



Editorial

Demographic and health surveys and its quality in India[☆]

Globally, national surveys play vital roles to provide required data for planning and formulating policy at national and sub-national levels and among these, health surveys play a major role in understanding health scenario and shaping the health programmes and policies in the country. In India, surveys are undertaken regularly and two of such large-scale health surveys in India are National Family Health Surveys (IIPS and MoHFW, 1993–2022)), Longitudinal Survey of Ageing in India (LASI: IIPS, NPHCE, MoHFW, HSPH, and USC, 2020) – both carried out by the International Institute for Population Sciences (IIPS), Mumbai. In addition, National Sample Survey Organisation (NSSO), in some of its rounds, also provide data on morbidity and health expenditure (Ministry of Statistics and Programme Implementation (MoSPI), 2019).

In India, with ever increasing demand for data at sub-state units as well as for various sections and sub-sections of the population, the sample size of the surveys has been increasing in every round, to provide critical health data for local-level planning. The length of the interview tool has also increased with each subsequent rounds of the survey. With the increase in the sample size and lengthening of the questionnaire, ensuring high-quality data remains a major challenge and requires constant feedback from research and analysis (James & Rajan, 2004; Rajan & James, 2004).

Various measures are often taken to ensure high quality standards of surveys as they adopt various strategies to address issues to minimise the non-sampling errors – a major source of errors in scientifically designed surveys. Standardised recruitment and training, uniform field work implementation protocols, technology-intensive tools to monitor and supervise the fieldwork are the examples of ways surveys attempt to control non-sampling errors. Still, errors remain, and researchers around the world have raised the issues related to the data quality of large-scale surveys from time to time.

The data quality issues pertaining to the large-scale surveys can be handled by prior understanding of the survey procedure and quality of data in similar surveys (or that of earlier round of the same survey). It is thus imperative to query the quality issues of the existing survey data to inform future surveys. There is a need to understand the quality of the information gathered from the different surveys to identify the areas that are vulnerable to errors and necessitate higher attention and caution. Regular research studies are also necessary to understand the mechanisms that work in a specific context to obtain high quality data from the field. These will act as feedback mechanisms on various aspects of the survey – recruitment, training, instruments, strategies of data

collection, and monitoring mechanisms. A set of guidelines on how to ensure data quality in large scale surveys can be found in ICMR-NIMS, 2021.

The National Data Quality Forum (NDQF: <https://ndqf.in>), a collaborative platform led by the Indian Council for Medical Research-National Institute of Medical Statistics (ICMR-NIMS) and Population Council, India was established to help improve the quality of the health and demographic data ecosystem in India that includes both survey and administrative data. The Forum specifically aims to educate and deepen interest among producers and users of data on quality and congregate them to a single platform to facilitate information sharing as well as equip them with appropriate solutions, guidance, strategies, and tools to contribute to the improvement in data quality. The NDQF has been supporting research on data quality that informs these exchanges between data producers and users.

The specific aim of this special issue is to bring out key findings pertaining to data quality issues related to diversified set of indicators drawn from large scale surveys of India, such as, National Family Health Survey (NFHS), Longitudinal Ageing Study of India (LASI), Indian Human Development Survey (IHDS) and health data from the National Sample survey (NSS). This collection contains 11 research articles and covers data quality issues related to survey implementation, collection, and use of economic data in large scale surveys, and issues related to inter and intra discordance between indicators.

Three papers dealt with survey implementation related issues as well as interviewer-level bias and found that interviewer-level bias does exist in some of the critical variables in National Family Health Surveys, such as, enumeration of total children ever born to a female respondent and that born in the 5-years preceding the survey, and mothers who did not receive ante-natal care. Two other papers dealt with interviewer workload and found that variable workload in a primary sampling unit (PSU) does increase variability in key indicators, including anthropometric indicators. The authors recommended several remedies including an outlier-based approach in field monitoring to capture interviewer-bias at the time of fieldwork, expanding the duration of survey or increasing number of team members in PSUs with more than average number of eligible respondents, taking a modular approach in designing questionnaires, and modified field implementation plans as per local need.

A paper that investigated intra-cluster correlation (ICC) by socio-demographic and behavioral characteristics of respondents concluded that survey design (particularly, sample size estimation) should be

[☆] Despite a focused approach to ensure good quality data in surveys in the past few decades, there is a dearth of research on quality of survey data in India and elsewhere. It is thus important to regularly bring out high-quality research that deliberates on data quality issues and innovative methodologies adapted to ensure quality data. The NDQF planned this special issue as a leading collection of research work on data quality of large-scale surveys in India.

determined by considering ICCs obtained in similar surveys to reduce loss of precision of the estimates of key indicators. Two other papers that dealt with survey procedures examined the difference between self-reported morbidity and that reported by standard tests, and presence of a third person on the reporting on tobacco consumption among women. Recommendations from these papers form important guidelines for future surveys collecting sensitive information.

Another paper argued against using wealth index, a current practice in national health surveys, for predicting economic variations in health and health care utilization, instead, recommended including an abridged version of a consumption schedule in the survey tools for this purpose. Finally, a paper that compared estimates drawn from large-scale surveys and health management information system (HMIS) drew conclusion that with further improvements, HMIS data can be a treasure trove for estimating selected vital statistics of the population. The findings from these 11 papers will pave way for new approaches and innovative strategies to ensure a quality data in the subsequent rounds or for any upcoming large-scale survey in India or elsewhere.

These papers also bring out the importance of data quality analysis before data are subjected to policy analysis. These analytics should be a routine process in any large-scale surveys, which is often missing, and conclusions are drawn without understanding the data limitations. The editors recommend and encourage data quality analytics to become a standard practice during and post data collection and before publication of the data by the data producers, along with concentrated efforts by researchers to bring out data quality issues in published data.

References

- International Institute for Population Sciences (IIPS). (2020). National programme for health care of elderly (NPHCE), MoHFW. In T. H. Harvard (Ed.), *Longitudinal ageing Study in India (LASI) wave 1, 2017-18, India report, international Institute for population Sciences*. Mumbai: Chan School of Public Health (HSPH) and the University of Southern California (USC). Also available at: <https://www.iipsindia.ac.in/lasi>. (Accessed 16 August 2023), 14:10.
- ICMR-National Institute of Medical Statistics. (2021). *National guidelines for data quality in surveys*. New Delhi: ICMR-NIMS. Also available at: <https://ndqf.in/guidelines-tools/>. (Accessed 16 August 2023), 14:27.
- International Institute for Population Sciences (IIPS) and Ministry of Health and Family Welfare (MoHFW). (1993). Various rounds of NFHS. Available at: <https://www.iipsindia.ac.in/content/nfhs-project>. (Accessed 16 August 2023), 14:08.
- James, K. S., & Rajan, S. I. (2004). Respondents and quality of survey data. *Economic and Political Weekly*, 659–663.
- Ministry of Statistics and Programme Implementation (MoSPI). (2019). *Key indicators of social consumption in India: Health*. New Delhi: MoHFW: MoSPI. Government of India. NSS 75th Round.
- Rajan, S. I., & James, K. S. (2004). Second national family health survey: Emerging issues. *Economic and Political Weekly*, 647–651.

Rajib Acharya^{a,*}, K.S. James^b, Sri Kant Singh^b, Niranjan Saggurti^a
^a Population Council, New Delhi, India

^b International Institute for Population Sciences, Mumbai, India

* Corresponding author.

E-mail address: RACHARYA@popcouncil.org (R. Acharya).