Demography India
A Journal of Indian Association of Study of Population
Journal Homepage: https://demographyindia.iasp.ac.in/

Evaluation of Health Management Information System Data Quality in Haryana

Bindiya Kumari[1]*, Suresh Sharma [2] and Madhu Ahlawat [3]

Keywords

Consistency, Data Quality, Health, HMIS, NFHS, Outliers, Service Delivery Indicators, Validation errors

Abstract

Background: High data quality and quality assessment of HMIS data are imperative for effective public health interventions, programme changes, and strengthening national plans. This desk research assesses HMIS data quality in Haryana across four key dimensions: Completeness, Accuracy, Internal consistency, and External consistency.

Methods: WHO Data Quality Framework is employed to gauge data quality dimensions, using HMIS portal data for 2015–16 (Base year) and 2019–20 (Reference Year). External consistency approach was measured by comparing HMIS data with NFHS- 4 and NFHS-5 estimates.

Results: The result revealed a notably high reporting rate at the upper health tier in Haryana. Validation errors (VE) decreased compared to baseline, though outliers increased particularly in child immunization. The highest outliers occurred in March, financial year's end, while mostly validation errors appeared in April, marking the year's start. External approach demonstrated commendable consistency in delivery care indicators, affirming HMIS data reliability for Institutional births and C-section. However, discrepancy emerged in IFA pills intake by pregnant women, except in Hisar district.

Conclusions: The study asserts substantial scope for improvement in outliers and VE across districts in Haryana. Key indicators with validity issues requiring immediate attention encompass- child immunization, newborn weighted at birth, and IFA tablets to PW. The study concludes a decline in completeness rate of health facility reporting, especially at PHCs level, requires urgent attention. Regular training, monitoring, and evaluation at concerned levels are essential for ensuring data quality.

*Corresponding Author

[1]*Field Investigator at Population Research Centre, Institute of Economic Growth, University of Delhi. Email: bindiya.prc@iegindia.org; kinhabindiya@gmail.com ; ORCID: 0009-0005-8009-6536*

[2]*Professor and Head at Population Research Centre, Institute of Economic Growth, University of Delhi. Email: suresh@iegindia.org*

[3]*Professor and Head at Department of Economics, BMU Rohtak, Haryana. Email: 28madhuahlawat@gmail.com*

Introduction

The Health Management Information System was launched in October 2008 with the initial objective to upload district wise consolidated figures, and later facility-based reporting initiated in 2011 (Ojha, D. K., 2023) Developed by the Ministry of Health and Family Welfare (MoHFW), GOI, HMIS is a web-based platform designed to monitor the National Health Mission and other Health programs, and to provide inputs for policy formulation and appropriate program interventions (HMIS Analytical Report, 2019-20). Currently, around 2.08 lakh health facilities, including 8,885 private ones, across all States/UTs are uploading monthly facility-wise service delivery data and annual infrastructure-related data on the HMIS web portal (HMIS Analytical Report, 2019-20). HMIS captures comprehensive facility-wise information on service delivery across various themes reproductive, maternal and child health, immunization, family planning, morbidity and mortality, OPD, IPD services, and surgeries.

HMIS is a system that integrates data collection, processing, reporting, and utilization information to support decision-making, enhance health service effectiveness and improve efficiency through better management at regional and global levels (Krishnan et al., 2010; Teklegiorgis et al., 2016; Githinji, et al., 2017). According to WHO, it helps in gathering, analysing and improving health system performance. Assessing HMIS data quality is extremely essential for key program changes and further strengthening national health plans, as existing studies have showed numerous quality-problems in practices, including data entry errors due, missing values, outliers, compilation/ computation errors, incomplete or untimely reporting (Maiga, A. et al., 2019; Dehury, R. K., & Chatterjee, S. C., 2018; James, K. S. & Prabhuswamy, P., 2016; WHO, 2003). These problems can undermine the reliability of data used for public health interventions and policy decisions. Therefore, it is essential data quality assessment should always be undertaken to understand how much confidence can be placed in the health data reported. This study is motivated by the need to address these data quality concerns to ensure that HMIS data is accurate and reliable.

Among the Northern States, Haryana depicted a steep fall in facility reporting rate from 98% in 2015-16 to 90% in 2019-20 (HMIS reports). This study focuses on Haryana to assess whether the decline in reporting rate is concurrent with any decline in the quality of data reporting. This paper examines the HMIS data quality of key health service delivery indicators for the state of Haryana, focusing on completeness, accuracy, internal consistency, and external consistency. To achieve this, our study employs the WHO data quality framework, which provides structured guidelines for evaluating various dimensions of data quality. The WHO framework for data quality review includes several comprehensive guidelines and modules such as the Immunization Data Quality Audit (DQA) Procedure (World Health Organization, 2003), the Data Quality Review Framework and Metrics (WHO, 2017), the Desk Review of Data Quality (WHO, 2017), and the Data Verification and System Assessment (WHO, 2017).

Data Source and Methods

This study utilizes HMIS portal data over two periods, 2015-16 (Base Year) and 2019-20 (Reference Year), to assess data quality. The WHO data quality framework is employed

to gauge the data quality dimensions: Completeness, Statistical Outlier and Data Validation, Internal and External consistency. The motivation for choosing the WHO methodology stems from its comprehensive and widely accepted framework for assessing data quality. This approach provides structured guidelines for evaluating data quality aspects such as completeness, accuracy, and consistency, which aligning well with our study's objectives.

The time frame of 2015-16 and 2019-20 was chosen for a meaningful comparative analysis over a significant period that captures recent changes and trends in data quality. Additionally, these years align with the availability of data from the National Family Health Survey, specifically round NFHS-4 (2015-16) and NFHS-5 (2019-20). Utilizing these rounds ensures external consistency in our analysis. As NFHS round 6 is still ongoing, the consistency in time frames across different quality dimensions is essential for the reliability and comparability of analysis/results.

To examine the completeness of data reporting, we calculated the facility completeness rate, completeness of data item, and accounted for blanks and zero values. Data for the study has been obtained from the standard reports on the HMIS Portal, including 'Data Reporting Status,' 'Monthly Data Item-wise Report,' and 'Performance of Key HMIS Indicators. Completeness of Facility reporting rate calculated by using following formula:

Facility Completeness rate=

Actual Reports/ Expected Reports* 100

Statistical Outliers and Validation checks have been identified using HMIS data outlier and validation reports, specifically from the 'Data Quality – Probable Validation and Outliers' and 'Validation Summary' files. In statistical terms, if the value lies 1.5 Standard Deviations away from the data Inter Quartile Range (IQR) it is identified as an outlier. Validation rule is measured by comparing values of two or more data elements. For example, the number of pregnant women given IFA Tablets should be ≤ the total number of PW registered for ANC.

Internal consistency was carried out over time (the most recent year compared with the mean value of the same indicator for the previous three years) as well as consistency for between selected indicators values calculated. The temporal consistency of selected few indicators, was measured against state values of those indicators by using following formula for each district:

$$\text{PercentageDifference}=\left|\frac{District\ Ratio-State\ Ratio}{State\ ratio}\right|*100$$

External consistency approach was calculated by comparing with household survey NFHS-4 (2015-16) and NFHS-5 (2019-21) estimates. If HMIS/NFHS ratio is 1, coverage rates are exactly same. If ratio > 1, HMIS coverage is higher; if < 1, NFHS coverage is higher.

Results

Dimension 1: Data Completeness

Completeness of Reporting

Overall, completeness of facility reporting rate in Haryana decreased from 98 percent in the base year to around 90 percent in the reference year (See Figure 1). Notably, upper health tiers exhibited better reporting rates in both years compared to lower tier in both years. DH maintained a 100 percent completeness rate throughout, while in CHC is also high hovered around 95 percent albeit slightly lower compared to base year. Completeness rate at PHC remains lowest in both years, underscoring the urgent need for focused attention on PHCs, especially considering the substantial decrease observed from 95 percent to 76 percent.

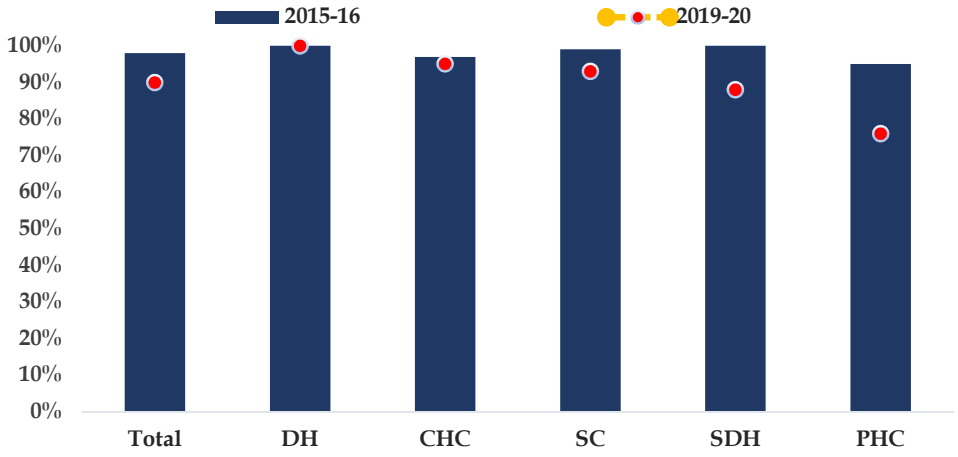

Figure 1 Decreasing trends in Completeness of Facility Reporting Rate in Haryana

The reporting rates for active health facilities also experienced a decline from 100 percent to 94 percent in reference year (see Table 1). PHC and SCs notably decreased their reporting rates compared to base year. The reporting level might be affected due to more no. of notional facilities/ non-active/ temporary close facilities in reference year compared to base. In 2020, the Ministry of Health and Family welfare (MoHFW) eliminated the notional facilities reporting from HMIS. This shift to reporting only physical entities on HMIS Portal, may have affected the reporting. New facility created rate may have further impacted data reporting, with lesser chances of reporting by these facilities. Moreover, COVID-19 may have affected reporting, with numerous facilities deactivated, some upgraded or activated during this time. Majority of facilities have undergone mergers. Merging of facilities is predominantly observed at the SCs & PHCs level. Appendix table 1 provides detailed information on healthcare infrastructure changes in Haryana.

Table 1 Calculation of facility reporting rate in Haryana: Base Vs Reference Year

Facilities	2015-16 (BY)				2019-20 (RY)			
	Total Facility	Expected Reports	Actual Reports	FCRR	Total Facility	Expected Reports	Actual Reports	FCRR
SC	2844	34128	33747	99%	2875	34500	31997	93%
PHC	622	7464	7065	95%	700	8400	6361	76%
CHC	158	1896	1830	97%	149	1788	1704	95%
SDH	24	288	288	100%	25	300	264	88%
DH	27	324	324	100%	29	348	348	100%
Total	3675	44100	43254	98%	3778	45336	40674	90%
Reporting against Active Facilities								
SC	2812	33744	33747	100%	2798	33576	31997	95%
PHC	591	7092	7065	100%	622	7464	6361	85%
CHC	152	1824	1830	100%	140	1680	1704	101%
SDH	23	276	288	104%	23	276	264	96%
DH	27	324	324	100%	29	348	348	100%
Total	3605	43260	43254	100%	3612	43344	40674	94%

Source: Authors' calculation from HMIS standard Report, 2015-16 and 2019-20
*Expected no. of reports = Total no. of Health Facilities*12*
Actual no. of Reports= Sum of the all 12 months (April- March) Reporting
FCRR stands for Facility Completeness Reporting Rate

Table 2 examined district-wise facility reporting rates in Haryana. Only Panchkula district had a reporting rate below 75 percent in the reference year, compared to none in base year. Panchkula's completeness rate was noted more than 100 percent in base year, indicating double entry or possibility of any error. Faridabad, Fatehabad, Gurgaon, Kaithal, Karnal, Kurukshetra, Mahendragarh, and Sonipat achieved 100 percent facility reporting completeness in base year, while none of districts maintained this level in the reference year.

Completeness of Data Item

Figure 2 (A&B) displayed the completeness of data item for 2015-16 and 2019-20. Overall, Haryana performs better with regards to data item reporting compared to baseline periods.

Table 2 District-wise Completeness of Facility Reporting: 2015-16 vs 2019-20

State/ Districts	2015-16			2019-20		
	Expected Reports	Actual Reports	Facility Completeness Rate	Expected Reports	Actual Reports	Facility Completeness Rate
Haryana	44100	43254	98%	45684	40674	89%
Ambala	1860	1728	93%	1944	1680	86%
Bhiwani	3480	3462	99%	2364	2154	91%
Charkhi Dadri	-	-	-	1164	1107	95%
Faridabad	1536	1536	100%	1548	1265	82%
Fatehabad	1980	1980	100%	2136	1951	91%
Gurgaon	1872	1866	100%	1944	1464	75%
Hisar	3300	3234	98%	3288	3016	92%
Jhajjar	2316	2292	99%	2316	2124	92%
Jind	2544	2508	99%	2616	2412	92%
Kaithal	2172	2172	100%	2184	2085	95%
Karnal	2364	2364	100%	2400	2213	92%
Kurukshetra	1872	1872	100%	2016	1746	87%
Mahendragarh	2112	2112	100%	2112	1978	94%
Mewat	1344	1310	97%	1524	1419	93%
Palwal	1548	1524	98%	1596	1492	93%
Panchkula	1116	1140	102%	1224	888	73%
Panipat	1656	1500	91%	1728	1440	83%
Rewari	1764	1752	99%	1788	1657	93%
Rohtak	2088	1889	90%	2244	1863	83%
Sirsa	2460	2333	95%	2736	2399	88%
Sonipat	2760	2760	100%	2844	2603	92%
Yamuna Nagar	1956	1920	98%	1968	1718	87%

Results	**2015-16**	**2019-20**
State Level Facility reporting completeness rate	98%	89%
No (%) of district with facility Reporting completeness rate < 75%	0%	1 (5%)
75-85%	0%	4 (18%)
85-90%	1 (5%)	4 (18%)
With 100%	8 (36%)	0%
> 100%	1 (5%)	0%

*Source: HMIS portal *Total Facility (Public only) as of Feb 2016 #Total Facility as per May 2020*
*Expected no. of reports = Total no. of Health Facilities*12*
Actual no. of Reports= Sum of the all 12 months (April- March) Reporting

Data reporting completeness rate increased to 94 percent from 92 percent in the base year. Zero reporting decreased over time, indicates improvement in data reporting. Zero reporting in data elements is noted high with the 23 percent in Charkhi Dadri, requiring regular monitoring. Blank reporting observed more in Rewari and Karnal districts.

Dimension 2: Accuracy/ Statistical Outlier and Data Validation

Overall, the number of probable Outliers is more than the Validation errors in Haryana. Figure 3 depicts that validation errors has decreased from 31 to 28 compared to baseline period in Haryana. While increasing trends among Outliers in HMIS data can be observed during same time period. Of the total outliers, majority of them were found in three domains- Child Immunization, Childhood diseases, Laboratory tests. Majorly Validity Issues noted in immunization (OPV, BCG), New-born weighted at birth, no. of pregnant women given 100 IFA tablets/ 360 calcium, and AFHC. The maximum Validation error noted in Kaithal district, while the minimum was seen in Mewat (HMIS state Report – Validation and Outliers, 2019-20).

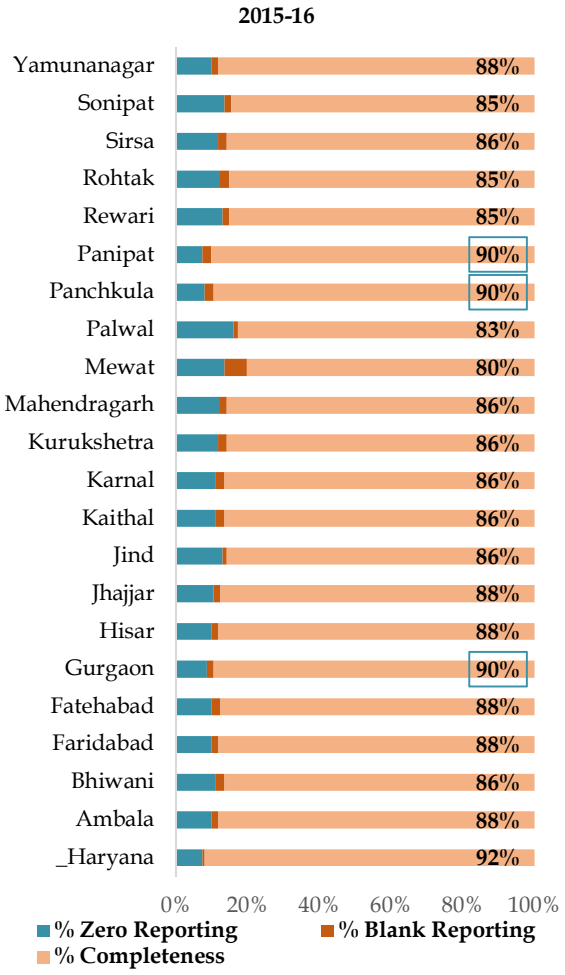

Figure 2 A

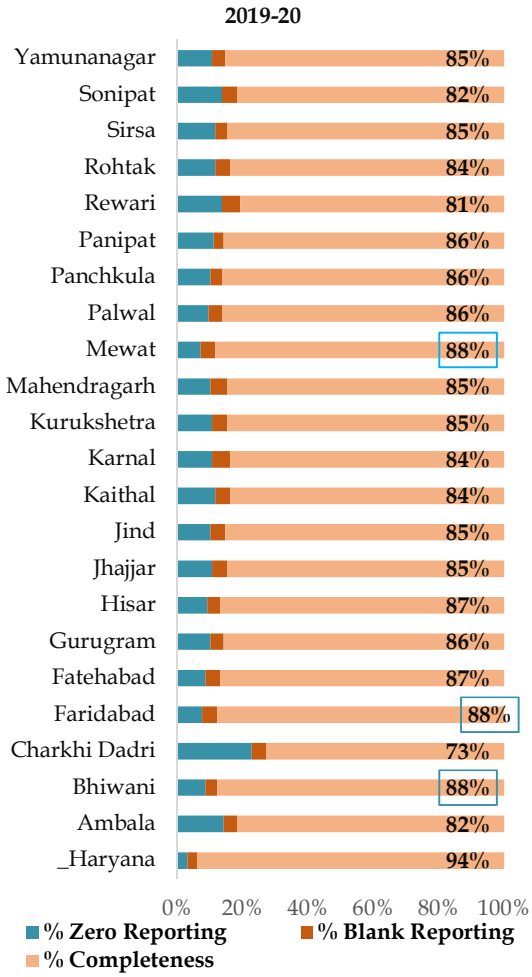

Figure 2 B

Figure 2 Data Completeness Rate, Zero and Blank Reporting against the Total Data Elements in a reporting in Haryana

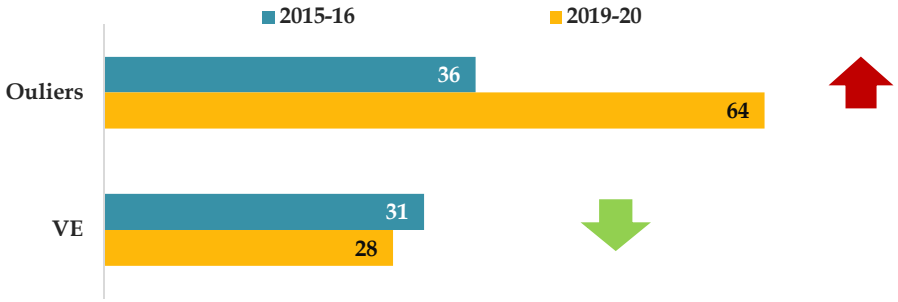

Figure 3 Trends in Outliers and Validation Error in Haryana (In No.)

Figure 4 depicts monthly data of probable outliers and validation error in Haryana over the years. Highest validation error is noted in April and October in both the years, indicate the need for staff training of HMIS data reporting format in the starting of financial year. The highest numbers of outliers noted in the month of March with share of more than 20 percent total outliers in state likely due to complete financial year data compilation. However, December consistently has the lowest outliers.

Dimension 3: Internal Consistency/ Temporal Consistency

To measure the overtime consistency, the most recent year (2019-20) compared with the mean value of MCH indicator for previous three years (2016-17, 2017-18, 2018-19). The temporal consistency of selected few indicators, was measured against state values of those indicators for each district.

Overall, ANC registration and no. of deliveries conducted seems more consistently than ANC1 and IFA tablet consumed by PW. Table 3 shows state ratio of 0.99, indicates that the ANC registration for the current year is 1 percent lesser than the mean of the past three years. Only Mewat district had an ANC registration ratio was higher than 33 percent of the state ratio.

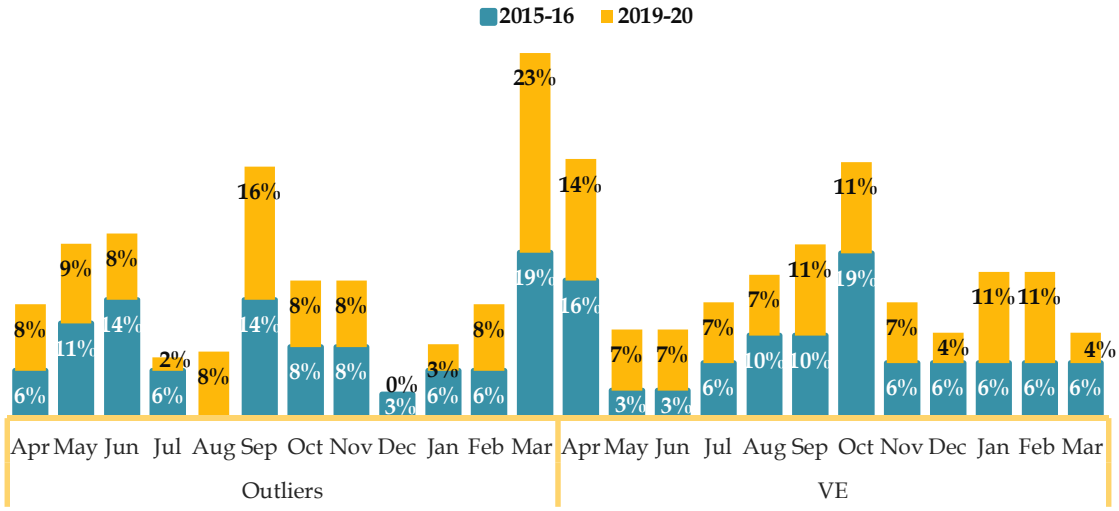

Figure 4 Monthly Trends in Outlier and Validation Error in Haryana

There is need to determine why driving this difference in Mewat. Fatehabad, Gurgaon, and Hisar are closely align with the state rate. ANC registration in the first trimester, showed a 7 percent greater than the mean of ANC 1 for the past three years. Dadri, and Mewat exceeds the state ratio with more than 33 percent. Faridabad, and Fatehabad are close to state ratio.

The state ratio of 1.19 indicates, IFA tablet provided to pregnant women for current year is 19 percent greater than the mean of the past three years. There are 3 districts- Charkhi Dadri, Hisar and Mewat had a ratio more than 33 percent of the state ratio. State rate is 119 percent- Fatehabad, Kurukshetra and Rohtak are close to this value, while Kaithal and Gurgaon also performing well relative to state level. The state ratio was 1.00 which means that the public deliveries for the current year is similar ratio than the mean of deliveries for the past three years. Thus, the data seems pretty good for deliveries conducted in the state as the data noticed without any fluctuations, except Charkhi Dadri District.

Overall, temporal consistency for selected MCH indicators reveal positive results establishing that consistency among data items across periods is maintained without any major fluctuations except Mewat and Charkhi Dadri. Additionally, Hisar district showed divergent score in IFA tablet consumed by PW.

Table 3 Temporal Consistency of selected Key HMIS Indicators for Haryana State

State/Districts	Current year/ Mean of last 3 years				> 33% Difference between the State and District Ratio			
	ANC Reg.	ANC 1	IFA Tablet	Deliveries	ANC Reg.	ANC 1st	IFA Tablet	Deliveries
Haryana	0.99	1.07	1.19	1.00				
Ambala	1.05	1.10	1.27	1.17	6%	3%	7%	17%
Bhiwani	0.82	0.95	1.30	0.80	17%	11%	9%	20%
Charkhi Dadri	1.30	1.50	1.71	1.74	31%	40%	44%	74%
Faridabad	1.02	1.07	1.07	0.89	3%	1%	10%	11%
Fatehabad	0.99	1.05	1.20	1.07	1%	2%	1%	7%
Gurgaon	0.97	1.05	1.14	1.05	2%	3%	4%	5%
Hisar	0.98	0.99	0.62	0.95	2%	8%	48%	5%
Jhajjar	0.94	0.95	1.06	0.92	5%	11%	11%	8%
Jind	0.97	0.98	1.42	0.92	3%	8%	19%	8%
Kaithal	0.89	0.96	1.15	0.88	10%	10%	3%	12%
Karnal	1.02	1.04	0.95	1.11	3%	3%	20%	11%
Kurukshetra	0.95	1.00	1.16	0.77	4%	7%	3%	23%
Mahendragarh	0.95	1.02	0.80	0.75	5%	5%	33%	25%
Mewat	1.40	2.46	2.13	1.17	41%	129%	79%	17%
Palwal	1.05	1.15	1.57	0.96	5%	8%	32%	4%
Panchkula	0.96	1.02	1.05	1.01	3%	5%	12%	1%
Panipat	0.81	0.89	1.07	0.97	18%	17%	10%	3%
Rewari	0.96	1.01	0.87	1.06	4%	6%	27%	6%
Rohtak	1.05	1.11	1.21	1.09	6%	3%	2%	9%
Sirsa	0.88	0.96	0.87	0.91	11%	10%	27%	9%
Sonipat	0.95	1.02	1.47	0.98	4%	5%	24%	2%
Yamuna Nagar	0.94	0.97	1.07	1.12	6%	9%	10%	12%

Source: HMIS standard Report, 2016-17, 2017-18, 2018-19 and 2019-20
>State Deviation (33 percent) considered to be Inconsistent data in the indicator over time

Dimension 4: External Consistency

External consistency refers to the level of agreement between two data sources measuring the same health indicator, often between an administrative dataset and a population-based sample survey for the same period (WHO, 2017). This approach was calculated by comparing independent external source of data i.e., NFHS-4 and 5 estimates with HMIS. Overall, external consistency was highly noted with the delivery care indicators in the state depicted in Table 4. On the other hand, when HMIS value compare with the household survey, result shows highly discrepancy in the intake of IFA pills/tablets by the PW for more than 180 days, which needs to be addressed.

Table 5 provides a comprehensive overview of district-wise external consistency in ANC 4, IFA tablet, Institutional births and C-section related indicators. In the base year, HMIS coverage rate has been recorded higher than NFHS for ANC4 by over 33 percent in 13 districts of Haryana. Only Yamunanagar has HMIS coverage rate equal to NFHS-5. While in the reference year, total 12 districts of Haryana recorded higher HMIS coverage rate with more than 33 percent difference- Ambala, Bhiwani, Fatehabad, Jhajjar, Jind, Karnal, Kurukshetra, Mahendragarh, Mewat, Palwal, Sirsa, Sonipat. None of the district have equal HMIS coverage rate of ANC 4 to the NFHS-5. With regards to IFA consumed by PW, the HMIS coverage rate recorded higher than NFHS with a difference of more than 33 percent in total 15 districts of Haryana in base year- Faridabad, Fatehabad, Gurgaon, Hisar, Jhajjar, Jind, Kurukshetra, Mahendragarh, Mewat, Palwal, Panipat, Rewari, Rohtak, Sirsa, Sonipat. In reference year, HMIS coverage rates were recorded higher in almost all the districts of Haryana for ANC 4 and IFA consumption by pregnant women. Notably, districts with significant difference (more than 33 percent) between HMIS and NFHS are flagged for attention. None of the district were observed in Haryana where NFHS coverage rate noted equal to NFHS-5.

In the base year, the HMIS coverage rate recorded higher than NFHS (with a difference of more than 33 percent) in Faridabad, Mewat, Palwal, Panipat districts of Haryana for Institutional deliveries at

Table 4 External Consistency in Haryana State: HMIS Vs NFHS across selected Indicators

Indicators	2015-16		2019-20	
	Ratio of HMIS/NFHS	Absolute % Diff.	Ratio of HMIS/ NFHS	Absolute % Diff.
Mothers who had ANC checkup in 1st Trimester	0.98	2%	0.89	11%
Mothers who had at least 4 ANC visits	1.47	47%	1.27	27%
PW who consumed IFA for 180 days or more	1.73	73%	2.32	132%
Institutional births	1.13	13%	1.01	1%
Institutional births in public facility	1.20	20%	1.02	2%
Births delivered by caesarean section	1.46	46%	0.97	3%
Births in a public health facility delivered by caesarean section	1.10	10%	1.14	14%
Mothers who received post-partum check-up from doctor/ Nurse/ ANM/ other	0.91	9%	1.03	3%

Source: HMIS standard Report, 2015-16 and 2019-20; NFHS 4 & 5
Note: Absolute % Difference indicates between HMIS & NFHS coverage rate

public health facilities (Table 5). Currently, the state ratio of 1.02 shows that the two denominator values are fairly similar to each other, with approximately 2 percent difference between them. When we compared the HMIS data with NFHS-5, only Palwal district of Haryana recorded higher HMIS coverage with more than 33 percent difference. None of the district had HMIS (2019-20) coverage rate equal to NFHS-5. The table 4 also shows the absolute percentage difference between HMIS and NFHS estimates for Institutional births as well as C-section from 2015-16 to 2019-20.

In the base year, the HMIS coverage rate exceeded NFHS-4 rate by over 33 percent in total 6 districts for C—Section at Public Facilities, namely Ambala, Fatehabad, Panchkula, Panipat, Rohtak and Yamuna Nagar (Table 6).

Kurukshetra is the only district which has HMIS coverage rate equal to NFHS-4 in C-sec delivery. In Jind and Palwal district, NFHS coverage rate noted higher than HMIS with more than 33 percent difference.

Table 5 District-wise external consistency in ANC & Delivery Care related indicators

State/Districts	2015-16 (HMIS/ NFHS ratio, Absolute Diff.)			2019-20 (HMIS/ NFHS ratio, Absolute Diff.)		
	ANC 4	IFA Tablets	Public Inst. births	ANC 4	IFA Tablets	Public Inst. Births
Haryana	1.47 (47%)	1.73 (73%)	1.20 (20%)	1.27 (27%)	2.32 (132%)	1.02 (2%)
Ambala	1.06 (6%)	1.33 (33%)	0.78 (22%)	1.57 (57%)	3.04 (204%)	1.18 (18%)
Bhiwani	-	-	-	1.52 (52%)	3.39 (239%)	0.87 (13%)
Charkhi Dadri	-	-	-	1.31 (31%)	2.41 (141%)	0.91 (9%)
Faridabad	1.71 (71%)	2.35 (135%)	1.57 (57%)	1.26 (26%)	2.18 (118%)	1.15 (15%)
Fatehabad	1.19 (19%)	1.34 (34%)	0.97 (3%)	1.37 (37%)	1.89 (89%)	1.05 (5%)
Gurgaon	1.69 (69%)	2.74 (174%)	1.30 (30%)	1.13 (13%)	3.92 (292%)	0.80 (20%)
Hisar	1.91 (91%)	1.88 (88%)	1.11 (11%)	1.09 (9%)	0.85 (15%)	0.85 (15%)
Jhajjar	1.34 (34%)	2.16 (116%)	1.11 (11%)	1.60 (60%)	2.40 (140%)	1.06 (6%)
Jind	1.33 (33%)	1.70 (70%)	0.91 (9%)	1.60 (60%)	3.13 (213%)	0.75 (25%)
Kaithal	1.29 (29%)	1.19 (19%)	1.07 (7%)	1.26 (26%)	1.95 (95%)	0.80 (20%)
Karnal	1.34 (34%)	0.88 (12%)	0.97 (3%)	1.65 (65%)	1.36 (36%)	1.02 (2%)
Kurukshetra	1.14 (14%)	1.38 (38%)	0.95 (5%)	1.36 (36%)	1.65 (65%)	0.67 (33%)
Mahendragarh	1.70 (70%)	2.05 (105%)	1.03 (3%)	1.73 (73%)	3.24 (224%)	1.04 (4%)
Mewat	5.78 (478%)	6.69 (569%)	4.22 (322%)	1.46 (46%)	12.56(1156%)	1.32 (32%)
Palwal	2.99 (199%)	6.69 (569%)	2.35 (135%)	1.63 (63%)	4.92 (392)	1.37 (37%)
Panchkula	0.79 (21%)	1.11 (11%)	1.04 (4%)	0.88 (12%)	1.54 (54%)	1.06 (6%)
Panipat	1.39 (39%)	1.40 (40%)	1.35 (35%)	1.30 (30%)	2.03 (103%)	1.19 (19%)
Rewari	2.60 (160%)	2.74 (174%)	0.97 (3%)	1.18 (18%)	2.06 (106%)	0.78 (22%)
Rohtak	1.57 (57%)	2.38 (138%)	0.97 (3%)	1.24 (24%)	2.55 (155%)	1.12 12%)
Sirsa	1.39 (39%)	1.64 (64%)	1.33 (33%)	1.35 (35%)	1.77 (77%)	0.86 (14%)
Sonipat	1.87 (87%)	1.79 (79%)	1.27 (27%)	1.34 (34%)	3.29 (229%)	1.06 (6%)
Yamunanagar	1.00 (0%)	0.98 (2%)	0.80 (20%)	1.20 (20%)	1.88 (88%)	0.96 (4%)

Results:	**2015-16**	**2019-20**
District with ANC 4 consistency ratio below 0.67 *	0	0
District with ANC 4 consistency ratio above 1.33**	13	12
District with ANC 4 consistency ratio is 1 ***	1	0
District with IFA tablet consistency ratio below 0.67*	0	0
District with IFA tablet consistency ratio above 1.33**	15	21
District with IFA tablet consistency ratio is 1 ***	0	0
District with Institutional births at public consistency ratio below 0.67 *	0	0
District with Institutional births at public consistency ratio above 1.33**	4	1
District with Institutional births at Public Facility consistency ratio is 1 ***	0	0

Source: HMIS standard Report, 2015-16 and 2019-20; NFHS-4 and NFHS-5

NFHS Coverage rate higher; **HMIS Coverage rate higher; ***HMIS Coverage rate= NFHS Coverage

Table 6 External consistency in C-section births at Public Facility in Haryana

State/Districts	2015-16 Ratio of HMIS/ NFHS-4		2019-20 Ratio of HMIS/ NFHS-5	
	C-Sec at Public	Absolute % Diff.	C-Sec at Public	Absolute % Diff.
Haryana	1.10	10%	1.14	14%
Ambala	1.81	81%	0.86	14%
Bhiwani	-	-	1.04	4%
Charkhi Dadri	-	-	0.07	7%
Faridabad	0.69	31%	1.39	39%
Fatehabad	1.80	80%	1.84	84%
Gurgaon	1.09	9%	1.22	22%
Hisar	1.06	6%	1.16	16%
Jhajjar	1.23	23%	0.66	34%
Jind	0.56	44%	0.62	38%
Kaithal	1.27	27%	0.84	16%
Karnal	0.88	12%	2.28	128%
Kurukshetra	1.00	0%	1.04	4%
Mahendragarh	1.24	24%	1.25	25%
Mewat	0.83	17%	1.29	29%
Palwal	0.00	-	1.09	9%
Panchkula	1.86	86%	0.81	19%
Panipat	1.49	49%	0.40	60%
Rewari	1.07	7%	1.10	10%
Rohtak	2.31	131%	2.09	109%
Sirsa	0.80	20%	1.09	9%
Sonipat	1.03	3%	1.33	33%
Yamunanagar	1.47	47%	1.05	5%
Results:			**2015-16**	**2019-20**
District with C-section at Public consistency ratio below 0.67*			2	4
District with C-section at Public consistency ratio above 1.33**			6	4
District with C-section at Public Facility consistency ratio is 1 ***			1	0

Source: HMIS standard Report, 2015-16 and 2019-20; NFHS-4 and NFHS-5
NFHS Coverage rate higher ** HMIS Coverage rate higher; * HMIS Coverage rate= NFHS Coverage*

In comparison with NFHS-5, 4 districts of Haryana had higher HMIS coverage rate by over 33 percent, namely, Faridabad, Fatehabad, Karnal and Rohtak. Charkhi Dadri, Jhajjar, Jind and Panipat district in Haryana where NFHS coverage rate noted higher than the HMIS by over 33 percent. No district has the equal HMIS coverage rate to the NFHS-5 in C-section delivery conducted at public institutions.

Discussion

The study found significant improvements in the data quality in Haryana, in the domains of Completeness of Data item, Validation Error and External consistency (ANC & Institutional deliveries). A study in Karnataka (2016), found remarkable improvement in HMIS data coverage and quality over the study period, 2012-13 & 2013-14 (James, K. S. and Prabhuswamy, P., 2016). Existing literature documented data quality may be improved significantly due to extensive used of HMIS data in State Health Index by NITI Aayog since its inception in 2017-18, HMIS augmentation, revamping and development of new HMIS, physical verification exercise of HMIS data by PRCs etc (Aayog, N., 2020 & 2021). Maiga, A. et al. showed, with the introduction of web-based digital platforms for health facility data analysis at district level leads to gradual

improvements in data quality. HMIS 2.0 was develop in December 2020 with the WHO support, featuring - consistency checks are inbuilt for data quality, such as compare option, inter validation checks, random checks of register (HMIS Annual report, 2020-21 & 2021-22).

The findings of Maiga, A. et al., (2019) revealed high levels of completeness of facility reporting, similarly our study also showed high reporting rate at DH and CHC level. However, our findings not corroborated completely with this. The present study results revealed significant gap- completeness of facility reporting rate in HMIS data, particularly for PHCs, across all districts in Haryana. Completeness of reporting rate somewhere influenced by changes such as mergers, upgrades, renaming and activations in healthcare infrastructure, highlights the potential impacts of these modifications on overall reporting integrity.

The research conducted by Sharma, Atul et al., (2016) in Haryana revealed high levels of completeness of reporting except for contraception and vaccine administration. In line with these findings, our analysis indicates Haryana performs better in data item reporting, however face challenges like blanks reporting in Rewari and Karnal, and a high percentage of zero reporting in data elements in Charkhi Dadri district, indicating a need for regular monitoring. The study by James, K S and others, provides a useful framework for understanding the types of errors encountered in reporting- procedural, lack of clarity, under-standing and over/ under reporting (James, K. S. and Prabhuswamy, P., 2016). These errors contribute to inaccuracies in reporting.

Our study findings corroborate with previous observations regarding persistent data quality issues, particularly the presence of extreme outliers identified in the 2019 study (Maiga, A. et al., 2019). The state experienced a substantial increase (approximately 78%) in probable outliers, from 36 in base to 64 in reference year. The highest outliers noted in the last month of the financial year i.e., March with the share of around 20 percent of total outliers in state. Trend aligns with the WHO study conducted in 2003 examined the consistency and reliability of the reported values over the period 1991 to 1996, which identified 30% of the reported values as "outliers". The analysis further reveals that despite a reduction in total validation errors, challenges persist in Haryana, with Validation errors predominantly occurring in April, the financial year's outset. Majorly Validity Issues noted in Immunization (OPV, BCG), New-born Weighted at Birth, No. of PW given 100 IFA tablet/ 360 Calcium, and AFHC.

A 2019 study noted inconsistency in reported data over time and between indicators (Maiga, A. et al., 2019). However, our findings on temporal consistency for MCH indictors demonstrate that consistency among data items across periods is maintained without any major fluctuations, except Mewat and Charkhi Dadri districts. ANC registration and deliveries showed more consistency than ANC registration in first trimester and IFA tablet consumed by pregnant women. External consistency was highly noted in delivery care indicators in the state; however, discrepancies were found regarding the intake of IFA pills/tablets by PW for more than 100 days, requiring attention.

Studies have shown that digital platforms play a crucial role in enhancing data quality (Jha, U. M. and Arora, R., 2022; Maiga, A. et

al., 2019). Training and support supervision of HMIS focal persons is required to strengthen quality assurance of HMIS (Kagoya, H. R., & Kibuule, D., 2018). Addressing HMIS data quality issues requires regular monitoring and validation of data by state/ district officials, training program in the beginning as well as outset of financial year, targeted interventions with validity issues.

There are few limitations to our study. Long term trends of HMIS data quality might not be visible in our study which covers a short span of time. This study captures the HMIS data quality issues solely on four key dimensions i.e., Data completeness, Accuracy, Internal and External consistency. In the present study, district-level analysis of VE and outliers was not included. The present analysis binds the results comprehensively, and it was felt that a district-level purview of VE and outliers – since it's a lengthy analysis – can be taken up in a follow-up study in the future. The strength of Our study is elevated by the inclusion of insights from expert perspectives i.e., specifically from a state level MIS Expert. This approach significantly enhances the clarity and depth of our research findings.

Conclusion and Policy Implications

The study was taken up with the aim to assess the status of data reporting and data quality of key HMIS health service delivery indicators in Haryana State. The study found significant improvements in the data quality of Haryana, in the domains of Completeness of Data item, Validation Error and External consistency (ANC & Institutional deliveries). The completeness of facility reporting rate noted high at upper health tier. However, result from analysis crystal clear a decline in completeness rate of health facility reporting in Haryana, especially for PHCs, compared to the baseline period. Reporting of blanks in the data reporting format observed more in Rewari and Karnal district as compared to other districts in state. Zero reporting in the data elements is noted high in the Charkhi Dadri district, which needs to be seriously addressed. Overall result call for the necessity to improve completeness of their reporting through the system in order to improve HMIS data quality, which is critical to monitor the Sustainable Development Goals (SDGs). There are also needs to monitor the zero reporting and reporting of blanks in the data reporting format. Data entry is typically seen as low operational priority. Incentivize quality data reporting may encourage the DEOs to prioritize data reporting and to improve the overall quality of health data in the state.

The study asserts that the state has a definite scope for improvement across all districts with respect to outliers as well as validation errors in HMIS data. The finding reveals increasing trends in outliers, majorly in child immunization, childhood diseases and laboratory tests. The highest number of outliers was occurred in March, while the validation error was predominantly appeared in April, the financial year's outset. To address the outliers and validity issue, more training and capacity building program in the beginning as well as outset of the financial year should be conducted for Data Entry Operators (DEOs) to enhance their data reporting skill. The state government should focus on reducing the number of outliers in HMIS data, especially in the areas of Child Immunization, Childhood diseases, and Laboratory tests. Further, there is need to be targeted interventions for areas with validity issues,

such as Immunization (OPV, BCG), New-born Weighted at Birth, No. of pregnant women given 100 IFA tablet/ 360 Calcium, and AFHC.

Temporal/ Internal consistency for selected indictors (ANC registration and No. of deliveries) reveal positive results, it is maintained without any major fluctuations across all districts except Mewat and Charkhi Dadri. IFA tablet consumed by pregnant women is noticed greater differences in Charkhi Dadri, Hisar and Mewat district from the state proportion. There is need to determine why deliveries data for the Charkhi Dadri district are notably higher. External consistency approach showed commendable level of consistency in the delivery care indicators. HMIS data is completely reliable for the Institutional births at public facility and C-section deliveries. However, concerning discrepancy noted in the intake of IFA pills by pregnant women across all districts except Hisar. The findings suggest that state should address the discrepancy in said indicator, to ensure the external consistency in HMIS data.

To sum up, there has been an overall improvement in the quality of HMIS data in certain domains in Haryana except few districts. The persistent data quality issues included- presence of extreme outliers, lack of external consistency in IFA supplement intake by PW. To address these issues, policy suggestion includes regular monitoring & validation of data by state/ district officials, training/ capacity building program, targeted interventions with validity issues, incentivizing data quality reporting etc. Furthermore, implement robust measures, including standardized reporting protocols, to ensure that changes in healthcare infrastructure, such as mergers, upgrades, and renaming do not adversely impact data reporting. Overall, the state official should have regular interaction with the targeted districts team, which are not performing well in-align in HMIS data quality.

References

Aayog, N. (2020) Healthy States Progressive India: Report on the Ranks of States and Union Territories, Health Index Round IV, 2019–20 NITI Aayog. Government of India, New Delhi.

Aayog, N. (2021) Best practices in the performance of District Hospitals. NITTI Aayog.http://www.planningcommission.gov.in/sites/default/files/2023 03/District_Hospital_Report_for_digital_publication.pdf

Dehury, R. K., & Chatterjee, S. C. (2018). Assessment of health management information system for monitoring of maternal health in Jaleswar Block of Balasore District, Odisha, India. *Indian Journal of Public Health*, *62*(4), 259.

Githinji, S. et al. (2017). Completeness of malaria indicator data reporting via the District Health Information Software 2 in Kenya, 2011–2015. *Malaria journal*, *16*(1), 1-11.

https://apps.who.int/iris/handle/10665/259226

James, K S and Prabhuswamy, P (2016). Quality of HMIS data at the district level in Karnataka, Compendium of studies conducted by PRCs, ISEC, Bengaluru.

Jha, Ugra Mohan and Arora, Ritu (2022). Health Management Information System: A Journey towards Public Health Data Ecosystem in India, Compendium of studies conducted by PRCs, Gokhale Institute, Pune, ISBN 978-81-85835-95-0, 241-249.

Kagoya, H. R., & Kibuule, D. (2018). Quality assurance of health management information system in Kayunga district, Uganda. *African Evaluation Journal*, *6*(2), 1-11.

Krishnan et al. (2010). Evaluation of Computerized Health Management

Information System for Primary Health care in Rural India. *BMC Health Services Research,* 10 (310), 1-13.

Maiga, A. et al. (2019). Generating statistics from health facility data: the state of routine health information systems in eastern and southern Africa. *BMJ Global Health*, *4*(5), e001849.

Ministry of Health & Family Welfare (2019). HMIS Analytical report, MoHFW, GOI.

Ministry of Health & Family Welfare (2020-21). HMIS Annual report, MoHFW, GOI.

Ministry of Health & Family Welfare (2021-22). HMIS Annual report, MoHFW, GOI.

Ojha, Dheeraj Kumar (2023). Health Management Information System: A Journey of digital health in India. *Demography India*, ISSN 0970-454X, 52 (2), 1-6.

Sharma, A., Rana, S. K., Prinja, S., & Kumar, R. (2016). Quality of health management information system for maternal & child health care in Haryana state, India. *PLoS One*, *11*(2), e0148449.

Teklegiorgis, K., Tadesse, K., Terefe, W., & Mirutse, G. (2016). Level of data quality from Health Management Information Systems in a resource limited setting and its associated factors, eastern Ethiopia. *South African Journal of Information Management*, 18(1), 1-8.

World Health Organization. (2003). *The immunization data quality audit (DQA) procedure* (No. WHO/V&B/03.19). World Health Organization.

World Health Organization. (2017). Data quality review: module 1: framework and metrics.

World Health Organization. (2017). Data quality review: module 2: desk review of data quality. https://apps.who.int/iris/handle/10665/259225

World Health Organization. (2017). Data quality review: module 3: data verification and system assessment.

Appendix

Table 1 Healthcare infrastructure Transformation in Haryana (2019-20)

Facilities		2019*	2020 *	Facilities Created/Activated/ Upgraded/ Deactivate/ Merge/ Down-grade
Total Facilities	SC	2875	2886	11
	PHC	700	715	15
	CHC	149	152	3
	SDH	25	25	0
	DH	29	29	0
	Total	3778	3807	29
Total Active Facilities	SC	2798	2664	-134
	PHC	622	526	-96
	CHC	140	142	2
	SDH	23	22	-1
	DH	29	29	0
	Total	3612	3383	229

**Shows facilities as on April 2019 & May 2020, HMIS # +Ve number indicate additions to the facilities, whereas -Ve numbers indicate a reduction in number of facilities.*