Study to Assess the Data Quality of Health Management Information System of a District from Central Maharashtra.

Rahul Digambarrao Anerao[1], Gangadhar G. Parge[2]

[1]District Training Centre, Latur, Maharashtra, India.
[2]District Health Officer, Latur, Maharashtra, India.

ABSTRACT

Background: Good governance, transparency and accountability have become the mantra of development, and consequently more attention is given to strengthening evidence-based decision-making and information systems. The launching of a national portal-based Health Management Information Systems (HMIS) by Government of India in 2008 was a bold and innovative step in this regard. Poor data quality can take many shapes; not just incorrect figures, but a lack of completeness or the data being too old (for meaningful use). **Methods:** This study design is a case study method. In this study realistic evaluation approach and employs case study design was used. The case here is of HMIS for Reproductive and Child Health programme within a district in Central Maharashtra. Multistage sampling was used to identify facilities covering all levels of health facilities. The case is a single case without a comparator. The district is being considered as a single unit for the case study. **Results:** Data quality mainly comprises completeness, correctness and timeliness. Data quality is multidimensional, and involves data management, modeling and analysis, quality control and assurance, storage and presentation. Inadequate or deficient validation check will lead to loss or compromised data accuracy and quality. Hence the correctness of data is lost. It will lead to ill informed decisions. Hence extra efforts, investments and recurrent costs are definitely required to ensure data quality in the district and facilities. **Conclusion:** Every action in the health system is translated into data. Thus, there is a lot of volume of data that is generated everyday. The current systems neither tackle volume nor increase the value of the information. There is a need to move from volume to value.

Keywords: Data Quality, Health Management Information Systems.

INTRODUCTION

Is the data complete? Is it collected on time? Is it correct? These are questions that need to be asked when analyzing data. Poor data quality can take many shapes; not just incorrect figures, but a lack of completeness or the data being too old (for meaningful use).

Data quality is a broad term most frequently used in any information system. Data are of high quality if they are fit for their intended uses in operations, decision-making, and planning. Data are fit for use if they are free of defects and possess desired features (Redman 2001). Data quality is a term that refers to the reliability and validity of data collected in the HMIS. It is measured by the extent to which the data in the system reflects actual information in the real world. The quality of data is determined by assessing certain characteristics such as timeliness, completeness, and accuracy.

Good governance, transparency and accountability have become the mantra of development, and consequently more attention is given to strengthening evidence-based decision-making and information systems. This trend is reinforced in the health sector by emerging infectious diseases and environmental disasters, which need timely information for appropriate action. (Anwer Aqil et al. 2009).

In case of health system, right information, at right time and right form can be associated with many crucial and important decisions made to benefit health of individuals and community.

According to Hurtubise (1984) the term Information System is described as "a system that provides specific information support to the decision-making process at each level of an organization". So the ultimate goal of information system is not only to collect information but to improve action planning. A Health Management Information System (HMIS) is defined by the World Health Organization as an information system specially designed to assist in the

Name & Address of Corresponding Author
Dr. Rahul Digambarrao Anerao
District Training Centre,
Latur,
Maharashtra,
India.

management and planning of health programs, as opposed to delivery of care. It can be summarized as integrated effort of collecting, processing, reporting and using health information and knowledge obtained through it to influence policy-making, program action and research". Health information system deals with the morbidity and mortality patterns of populations, causative analysis and the scope and effectiveness of public health interventions.

One of the major barriers in the implementing effective health care delivery particularly in developing countries is the unavailability of adequate information for efficient decision-making. Effect of poor data quality on decision-making can have big social and economic consequences. It can affect customer/user satisfaction level, operational costs, and effective planning and decision-making. (Redman T 1998, Wang RW et al1996). It is necessary to re-structure the health information system for better planning and effective monitoring of health services (Campbell, 1997; Lippeveld et al. 2000). Health sector reforms have also increased the need for standardized information with high quality data.

Health information systems in developing countries have been shaped by political, administrative, economic and donor pressures. Improvement of the quality and accuracy of data coming from developing countries have been promoted since the 1990s by augmenting the routine health information system with the help of information technology.

Health information system in India includes system reporting demographic information, system reporting information regarding health services and health research. Demographic information system has been existing since many years and includes systems like census, vital registration etc. Health surveys include national family and health survey, district level household survey etc. These surveys provide policy makes with essential information required for effective planning.

To strengthen use of data for local analysis, display of information, feedback forms and use, NHSRC has introduced a district level tool – the DHIS-2 applications: as of now it is functional in Maharashtra state also. This application also allows customization to include new indicators and integration with IDSP, diseases control programs and functions such as HR Management and Hospital Management. DHIS 2 helps governments and health organizations to manage their operations more effectively, monitor processes and improve communication. DHIS 2 is typically used for data management and analysis purposes, for health program monitoring and evaluation, as facility registries and service availability mapping, for logistics management and for mobile tracking of pregnant mothers in rural communities. It has been mandatory for all districts of Maharashtra to use DHIS-2 for providing data from March 2014. The Mother and Child Tracking System (MCTS) is also in place which provide reliable data for effective decision-making through name-based tracking of each client.

Two terms are commonly used to describe the data quality measurements. These are Quality Assurance and Quality Control. Redman (2001) defines Quality Assurance as 'those activities that are designed to produce defect-free information products to meet the most important needs of the most important customers, at the lowest possible cost'. However, Taulbee (1996) makes the distinction between Quality Control and Quality Assurance. Quality Control as a judgment of quality based on internal standards, processes and procedures established to control and monitor quality; and Quality Assurance as a judgment of quality based on standards external to the process and is the reviewing of the activities and quality control processes to insure that the final products meet predetermined standards of quality. However both the terms are used synonymously in practice of data quality management.

Thus this study is planned to determine the extent of the data quality problems and explore the factors that contribute to poor data quality in selected health facilities in Beed district of Maharashtra.

Review of literature

Debates abound at different forums regarding which data source is preferable for developing and tracking health system performance, documenting best practices or effectiveness of interventions, and identifying gaps in implementation. It is argued that household and facility surveys yield better quality information than self-reported health management information systems (HMIS) because of more objectivity and less bias. Some perceive HMIS to be costly, producing low quality and mostly irrelevant information [Mutemwa 2006] thereby contributing less to decision-making processes. But it should be noted that each method of data collection serves a different purpose and has its own strengths and weaknesses. Further, there is no evidence that a third party survey assures better accountability or improvement in health system performance [T. Sundararaman 2011]. Performance remains an organizational issue and needs to be dealt with as such. The HMIS allows organizational members to track their progress routinely in meeting organizational objectives, including patient management objectives, for which data cannot be collected otherwise (Lippeveld et al. 2000). Health system managers have no substitute for routine information in terms of monitoring progress towards achieving service coverage objectives and managing associated support services (e.g. logistics, human resources, finance) for their local target populations. Thus, the focus of debate should shift from abandoning HMIS over other sources of data to showing how to improve HMIS.

HMIS evaluation studies conducted in different countries showed increasing evidence that, routine information systems were not producing the intended results. Studies showed that data quality was poor in Mozambique and Kenya (Mavimbe et al. 2005; Odhiambo- Otieno 2005a), while use of information for planning and decision-making was found to be weak in Brazil and South Korea (Chae et al. 1994; da Silva and Laprega 2005). Many factors contributed to under-performing information systems, such as difficulty in calculating indicators because of poor choices for denominators in DR Congo (Mapatano and Piripiri 2005) and inadequacies in computerization, data flow, human and capital resources, and low management support in Kenya (Odhiambo-Otieno 2005a). Nsubuga et al. (2002) in Tanzania found weaknesses in the areas of standardized case definitions, quality of reporting, analysis, supervision and feedback. Rotich et al. (2003) and Kamadjeu et al. (2005) noted that user involvement, the choice of a standardized terminology, a pre-existing culture of data collection and leadership remain crucial issues for HMIS financial and technical sustainability. Another problem in strengthening information systems was the scarcity of structured evaluations for best practices in information systems (Mitchell and Sullivan 2001; Tomasi et al. 2004). Odhiambo-Otieno (2005) suggests that lack of evaluation of district-level HMIS has been partly due to the lack of defined criteria for evaluating information systems.

Aim

- To assess the data quality of the existing Health Management Information System (HMIS) of National Rural Health Mission (NRHM) at district level.

Objective

1. To assess the quality of RCH-NRHM data in relation to completeness.
2. To check the timeliness of RCH-NRHM data.
3. To assess the accuracy of the RCH-NRHM data
4. To study, how this NRHM-HMIS processes affecting data quality?

MATERIALS AND METHODS

The HMIS has several subsystems. In districts of Maharashtra it includes vital registration, District Level Household Surveys, Information systems of national health programs (such as leprosy, TB, HIV control programs), information systems of general administration (such as human resource, salaries of staff). In this study, focus was be on maternal health care and child health care componenets within the Reproductive and Child Health Program.

There are many things which create hurdles in the smooth functioning of HMIS. Some of these are non user friendliness of grass root level workers about computer and internet, infrequency of internet

facilities, lack of supportive supervision, irregularities in timely conduction of data validation committee meetings etc. This study was planned to assess the data quality of the existing Health Management Information system of Beed district in Maharashtra. This study focuses on RCH componenets of HMIS system i.e. DHIS2 and MCTS.

Study design

This study design is a case study method. The case here is NRHM-HMIS which includes maternal and child health care under HMIS and Mother and Child Tracking System (MCTS). Both quantitative and qualitative methods were used as part of the case study. The case is a single case without a comparator. The district is being considered as a single unit for the case study. The study setting of the case is described in subsequent paragraph.

Study setting

It has been decided to study health information system of Beed district in Maharashtra purposely. Maharashtra is divided into 36 districts, which are grouped into six divisions. One of these divisions, namely Marathwada has 8 districts, one of which is Beed. Beed district is among ten poor performing districts as per HMIS report of Maharashtra. For these reasons, Beed district has been selected as a case.

Beed district is divided into eleven blocks and has a total of 362 health institutes including one District Hospital (DH), one Women Hospital (WH), 11 Rural Hospitals (RH), 2 Sub-District Hospitals (SDH), 50 Primary Health Centres (PHC) and 280 Sub Centers (SC).

Permission to undertake the study was obtained from the Director of Health Services, Maharashtra and the District Health Officer of the selected district (Beed) before commencing the study.

Ethical clearance

Proposed study was placed before the Ethical committee of Government Medical College, Ambajogai. After receiving written approval letter from the authority, this study was initiated.

The information sheet and the consent form was given to study participants before starting of data collection. This was to ensure the participants to understand the purpose of the study and the purpose of their participation in the study. Under no circumstances the identity of study participants, the

facility where they work has been revealed during and after the study.

Sample Size

There are 362 health facilities in Beed district. Sample size based on the HMIS indicators for 80% data quality and 60% use of information is achieved or not, a sample size of 19 facility was needed (prism theory page no 88-90). It is based on the 95% confidence, and a margin of error of 10%. It not only provide an average statistic of the variable, but it also provides information about whether lots (districts, supervisory areas, divisions of organization) are below or above the average. Based on this, 19 facilities including Women Hospital, Rural hospital, Primary Health Centers and Sub centers were included in the study.

Sampling procedure

Multistage sampling was used to identify facilities covering all levels of public health facilities. Three blocks were selected with good, average and poor performance with respect to HMIS. In each block, there is a rural hospital and averagely 4-5 PHCs. Under each PHC there are several sub-centers. For attaining total sample size of 19 facilities, a women hospital; three Taluka Health Office/ Rural Hospital; five Primary Health Centers and ten sub centers were selected.

Stage 1: selection of randomly three blocks from district.

Stage 2: selection of one rural hospital from selected block and selection of five PHC randomly from same selected blocks.

Stage 3: Selection of two sub centers from each selected PHC.

Stage 4: Selection of healthcare workers involved in HMIS.

Selection of Participants

The district level officers who were part of the study include District Health Officer, District Reproductive and Child Health Officer (DRCHO), Statistical Officer, Monitoring and Evaluation Officer, Technical Assistant, and statistical assistants.

At the Block level the Taluka Health Officer of the block (THO), Health Supervisor (HS), Block Nursing Officer (BNO) and Data Entry Operator (DEO) were interviewed with appropriate tools.

At the PHC level the appropriate tools were administered to the Medical Officer-in-Charge, Supervisor Female (LHV) and Health Assistant (HA).

At the sub-centre level the appropriate tools were administered to the Auxiliary Nurse Midwifery (ANM) of that respective Sub Center. The participants were all Government officers occupying designated posts at the time of the study and in charge of activities related to Health Management Information System. When a designated post was vacant or the person was on long leave, efforts were made to include the person in charge of the related activities subject to his/her consent.

Study instruments / Data collection tools

Being a case study method, multiple methods were needed and employed to collect data. The main data collection methods were semi-structured interviews of study participants (detailed above), observation & review of the records (reporting formats, recording registers, reports, electronic systems). The tool for semi structured data collection would include questions related to determinants which influence procedural, managerial and psychosocial factors. Since we were considered 19 facilities for our data collection, so at each facility level the main officers in charge of HMIS related activities were interviewed. Review and observation of facility records and information system infrastructure were done at each facility. For observation and review of records and registers, checklists were developed. To check for consistency of data across multiple records and registers, three indicators were chosen. These include total number of deliveries, total number of child registration and administration of measles vaccine to child. Pre□testing was performed and the tools were modified accordingly as per pilot study experience.

RESULTS

Two terms are commonly used to describe the data quality measurements. These are Quality Assurance and Quality Control. Taulbee (1996) makes the distinction between Quality Control and Quality Assurance. Quality Control as a judgment of quality based on internal standards, processes and procedures established to control and monitor quality; and Quality Assurance as a judgment of quality based on standards external to the process and is the reviewing of the activities and quality control processes to insure that the final products meet predetermined standards of quality. However both the terms are used synonymously in practice of data quality management.

Completeness of data

From the observations of DHIS2 web portal and HMIS training modules, it is observed that, data completeness is assessed for the following two entities -

a. Number of facilities reported against total facilities

b. Number of data elements reported against total data elements in a reporting form

It was observed that with respect to both these aspects, the HMIS data was complete. HMIS has a module where a facility is registered. After registration of the facility, user name and password is generated for the facility. It is only after that the system can record data pertaining to that facility. All the government healthcare facilities are registered within HMIS. Unless a particular data element is filled, the application does not move to next data

element which addresses the issue of incompleteness of monthly reporting, as depicted by an ANM,

"सगळे रकाने भरल्याशिवाय ते एस.सी. चा फॉर्म सबमिटच होत नाही..."

"until you fill all the columns, the SC form not get submitted"

Roopali (SC ANM)

Table 1: Showing facility wise deadlines for submission of data

Sr no	Facility type	Last date for data submission
1	SC	3rd of every month
2	PHC	5th of every month
3	THO, RH, SDH,WH	7th of every month
4	Whole district as a unit	11th of every month

The operating system of DHIS2 does not allow submitting the incomplete facility based forms. But the issue of delay in the data entry from daily diary into the respective registers could potentially result in incompleteness.

Data Timeliness

The average two days gap is maintained in the schedule so as for data validation and monitoring process at each level in the hierarchy like SC, PHC, THO etc. The time framework for different facilities is given in [Table 2].

(Source - DHIS 2 Standard Operating Procedure document. Attached in annexure x)

During the study, a prominent finding was that, all the visited facilities submit the data on their respective time. Generally by and large after 11th of every month at 11 am, any data entry after this crucial timeline stops automatically as the website gets closed. These crucial timings got its own advantages and disadvantages. During this particular window period data should get entered, moreover, indirectly it ensures timeliness but at the same time it indicates a lack of trust over the lower hierarchy. Well whatever may be the reason but the benefit of it is in the form of data from all facilities gets filled up on time.

Timeliness is very important quality component of HMIS. Timely processing and reporting of data facilitates timely availability of data for decision making. e.g. during monthly meetings for the filling of PHC information on HMIS web portal, all the sub center information is collected and gathered in the PHC form and further evaluated and strategic decisions are taken to address the problems. If out of 7 sub-Centers 4 do not submit report on time, then it will be difficult for the MO to assess the performance and develop aplan for PHC in particular and of sub-Centers in general. Adherence of this schedule can improve data timeliness and overall data quality. For each level of information flows, that is, from SC facility to the district, from the district to the national level, there is a deadline for

sending reports from one level to another (shown in table).

In case of MCTS, as it runs on 'mission mode project', the real time data entry is the most vital part. Hence it is compulsory to do the entry of every given maternal and child health service in to the MCTS software within the given stipulated time. When asked to one of the PHC LHV about this MCTS data entry, she told that,

"सत्र सपल की लगेच तीन दिवसानमध्ये एम. सी. टी. एस. ला एन्ट्री करावी लागते....नाही तर ते मग इयू.ओव्हर इयू दाखवत राहते ..."

"After completion of immunization session, we have to made data entry into the MCTS…otherwise it shows 'due' and 'over due'…"

Shobha (PHC LHV)

After providing MCH related services, the service entry is to be made into MCTS, otherwise it shows that service as 'due service' and 'over due' service in next month's work plan. But during the study it was observed that, sometimes there is considerable delay in data entry and services updation in the MCTS. In this case the letters were issued from state offices to districts about to complete all the service updations.

Data Accuracy

Data accuracy refers to the correctness of data collected in terms of actual number of services provided or health events organized. Inaccurate data will yield incorrect conclusions during analyses and interpretation. Small errors at facility level will cumulate into bigger mistakes since data from various providers/facilities are aggregated. During the study it is observed that, poor data accuracy is mainly because of following three factors:

1. Gaps in understanding of data definitions and data collection methods
2. Data recording and data entry errors
3. Misreporting

During the study it was observed that, data entry errors could be either due to typing errors (like wrong numbers typed in computer) or due to wrong box entry (like data entered in wrong box e.g., 'ANC registration' data entered in 'Registration in first trimester'). One of the LHV working at PHC highlighted the same issue as –

"शेवटची तारीख गाठण्याच्या घाईमध्ये कधी कधी एखादा क ा चुकीच्या जागी प तो. मग ी.एच.ओ ऑफिस मधले एम.अं इ. असिस्टंट ला कॉल करून विचारतात"

"In hurry and to achieve the deadline, sometimes one of the figures are entered in the wrong places…then M and E assistant from DHO office call for the verification of the same…"

Geeta (LHV from PHC)

During the study, it was significantly observed that, there is constant pressure of other national programmes apart from DHIS2/HMIS and so data

can be manipulated and entered in order to avoid the action by seniors/memo and so the data quality get hampered.

From the interactions with health workers at SC and PHC facilities, it is observed that, many a times they report incorrect values on the total number of patients attended, so as to compensate previous/ last months over reporting or to compensate any medicine stock etc.. Which results in poor data quality. Data quality is hampered because people are collecting data to show the work done, as well as the use of resources, such as medicines, without considering the purposes of data collection for health management.

Data Duplication

Regarding the HMIS data quality, one of the important observation find during the interaction with the M and E Officer is that, data duplication is a frequent problem and it does hamper data quality. In many cases, ANM also reports services that are delivered by facilities such as PHC, RH or Women Hospital. Data duplication leads to false higher coverage of services and inaccurate decision making. For example, if a pregnant woman delivers in the RH, the ANM should not report this delivery. She can record this delivery in her register because the pregnant woman is registered with her but she should not report it. If ANM reports this delivery and the RH also reports, this leads to duplication. When asked to district 'M and E officer' about this issue, he told that,

"त्याच कस आहे कि, हा डी.एच.आय एस. २ चा नंबर त्या एम.सी.टी.एस सोबत जुळतच नाही, थोडा तरी गप राहतोच. बहुतेक हे मातांच्या सासर माहेर जाण्या येण्यामुळे होत असेल"

"The thing is that, the DHIS2 number never matches with MCTS, some gap is remaining. This may be happening because of travelling of mother from in laws house to mother's house… "

Mangesh (District M and E officer)

In case of MCTS, as the data entry is name and 'mother/child MCTS ID number' based, there is less chance of data duplication. And this is the reason that, there is potential gap in the 'maternal and child health' related data from the two soft wares i.e.DHIS2 and MCTS.

Data validation

Validation rules are data quality check mechanism based on verification of the logic of relation between related data elements. In this computerized DHIS2/HMIS system Data validation can be done by visual scanning, system generated validation facility and also with the formation of data validation committee.

Visual Scanning

From the observations of the DHIS2 reports, it finds that, through simple but careful examination of complete report many errors can be identified and

rectified. Once data entry is complete and a report is ready it should be checked for missing values, calculation mistakes, abnormal figures etc. The following table below shows the common errors in data entry.

Photo 1: Showing a table from a PHC's HMIS report for visual scanning.

System Generated Validation Facility

This DHIS2 software is customized by adding some validation rules based on statistical logics. Hence when the data is not matched as per given validation rules, then it shows validation error in the system.

Validation rules

Validation is also a means to check accuracy. Generally validation is performed by comparing values of 2 (or more) data elements that are related. One (or more) data elements are placed on left side and other data element(s) are placed on right side with an operator separating both sides e.g. 'Early ANC registration' is a part of 'ANC registration' and it can equal to 'ANC registration' or it will be less than or equal to 'ANC registration' but it can't be greater than 'ANC registration'. This rule can be expressed as:

Table 2: Showing validation rule for ANC registration.

Validation rule	Left side	Operator	Right side
Early ANC registration is less than or equal to total ANC registration	Early ANC registration	≤ (less than or equal to)	Total ANC registration

Table 3: Showing data validation rules for RCH HMIS.

Data Validation Rules	
1	**Antenatal Care**
I	ANC registration should be equal or greater than TT1
II	Early ANC registration must be ≤ to ANC registration
2	**DELIVERIES**
I	Caesarean Deliveries must be ≤ to institutional deliveries
II	Deliveries discharged under 48 hours ≤ deliveries at facility
III	BCG given should be ≤ Deliveries
IV	OPV0 given should be ≤ Deliveries
V	Total deliveries should be equal to live births +

		still births
3		**IMMUNISATION**
I		BCG should be ≤ to live births
II		Immunization sessions planned should be greater than or equal to sessions held
III		Measles dose given should be greater than or equal to full immunization
4		**NEWBORNS**
I		Newborns breastfed within 1 hour are ≤ total live births
II		Newborns weighed at birth ≤ total live births
III		Newborns weighing less than 2.5 kg≤ total newborns weighed
5		**POST NATAL CARE**
I		Women receiving first (within 48 hour) post-partum checkup ≤ to total live births plus still births

Hence statistically, the rules guide the validations and show the wrong data in grey color. It is important to note that violation of a validation rule does not always indicate error. You need to know that sometimes inconsistent/unexpected values may be due to management issues like availability of vaccines or medicines in stock, disease outbreak, etc. Violation of validation rule prompts you to enquire and check/verify data until satisfactory answer is not found.

From the observations and discussion with the study participants during the process of data collection, it is found that this computerized DHIS 2/HMIS system has its own developed validation rules in place. Few are listed in the following table

Break of Validation rules
Based on the provided validation rules, the system shows validation errors. But one can still enter the data with explanatory remarks for that validation error. This is provision, but not necessarily implemented in every case, as officials may review validation rules so as to avoid entry of false data entry. In an interview, one ANM was asked whether all newborn receive BCG on first day of life. She replied that it is not possible for home deliveries. Even for some institutional deliveries, the woman would get discharged and then come back for BCG. When asked whether she encounters cases where a child is born in previous month receives BCG in current month and she responded affirmatively. During further conversation, she mentioned that the system cannot take a higher figure. Thus, they have to enter same figure for the number of newborns and number of newborns given BCG, if the BCG figures are higher. This is a notable example of accuracy getting affected by validations.

During the study it finds that, along with this system generated validation rules, the monthly report sent to DHO office is checked/verified by M and E asst. Mostly the report is verified for its completeness, internal consistency, any outlier in the report, limit checks and correctness of the report received though logical observation. If any error is found then it is cleaned in consultation with the respective ANM or LHV via telephone.

Data Validation Committees
Once a monthly report has been prepared at the facility level, data has to be verified before it is submitted to the next facility level. The meeting of the facility wise validation committee is mandatory procedure before actual data uploading on the HMIS server.

The primary data verification should be done by the Medical officer of PHC which would be very much useful. The study finds that there is very little involvement of medical officer for the data verification process.

"ते (वैद्यकीय अधिकारी) महिन्याचे रिपोर्टवर फक्त सही करतात; त्यातली माहिती कधीच पहात नाहीत".

"He (Medical officer in charge) only puts his signature on report; never look at the data"
Mr. Mangesh (M and E of DHO office.)

In the study it is observed that, data validation committee meetings at SC and PHC are not taking place regularly. Even at Taluka and District level offices also the meeting for validation of HMIS related data was not found regular. Because of work pressure and target oriented work culture, the higher officials not find enough time for conduction of these validation meetings.

Minutes of meetings' of this validation committee meetings at district level and taluka level were seen but not seen in PHCs visited during the study, as these meetings were not regular at these PHCs.

Data quality assessed during supervision
Although DHO, Deputy Director and above senior officials, regional coordinators, circle M and E visit the PHC or SC for monitoring and supervision, but during these visits verification of records, cross checking, spot checking of data are less focused.

As a LHV replied about the visit of DHO to their PHC –

"हो ते भेट देतात, हे सर्व पहातात(माहिती,रेकोर्ड).सर्व व्यवस्थित आहे का ते विचारतात. खर तर त्याच्याकडे वेळच नसतो."

"Yes, they visit, do see all this(data, records)..just asks me if everything is fine; actually they do not have time".
Vaishali (LHV from PHC)

So because of time constraint, data quality is not much assessed by higher officials during their during supervisory results. Hence although this computerized HMIS system provide insight and tools that help in improving data quality, the quality of data cannot be improved through information technology only. Processes, people and technology involved in a information system together

determines the data quality (Karr, Sanil, and Sacks, 2000).

DISCUSSION

The discussion that follows is based on the findings of the study on the assessment of quality of data generated by the health management information system in selected health facilities.

Data quality is multidimensional, and involves data management, modeling and analysis, quality control and assurance, storage and presentation. As stated by Chrisman (1991) data quality is related to use and cannot be assessed independently of the user. The information system merely acts as an 'enabler' that helps to create, store, retrieve and manipulate the data items. Although this provide insight and tools that help in improving data quality, the quality of data cannot be improved through information technology only. Processes, people and technology involved in a information system together determines the data quality (Karr, Sanil, and Sacks, 2000). So, the quality assurance mechanism including data validation techniques plays a significant role in determining data quality even if we have a strong information technology support.

It is the responsibility of all health workers to check if HMIS data is of good quality. These quality checks are at all stages of data collection and analysis and reporting. Once these checks have been done and errors have been identified, the in-charge has the responsibility to make the corrections. In some cases the HMIS focal person or the nurse makes the corrections. Once a report has been prepared it should be verified before sending to the district health office.
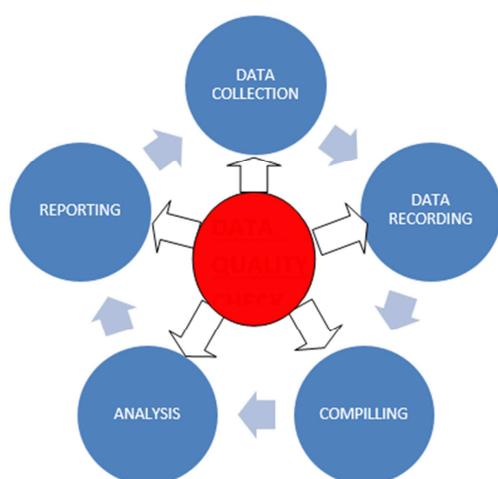
Diagram 1: Showing points of data quality check on the various steps of HMIS cycle.

Data quality mainly comprises completeness, correctness and timeliness. Quality of data means correct and meaningful data, which can be utilized for managerial, policy and planning issues or any other purposes it is intended for. The quality of data depends not only on attitudes of health workers but also on having necessary skills for doing things. The incompleteness of data raises doubts on the validity and reliability of data and its utilization in the management of health services. Secondly incompleteness of data will result in late report compilation and submission to the next level and thereby rendering the data not to be used by the different users. Mbananga during his study in South Africa has observed that, the data quality of health management information system in the study area as inaccurate, incomplete and inadequate in informing managers on all the decisions they needed to make (Mbananga et al., 2000).

On a monthly basis, reports from all facilities ought to be sent. In paper based system, the monitoring at state level would not be knowing how many and which facilities have not reported as the data was getting aggregated at PHC (all subcentre data was aggregated), at Block (all PHC data was aggregated), district (where all block data was aggregated). With HMIS, the data transmission changed to a central server system. This facilitated identification of the facilities that did not report. This led to reduction of this error. However, this does not indicate any positive change in other aspects notably accuracy. Rather, it is likely to get compromised.

In paper based system, certain fields not being entered was not uncommon. Blank spaces could mean service not available as well as service is available but not utilized. With HMIS, the data entry rules got created as the system allows entering data into a field only if the previous field has been completed. Thus, HMIS removed possibility of blanks. Thus, this error is prevented in HMIS. However, this would put pressure on the healthcare providers to enter some 'number' even when number is not known or not computed from existing registers. Thus, this rule affects completeness positively with certainty but it possibly could affect data accuracy negatively.

Accurate data is vital to the effective management of a health system or program. As the health workers record the data into their daily diary and then enter this data in to the various respective registers. This culture of multiple registers and the process of transcription of data from this daily diary to various registers can affect the data accuracy. Information is important because it, among other things, ensures the timely provision of effective services that in turn improves health care consumer satisfaction and it is vital for making informed decisions. Hence disparities in quality have variable impacts, ranging from the quality of care provided, timely delivery of services to clients to planning and organisational issues. Despite routine supervision conducted by the higher level

staff to the SC, PHC, THO poor quality of information continued. It could seem structural problems are some of the contributing factors to persistent poor data quality issues. Incompleteness and inaccuracy of data contribute to delays in preparing timely reports. These results in late reporting from health facilities.

Study conducted by Forster et al. (2008) shows that the training of staff and the time spent on the database by the health staff are positively associated with more complete data. Data Verification is the process of evaluating the completeness, correctness, and conformance /compliance of a specific data set against the method, procedural, requirements. The improvement in the process of data verification for better data quality and local use of data can be better achieved through the involvement of Medical officer and developing a sense of ownership at the block and district level staff. Greater involvement and support of MO at the PHC level is highly essential for clearing doubt and for data analysis and interpretation.

Inadequate or deficient validation check will lead to loss or compromised data accuracy and quality. Hence the correctness of data is lost. It will lead to ill informed decisions. Sometimes the missing data will lead to additional problems like incomplete reports, indicator calculations will be misleading. The study could not find any case of external verification that has been carried out in any of the study facility.

If the data validation is not there, then it has no meaning since it is just clustering of data. So the reports are inconclusive. Poor quality data adversely affects the quality of decision making, and even may lead to inappropriate decisions being taken if the data incorrectly assumed to be adequate quality. Kerr et al., (2007) have argued that - defining data quality and realizing the need for information that is free of defects and that possesses the right qualities for the task at hand remains a difficult issue. This is particularly so in the healthcare sector where the need for effective decision making is high.

From the study findings it can be suggested that, the monitoring and supervision activity should also focus on practice of looking at data, performing data validation, cross checking, record checking besides programmatic supervision. Further focus on data, record verification and other practices of data management by the supervisor at the reporting unit will also create a sense of ownership among the data collector. H/she may realize the importance of data and will take utmost care while not to repeat the same error.

Due to the culture of targets, the health staff in general fears 'zero' in their reports. The fact that zero can have different meaning is not considered. For example it could mean that there was no service given, there was a beneficiary but no means to provide service to her/him or there was no beneficiary and hence no service was given. However, this is not considered. This culture also leads to numbers deciding if a program is a success or failure. Also, it is seen that the service provider and the services being provided are equated. There is no differentiation. Due to this the onus of service delivery is completely on the individual. This again leads to fear of accountability and getting caught and hence false reporting. Thus the low quality of data prevails. This cycle continues. Thus health information systems thus end up being a mere representation of the authority.

Generally the accountability and sincerity is based on by the quantity of services rather than quality of services provided by those employees. The notion that the field level workers are responsible for the low quality of data is also incorrect. But the top management completely overlooks the possibility that the middle level management or they themselves could be responsible for the data quality. The ordering of mandatory reporting into HMIS will not solve the problem of inaccurate reporting in the sub centers. The normative pressures do not aid in the getting near to the root cause but causes it steer further where they are overlooked.

In a database, the data have no actual quality or value unless someone uses the data to do something useful (Dalcin 2004). The best incentive to the data producers/collectors is to ensure that the data collected are useful to them(Shrestha andBodart 2000).If there is no data analysis and use of data at local level, this may affect the moral of the people involved in data collection, recording and reporting. Which further decrease their interest to collect, record and report of the data. It becomes a continuous cycle finally resulting in poor quality of data collected and reported. So, the staff associated with data collection, recording, reporting, processing and data analysis should have clear-cut understanding of the importance of the data elements and local use of data.

By the findings and observations of the study, it can be said that data quality issues are due to various reasons. The data quality of the computerized MCH information systems is influenced by factors like incomplete training sets, top down solutions and decisions from the top-level management, lack of technology, normative pressures to comply with the targets, etc.

CONCLUSION

The DHIS2 introduced in the Maharashtra state in 2012, seems to have all the necessary essentials needed to operate an effective routine health information system. The system is providing data to stakeholders for use in managing health services.

However there is need to ensure data is of good quality. Since quality does not come automatically, deliberate effort is required to continue with quality assurance of the system.

Extra efforts, investments and recurrent costs are definitely required to ensure data quality in the district and facilities. However to improve data quality of routine health data requires innovative interventions that are cost effective to motivate the health workers to have quality data. The introduction of the incentive scheme that encourages health workers and facilities to improve data quality, introduction and intensification of innovative ways of providing supervision to health facilities and providing on the job practical sessions, quarterly and annual peer reviews at zone level and providing feedback by the DHO to the reporting units are some of the cost effective interventions that could contribute to data quality improvements. Further more implementation of advocacy activities on the importance of having quality data and its use targeting policy makers and other stakeholders could also be given prominence on the intervention list to be implemented. It is the responsibility of the district and facility management teams to take a leading role to ensure that all reporting units produce quality data that is fit for use. This therefore requires empowering the district health and facility management teams by defining clearly their roles and responsibilities and equipping them with the necessary skills, defining appropriate procedures and providing logistical support.

Despite some limitations such as high workload, the health workers are enthusiastic about the health management information system that has been established. What needs to done is to encourage them to do it. The proposed innovative interventions would help in improving the situation. The inclusion of the proposed HMIS interventions in the district implementation plans (DIPs) and sufficiently funding such activities would go a long in improving the routine health management information system at the facility, district and country levels.

REFERENCES

1. Aqil A, Lippeveld T, Yokoyama R. Yunnan Baseline HMIS Report, MEASURE Evaluation, CDC Yunnan, USAID; Aqil A, Lippeveld T, Yokoyama R. 2007 Guangxi Baseline HMIS Report, MEASURE Evaluation, CDC Guangxi, USAID. 2007.
2. Bernardi, R. (2011). IT Enactment of new Public Management: the Case Study of Health Information Sys-tems in Kenya. Leading Issues in E-Government, 1, 51.
3. Campbell B, Adjei S, Heywood A. From data to decision making in health: the evolution of a health management information system.Amsterdam: Royal Tropical Institute; 1996.
4. Chae YM, Kim SI, Lee BH, Choi SH, Kim IS. Implementing health management information systems: measuring success in Korea's health centers. International Journal of Health Planning and Management. 1994;341 – 8.
5. Evans T, Stansfield S. Health information in the new millennium: a gathering storm? Bull World Health Organ. 2003;81(12):856.
6. Health Metric Network (HMN), WHO, 2008, Framework and Standards for Country Health Information Systems, Second edition, June 2008, ISBN 978 92 4 159594 0, WHO Library Cataloguing-in-Publication Data, (NLM classification: W 26.5).
7. Hurtubise R (1984), Managing information system: concept and tools, west Hartford,CT, Kumarian press; 1-168.
8. Lippeveld T, Sauerborn R, and Bodart C. (eds.)2000. Design and implementation of health information systems, Geneva: WHO.
9. Lippeveld T, Aqil A, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy and Planning. 2009;24:217 – 228.
10. Freund, P.J., Kalumba, K. (1985). Monitoring and evaluation of primary health care in rural Zambia. A comparative study. Scand J Soc Med 13, 137–146.
11. Mapatano MA, Piripiri L. 2005. Some common errors in health information system report (DR Congo). Sante´ Publique 17: 551–8.
12. Mavimbe JC, Braa J, Bjune G. Assessing immunization data quality from routine reports in Mozambique. BMC Public Health. 2005;5:108.
13. Murray, C.J.L., Shengelia, B., Gupta, N., Moussavi, S., Tandon, A., Thieren, M. (2003). Validity of reported vaccination coverage in 45 countries. Lancet 362, 1022–1027.
14. Mitchell E, Sullivan FA. 2001. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980–97. British Medical Journal 322: 279–82.
15. Odhiambo-Otieno GW. 2005a. Evaluation of existing district health management information systems a case study of the district health systems in Kenya. International Journal of Medical Informatics 74: 733–44.
16. Rotich JK, Hannan TJ, Smith FE, Bii J, Odero WW, Vu N, Mamlin BW, Mamlin JJ, Einterz RM, Tierney WM: Installing and implementing a computer based patient record system in sub-Saharan Africa: the Mosoriot Medical Record System.
17. Redman, T.C. 1996. Data Quality for the Information Age. Artech House, Inc.
18. Shrestha, Laura. & Bodart, Claude. 2000. 'Data transmission, data processing & data quality', in Lippeveld, Theo., Sauerborn, Rainer. & Bodart, Claude.2000. Design and implementation of health information system. WHO Library Cataloguing in Publication data, Geneva, ISBN 92 4 1561998, pp.128-145
19. Sundararaman., Gupta, Pankaj., Mishra, Amit., Vasisht, Itisha., Kauser, Alia., Mairembam, Dilip Singh.2012.Designing an information technology system in public health : observations from India. BMC proceedings, Oral presentation, 2nd National conference on Bringing Evidence into public policy(EPHP 2012), Bangalore, India, 5-6 October 2012.
20. Sahay, S. (2011). Assessment of the ability of the Health Management Information System in India to use information for action (Doctoral dissertation, School of Public Health, University of the Western Cape).
21. S. Sahay, T. Sundararaman, J. Braa (2017). Public Health Informatics: Designing for Change- A Developing Country Prespective, Oxford University Press
22. T sundaram. Indian approaches to retaining skilled health workers in rural areas; Bulletin World Health Organization 2011;89:73–77
23. Taulbee, S.M. 1996. 'Implementing data quality systems in biomedical records,' pp. 47-75 in Handbook of data recording, maintenance, and management for the biomedical sciences. Boca Raton: CRC Press.

24. Van Etten E, Baltussen R, Bijmakers L, Niessen L. 2006. Advancing the Mexico agenda for the health systems research- from clinical efficacy to population health. Tropical Medicine and International Health 8: 1145–6.
25. Wang, R W.; Strong, D M. Beyond Accuracy: What Data Quality Means to Data Consumers Journal of Management Information Systems, Spring 1996, Vol. 12 Issue 4, p5, 30p
26. https://arogya.maharashtra.gov.in/1105/Divisional-Level
27. http://indianexpress.com/article/opinion/columns/neha-national-e-health-authority-4393612/ (last seen November 25, 2016)
28. https://orgtheory.wordpress.com/2017/01/02/the-not-so-new-institutionalism/ (last seen January 7th, 2017)

How to cite this article: AneraoRD, Parge GG. Study to Assess the Data Quality of Health Management Information System of a District from Central Maharashtra. Ann. Int. Med. Den. Res. 2018; 4(4):CM08-CM18.

Source of Support: Nil, **Conflict of Interest:** None declared