A Report of Jan Swasthya Abhiyan online Panel Discussion: Towards transparent health data ecosystem; Challenges and Prospects in India (9th November 2024)

This report summarises the discussions from the workshop on 'Creating a Transparent Health Data Ecosystem in India,' organised by Jan Swasthya Abhiyan (JSA). The event addressed critical challenges in India's health data landscape, including data integrity, accessibility, and its implications for public health policymaking.

The workshop highlighted key issues such as delays in the Census, diminishing autonomy of data institutions, suppression of independent research, privacy concerns surrounding digital health initiatives, and barriers to accessing government-collected data.

Featuring expert presentations by **Professor K.S. James**[1], **Ms Rema Nagrajan**[2], and **Professor Girija Vaidyanathan**[3], the workshop explored solutions such as decentralising data systems, enhancing data transparency, improving routine health data collection, and fostering collaboration between stakeholders. These discussions aim to inform actionable strategies for building a robust and transparent health data ecosystem in India.

Pre-workshop JSA Note

India's health data ecosystem is facing multiple challenges that threaten the integrity, quality, and accessibility of vital information needed for effective policymaking and health interventions. Though non exhaustive below are some the key issues that need urgent attention of Jan Swasthya Abhiyan Constituents.

a. Concerns about Health Data in India

- ***Statutory institutions like the Census:*** The decennial Census, which is a critical source of population information, has not been conducted for over three years. This delay disrupts many essential planning and development activities, as the Census provides the foundational data for other surveys, including those related to health and the economy. Without current Census data, many policies and programmes, particularly for urban and rural populations, are being developed using outdated 2011 data, which fails to reflect significant changes in demographics and living conditions over the past decade.

---

[1] Prof. K.S. James is the senior visiting scholar at Newcomb Institute, Tulane University, USA, Ex-Director IIPS Mumbai.

[2] Ms. Rema Nagrajan is a renowned journalist and data expert.

[3] Prof. Girija Vaidyanathan is a former chief secretary, Govt of Tamil Nadu, and Professor of practice, humanities and social sciences dept, IIT- Chennai.

- ***Autonomy of data-gathering institutions:*** Institutions such as the National Family Health Survey (NFHS) and National Sample Survey (NSS), which provide crucial health and social data, are facing reduced independence. These organisations, built through decades of collaboration and expertise, are now subject to ***government*** interference, potentially compromising the reliability and credibility of the data they produce. This undermines trust in the data used for health planning, policy development, and academic research.

b. Suppression of Independent Research

- ***Dismissal of independent studies:*** Research findings from reputable organisations such as the World Health Organisation (WHO), and independent researchers both ***within*** India and abroad, have been dismissed or refuted by the government when the findings contradict official narratives or portray the government in an unfavourable light. This suppression of evidence-based findings undermines public trust and can lead to policy decisions that are not grounded in reality.

- ***Suppression of researchers:*** There have been reported instances where researchers presenting critical findings, especially those challenging government claims, have faced backlash. This creates an atmosphere where academics and ***independent*** researchers may be discouraged from pursuing or publishing studies that could be deemed controversial, thus limiting open scientific discourse and inquiry.

c. Atal Bihari Vajpayee Digital Mission (ABDP)

- ***Creation of a national health database:*** The ABDP seeks to create a comprehensive health database for every Indian citizen, health provider, and institution. Whilst this ambitious project could significantly improve healthcare delivery and ***planning***, it raises serious concerns regarding data privacy, security, and the potential misuse of sensitive personal health information. Without adequate safeguards, there is a risk of unauthorised access, misuse by private entities, or even government surveillance.

- ***Concerns over data safety:*** There are fears that the collected health data could be used for purposes beyond healthcare, such as profiling or discrimination, without individuals' proper consent. This has raised alarms among privacy advocates and ***civil*** society organisations, especially in a country like India, where data protection laws are still developing.

d. Diminishing Space for Independent Inquiry

- ***Reduced freedom for academia and civil society:*** The space for independent, critical thinking and honest scientific inquiry is shrinking in India. Researchers, universities, and civil society organisations are facing increasing challenges when their work presents an alternative view or criticises government actions. This restriction affects the quality of public discourse and stifles innovation in addressing the country's health and social challenges.

- ***Impact on policy and ground realities:*** With restrictions on independent inquiry, the government may not receive accurate feedback about real-world conditions. As a result, policies may be formed without a comprehensive understanding of ground-level challenges, potentially leading to ineffective or misaligned health interventions.

e. Impact on Evidence-Based Policymaking

- ***Delay in the Census:*** India remains one of the few countries unable to conduct its decennial Census on time, delaying crucial population and demographic data. This delay impacts all data-dependent sectors, including health, as other surveys and planning processes rely on updated Census data for accurate sampling frames and needs assessments.

- ***Outdated data affecting policy:*** Current programmes for the urban poor, for example, are based on the rural and urban population ratios from the 2011 Census. This does not reflect the significant urbanisation and rural-to-urban migration that has occurred over the past decade, leading to policies that are increasingly disconnected from present realities.

f. Censorship of Data Institutions

- ***Interference with the National Sample Survey Office (NSSO):*** The National Sample Survey Office, once highly respected for its primary surveys, has faced censorship from the government, particularly when data does not align with favourable government narratives. A notable example is the withdrawal of the 76th round of the NSS on consumption expenditure, reportedly because the results were unfavourable to the government. Such actions damage the integrity of India's statistical institutions.

- ***Reduced credibility of periodic surveys:*** With increasing interference in data collection and reporting, the credibility of India's premier data institutions is eroding. This leads to doubts about the reliability of health and economic data, both nationally and internationally, affecting the formulation of health policies and programmes.

g. Government's Control Over the NFHS

- ***Shift from independent international oversight:*** The National Family Health Survey (NFHS), previously overseen by the internationally reputable Demographic and Health Survey (DHS), is now under greater government control. This shift raises concerns about potential data manipulation or bias, as the NFHS provides vital health indicators used by policymakers and international bodies to assess the health status of India's population.

h. Lack of Improvement in Routine Health Data

- ***Weak systems for recording health data:*** Routine health data systems, such as the Health Management Information System (HMIS), are of poor quality, and the government has made limited efforts to improve them. During critical times, such as the COVID-19 pandemic, essential systems like the Integrated Disease Surveillance Project (IDSP) were not fully operational, exposing the gaps in data management.

- ***Inconsistent data on key diseases:*** Important data on diseases such as tuberculosis, dengue, and malaria are not consistently or accurately recorded, hampering efforts to monitor and control these diseases. In many states, even basic birth and death registration is far from universal, complicating health planning and policy formulation.

i. Underutilisation of Data

- ***Government data not used for policymaking:*** Despite significant resources being spent on data collection, government data is often not effectively utilised for policymaking due to known but unaddressed issues with data quality. This undermines the utility of the data in shaping health policies and actions. Frequent changes in data systems: Software and data portals are frequently altered, making it difficult for officials at the ground level to keep up with updates. Furthermore, feedback from these officials is rarely sought or utilised to improve the systems, resulting in persistent inefficiencies and data collection issues.

j. Lack of Access to Research Data

- ***Barriers to independent research:*** Increasing obstacles are being placed on researchers seeking access to government data for independent studies. The data collected by the National Health Authority, for example, remains inaccessible to the research community, limiting opportunities for independent analysis that could inform better health policies and programmes. Non-Compliance with the Clinical Establishments Act Limited registration and reporting by private facilities: Despite the enactment of the Clinical Establishments Act in 2010, many states have not ratified it. As a result, the majority of private healthcare facilities remain unregistered, and there is minimal compliance with mandatory data reporting requirements. This means that crucial health data from private providers, who serve a large portion of the population, is absent from national health datasets. This lack of comprehensive data hinders effective health policy planning and monitoring. These issues collectively present serious challenges to the creation of an accurate, transparent, and accountable health data ecosystem in India.

To discuss all the above issues, Jan Swasthya Abhiyan successfully organised a special workshop and panel discussion to address the critical challenges facing India's health data architecture. In light of the growing concerns around data integrity, privacy, and the suppression of independent research, the workshop brought together some of the most distinguished voices in the field to discuss actionable strategies for moving forward.

Summary of the discussion-

1. Prof. K S James

- ***Importance of regular data collection:*** He highlighted the need for regular health data collection, suggesting that certain areas may require monthly data updates to effectively address developmental goals and combat data deprivation.

- ***Significance of census and survey data:*** Prof. James stressed that census data is vital for understanding local needs and assessing development programs, providing essential contextual information. But, pointed out that India's reliance on survey data, especially for health metrics, is excessive and often suffers from quality issues, with inconsistencies across states affecting reliability.

- ***Strengthening administrative data systems:*** There is a strong advocacy for enhancing administrative data systems as they are more cost-effective and can provide continuous updates compared to sporadic surveys. Public access to this data is crucial for transparency and improvement through public scrutiny.

- ***Decentralization of data systems:*** Prof. James argued for the decentralization of data collection, enabling local governments to take ownership and ensure data quality at state and district levels. This would enhance both relevance and reliability of the data collected.

- ***Data quality and integrity:*** He emphasized the importance of maintaining data integrity through consistent checks, ensuring that both survey and administrative data align with Sustainable Development Goals. The need for rigorous processes for data cleaning, validation, and ethical considerations in data collection was also highlighted.

- ***Public domain access to data:*** Advocating for making data publicly accessible, he noted that this fosters transparency, accountability, and critical evaluation, which can lead to improved data quality and informed public discourse.

- ***Contextualized data needs:*** Prof. James called for an emphasis on understanding the context in which data is collected, advocating for disaggregated data to identify disparities effectively and target interventions relevant to specific populations and geographic areas.

2. Ms. Rema Nagrajan

- ***Challenges in accessing government data:*** Ms. Nagrajan highlights the difficulties journalists face in accessing and interpreting government data due to a lack of expertise and increasing government restrictions on experts sharing information with the media.

- ***Climate of fear:*** She notes a growing climate of fear among researchers and analysts, who are discouraged from engaging with the media due to government crackdowns on independent analysis. This is critical as often journalists rely on external experts to understand the data.

- ***Limited public access:*** While the government collects extensive data on citizens, it provides little access to useful government data for the public, even on fundamental resources like doctor or hospital registries, which impedes public accountability and safety.

- ***Data dumping issues:*** The government releases large amounts of data without organizing it in ways that allow for meaningful comparisons or accountability checks. Dashboards often lack historical data, and data inconsistencies make it challenging for journalists to hold the government accountable.

- ***Bureaucratic barriers:*** Journalists encounter bureaucratic obstacles when seeking data through the Right to Information Act, often receiving responses that deny access or state that data is unavailable.

- ***Inconsistent medical data:*** Poor administration of medical data, such as death certifications, limits accurate health statistics and public health accountability.

- ***Frequent website changes:*** Government websites frequently change without notice, complicating the process of locating data and limiting public access.

- ***Lack of Accountability in Reporting Standards:*** Changes to data collection methods (e.g., NFHS no longer including malnutrition surveys) are sometimes made without explanation, reducing the ability to compare new and past data, which undermines government accountability.

- ***Delayed accountability:*** Many government targets are set for distant future dates, reducing immediate accountability and delaying progress on pressing issues.

3. Prof. Girija Vaidyanathan

- ***Perspective on Health Data in Policy:*** Prof. Vaidyanathan, drawing on her experience as a policymaker, discussed challenges in how health data is perceived and used by policymakers. She emphasized that the bureaucratic approach often doesn’t align with the ideal policy cycle (problem identification, evidence review, solution design, implementation, and evaluation).

- ***Limited access:*** Access to quality health data is restricted, even within governmental departments, leading to centralized data that is not effectively shared with field functionaries who collect it.

- ***Quality improvement:*** She emphasized that sharing data with those who generate it can enhance its quality, citing successful initiatives in Tamil Nadu where feedback mechanisms improved data accuracy. Under the DANIDA Tamil Nadu Area Health Care Project, Primary Health Centre staff used an early feedback loop that could review and discuss data, significantly improving data accuracy.

- ***Incentive misalignment:*** Field-level staff often face punitive measures for negative data outcomes, resulting in data distortion, e.g., clustering of maternal BP readings within “safe” ranges to avoid scrutiny. Aligning incentives to encourage accurate data reporting and providing constructive feedback can enhance data quality.
- ***Bureaucratic constraints:*** Short tenures: Policymakers often do not remain long enough in their positions to see through, affecting continuity in decision-making.
- ***Frequent changes in platforms:*** Constant updates to data platforms disrupt continuity and make it challenging to utilize historical data effectively State-level innovations challenges: State-level health data initiatives are hampered by bureaucratic reluctance to adopt central or locally relevant data initiatives.
- ***Localized data collection:*** There is a strong need for state-level governments to conduct their own studies (e.g., morbidity studies) to gather district-specific insights instead of relying solely on national-level data.

Suggestions for improvement-

- ***Feedback mechanisms:*** Establishing continuous feedback loops for data collectors can enhance accountability and quality.
- ***Aligning incentives:*** Creating a supportive environment for accurate reporting and error acknowledgment can prevent manipulation of data.
- ***Empowering local initiatives:*** Encouraging district-level data collection initiatives can better address local health needs, drawing on successful examples from Tamil Nadu.

Comments, question and answer session-

Key Questions and Comments

1. **Incorporating Nutrition and Coverage Data**

   A participant highlighted the need to include data from the National Nutrition Bureau and UNICEF's coverage evaluation surveys in policy analysis. The example shared underscored how evidence-based interventions can drive policy changes, while also advocating for flexibility in implementation to account for evolving needs. This was presented as a comment without a direct response.

2. **Discrepancies in Insurance Data**

   Concerns were raised regarding inconsistencies in data reported by the Employees' State Insurance Corporation, particularly on pneumoconiosis cases among family members with no direct exposure. This observation was shared as a comment.

3. **Data on Substance Use and Drug Abuse**

   Another participant highlighted the absence of national and state-level data on substance use and drug abuse, emphasising the significant impact this has on younger generations. This question was later addressed along with another.

4. **Comparability of Data Frameworks**

   A participant inquired about the impact of changes in data collection frameworks on the comparability of data over time, noting that this could pose a significant challenge for statisticians and demographers.

Responses from Panellists-

- **On Substance Use and Comparability**

  Prof. K.S. James explained that certain traditional datasets, including those by UNICEF, are no longer publicly accessible without government permissions. He acknowledged the limitations in data quality related to substance use, which often surfaces only through criminal records, highlighting the need for better mechanisms to capture such data.

- **On NSSO's Focus on Health Data**

  A participant questioned the National Sample Survey Office's limited focus on health data. Prof. James responded by tracing the historical collection schedule of health-related data, which initially occurred every 10 years but was later adjusted to a five-year cycle. He noted ongoing debates regarding the inclusion of data such as out-of-pocket expenditure in surveys like the NFHS but identified logistical challenges as a barrier.

  Prof. Girija Vaidyanathan supplemented this response, recommending the use of alternative data sources, such as MCC data or state-level studies similar to NSSO surveys. She emphasised that states could collect health utilisation data at a relatively low cost, thereby addressing gaps in national datasets.

- **Challenges Faced by Researchers**

  Participants voiced concerns about difficulties faced by young researchers, particularly due to non-cooperation from official bodies and ambiguities in permissions. Prof. Girija suggested fostering partnerships, such as Memorandums of Understanding (MOUs) with state governments, as a solution. She stressed the importance of trust-building to ensure smoother field access for researchers.

- **On Manipulated Data from District Hospitals**

  Another query pertained to data manipulation at district hospitals and the challenges this posed for public university researchers. Prof. Girija acknowledged these difficulties, advocating for a case-by-case approach to resolve such issues. She highlighted the value of webinars like this as platforms for raising awareness and seeking solutions.

Concluding Observations-

- **Ms. Rema Nagrajan** called for a reciprocal approach to data sharing. She emphasised that if governments expect data contributions from the public, they should also reciprocate by making government data more accessible, fostering accountability and transparency.

- **Prof. K.S. James** reiterated the importance of integrating data transparency into everyday practices, not just within government operations but also among academics and researchers. He urged universities and scholars to make their data publicly available, fostering a culture of openness and shared responsibility.

- **Prof. Girija Vaidyanathan** emphasised that the challenges within the health data ecosystem cannot be addressed solely by the government. She highlighted the defensiveness surrounding data sharing, stemming from perceptions of vulnerability. Prof. Girija advocated for shifting this mindset and promoting a collaborative approach, where data challenges are seen as shared responsibilities rather than solely governmental shortcomings.

**