



Munich Personal RePEc Archive

Health Information Systems in India: Challenges and way forward

Tripathi, Shruti and Sharma, Rachna and Nagarajan,
Shyama

29 May 2018

Online at <https://mpra.ub.uni-muenchen.de/87067/>
MPRA Paper No. 87067, posted 03 Jun 2018 01:20 UTC

Health Information Systems in India Challenges and way forward

Dr. Shyama Nagarajan ^{*}, Dr. Shruti Tripathi [†], Rachna Sharma [‡]

May 29, 2018

Abstract

Universal Health Coverage (UHC) now endorsed by all UN member state as a part of their sustainable development goals, ushered in an era of improved governance and further development of health care systems. For the seamless functioning of health care systems and for their future development and sustainability, complete and reliable information is essential. The role of Health Information System (HIS) is to ensure recording, analysis, dissemination and use of reliable and timely data by decision-makers at all levels of the health system. Information is used in a wide range of situations: when developing national strategies and plans; when monitoring progress against national priorities; or when responding to public health emergencies. Information is also needed for greater accountability for results. Given India's commitment to UHC, this paper explores the current HIS of India and understands its completeness and usefulness given the international standards. Article further describes various issues pertaining to data collection, definition, analysis and dissemination at national, sub-national and institutional level, where do India lags and what can be done to develop a seamless HIS for India.

^{*}Managing Director, SahaManthran Pvt. Ltd., Haryana

[†]Economist, National Institute of Public Finance and Policy, New Delhi

[‡]Consultant, National Institute of Public Finance and Policy, New Delhi

1 Introduction

Complete and reliable information is essential for the seamless functioning of health care systems and for their future development and sustainability. With the countries striving to achieve UHC) as part of their Sustainable Development Goal (SDG), existence of a functional and integrated HIS which is able to connect with other information systems, such as civil registration and vital statistics, community health information systems, and various other information system has become the key requirement. Information is used in a wide range of situations: when developing national strategies and plans; when monitoring progress against national priorities; or when responding to public health emergencies. Information is also needed for greater accountability for results. The role of a health information system is to ensure the production, analysis, dissemination and use of reliable and timely data by decision-makers at all levels of the health system and pave the path for effective policy making.

However, evidences reflect that India is significantly compromised in terms of its data quality and quantity (in terms of periodicity and coverage). A study by Mikkelsen and et. al. (2015), reported that India fares very low (<0.25) in terms of the Vital Statistics Performance Index (VSPi), a composite index, that comments on timely generation of mortality and birth data. The report demonstrated that out of 148 countries, India was laid in the group of 39 worst performing countries. Survey of data sources conducted by Pandey *et al.* (2010) pointed out that there is a visible discrepancy between the type of information provided by the existing health information system and what is required by the public health planner. While number of studies have (i) evaluated the Health Management Information System (HMIS) conceived under National Rural Health Mission (NRHM) and its implementation challenges Madon *et al.* (2007), Bodavala (1998); (ii) elaborated on types of data management systems and data quality issues Husain *et al.* (2012); (iii) studied data related challenges with regard to cause of death, clinical audits and infection control audits(Mahapatra and Chalapati Rao (2001)); (iv) brought out the challenges to inter-linkages between health sensitive indicators impacting health (Chalasani (2010)). In our knowledge there are no existing studies that have comprehensively reviewed the HIS at a systemic level in India.

Therefore, the objective of this study is to bring out the data quality issues in HIS in India and also to question availability of such data for making policy decisions. We provide evidence both at a national and sub-national level with Andhra Pradesh (AP) as case study and explore possible solutions for India

from international experience. AP has been chosen because of its relative tech savvy nature, serving as benchmark for other states in HIS. Therefore, challenges that would get reflected in AP would be the minimum that a state would face while implementing a complete system for health information.

In the next section we discuss the main attributes of a well functioning HIS giving evidence from other countries. Section 3 discusses approach and methodology of the paper, section 4 gives the current structure of HIS in India. Section 5 and 6 focus on challenges in HIS at national and sub-national level and finally section 7 gives out the conclusion and way forward.

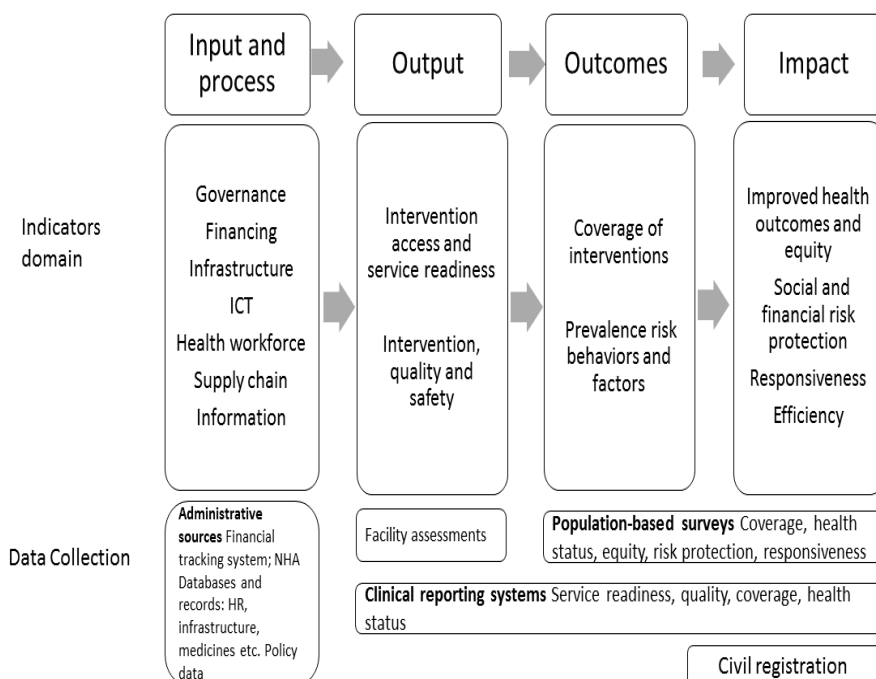
2 Features of strong HIS: International evidence

According to World Health Organisation (WHO) (WHO (2008)), HIS underpins decision-making and has four key functions: (i) data generation, (ii) compilation, (iii) analysis and synthesis, and (iv) communication and use. The HIS therefore, collects data from health and other relevant sectors; runs quality checks; ensures relevance and timeliness of data, analyses and converts the data into information for decision making.

For relevant data generation the WHO's HIS framework (Figure 1) brings together indicators and data sources across the results chain in its entirety, i.e. from "inputs/processes", "outputs", and "outcomes", to "impact". It is designed to address monitoring and evaluation needs for different users and multiple purposes at different levels of decision making, such as:

- **Individual** level data noted in the medical record provides information on patient's profile, health care needs and treatments and support continuum of clinical care at individual level.
- **Health facility** level data, are aggregated at hospital/clinic and at administrative levels (district, state and national levels) on procurement, financing, equipment and manpower to determine input required for operationalisation of services for the population as a whole.
- **Population** level data are essential for public health data at community level. Herein the information generated directly from the household sample/population, enables the decision makers to understand the people's need, health seeking behaviour and practices that are critical determinants for planning.

Figure 1: WHO framework for monitoring of health systems



Source: Monitoring the building blocks of health systems: a handbook of indicators and their measurement strength, WHO(2010)

- **Public health surveillance** brings together information from facilities and communities with focus on defining problems and providing timely information for urgent action to tackle larger public health issues, especially during epidemics.

This WHO framework has been used to map the globally agreed 100 core health indicators (WHO *et al.* (2015)) of Sustainable Development on the results chain in terms of the requisite health inputs and processes (for example: health workforce and infrastructure), defined outputs (interventions and available services), expected outcomes (coverage) and impact (morbidity and mortality).

After generation and compilation of data, it goes through quality check and analysis. For such checks, most international organisations and countries have developed data quality assessment framework outlining the various dimensions of quality measurement. For example, the International Monetary Fund (IMF) *Data Quality Assessment Framework* takes a holistic view in-

cluding governance of statistical systems. It covers five dimensions of quality in core statistical processes and products: assurance of integrity, methodological soundness, accuracy and reliability, serviceability and accessibility. The framework adopted by European Statistical System focuses more on statistical outputs and defines the quality of statistics with reference to six criteria: relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability and coherence. The Organization for Economic Cooperation and Development (OECD) framework views quality as a multifaceted concept; and the Eurostat approaches data quality from users perspective.

A survey undertaken in 25 OECD countries on strengthening HIS (OECD (2013)), reported that most OECD countries have a infrastructure both at national and sub national level to support HIS. They have the legal authority to collect identifiable personal health data ensuring data privacy and to monitor status of public health and undertake research by analysing the data collected . The evidence from participating countries show provision of:

- **Nationalised databases** reporting vital statistics and health care quality
- **Record linkage** across different databases using unique identifier for patients
- **Sub-national infrastructure for data linkage** for advance research as this brings access to greater details to individual's data than is available nationally
- **Legal framework for data identification and linkages.** The country specific laws adopt OECD guidelines that emphasise data collections to be respectful of the protection of personal privacy.

OECD countries also have a well laid system of **electronic health records (EHR)**. The longitudinal electronic record of individual patients virtually links together multiple electronic medical and non medical record systems that is interoperable across health care setting and provides historical data of patient's contact with the health care system. To support standardisation and to allow for cross reference a concept of **minimum data set** has been created. The minimum dataset would contain patient identifiers, such as a unique patient identifying number and a set of patient characteristics, namely: patient demographics, clinically relevant diagnostic concerns, such as chronic conditions, allergies; and unique identifiers for health care providers. This data is then embedded in a **smart card** and issued to the patient. These cards ensure access to accurate health records in a secured environment on-line.

Overall international evidence show that the main attributes of a well functioning HIS include: (i) generation of data at every step of result chain, (ii) Capacity to detect, investigate, communicate and contain events that threaten public health security, (iii) ability to synthesize information and apply the knowledge of financing, planning, and implementation (Lippeveld *et al.* (2000)).

3 Approach and Methodology

Health information systems, in India, have evolved in a haphazard and fragmented way as a result of administrative, economic and policy related decision. The responsibility for health data is often divided among different ministries or institutions, and coordination is difficult due to financial and administrative constraints. The goal of a health information system is often narrowly defined as the production of good-quality data. However, the ultimate goal is more than this - it is to produce relevant information that health system stakeholders can use for making transparent and evidence-based decisions for health system interventions. Little effort is made to present the information in formats that are relevant and comprehensible to policymakers, the public or those working in sectors other than health. Therefore, in the following sections we will explore the current HIS in India and understand where India stands given the international standards and where do India lags and what can be done to develop a seamless HIS for India.

In order to study the data quality issues in India, it is important for us to understand: (i) what are the different types of data that are being collated? (ii) what are the different sources from where data is generated? (iii) who analyses and reports it? As the data is being reviewed from the user's perspective, this is done by studying the data generation and reporting system as the exist currently in India. Then we look at data related challenges and concerns by comparing different sources for standard variables, listing gaps and shortage in data availability, various methodical and definitional inconsistencies. This will be supported with the analysis of data available in the public domain both at national and sub national level. **To study the data challenges at a sub-national level, this paper will undertake internal data validations and external consistency checks using various data sources of state of AP .**

4 Health information system in India

The health data in India gets generated from multiple sources (CSO (2015)). The main sources of health information are the central and state agencies. Some of the common data sources, generating **health-specific and health-sensitive** indicators¹, along with the ministries under which they are maintained are provided in Table 1. The Office of Registrar General and Census Commissioner under MHA reports on vital statistics, live births and death related data collected and disseminated through reports of Sample Registration System (SRS), Civil Registration System (CRS) and population Census. The Directorate General of Health Services (DGHS) under MoHFW is responsible for various sample surveys as well as administrative records. Data related to health resources like number of doctors, nurses; number of facilities and expenditure on health and, health insurance etc. gets recorded in National Health Profile (NHP) and Rural Health Statistics (RHS). There are four major health related surveys, conducted regularly in India at district level to monitor the performance of the government's various health interventions, including those under the National Rural Health Mission (NRHM): Annual Health Survey (AHS), District-level Household Survey (DLHS)(now merged with National Family Health Survey (NFHS)), NFHS and Coverage Evaluation Survey. They disseminate data on fertility, mother and child-care, family planning practices, mortality, disability, marriage etc. Another survey is carried out by National Sample Survey Organisation (NSSO) under Ministry of Statistics and Programme Implementation (MOSPI) called Consumer Expenditure Surveys, they are also sample survey held quinquennially. These surveys can be important source of information on affordability of health services including information on out of pocket expenditures for analysis on impoverishment and catastrophic expenditures. Health specific expenditure related surveys by NSSO have been conducted in 1986-87, 1995-96, 2004 and 2014; providing information on acute and chronic illness, hospitalisation, expenditure on medicines and treatment, source of financing, socio-economic status of individuals, demographic and educational profile of individuals, availability of drinking water and sanitation. The data are available by region, district, states, social strata and income quintiles. Table 2 summarises the features of various sources of health information like their

¹Health-specific data are those that get directly reported by the Ministry of Health and Family Welfare (MoHFW) and provide indicators that are specific to health. Health-sensitive data are those that get reported by other ministries, such as the Women and Child development, Ministry of Home Affairs (MHA) through NCRB, and provide indicators that impact health indirectly-lancet 2013, India

periodicity, responsible authority, usability of these key sources.

Table 1: Data sources for health information system

1. MHA	2. MoHFW	3. MOSPI
<ul style="list-style-type: none"> • Census • Sample registration system • Civil registration system 	<ul style="list-style-type: none"> • National family health survey • HMIS, National health mission • Annual health survey • Coverage evaluation survey • Concurrent monitoring • Rural Health Survey • National health profile • National health accounts • Integrated disease surveillance program 	<ul style="list-style-type: none"> • Consumer expenditure survey, NSSO • Health specific survey, NSSO

5 Challenges to HIS in India

Evaluation of indicators across data sources and examination of linkages of health data with reporting systems presents a major shortfall in India's HIS capacity to provide useful information. In this section we will discuss various issues pertaining to data collection, definition, analysis and dissemination at national, sub-national and institutional level. Overall we find that, data that is collected is incomplete or of poor quality; there is duplication and fragmentation of data across reporting systems that have been developed by different users.

5.1 Issues at national level

At the national level there's no independent central resource centre to provide stewardship to the health information and data management in the country. There's no unified body at a national level to provide guidance around: (i) types of data that can be collected at various level; (ii) responsibility for collection; (iii) legal and policy framework ; (iv) use of standardised definitions; (v) quality control and validation mechanism. Different ministries produce different data sets to fulfill their program specific needs (as shown in Table 2). There are four main issues relating to data at national level:

- **Different sources, divergent figures:** Variety of data sources in India have reported divergent population demography figures which leads to confusion in interpretation (Table 3). For example, while SRS, 2016 report sex ratio at birth for India as 898, NFHS-4, 2015 reports the same variable at 923 for the year 2015. Similarly, there is divergence between infant mortality rate presented by the two data sets. The data sources like NFHS and HMIS(NHM) provide the similar data on maternal and child health, that is, information on maternity care, delivery care, immunization, child hood diseases, family planning among others. However, they are incomparable owing to the following: while NFHS is a survey data from users point of view and done at irregular period, HMIS data is service statistics from PHCs, CHCs and hospitals collected annually. The recall period for questions related to maternity and delivery care is 5 years and that for child hood disease in 2 weeks in the case of NFHS, HMIS is actual data collected from the source of registration of mothers and children. Therefore, ability to compare or cross check an indicator across two data source it not possible.

Table 2: Summarising differences across data sources*

Information system	Responsible authority	Mechanism of data collection	Intent/ Usability	Periodicity
Sample registration system	ORGI	Population Based survey of the usual residents	vital events & mortality data, MDGs-outcome based data	Annual
Civil registration system	ORGI	Administrative data-compulsory & permanent reporting	Legal requirement under UN, enacted through a central birth & death Registration Act-69. Provides Medical Certification for Cause of Death. Provides basis for planning of infrastructure and services	Continuous
Integrated child development scheme	DGHS, Mo-HFW	Decentralised state based surveillance program	Picks early warning signs for an impending outbreak of specific diseases in specific states. District based surveillance of communicable/NCD.	Continuous
National family health survey	DGHS, Mo-HFW	Large scale district house hold survey	Covers MCH service utilisation, adolescent reproductive health, high risk behaviour, immunisation, NCDs	Every 5 years
Rural health statistics	DGHS, Mo-HFW	District house hold survey. generates evidence by undertaking requisite blood test and other lab investigations	Report on decentralised planning and coverage of RCH services.	Not defined
Annual health survey	Independent survey supervised by ORGI	Independent population survey in EAG states	Report on outcome/impact of schemes under NHM.	Annual(covering select states)
Concurrent monitoring	DGHS, Mo-HFW	Patient satisfaction survey	To report on patients opinion on NRHM course strategy	Not defined
Coverage evaluation survey	UNICEF and MoHFW	Population based survey	Evaluate the impact of NRHM strategy	Not defined
HMIS	MoHFW	Web based administrative reporting	Report on input and outcome of NHM strategy	Monthly
Central bureau of health intelligence	DGHS, Mo-HFW	Different web based formats	To report on incidence/prevalence of disease, health risk, performance, resources.	Annual
NSSO	CSO, MOSPI	Large scale population based survey	Reports on consumption and utilisation on different aspects every year essential statistics socio economic, demographic, agriculture etc.	Multiple rounds
National nutrition monitoring bureau	WCD	House hold survey of 10 states	to report on nutritional intake and nutritional status	Not defined
Human resource information system	DGHS, Mo-HFW	Web based employee data base	Reporting on efficient utilisation of HR, establishing transparent, transfer and evaluation system in 14 states	continuous

*Note: this apart there are program specific data reporting by various national programs, such as the

National Blindness Control Program

Table 3: Vital Statistics from different sources - All India

	CRS	SRS	NFHS -4
	2015	2016	2015
IMR		34	41
NMR		24	
U5IMR		39	50
BR	20.8	20.4	
DR	6.5	3.1	
Sex Ratio at birth	881	898	923

Source: SRS report 2016, CRS report 2015, NFHS -4 India fact sheet

- **Non-standardised decentralised procurement of nationalised Human Resource Information System (HRIS):** Currently, most human resources information system (HRIS) information is neither complete nor up-to-date and is often found in disparate paper files. The Government of India established human resources for health (HRH) as an important policy initiative in its 12th five-year plan. It was agreed that the most effective method to bring together such data is in a web-based information system. The review of states' HRIS system by Shukla *et al.* (2014) showed that states have either a paper-based system or one with some data in electronic spreadsheets. Some systems have only 10 data elements (Haryana) and some systems have more than 200 fields (Bihar and Jharkhand). Moreover, the protocols for data quality and updating also span a wide spectrum, from few checks to very sophisticated ones to assure quality and accuracy. This is because of lack of standard mechanism for procurement of such specialised softwares at state level. Also, as National Informatics Centre remained constrained in terms of their ability to incur any additional cost, other than for staff time and training; states were required to seek continuous external funding for its upkeep.
- **Lack of centralised mechanism for linking data across sources:** Various ministries produce different data sets, such as the MoHFW, WCD (indicators like violence against women, nutrition in children), Ministry of road transport and highways (road accidents, deaths). However, in the absence of a central data resource centre or data stewardship the system is unable to draw correlations across data sources and estimate the degree of impact of different social and environmental factors on health of people to support informed planing and improve implementation process. For instance, an article by Ackerson and Subramanian (2008) on analysis of NFHS-2 (1998-99) data indicated strong association between domestic violence on women with anaemia of chil-

dren.

- **Data dissemination and utilisation challenges:** There are gaps in dissemination and use of data at a national level, in terms of timeliness of dissemination. For instance, CRS data on population reporting are available only till 2015 making it impossible to serve as basis of policy recommendation. Also, there is a delay in compilation and synthesis of data, for example, NFHS -4 for 2015-16 could bring out All-India data only in December 2017, owing to lack of processed data from several states. Again, in terms of meeting international obligation to data reporting on health indicators for, example, SAARC development goals, WHO health indicators or India stands behind. Same is the case for expenditure on health data, National health accounts which disseminates this information was last published in 2017 with assimilated account for 2014-15.

5.2 Issues at a sub-national management level

At a sub-national level there's no state level data resource centre to coordinate and collaborate with national data resource centre for steering decentralised data collection and collation across sources. The data that is generated at state level lacks any information on private sector where about 70 percent of population seek treatment (NSSO (2014)).

- **Lack of private sector data** The data both at a national and sub-national level is compromised in terms of private sector data. The strength of the private sector is illustrated by the fact that it controls 80 per cent of doctors, 26 per cent of nurses, 49 per cent of beds and 78 per cent of ambulatory services (Planning Commission (2012)). In treating the in-patients, private institutions dominated both the rural (58%) and urban areas (68%) (NSSO (2014)). However, in case of both service records and administrative records private sector is not included. Therefore, there needs to be better mechanism for capturing this source of information for completeness of data. For example, the HMIS data that serves as the backbone for monitoring results of the National Health Mission comprising of the urban and rural sub-missions, needs extensive reform to accommodate the private health-care delivery system and surveys like NFHS are also restricted in terms of informing only about utilization of mother and child care at public facilities. Even the data on incidence of communicable and non-communicable diseases and data on cause of death is limited due to very marginal coverage

of private sector. Even in the case of national health accounts (most recent report was published in 2017 presenting account for 2014-15) which gives out expenditure on health in both public and private sector, list of health care providers and related capital expenditures in the private sector, is not exhaustive due to non-availability of disaggregate data.

- **Limited disaggregation of data:** Disaggregated data on coverage for specific population groups residing in remote areas, referring to vulnerable and marginalised population, especially those working in informal sectors are limited. Details around preventive, promotive, rehabilitative and palliative services are insufficient. Similarly, disaggregated health insurance data by socio-economic categories under different schemes are not available in a timely manner to evaluate impact on the vulnerable population. For example, in case of SRS information on MMR, IMR is available only for bigger states and that too combined for rural and urban. It does not provide data below state level. Like the SRS, NFHS was not providing estimates below the State level till the third round. However, NFHS- 4 provides estimates at district level for most indicators

5.3 Issues in data compilation and analysis

- **Lack of training on probing skills:** A lot of success of any survey depends on probing and interviewing skill of field officers conducting surveys. While both CSO and NSSO has a training divisions for their field officers, but they lack in various areas. There is a marked lack in infrastructural facilities, there is no systematic calendar or arrangement for training and there is often no linkage between the kind of training received by an officer and his or her assignment. As a result, for example, while both Consumer Expenditure Survey and morbidity and health care surveys both conducted by NSSO and collect information on health expenditures, it has been shown that the way questions are put forth to households, the spending reported by both surveys lead to very different results(Garg and Karan (2009)).
- **Shortage of staff:** Primary reasons for poor data quality is the shortage of qualified personnel specially the nurses, ANMs, data entry operator who are responsible for data inputs. As per the report of Economist (2008) only 30 per cent of nursing positions in rural hospitals are filled and a single ANM covers over five villages. Data entry at the sub-centre

level is by ANMs writing into physical registers. There are bound to be errors at this level because ANMs record data in handmade registers which are very badly designed. These registers sometimes do not have enough space available to write. Therefore for the over worked ANMs spending hours in data entry along with touring the villages and registering and providing medicines to pregnant women, quality of data while fulfilling reporting obligation becomes secondary and often subject of error and omission. At the PHC level, the Data Entry Operator is responsible for entering data for District-HIS. Alongside, she is responsible for fulfilling several other reporting requirements too. For example, there is another health information system called Mother and Child Tracking System (MCTS). This too has its parallel reporting requirements and the Data Entry Operator has to report data for MCTS too. Similarly, the Data Entry Operator has to undertake data entry of immunisation report, vaccine and logistics, release and logbook data.

5.4 Intrinsic data quality issues

- **Methodical issues.** There is little standardization in amount of information collected and definitions of indicators across states, making the data being reported of little practical use. For instance, under CRS many states report births according to the date of registration instead of the date on which the the birth takes place. Thus, making data incomparable across states.
- **Data traingulation issues.** Before usage of data for research or policy formation data fidelity should be assured by triangulation with data from periodic surveys and community based monitoring. In SRS, quality is assured, as there is a continuous enumeration of births and deaths in selected sample units by resident part time enumerators, and an independent survey every six months by SRS supervisors. The data obtained by these two independent functionaries are matched, re-verified and thereafter an unduplicated count of births and deaths is obtained. Further the SRS is also compared with the Civil registration and vital statistics (CRVS) (Garg and Karan (2009)). However, civil registration is still not complete and a better process needs to be established for states lagging behind. Various surveys like Census, National Sample Survey (NSS) and NFHS collect valuable information on demography, expenditure and utilisation of services. However, the irregular time difference between any two surveys and the time lag with which reports

are made available on the website differ across the sources. Therefore, the triangulation of statistics at disaggregated across sections of the society becomes difficult.

- **Inconsistent data definitions across survey rounds:** There exist differences in concepts and definitions with respect to some important parameters in various survey rounds, across organisations, making the results incomparable. For instance, the results of NSS 71st round is not strictly comparable with the results of NSS 60th round. Few of these differences are listed in Table 4. In the 60th round and earlier surveys on health, “persons with disabilities” were regarded as ailing persons. In the 70th round, “pre-existing disabilities” have not been recorded as ailments, unless they were under treatment for over a month during the reference period, in which case they were considered as chronic ailments. In the earlier surveys, for each person aged 60 years or more, (up to three) ailments existing on the date of survey and the nature of treatment of such ailments, were recorded in addition to information on ailments suffered during the reference period of last 15 days. In the 71st round, such information on ailments as on the date of survey were not collected.

Table 4: Difference across rounds of NSSO: 71st round compared with 60th round

Heads	60th Round	71st Round
Definition of Household	As used by NSSO in the consumption expenditure surveys	few exceptions for household members and included <ul style="list-style-type: none"> • students residing in hostels irrespective of period of absence • any women undergone childbirth in last one year and incurred some cost irrespective of her place of residence • child less than 1 year is member of the household to which its mother belongs
Disability	Persons with pre-existing disability regarded as ailing persons	<ul style="list-style-type: none"> • Persons with pre-existing disability under treatment for a specified period during the reference period ² was classified as chronic ailment, else not recorded as ailment • disability acquired within the reference period is recorded as ailment
Medical treatment	Self-medication or use of medicines on advice of chemist - not considered as medical treatment	all such treatments considered as medical treatment
Child birth	Collected consolidated expenditure incurred pre and post and during childbirth	<ul style="list-style-type: none"> • Is coded as a dummy ailment to record detailed expenditure incurred • Not included in estimating proportion of ailing persons (PAP)
Persons aged 60+	Collected the information of ailment <ul style="list-style-type: none"> • reported during reference period (last 15 days) • reported on date of survey and nature of treatment 	<ul style="list-style-type: none"> • Information of ailments reported on date of survey not collected (for any age group) • however, information on economic dependence; • living arrangement; and • State of health
List of ailments		updated (20 new ailments added to the list used for the 60th round)
Nature of treatment	Not differentiated	recorded as separate categories

6 Case Study: Andhra Pradesh

6.1 HIS in Andhra Pradesh

Before analysing the health data in AP, there is a need to understand the system through which that information stems. The Figure 2 illustrates the health information flows in AP. Secretary of health heads the department of health. The following directors - Commissioner Family Welfare, AP Vaidya Vidhan Parishad (APVVP); Director of Health Services, Medical Education and First Referral and Primary Health Care projects- report to the secretary and provide the relevant data. Respective project/ program director are responsible to furnish reports to the Directors, timely on data both program (input, output, outcome) and administrative.

Figure 2: Information flow from various sources

Secretary Health and Family Welfare				
Commissioner family welfare <ul style="list-style-type: none"> Deputy Director Statistics 	Director health services <ul style="list-style-type: none"> Deputy Director planning Deputy Director Vital statistics (CRS, Municipal registration) Missions: TB, Malaria, Leprosy, NPCB 	Commissioner APVVP <ul style="list-style-type: none"> Officer Statistics 	Director medical education <ul style="list-style-type: none"> Section officer statistics 	Director AP first referral and PHC projects <ul style="list-style-type: none"> Deputy Director MIS (first referral project)

Source: Evaluation of HMIS in India: Need for computerized database (Bodavala (2012))

For our analysis, we analysed various data sources providing information on AP - both reported centrally and at the state level, and performed various checks for internal and external validation. Often errors are made during data entry, which give rise to absurd data. To cross check and look for inconsistencies is called a method of internal validation. For internal validation we used data from HMIS, a set of simple validation checks based on logical relationship between data elements was used to analyse the data. For exter-

nal validation, that is checking whether same data elements across different sources are in sync with each other, we looked at vital statistics at state level for AP, mother and child health, facility level information and workforce level information for the comparable years. The results of this exercises put a question mark on the usability and trust worthiness of the data.

6.2 Inconsistency due to manual feeding

While entering information manually, often errors are made during typing, which give rise to incorrect data. Given below (Table 5) are a set of simple validation checks that are based on logical relationship between data elements in HMIS as reported by National Health Mission (NHM), which testifies the above inference of incorrect data entries. We find that in the cases listed below the criterion of validations checks are not met. For example, number of infants reported to have received BCG should be less than total number of deliveries as well as total number of live births reported. However, in both the cases the number of infants reported to receive BCG were higher than the number of births. Similarly, number of complicated pregnancies treated with particular form of IV antibiotics and oxytocics should be less than total number of obstetrics complications reported. Data shows that number of complicated pregnancies treated are much higher than the total complicated pregnancies reported.

Table 5: Internal Inconsistency in AP HMIS

Validation Criteria	2015-16	2016-17
1. BCG given should be \leq number of deliveries		
Infants received BCG	834135	778563
Total reported deliveries	767534	737741
Ratio in %	110	108
2. Total deliveries should be \equiv live birth + still birth		
Total Deliveries	767534	737741
Total Live + still birth	814276	755359
GAP(%)	6%	3%
3. No. of complicated pregnancies treated with IV Antibiotic \leq No. of pregnant women with Obstetric Complications		
Number of complicated pregnancies treated with IV Antibiotics	71735	104525
Total Complicated Pregnancies	63520	71678
Ratio in %	113	145
4. No. of complicated pregnancies treated with IV Oxytocics \leq No. of pregnant women with Obstetric Complications		
Number of complicated pregnancies treated with IV Oxytocics	96792	133053
Total Complicated Pregnancies	63520	71678
Ratio in %	152	185

6.3 Discordance between data sources

1. **Vital Statistics** We reviewed the vital statistics given by CRS, SRS and NFHS simultaneously, for the most recent year available. We found that while CRS is supposed to collate administrative data at sub national level on birth and death, the recent 2014 CRS report has published birth and death rate numbers using SRS data. Therefore, adding CRS to our analysis doesn't add value. SRS in it's 2016 report has published data for 2014 survey and NFHS-4 survey publishes the numbers for the year 2015. However, both the data sets give the estimation of vital statistics on the basis of a sample collected, which may vary from the numbers calculated using the population. Comparing the two databases in Table 6 we find the value of the indicators across data sources do not match for comparable indicators. For other vital statistics like Birth and Death rate there is no way to cross check the correctness of any given number, as this information is not available across sources.

Table 6: External Validation - AP Vital Statistics

	SRS*	NFHS-4
	2016	2015
IMR (urban)	26	20
U5IMR	37	41
BR	17	
DR	7.3	
Sex Ratio(urban)	885	1010
Delivery conducted by skilled health personnel(urban)	0.1	2
Institutional birth in public facilities	51.3	38.3

* SRS: given for period 2013-15

2. **Mother and Child health:** Comparison between NFHS -4 and NRHM data for the year 2015-16 for AP for some comparable indicators, show wide diversion between the two data sets. For instance percentage of total cesarean deliveries, cesarean deliveries in private facilities and cesarean deliveries in public facilities; all the three data points were incomparable across the two data sources (see Table 7). Even the reported sex ratio at birth by NFHS (914)and NRHM (959) were very different. It is difficult to say which one is over/under estimated over the other, but the difference across the two leading data sets is alarming.

Table 7: External Validation - NFHS-4 State fact sheet and NRHM key HMIS indicators for AP

NFHS-4,2015-16		NRHM 2015-16	
Definition	Value	Definition	Value
Institutional deliveries to total reported deliveries	91.6	Institutional births (%)	98.2
Births delivered by caesarean section (%)	40.1	C-section deliveries to reported institutional deliveries (%)	33.9
Births in a private health facility delivered by caesarean section (%)	57	C-sections conducted at private facilities to Deliveries conducted at private facilities (%)	41.9
Births in a public health facility delivered by caesarean section (%)	25.5	C-sections conducted at public facilities to Deliveries conducted at public facilities (%)	22.8
Institutional births in public facility (%)	38.3	Deliveries conducted at Public Institutions to Total Institutional Deliveries (%)	41.9
Sex ratio at birth for children born in the last five years (females per 1,000 males)	914	Sex Ratio at birth (Female Live Births/ Male Births *1000)	959

3. **Health care infrastructure.** To understand the extent of mis-reporting and non comparability across data sources we also looked at data on health care human resources and infrastructure. We compared the data availability of various public health care facilities, its penetration and availability of human resources (See Table 8 and 9). To achieve this we compared the last available DLHS survey, which was fourth round survey for the year 2012-13 and RHS for the year 2012-13. Even here we found that the numbers reflected in both the data sets were widely different.

Table 8: Facility level information - 2012-13

Variables	DLHS-4	RHS
Average population covered by health facility, Sub Center	5134	4501
Average population covered by health facility, PHC	40945	32979
Average population covered by health facility, CHC	106273	193020
Total number of Sub Center	468	12522
Total number of PHCs	361	1709
Total number of CHCs	156	292
Total number of Sub-Divisional Hospital	38	61
Total number of District Hospital	17	17
Number of CHCs having new born care services	106	240

Table 9: Workforce level information - 2012-13

Variables	DLHS-4	RHS
Percentage of Sub center with ANMs	90.4	90
Percentage of PHC with Medical officer	90	86
Percentage of PHC with Lady medical officer	41	54
Percentage of PHC with Ayush Doctor	21	93
Percentage of PHC with Pharmacists	53	79

These data differences across various data sources reflect that there are serious errors in the state on data entry, data aggregation, data compilation and validation. These errors stem from systemic issues like existence of multiple registers, misinterpretation, duplication of data, lack of written guidelines and procedures, shortage of staff etc. Errors occurring when validating data elements across data sources show that there is a need for standardisation of data definitions across sources or element's nomenclature across source be different.

7 Conclusion and Way forward

A review of international HIS shows that, the system should have following characteristics: (i) be both population and community based and should include facility-based information. (ii) It should link all service providers, laboratories and manufacturing units, so that it is able to provide information needed to monitor disease burden, mortality and subsequently support decision-making and resource allocation. (iii) To ensure quality, the data needs to be tallied with periodic surveys and community based monitoring. (iv) Requisite safeguards for protection of privacy rights will have to be put in place to ensure anonymity. To facilitate the transition of HIS in India from its present state to a level where it is compliant with international standards, there is a need to review the information processes and existing systems; highlight the existing flaws in the system and suggest ways to improve upon them. This paper makes an attempt to plug this gap by reviewing the existing health information sources at national and sub-national level in India and understanding the challenges faced by them. This section focuses on our learning and tries to give ways around those challenges.

- **Strengthening data at a national level:**

1. *Nationalised HIS infrastructure:* National databases with individual level records should be made available across the spectrum of

health care administration linked to unique ID. Mobile devices can serve as enablers, which can be rolled out across the country in a phased manner with a uniform guideline to avoid mismatch across IT platforms at sub-national levels and to reduce the additional expense of adding a bridge to pull and push data across sources. This would also help in timely dissemination of relevant data for various analysis and decision making purposes.

2. *Legal framework to ensure data privacy:* The country should lay down regulations to ensure people's privacy in data accessibility and usability across various environment and stakeholder, especially for research purposes to prevent misuse. Informed consent is the pillar for protecting individual's autonomy thereby build professional codes of practice to obtain consent. Where there are multiple data custodians, there should be legal and information custodian policy frameworks with accreditation and certification process to provide for the safe sharing identifiable personal health data.
3. *Integrating data bases:* It is important to link data bases across operations of an individual to draw effective correlations, avoid risk factors and take decisions to impact people's health positively and thereby improve overall health outcomes.

- **Strengthening data at a sub-national and institutional level:**

To use data seamlessly, data needs to be of high quality so that data users are confident that the data they are analysing are accurate, complete, and timely. Therefore, there is a need for:

1. *Standardisation of data definitions:* in order to ensure that users understand variables in uniform way, the data sets needs to be defined at each level, standardised definition for each data element should be used, data elements should also be made internationally comparable.
2. *Adequacy of human resource:* Given the importance of data in decision making it is essential that separate workforce outside of the operational system should be made available to collect, collate, aggregate and validate data from the administrative system of the health sector. The job description of such staff should clearly align their roles to ensure development of robust HIS at sub national and national level
3. *Capacity building of staff:* The lack of interaction between individ-

uals who design and manage the data manual and individuals who provide data input and professionals who use data in program improvement. To improve sustainability and use of data in decision making, individual competency to demand and analyse data must exist at all levels of the health system. Competencies include skills in data analysis, interpretation, synthesis, presentation, and the development of evidence-based programmatic recommendations.

4. *Improve data availability:* The best way to capture data is to ensure routine capturing of disaggregated data, without duplication of forms and formats to reduce the efforts of the data producers. Secondly, creating mechanisms for mandatory disclosure of public health data by private sector.
5. *Monitoring and communication:* In order for stakeholders and decision makers to use data in decision making, they need to place value on data (Lavis *et al.* (2006)). This value can be built through a positive experience using information to support a decision, through training or through exposure to positive messages about the benefits of using data in the decision-making process (Foreit *et al.* (2006)).

References

- Ackerson LK, Subramanian S (2008). “Domestic violence and chronic malnutrition among women and children in India.” *American journal of epidemiology*, **167**(10), 1188–1196.
- Bodavala R (1998). “Evaluation of Health Management Information System in India Need for Computerized Databases in HMIS.” *Ranganayakulu Bodalava, Takami Fellow in International Health, Harvard School of Public Health*, **665**.
- Chalasanani S (2010). “The Changing Relationship between Household Wealth and Child Survival in India.” Demographir and Health Research, Working Paper No. 69.
- CSO (2015). “Manual on Health Statistics in India.”
- Economist T (2008). “A special report on India: Creaking, groaning: Infrastructure is India’s biggest handicap.” *The Economist*, **11**.
- Foreit K, Moreland S, LaFond A (2006). “Data demand and information use in the health sector: conceptual framework.” *Chapel Hill, NC: MEASURE Evaluation, Carolina Population Center*.
- Garg CC, Karan AK (2009). “Reducing out-of-pocket expenditures to reduce poverty: a disaggregated analysis at rural-urban and state level in India.” *Health policy and planning*, **24**(2), 116–128.
- Husain Z, Saikia N, Bora R (2012). “Opportunities and challenges of health management information system in India: A case study of Uttarakhand.” (40014).
- Lavis JN, Lomas J, Hamid M, Sewankambo NK (2006). “Assessing country-level efforts to link research to action.” *Bulletin of the World Health Organization*, **84**(8), 620–628.
- Lippeveld T, Sauerborn R, Bodart C, Organization WH, *et al.* (2000). “Design and implementation of health information systems.”
- Madon S, Sahay S, Sudan R (2007). “E-government policy and health information systems implementation in Andhra Pradesh, India: need for articulation of linkages between the macro and the micro.” *The Information Society*, **23**(5), 327–344.
- Mahapatra P, Chalapati Rao P (2001). “Case of death reporting systems in India: a performance analysis.” *National Medical Journal of India*, **14**, 154–62.
- Mikkelsen L, *et al* (2015). “A global assessment of civil registration and vital statistics systems: monitoring data quality and progress.” *The Lancet*, **386**(10001), 1395–1406.

- NSSO (2014). "Health in India." *Technical report*, National Sample Survey Organisation.
- OECD (2013). "Strengthening Health Information Infrastructure for Health Care Quality Governance." *OECD Health Policy Studies*. doi:<http://dx.doi.org/10.1787/9789264193505-en>.
- Pandey A, Roy N, Bhawsar R, Mishra R (2010). "Health information system in India: issues of data availability and quality." *Demography India*, **39**(1), 111–8.
- Planning Commission (2012). "Report of the steering committee on Health for the 12th five year plan." *Technical report*, Health Division.
- Shukla M, Verma S, Narayanan M, Potenziani D (2014). "Human Resources Information Systems (HRIS): A Review across States of India."
- WHO (2008). "Framework and standards for country health information systems." URL <http://www.healthmetricsnetwork.org>.
- WHO, *et al.* (2015). "Global reference list of 100 core health indicators."