
A snapshot of data quality issues in Indonesian community health

Maulidyati Aisyah

Faculty of Economics,
Islamic University of Indonesia,
Depok Sleman Yogyakarta 55283, Indonesia
Email: maulidyati@uii.ac.id

Sophie Cockcroft*

University of Queensland Business School,
Queensland 4072, Australia
Email: Sophie@business.uq.edu.au
*Corresponding author

Abstract: Healthcare services in Indonesia remain ‘poor’ by international standards. At the heart of the problem are systemic data quality issues. Little work has been done on health data quality in rural settings in this region. In this work, an exploratory study of data quality within a health centre (HC) in rural Indonesia is carried out with reference to two well-known sets of qualitative data quality measures AIMQ and PRISM. The research aims first to uncover data quality issues within a typical health facility in rural Indonesia, and second to discover whether these problems relate to operational issues. The research uses an inductive, qualitative case study approach using the following methodology. Key data quality issues are identified in the literature; these are used as a framework from which to develop seed questions for data collection via semi structured interviews. The full interview transcripts are analysed manually and using the text mining software Leximancer. Issues relating to data validation and integration are identified. Suggestions are put forward for development action both locally and nationally. This work provides a snapshot of the state of play in a typical rural health facility in Indonesia.

Keywords: health information systems; HIS; developing countries; PRISM; AIMQ; Leximancer; case study; data quality; Indonesia.

Reference to this paper should be made as follows: Aisyah, M. and Cockcroft, S. (2014) ‘A snapshot of data quality issues in Indonesian community health’, *Int. J. Networking and Virtual Organisations*, Vol. 14, No. 3, pp.280–297.

Biographical notes: Maulidyati Aisyah earned her Master degree in 2013 from the UQ Business School in the field study of information systems. Currently she is a Lecturer in Accounting Department, Faculty of Economics, Islamic University of Indonesia. She teaches management information systems, database management systems, and accounting management. She is a member of Indonesian Institute of Accountant and Information Systems cluster teaching groups. Her research interest is on accounting information systems, particularly its application in database management.

Sophie Cockcroft teaches systems analysis and design and business intelligence in the UQ Business School. She completed her PhD at the University of Otago, New Zealand, specialising in the application of conceptual data modelling in spatial information system (IS) design. Her current research interests relate to conceptual modelling in health IS, and in particular its application in larger health infrastructures. Related to this has been work relating to the effect of national culture on IS in general, and health IS in particular. She is a member of the Individual and Behavioural Information Systems, IT Artefact Research, Organisational Information Systems research groups.

This paper is a revised and expanded version of a paper entitled 'Data quality in community health: an Indonesian study' presented at ACIS 2013, Melbourne, December 2013.

1 Introduction

Despite much local development work, healthcare services in Indonesia have been assessed as 'poor' by international standards (OECD, 2010). Previous studies (Husada and Nguyen, 2012) have identified low health literacy, patient doctor ratio, incorrect diagnosis and long ordering times as motivation for providing better health information systems (HIS).

Health service delivery has become a major concern for the Indonesian Government. The Indonesian Ministry of Health (MoH) has established almost ten thousand Pusat Kesehatan Masyarakat/primary health centres (PUSKESMAS/PHC) in various districts all over the country (Husada and Nguyen, 2012). The main purpose of these PHCs is to help rural people to access health services. In 2002, the ministry enacted an order which instructed PHCs to adopt information systems called Sistem Informasi Kesehatan/Health Information Systems (SIK/HIS). Since then, various software applications have been developed and implemented, for instance, simpus, e-puskesmas, and simkes (Liebscher and Hui, 2007; E-Puskesmas, 2011). Unfortunately, evaluation by MoH stated that the overall status of HIS implementation is 'present but not adequate' based on WHO assessment tools. This situation led the Ministry to enact a decree that encompasses HIS Roadmap in 2012 which mandate all health agents to submit minimum standard report on schedule (The Ministry of Health, 2012). The current study examines a health facility that is complementary to the PHCs, and is privately funded. The funding body, Badan Amil Zakat Nasional (BAZNAS), based on Islamic Philanthropy, and endorsed by the Indonesian Government, runs five health centres (HCs) around Indonesia.

The focus of BAZNAS has been to provide programmes to meet life essentials such as health and education. The HC in this study, represents one of the fund distribution channels, and delivers free health services using a membership system. This approach is known as no-cashier healthcare. HC also provides 'outdoor', or roving health services. The focus of the initiative has been management of chronic non-communicable disease, and child and maternal health including immunisation and family planning. For BAZNAS, in delivering these programmes, information systems issues are secondary to those of health service delivery. Despite this focus, the HC is obliged to gather and manage data pertaining to mandated reports. The work presented here is a small

facility-based study concerning the data quality of a particular HC and the impacts this has on their day to day work.

The research questions examined here are:

- RQ1 ‘What is the current situation with respect to data quality in a typical rural healthcare setting in Indonesia?’
- RQ2 ‘What specific operational issues arise from data quality problems?’
- RQ3 ‘What are the possible remedies both locally and nationally from a data management and policy viewpoint?’

It should be noted that whilst the items described in Table 1 provide a snapshot, it is the richer qualitative data captured through the interview data that enabled us to address RQ2 and RQ3. This paper proceeds as follows: first the background to the project and data quality is given, then the methodology is described finally the results, discussion and conclusions are presented.

2 Background

2.1 The health facility

The health facility under investigation is located in the province of Yogyakarta in Java Indonesia. It is funded by BAZNAS and has an obligation to produce quarterly reports to BAZNAS head office on chronic non-communicable disease, child and maternal health including immunisation and family planning. Further periodic reports are sent to the Regional Health Department (RHD) and ad hoc recommendations are made to the MoH. The management team meet once a month to discuss work progress, and review and submit these reports. There is an additional monthly nurses meeting. The information system has evolved to assist in gathering this data and producing reports. The system incorporates membership and visit details – see Figure 1.

Figure 1 Data structure (see online version for colours)

1. PATIENT REGISTRATION	3. PATIENT VISIT REPORT
Medical record number	Visit date
Head of family's name	Patient's name
Patient's name	Medical record number
Age and sex	Diagnosis
Patient's relation to head of family	Sex (which filled with age)
Address	Child category (filled with age)
Clinic	Identification whether the patient is a member or not
Notes (if any)	Identification whether the patient is examined in ER
	Identification whether the patient is examined in Clinic
2. DISTRICT DATABASE	Nurse's Shift
Head of family's name	Notes (if any)
Member number	
Dependent	
Address	
Card Expired Date	

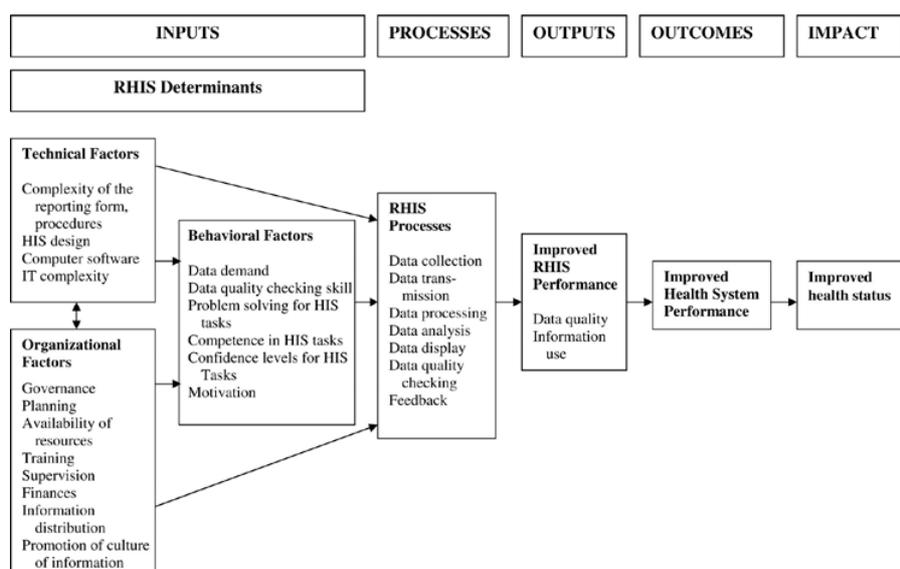
Note: Table titles as given in HC spread sheets.

2.2 Quality measures

The predominant emerging standard, for evaluating routine health information systems (RHIS) in developing countries, is the performance of routine information system management (PRISM) framework. Given the domain of the present study it would be remiss not to work in cognisance of this existing framework. PRISM tools, however, are designed to be used at the district or local government level, and many of the tools available within this framework do not apply to BAZNAS. The framework assumes a hierarchical reporting structure, and access by researchers at all levels (Figure 2). Previous studies have used PRISM framework (Aqil et al., 2009; Krishnan et al., 2010; Battulga et al., 2007; Hotchkiss et al., 2010; MEASURE Evaluation, 2010a). Aqil et al. (2009) summarised studies applying PRISM framework and its tools in Uganda, Pakistan, China, and Mexico.

Whilst many previous PRISM studies have a much larger regional or national scope, our work focuses at the facility level and the richness derives from more in depth interviews with the personnel on the ground. PRISM was used only to develop seed questions.

Figure 2 PRISM framework



Source: Aqil et al. (2009)

In the HC studied here, technical and organisational factors have received little attention. This shortcoming was not apparent at the outset of the study but has emerged as a point for recommendation similar to the case presented by Aung and Whittaker (2013). Thus, with respect to Figure 1 our work focuses predominantly on the process section and specifically data quality checking

As an idea of scale, within the PRISM framework is a recommendation that 12 to 19 facilities within a district (Aqil et al., 2009) are sampled. The framework does, however, provide data quality checklists, in the form of questions with largely yes/no responses, which can assess the data quality of a *single* health facility.

To expand the data quality framework we also use measures from the AIMQ framework (Lee et al., 2002). AIMQ is a validated, multidimensional information quality assessment framework that has found it application largely in mainstream management information systems studies.

Table 1 Measurement items drawn from PRISM and AIMQ

<i>Concept</i>		<i>Measurement item</i>	<i>Source</i>
Completeness	CP1	How many data items does the facility need to report on in the rural health information systems (RHIS) monthly report?	PRISM
	CP2	Is an official record of management meetings maintained?	
	CA1	This information includes all necessary values.	AIMQ
	CA2	This information is incomplete.	
	CA3	This information is complete.	
	CA4	This information is sufficiently complete for our needs.	
	CA5	This information covers the needs of our tasks.	
Relevancy	CA6	This information has sufficient breadth and depth for our task.	
	RP2	Does the facility's database provide comparisons among different service types?	PRISM
	RP3	During the last three months, did the facility receive any feedback report from higher authority on their performance.	
	RP4	Have any action-oriented decisions been made based on management reports?	
	RP5	Does the facility have routine meetings for reviewing managerial or administrative matters.	
	RP6	Does facility receive annual/monthly planned targets based on the report.	
	RA1	This information is useful to our work.	AIMQ
	RA2	This information is relevant to our work.	
	RA3	This information is appropriate for our work.	
	RA4	This information is applicable to our work.	
Timeliness	TP1	Observe whether report is submitted by the deadline.	PRISM
	TA1	This information is sufficiently current for our work.	AIMQ
	TA2	This information is not sufficiently timely.	
	TA3	This information is not sufficiently current for our work.	
	TA4	This information is sufficiently timely.	
Accuracy	TA5	This information is sufficiently up-to-date for our work.	
	AP1	Did you receive a directive in the last three months from the management to check the accuracy of data at least once in three months?	PRISM
	AA1	This information is correct.	AIMQ
	AA2	This information is incorrect.	
	AA3	This information is accurate.	
	AA4	This information is reliable.	

Note: AIMQ items generally require true/false responses.

This study focuses on the dimensions commonly referred to in data quality research (Wand and Wang, 1996; Pipino et al., 2002) these include completeness, accuracy, relevancy and timeliness. Aqil et al. (2009) assessed each data quality dimension within PRISM, as follows:

- a completeness – ensure that the all data elements are filled
- b accuracy – compare what is recorded and reported by the facility
- c relevance – compare between data collected and management needs
- d timeliness – observe whether report is submitted by the deadline.

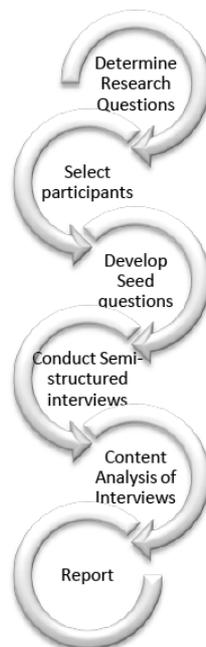
Both PRISM and the AIMQ methodology have derived a set of validated measurement items. The use of these items in our interview protocol adds to the body of knowledge from studies using these measures. The aggregated set is illustrated in Table 1. It will be noted that the PRISM questions are more procedural in nature.

3 Research methodology

3.1 Methods

Figure 3 illustrates the research process following the case study methodology of Yin (1989). Ethics approval for the study was obtained from the relevant ethics committee. A qualitative research design using in-depth semi-structured interviews was selected because it allows for a richer exploration of interview topics and the introduction where necessary of peripheral issues

Figure 3 The research process



The process included the systematic collection, organisation and interpretation of textual materials using content analysis.

3.2 Site, participants, and data collection

A survey in the form of a semi-structured interview was selected as the data collection method, in order to gain a deeper understanding, and provide a richer set of data for analysis. This approach follows Dul and Hak (2008). The interviews were performed in the native language of the participants. Nine interview participants were identified through the organisational gatekeeper. These individuals were individually approached in person while on the job, and all agreed to participate in recorded interviews of approximately 1 hour duration over the phone at their workplace, about their perception of data quality. The participants were the staff of the healthcare centre which consists of one manager, five nurses, two front officers (FO), and one surveyor (the surveyor is the person who runs the 'outdoor' operation and recruits members). None of the participants have an IS background and only two have had extensive experience of HIS as end-users.

3.3 Procedure

For the full validated questionnaires, the reader is referred to PRISM (MEASURE Evaluation, 2010b) and AIMQ (Lee et al., 2002) Table 1 shows the data items under review. In order to cover the dimensions addressed by the two aforementioned studies, but still attain the richness of a semi-structured interview a list of seed questions were developed to encourage participants to elaborate whilst at the same time addressing the measured items.

Each interview was recorded digitally and transcribed. These interview transcripts were independently translated by two native speakers and then checked for commonality. Each interview was placed in an individual file, and then the files were grouped in folders according to the four roles within the HC (surveyor, nurse, front office staff, manager, and surveyor). Within Leximancer a facility known as user tagging was employed. User tags were added to make it easier to see which remarks related to which seeding questions, and to which participants.

3.4 Analysis

The text was analysed using Leximancer (Smith, 2005). Leximancer is a research tool used for determining the presence of words or concepts in collections of textual documents (Smith, 2005). One of Leximancer's strengths is its ability to conduct both conceptual and relational analysis. Leximancer is programmed to develop a corpus of specific terms from within a text and to create a dictionary from these terms. Further, Leximancer has ability to make inferences from the concepts within the text establishing relationships between them. Concepts which tend to appear together in the text are extracted. Terms extracted in Leximancer are weighted based on the frequency of occurrence in sentences. Once the weight reaches a threshold, the terms are defined as a concept. Besides providing quantitative means for analysing text, Leximancer also allows the information drawn to be presented in a concept map. It is a map where co-occurrence

concepts will be put near to each other and clustered into higher-level ‘themes’. As part of the data preparation certain modifications were made as per Table 2.

Table 2 Modifications made to Leximancer default settings to automatically analyse interviews with HC employees

<i>Modification</i>	<i>Concepts</i>
Automatically identified concepts manually removed	Example, further, provide
Concepts merged	Patient, patients nurse, nurses
User defined tags added	Complete/completeness, relevancy, timeliness, and accuracy, medical record*

Notes: *In the initial analysis, it was found that Leximancer was not distinguishing adequately between the noun and the verb use of the word ‘record’. Thus, all occurrences of medical record were concatenated in the original file as medical_record.

4 Results

The results of analysing interview data are presented first via a Leximancer analysis of the raw interview data and second using the qualitative assessment measures (AIMQ and PRISM).

It emerged from discussions with key personnel that there is an implicit data structure which has evolved organically and with little input from information systems literate staff. This structure is illustrated in Figure 1.

The focus of the HC and the wider mandate of BAZNAS has been to provide healthcare to less advantaged people in rural communities in Indonesia. This service includes an ‘outdoor’ or on the road service. The focus on IT has only stretched as far as getting the appointment records into digital form and maintaining consistency between paper and electronic records. We see this in remarks such as:

“...what written manually is exactly the same as what we input into computer. Our first process is we search patient Medical record on the cabinet then we input name and address manually. If we compare to hospital, once we input Medical record number we can retrieve all patient’s data. Unfortunately we still input the data manually because there is no good system running so far.”
FrontOfficer1

Whilst it is beyond the scope of this work to fully explore IS maturity in the organisation, this approach would be defined, in terms of IS maturity, as automation (Zumpe and Ihme, 2006).

The record that HC stores, has evolved in an ad hoc fashion around three spreadsheets. About 1,000 visits are recorded per month. The structure of this record, illustrated in Figure 1, which has served the HC for the first two of years of operation, is starting to cause problems. This is revealed in the interview data. Specifically, the spreadsheets are used to store age instead of date of birth, so this has to be recalculated on every visit, and is inaccurate for reports. The membership card is fraught with difficulty

since there are many family members on one card, the card does not have a photograph, and procedurally there is no way of updating a member's marital or change in name, or the fact that they have become adult. Medical record number has to be reconciled across two spreadsheets; the gender of children is not stored.

4.1 Leximancer results

Figure 4 shows the breakdown of themes showing speaker tags only. File and folder tags identifying roles are omitted because they cause unnecessary crowding. There is a clear division of concerns between the front office staff, surveyor and managers, tagged in the top of the diagram and the nursing staff tagged to the bottom. With the front office staff being more concerned with card and membership data and the nursing staff being more concerned with patient demographic data required for reporting. Themes are heat-mapped hot colours denote more important themes (in this case health, data, and report) and green or blue shades the less important ones – in Figure 2 the terms meeting, computer, and case take less prominence.

Figure 4 Leximancer output (see online version for colours)

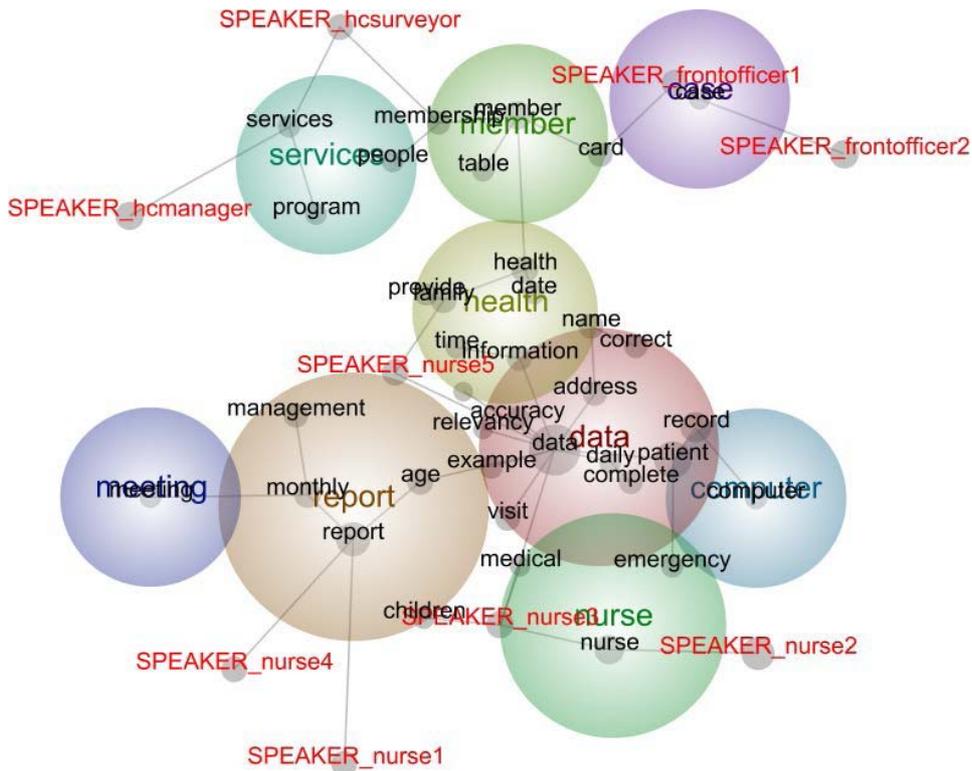
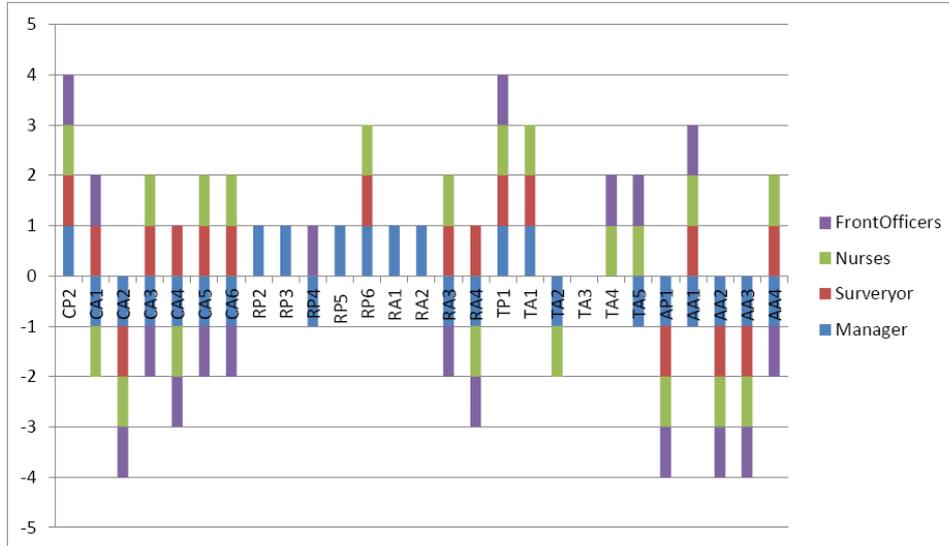


Table 3 gives a breakdown of the type of issues discovered in relation to each to the Leximancer themes.

Table 3 Summary of Leximancer themes and major concerns emerging from interview data

	Related concepts	Most frequent role (Figure 2)	Major concern	Sample quotes
Data	Data, patient record	All	Various, from membership validity at the higher levels (manager and front officer), to problems of health record numbers in relation to service provision at the nursing level.	"Patients do not really care about the accuracy of their data written in the card because what important for them is to get the card and be able to get free health service." Front officer 1
Report	Report, monthly, age, visit management	Nurse 5	Making sure data is accurate to produce necessary reports particularly with respect to age, needing to make monthly reports to monitor specific diseases such as diabetes and blood pressure.	"After couple of months opening we improved our data by differentiating children in to separate group. The cut of is at age 16. Patient under 16 years old categorized as children." Nurse 3
Health	Health, family, name, information	Nurse 5	Ensuring data allows smooth healthcare delivery.	"So we check to the pharmacy whether drugs disbursement for family planning is conform to our record or not."
Member	Member, card	Front officer 1, Surveyor	Managing membership data problems of incorrect or missing member numbers, accuracy of data on card.	"For membership, the accuracy is still lacking because there are many factors were not recorded which by that we can provide better tools in assessing the eligibility for a person to be classified as poor people. There are several items that should be included in the table beside occupation and income, and this is include expired date of the card." HC Manager
Nurse	Nurse	Nurse 2, Nurse 3	Anomalies that occur between the nursing station and front desk. For example, a person asks to the nurse to see their child as well.	"We realised this at the examination, the nurse came to Front Office and told that the name is same but the member card number is different." Nurse 3
Services	Services	HC Manager, Surveyor	Making sure the right people get access to services and that the correct information is stored about them.	"The management expect that our member also possess other health allowance from government such as community health allowance, regional health allowance. Since we do not have this information, we found difficulties in assessing whether particular member already covered by this type of health coverage or not."
Computer	Computer	Nurses 2,3 5 and Front office 1	How electronic and paper data are kept current, back up procedures. In particular, in relation to a patient visit booking, and family planning services.	"Daily we check computer's data and book's data, so we do checking in every input to computer. So we never found dissimilarity." Nurse 5
Meeting	Meeting	All	Discussions of what goes on in monthly evaluation meeting. Meeting performance targets.	"Yes, before monthly meeting we have internal meeting to discuss about the challenges." Nurse 3
Case	Case	Front officer 1 and 2	Discussing cases where accuracy had been compromised. Suspicions were raised in the 'case' of someone being married but still operating on a parent's card, or someone arriving at the clinic in a car.	"If [use of wrong number] was found when the patient had been diagnosed." Front Officer 1

Figure 5 Summary of responses to survey questions in Table 1 (see online version for colours)



4.2 Quality measure results

The answers given to the yes/no questions in Table 1 are illustrated in Figure 3. Where a bar appears above the x axis the survey group has said ‘yes’ to the question, when it appears below they have said No and when it is missing there is missing data. Three of the survey items CA2, TA2, and AA2 relating to completeness, timeliness and accuracy, were stated negatively thus a negative bar represents a positive result. It can be seen that participants were unanimous in six of the responses TP1 – the fact that reports are submitted on time, CP2 –an official record of meetings is kept, CA2 – that information is *not* incomplete, AP1 – that they have not received any directives with respect to data accuracy, AA2 – that information is *not* incorrect and AA3 that the information is not accurate in some regard. In all the other dimensions, there was some difference on opinion and in many cases only one group responded positively. The item TA3 concerning timeliness, did not receive valid answers. The responses to RA4 are given below for illustrative purposes, in this case only the surveyor found the data reliable.

4.3 Reliability (RA4)

it is still not as expected. Since one of our obstacle is that the system is offline, manually input the data, counting changes on age will be a challenge. I believe that not all the nurse pay attention on the date of birth and age changes.
 Manager

Surveyor data is reliable in supporting health care program. It uses membership data I managed, for instance program for elderly people, program on certain region and other certain program. Surveyor

Yes, but in the examination we will check medical history through health record to deliver correct drugs to the patient. Nurse

There is no patient who informs us voluntarily if for example his child has married or worked which affect their membership status. I will be suspicious when a patient within marriage age but he still include in his parents card. Then I will confirm about his marital status and often the patient admit that he is married. FO

We now look at responses to specific questions from the AIMQ and PRISM frameworks shown in Table 1.

4.4 Completeness – issues with reporting

The manager is required to submit quarterly reports to BAZNAS on activity (indoor and outdoor) and financial aspects concerning monthly membership. The indoor report includes emergency treatments, dental visits, lab and pharmacy work. The outdoor report is concerned with such things as school health programmes, and nutritional advice in the community. Most nurses submit monthly visit reports and reports on the top ten diseases to BAZNAS management. One nurse has different reporting duties she is required to prepare a family planning and nutrition report monthly to the national body of population and family planning and them, in partnership with PHC a summary report is delivered to district management and the RHD.

Only two of the PRISM questions were relevant in this study (CP1 and CP2). In terms of the PRISM measurements, the manager mentioned that there are two reports submitted by BAZNAS HC in Yogyakarta to its headquarters in Jakarta. The manager keeps an official record for monthly evaluation meetings between management and staff key personnel.

When asked about the AIMQ indicators (CA1-6), the HC manager noted that she is required to record possession or otherwise by patients of a particular government allowance. This allowance is administered by the MoH and covers hospitalisation services.

Unfortunately, the current information system does not allow for the recording of this item. It was the manager's opinion that knowing this information might enable the HC to give recommendations to MoH about members who have not received the allowance.

The manager also noted that she occasionally found missing 'age' data from the table provided by nurses. This statement highlights an inconsistency because all nurses mentioned that they have filled all columns completely.

The surveyor's main reporting responsibility concerns membership, its monthly growth and total, in detail. Under the CA4 AIMQ measurement, he was the only participant who found the data sufficient.

Other contradictory answers were also found in nurses statements. One nurse stated that classifying children by broad age bands such as 'baby', 'toddler' was sufficient for their task. Another nurse also argued that her report is sufficiently complete for reporting requirements to the RHD. However, the consensus among the remaining nurses was that further age categorisation is necessary for children. Additionally, they highlighted the need to add personal information such as address and enable the system to identify new patient within the current month. This contradiction indicates differences in reporting requirements for the role of each nurse. Reports required by RHD need more data items which are not covered by the report prepared for BAZNAS. Thus, nurses with RHD reporting requirements have more stringent data needs. All nurses stated that no official

minutes are maintained for the monthly meeting with management, but they have one for internal nurses meeting.

The fourth key set of personnel is the FO. Their report covers patient registration on a daily basis which also includes records for non-member patients. Like the nurses, they are not required to manage the official record for monthly meeting. Although, the FOs argued that even if all the columns in the report are completely filled, the current system is insufficient to cover their needs. This can be viewed on their statements where give the opinion that it would be better to record date of birth instead of age, and remark that current systems could not accommodate any issues arising with respect to membership eligibility.

4.5 Accuracy

The only PRISM question that was applicable was API relating to accuracy or directing the staff to ensure details on member's data are accurate. Answers to the AIMQ questionnaire show that from the manager's perspective, HC's data is not accurate. However, in term of report, the manager expects that the FOs and nurses will ensure the number of patients on registration and visit reports are exactly the same. Data on age impacts accuracy since current systems cannot reflect any changes in 'age'. Another data accuracy issue is 'diagnoses' data input by nurses. Sometimes the manager found typos on the data due to nurses' misinterpretation on doctor's handwriting and none of the nurses realised that these typos had occurred.

The surveyor argued that membership data is reliable as this data has been used for supporting the other HC programmes. Nevertheless, he also mentioned that there are often differences in members' details between those held by the surveyor, nurses, and FO.

Accuracy of patient's age data was also of concern to the surveyor. He would prefer to record patient's date of birth rather than age because no standard had been setup for manual calculation of age.

The nurses' responses to the questionnaire varied. For certain data such as age, nurses, agreed that even though were calculated manually, patient's age data is accurate. The nurses also stated that input for diagnoses are accurate. However, as seen these are not in line with manager arguments above. Additionally, the nurses responded differently over issues relating to patient name. Some of them argued that the name is accurate but the other mentioned there are typos in the data. The nurses also commented on inaccuracy in marital status recoding.

FO were also concerned about typos when assessing accuracy on data especially member's name and number. One FO experienced problems due to the accuracy of membership numbers. This fault made the nurse use a wrong health record which eventually resulted in incorrect medication being given to the patient. There was no emergency situation arising from this on this occasion, however FO mentioned that this type of error had occurred two or three times whilst she had been working at the HC. Cases such as these are quite a significant issue regarding data accuracy. Additionally, the fact that number of patients registered daily did not always coincide with the visit reports, and their marital status was sometimes incorrect (i.e., some patients has passed themselves off as unmarried to remain on their parents' health card) had also become a concern for the FO.

4.6 *Relevancy*

The PRISM questions on relevancy RP2–RP6 are mainly applicable to the manager because they are situated at the organisational level. These include comparison among services provided and action-oriented decision made based on the reports. Therefore, interviews with other participants focussed mainly on the AIMQ questionnaire.

The manager explained that comparisons made were on each service against its target, rather than between the services. She added that informally she always make comparison on HC services against PHC in order to maintain their services. HC sets targets on performance including registration levels for itself upon approval from BAZNAS. One action-based decision by the HC was provision of dental services. BAZNAS reviewed the first three months report to grant an approval to provide the service permanently. From AIMQ perspective, the manager argued that current data was useful and relevant. However, the large amount of incomplete data, especially demographic data has limited its usefulness and relevancy.

The surveyor, who manages membership data, stated that this data is more applicable and useful for indoor (in house) services which have recruited a lot of members. Meanwhile for outdoor (mobile) services, member recruitment only started few months ago. Thus, the data is useful for the surveyor to the extent that he can triage the process of membership eligibility using this data. Another challenge for the application of data is that different regions have different needs.

The Nursing staff in general made favourable remarks about the relevancy of the data. As they report on every treatment given to the patients, this kind of data is clearly fulfilling all data relevancy items (useful, relevant, appropriate, and applicable) for deciding what treatment the patients should receive next. The relevancy of data was also in relation to dispensing drugs. The nurses agreed that their drug records should match those of the pharmacist. On the other hand, the FO did not find the current data relevant. She felt unsure on the usefulness of the data. This FO argued for a completely new system to be installed rather than to make an improvement to the current system. She explained her preference by saying that the computer system was not adding any benefit beyond putting handwritten records in digital form.

4.7 *Timeliness*

In measuring the timeliness dimension, PRISM only requires users to examine the submission date of the reports, and whether the deadline has been met or not, whereas, the AIMQ questionnaire focuses on the currency of the data.

The manager always submits the reports by the deadline since she is concerned with the disbursement of the next budget. In terms of content of data, she argued that the data is sufficient on a daily basis. Again, one issue she put forward concerned the manual calculation of ‘age’ data. Specifically, if this data item is not calculated one by one, it will remain the same and not be updated.

The surveyor was confident that his membership data are up to date, especially the field ‘member number’ which is used as a primary key for the tables. He also said that he usually submits the reports a few days before the deadline.

In questions relating to meeting the deadlines nurses’ responses depend highly on whether they need to report to external parties or not. The nurse who is not responsible to external report tends to delay the report whereas the one who needs to submit reports to

external parties cannot. Administering the AIMQ parts of the questionnaire also gave varying results. Once again the issue of 'age' was the most pressing. The chance to update 'age' data only occurs when a patient comes to HC. Thus, it is important for the nurses to finish the age calculation rather than postpone it. Nurses' answers also implied that they define currency of data as the fact that they perform input data daily.

One FO mentioned that she is never late in submitting the report, while another said that even sometimes she did not meet the deadline, it will not become a problem as long as the report is handed-in before the monthly meeting. However, they both agree that their data is quite up to date with respect to the current patient condition.

5 Discussion and conclusions

It is argued that all of the results are consistent among these countries where low data accuracy was explained by higher perceived confidence level in checking data accuracy, yet much lower knowledge of the methods available to check the data (Aqil et al., 2009).

5.1 Compliance with AIMQ and PRISM measurements

With respect to the first dimension, completeness, the HC has complied under PRISM measurements in terms of providing the required reports, but not in managing official meeting records. Meanwhile with respect to the AIMQ questionnaire, data in HC still has a long way to go in order to comply with completeness measures due to insufficient and missing values.

The next dimension is accuracy which clearly does not comply with PRISM since a management directive on data accuracy is missing. In the meantime, AIMQ questionnaire put HC in a low compliance as there are still many fields of data have been questioned for accuracy, with respect to effort from HC's staff to maintain data accuracy.

The third dimension is relevancy. By both PRISM and AIMQ measures it is concluded that the relevancy of data in HC is still partial. With respect to PRISM, HC has made use of data to take a strategic decision in providing a new service. However, analysis made on current data is limited due to data incompleteness. With respect to AIMQ, only the nurses identified the data as relevant. Further, it is also identified that only indoor services data is relevant because it has been managed since the HC was established.

The last dimension is timeliness. According to the PRISM measures, most of HC staff, including the manager, meet timeliness measurements in the sense that they manage to submit their reports within defined deadlines. Only few do not. In term of currency of the data as measured by AIMQ, the data is quite up-to-date, but age is obviously not because it is calculated manually.

In general, with reference to PRISM measurement and AIMQ questionnaire, it is concluded that data quality in HC is still far from complying with items in the data quality dimension. This might also be explained by the IS maturity level for HC. The utilisation of IS within HC exists to the extent that manual records are transferred into digital form and stored in a computer. Staff still needs to maintain the consistency between two records. As stated by Zumpe and Ihme (2006), the stage where human component are excluded from a process to some extent is called automation. This is the first level of e-business IS maturity model. This status goes some way to explaining the

HC's low data quality compliance to the measurements. Understanding the current stage of maturity helps to identify strategic action to be made in order to pursue the next maturity level (Liu et al., 2011). Development of a maturity model specific to primary healthcare in developing countries is a possible avenue for further research.

These results are consistent with a previous PRISM study (Aqil et al., 2009) in the sense that there is an element of overconfidence in the data, and a lack of knowledge on how to detect problems.

Specific problems within the HC were identified. Several staff found difficulties in reporting due to the many manual approaches applied to softcopy data such as in calculating and categorising age. Concerns about fraudulent membership were also revealed by the Leximancer analysis. Since many staff found irregularities in patient data, this content appeared in a number of themes in the Leximancer concept map. Repeated statements on data inaccuracy that effect on recording and reporting processes resulted in a number of concepts emerging in the 'data' theme

Moreover, the Leximancer analysis identified isolated error events specifically, when FO used the wrong medical record for a patient. This event is a significant one in a healthcare setting. Identifying such case might guide the management to put more attention to data management

Addressing RQ1, the current situation is a low level of IS maturity resulting in poor data quality and a lack of process for checking accuracy. Specific operational issues that have arisen with respect to RQ2 include non-availability of data for reports, difficulties with uniquely identifying patients and problems with the data structure resulting in update problems. RQ3 is addressed in the following section.

5.2 Recommendations

By considering the results on current study, several recommendations RQ3 are suggested to the BAZNAS HC as follows:

- In pursuing better data quality, HC should acquire a better IS. This can be done through software development. However, this development should be accompanied emphasis on including crucial features such as data checking and data validation to avoid the use of wrong medical records and fraud.
- Alternatively, the HC could buy a software package of the many readily available in the market. This option would be less resource intensive than a custom solution.
- As part of any new development HC could consider integration with data in the government-based PHC schemes.
- Establishing formal standard operating procedures (SOP) including data flow will help HC avoid many of the operational issues.
- Human resource training is necessary for staff, including management to improve IS literacy which may reduce errors.

Ultimately, this study presents a snapshot of the current state of data quality in small to medium sized health facilities, which serves as a benchmark for future development. A further contribution of this study is to the practice of using Leximancer in content analysis in a healthcare setting.

5.3 Limitations

During the research process, we documented some limitations:

- Interviews were conducted by phone. Thus, it is possible that messages were misinterpreted during the session due to the lack of physical cues.
- It is possible that some of the team interviewed were either defensive of their system, or thought interviewer was looking for a particular answer which could potentially lead to bias. This feature also lends itself to further research in terms of learning within an organisation (Aggelidis and Chatzoglou, 2012)
- Interviews were transcribed and analysed in different language. This might reduce the nuance of the words chosen.
- The research was limited to one type of healthcare provider in Indonesia.
- Limited exploration on IS maturity model in its relation to data quality.

5.4 Future work

- development of a maturity model specific to primary healthcare in developing countries
- user acceptance of health systems.

References

- Aggelidis, V.P. and Chatzoglou, P.D. (2012) 'Hospital information systems: measuring end user computing satisfaction (EUCS)', *J. Biomed. Inform.*, Vol. 45, No. 3, pp.566–579.
- Aqil, A., Lippeveld, T. and Hozumi, D. (2009) 'PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems', *Health Policy and Planning*, Vol. 24, No. 3, pp.217–228, %R 10.1093/heapol/czp010.
- Aung, E. and Whittaker, M. (2013) 'Preparing routine health information systems for immediate health responses to disasters', *Health Policy and Planning*, Vol. 28, No. 5, pp.495–507.
- Battulga, S., Ajnai, L., Purevbat, N., Tsevelmaa, B. and Ochirbat, B. (2007) *Establishment of Patients' Information Database in Mongolia*, Springer, Berlin, Heidelberg.
- Dul, J. and Hak, T. (2008) *Case Study Methodology in Business Research*, Elsevier Butterworth-Heinemann, Amsterdam.
- E-Puskesmas (2011) *e-Puskesmas Dalam Inaicta 2011*, e-Puskesmas [online] http://www.e-puskesmas.com/index.php?option=com_content&view=article&id=108:e-puskesmas-dalam-inaicta-2011&catid=40:berita&Itemid=53 (accessed 10 May 2013).
- Hotchkiss, D.R., Aqil, A., Lippeveld, T. and Mukooyo, E. (2010) 'Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda', *BMC Health Services Research*, July, Vol. 10, No. 188.
- Husada, R. and Nguyen, L. (2012) 'ICT alternative for primary care delivery in Indonesia: a proposal', *Proceedings of the 23rd Australasian Conference on Information Systems 2012, ACIS*, Geelong, Vic., pp.1–7.
- Krishnan, A. et al. (2010) 'Evaluation of computerized health management information system for primary health care in rural India', *BMC Health Services Research*, November, Vol. 10, No. 310.

- Lee, Y.W., Strong, D.M., Beverley, K.K. and Richard, Y.W. (2002) 'AIMQ: a methodology for information quality assessment', *Information and Management*, Vol. 40, No. 2, pp.133–146.
- Liebscher, C. and Hui, K. (2007) *SIMPUS Implementation Midterm Assessment: Findings Report*, IG Health, Aceh, Indonesia.
- Liu, C.F., Hwang, H.G. and Chang, H.C. (2011) 'E-healthcare maturity in Taiwan', *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, Vol. 17, No. 7, pp.569–573.
- MEASURE Evaluation (2010a) *Fourth International RHINO Workshop: Measuring and Improving RHIS Performance*, USAID, 8–1 March.
- MEASURE Evaluation (2010b) *PRISM Tools for Assessing, Monitoring, and Evaluating RHIS Performance*, USAID.
- OECD (2010) *Economic Surveys Indonesia*, OECD Publishing [online]
<http://dx.doi.org/10.1787/9789264000000-en> (accessed 16 April 2013).
- Pipino, L., Lee, Y.W. and Wang, R.Y. (2002) 'Data quality assessment', *Communications of the ACM*, Vol. 45, No. 4, pp.211–218.
- Smith, A.E. (2005) *Leximancer Manual*, Unpublished Manuscript, The University of Queensland, Australia.
- The Ministry of Health (2012) *Roadmap Sistem Informasi Kesehatan 2011–2014*, Jakarta.
- Wand, Y. and Wang, R.Y. (1996) 'Anchoring data quality dimensions in ontological foundations', *Communications of the ACM*, Vol. 39, No. 11, pp.86–95.
- Yin, R.K. (1989) *Case Study Research: Design and Methods*, Sage, Newbury Park, California.
- Zumpe, S. and Ihme, D. (2006) 'Information systems maturity in e-business organizations', *European Conference on Information Systems*, Vol. 14, IT University of Gotteborg.