'Registries are not only a tool for data collection they are for action': Cancer registration and gaps in data for health equity in six population-based registries in India

Amiya Bhatia, Cesar Victora, Jason Beckfield, Atul Budukh, Nancy Krieger

Amiya Bhatia. amb803@mail.harvard.edu; Harvard T.H. Chan School of Public Health, Department of Social and Behavioral Sciences, 677 Huntington Ave, Boston, MA, USA 02115 *[corresponding author]*

London School of Hygiene and Tropical Medicine, Department of Global Health and Development, 15-17 Tavistock Place, Saint Pancras, London WC1H 9SH

Cesar Gomes Victora. cvictora@gmail.com; International Center for Equity in Health, Federal University of Pelotas, Mal. Deodoro, 1160, 3d Floor, Zip Code: 96020-220, Pelotas (RS), Brazil.

Jason Beckfield. jbeckfie@wjh.harvard.edu; Harvard University, Department of Sociology, William James Hall, 33 Kirkland Street, Cambridge, MA 02138

Atul Budukh. atul.budukh@gmail.com; Centre for Cancer Epidemiology, Tata Memorial Center, Homi Bhabha National Institute, Mumbai, Maharashtra, India

Nancy Krieger. nkrieger@hsph.harvard.edu; Harvard T.H. Chan School of Public Health, Department of Social and Behavioral Sciences, 677 Huntington Ave, Boston, MA, USA 02115

Abbreviations

ASR; Age standardized rate
DAE; Department of Atomic Energy
ICMR; Indian Council of Medical Research
IARC; International Agency for Research on Cancer
GBD; Global Burden of Disease
GICR; Global Initiative for Cancer Registry Development
HBCR; Hospital-based cancer registry
LMICs; Low and middle-income countries
NCDs; Non-communicable diseases

NCRP; National Cancer Registry Program
PBCR; Population Based Cancer Registry
TMC; Tata Memorial Center

Novelty and impact statement: This study examines which sociodemographic information population-based cancer registries in India collect and use to make visible how poverty and inequality contribute to unequal cancer incidence. Findings show equity analyses of cancer incidence data in India are possible as most registries published estimates of cancer outcomes disaggregated by age, sex and geography, and several registries collected – but did not report – information on education, marital status, mother tongue, religion, income, and occupation.

Abstract

In India, population-based cancer registries (PBCRs) cover less than 15% of the urban and 1% of the rural population. This study examines practices of registration in PBCRs in India to understand efforts to include rural populations in registries, and efforts to measure social inequalities in cancer incidence. We selected a purposive sample of six PBCRs in Maharashtra, Kerala, Punjab and Mizoram and conducted semi-structured interviews with staff to understand approaches and challenges to cancer registration, and the sociodemographic information collected by PBCRs. We also conducted a review of peer-reviewed literature utilizing data from PBCRs in India. Findings show that in a context of poor access to cancer diagnosis and treatment, and weak death registration PBCRs have developed additional approaches to cancer registration, including conducting village and home visits to interview cancer patients in rural areas. Challenges included: PBCR funding and staff retention, abstraction of data in medical records, address verification, and responding to cancer stigma and patient migration. Most PBCRs published estimates of cancer outcomes disaggregated by age, sex and geography. Data on education, marital status, mother tongue and religion were collected, but rarely reported. Two PBCRs collected information on income and occupation and none collected information on caste. Most peer-reviewed studies using PBCR data did not publish estimates of social inequalities in cancer outcomes. Results indicate that collecting and reporting sociodemographic data collected by PBCRs is feasible. Improved PBCR coverage and data will enable India's cancer prevention and control programs to be guided by data on cancer inequities.

Introduction

The International Agency for Research on Cancer (IARC) estimates that only 1 in 5 low-and middle-income countries (LMICs) have the necessary data to drive cancer policy.[1] High quality population-based cancer registries (PBCRs) cover 1% of the population in Africa, 4% in Asia, and 6% Latin America.[2] Yet, more than half of the global cancer burden is in India, Russia and China, which include approximately 40% of the world's population, experience 46% of all new cancers worldwide, and account for 52% of cancer deaths globally.[3] Recognizing the need to address the invisibility of places and populations in cancer statistics and improve cancer registries in LMICs,[4-7] IARC established the Global Initiative for Cancer Registry Development (GICR) in 2011. A technical support hub was created in the Tata Memorial Center (TMC) in Mumbai, India to support countries developing or expanding PBCRs in South and East Asia.[5]

In a context where cancer registration is expanding in LMICs, this study examines the practices of population-based cancer registration in India among six PBCRs to understand approaches to, and challenges of, cancer registration, efforts to include rural populations in registries, and efforts to measure social inequalities in cancer outcomes.

India in particular faces unique challenges for cancer registration. Second to China, India has the world's largest population (1.34 billion) which is predominantly rural (66.4%). Access to affordable and equitable cancer care is limited[8-11] and low death registration coverage (67.40%) makes it challenging to estimate cancer mortality (Supplementary Table S1). India's network of PBCRs are estimated to cover less than 15% of the urban population[12] and less than 1% of the rural population[13] pointing to serious deficits in coverage. Important discrepancies exist in estimates of both cancer incidence and leading types of cancer, depending on which data are employed. For example, IARCs estimate of India's age-standardized incidence rate (ASR) (89.4/100,000)[14] is notably lower than that estimated by the by the Global Burden of Disease (106.6/100,000).[15, 16] Estimates of cancer cases are close to one million, but also differ, with non-overlapping 95% confidence intervals (Table 1). Moreover, whereas IARC suggests that lip and oral cavity, lung, stomach and colorectum are the leading cancer sites for males, GBD suggests lip and oral cavity, lung, other pharynx, and prostate cancer. Such inconsistencies raise concerns for cancer prevention and control in India and more widely, given that India's registries are relied upon to estimate cancer incidence in other Indian states,[17] and across South Asia, including in Bangladesh, Bhutan, Nepal, Pakistan and Myanmar (Supplementary Figure S1).[2, 14, 18]

Still another challenge in the current estimates of cancer presented in Table 1 is the absence of cancer incidence or case counts disaggregated by sociodemographic variables other than age and sex raising unanswered questions about how cancer incidence differs by income, ethnicity, religion and caste. Such data are crucial for documenting, monitoring, analyzing, and preventing cancer inequities,[19] which represent suffering and death which is unfair unjust and preventable[20-24] and most often experienced by historically marginalized populations.

[Table 1]

Methods

Sample selection

At the time of data collection in 2018, the National Cancer Registry Program (NCRP) of the Indian Council of Medical Research (ICMR) included 27 PBCRs and 7 hospital-based cancer registries (HBCRs).[25] At least 17 additional PBCRs had additionally been funded by the Department of Atomic Energy (DAE) and established by the Tata Memorial Center (TMC).

From among the NCRP and TMC registries, we selected a purposive sample of six PBCRs based on registry characteristics and discussions with colleagues at TMC and NCRP. To ensure we included registries functioning in a range of contexts, we selected: two rural PBCRs in India (in Sangrur, Punjab and Barshi, Maharashtra); the oldest PBCR (Mumbai, Maharashtra); the most recently formed urban PBCR (Chandigarh); the PBCR reporting the highest burden of cancer in India (Aizawl, Mizoram); and, a PBCR serving an urban and rural population (Thiruvananthapuram, Kerala).

Data collection

We developed a semi-structured interview guide and a registry profile form to understand: the characteristics of each registry; the data collection process; which sociodemographic variables were collected and reported; the strengths and challenges of cancer registration; and recommendations to improve registration.

The first author traveled to each registry over eight weeks in 2018 to conduct key informant interviews. In each registry, 3-7 people were interviewed and included the PI or Co-PI and at least one fieldworker. Additional interviews were conducted based on staff availability and the division of labor in the registry: these included data managers, social workers, coders, or physicians. Interviews were also conducted with other experts in cancer registration who were affiliated with these registries, including NCRP. Interviews were audio recorded when possible, or notes were taken. We sought and obtained verbal informed consent to name each registry in the analysis.

To gain insight into the use of India's PCBR data for population health monitoring and research, we also examined the most recent report from each registry and searched the Web of Science and PubMed for all peer-reviewed articles using the following terms: "population-based cancer

registry" or "registr*", "India", "cancer" as well as the locations of each registry in the sample. All studies that discussed the state of PBCRs in India or used PBCR data were included. For these studies, we examined which articles presented estimates of cancer disaggregated by sociodemographic variables. Literature reviews, abstracts, letters to the editor, multi-country studies, or studies that used data from hospital-based registries or tumor registries were excluded.

Analysis

We used the registry profile form and interview guide to develop a codebook and define themes for analysis a priori which included: PBCR characteristics; challenges of registration; funders and technical assistance; incidence and mortality data collection; reporting; data uses; recommendations to improve registration. The following additional themes, which emerged from the data, were added during data analysis: history of the PBCR; patient engagement; approach to registration; strengths of the PBCR; government schemes; links between hospital and population-based registries. Analysis was conducted using NVivo 12. Each registry had an opportunity to check if the information in the profile was accurate but did not analyze or interpret findings.

Records that met the inclusion criteria of the literature review were categorized into the following thematic groups based on a review of the title and abstract: cancer surveillance; time trends; projections; case-control or cohort studies; associational studies; costing studies; studies about PBCRs in India. For studies that reported data, we conducted full text searches and calculated the number and percentage of articles which reported estimates of cancer disaggregated by sociodemographic variables.

Findings

Context and data collection approaches of six Population Based Cancer Registries

Six PBCRs were included across four states: Maharashtra, Kerala, Punjab and Mizoram. Four were located in large cancer hospitals, the Mumbai PBCR was located in the Indian Cancer Society, and the Chandigarh PBCR in the School of Public Health in the Post Graduate Institute of Medical Education and Research. Registries were funded and supported by NCRP or TMC. Core staff ranged from 7-14 members and included: a program manager, 'social investigators' (fieldworkers), and a data manager. Each registry worked in a different context of population size and density, geography and cancer burden. Table 2 and the additional text in the Supplementary Materials describe the context, coverage, registration processes and cancer cases in each PBCR.

Each PBCR used methods of cancer registration recommended by IARC and ICMR and visited between 29 and 75 institutions (data sources) for case finding. These included hospitals,

pathology and cytology labs, CT scan centers, nursing homes, and the offices coordinating enrolment in government schemes for cancer treatment. For mortality data, PBCR staff either accessed death registration data online or visited the office of vital statistics, crematoriums and burial grounds. However, the history and context of each registry -- population size, density, geography and cancer burden -- shaped approaches to case finding (Table 2). For example, the PBCR in Mizoram explained, *"there are so many cancer patients every day…many homes have cancer"*, while the Sangrur PBCR underscored the long distances staff traveled to find cancer cases: *"If you cover…15,000-16,000 households you don't find many symptomatic patients so it's very difficult"*.

Notably in Sangrur and Barshi -- PBCRs working in rural areas -- staff routinely visited villages and households, and conducted case finding with health workers, and local leaders. Phone calls or home visits with patients were not only a way to confirm address and collect data, but offered an opportunity for PBCR fieldworkers to build relationships with patients, address stigma, encourage treatment, demystify hospital processes, and provide assistance in accessing government schemes for cancer. Given its large caseload, the Mumbai PBCR relied solely on information in patient records from each data source. Similarly, the Thiruvananthapuram and Chandigarh PBCRs relied predominantly on medical records for case finding, but telephoned patients to address any incomplete fields in the data collection form. The Aizawl PBCR interviewed patients in the hospital in which the registry was located and then relied on phone calls, local birth and death registrars and the Young Mizo Association, a state-wide community group, to collect additional information.

[Table 2]

The distribution and heterogeneity of cancer cases across the data sources PBCR staff relied on for case finding reflected the journeys of cancer patients through local health systems, and the quality of local and state-level health data. For example, in Mizoram, as one respondent explained, PBCR staff found most cancer cases in Aizawl as patients traveled to the state capital for diagnosis and treatment: *"we don't have pathology or diagnostics in sub-centers…every district has a hospital, but only Lunglei [another district] has a pathologist…so everyone comes to Aizawl"*. Similarly, in Thiruvananthapuram, because the PBCR was located in the Regional Cancer Center (RCC), the vast majority of cases came from here. Cancer cases in Barshi often came from cancer detection camps conducted across all 346 villages in the catchment area by the PBCR and cancer hospital. For mortality data, the Aizwal, Thiruvananthapuram and Mumbai PBCRs benefited from high coverage and quality of death registration, including online access in Thiruvananthapuram. In contrast, the registry in Sangrur, Punjab had developed processes to check crematoriums, burial grounds and religious sites, review data in the local death registrar's office and to conduct a verbal autopsy with relatives to find and identify cancer deaths.

Sociodemographic data collected by six Population Based Cancer Registries

All PBCRs published estimates of cancer incidence disaggregated by age and sex and most by geography - village, district, urban/rural (Table 3). Although most PBCRs collected information on education, marital status, mother tongue and religion, only one PBCR published estimates of cancer incidence disaggregated by all these outcomes. Only two PBCRs collected – but did not publish – information on income and occupation and no PBCR collected information on ethnicity, Scheduled Caste, Scheduled Tribe, or caste. The PBCRs which did not interview patients directly were limited to the sociodemographic information available in medical records.

[Table 3]

Among registries with direct patient contact, a minority of respondents felt it was easy and possible to collect sociodemographic variables while others expressed concern about the ramifications of asking about sociodemographic information. Education was considered to be easy to ask about. However, in Mizoram, pointing to the lack of variation in education, one respondent observed: *"it might be possible to look at education and cancer, but Mizoram is the second most educated state...so it is not a good criteria"*. Income was considered challenging to get accurate information about and data on caste – which is not collected – was also perceived as challenging: "*we can't ask caste...in Punjab the biggest political fights happen over caste, more than religion"*. Contrastingly, in Kerala, religion was considered to be more sensitive. Other respondents described how it was possible to *'figure out'* these variables based on other identifiers. For example, the patient's name, occupation, whether a spouse was listed or not were occasionally used to guess religion, mother tongue and marital status. One respondent summarized the status quo as: *"there are widespread socioeconomic inequalities in cancer...but they need to be further documented"* and another respondent discussed how controversial it might be for government stakeholders if registries published estimates of cancer disaggregated by caste.

Uses of cancer registry data

There was agreement about the benefit of cancer registries, especially in examining if cancer burden, leading sites, and cancer stages were changing. Uses of data within registries included: flagging cases to guide follow up with patients and preparing reports that were either issued annually or every 2-3 years (Table 2). No individual level or aggregated data outside of the data presented in reports were publicly available. Few registries conducted additional or exploratory analyses of data. All registries submitted data to NCRP or TMC and most submitted data to IARC for inclusion in CI5 and GLOBOCAN 2018.

When asked about uses of data outside the registry, several respondents discussed answering the state government's questions about cancer burden. In Punjab, registry staff described how the 2013 report showed cancer in Punjab is: *"lower than the national average and now we are trying to dispel the myth. People feel like that because they don't have access to care"*. There were very

few examples where registry data were used to inform treatment and control programs at the local or state level. One example included the introduction of endoscopy centers in response to stomach cancer incidence in Mizoram. Respondents agreed that registry data could be better used to guide treatment and control programs, and to understand sociodemographic inequalities in cancer care and outcomes.

The review of the literature revealed 230 out of 1290 peer-reviewed articles that were either about PBCRs or utilized PBCR data published between 1960 and 2019. Most articles were on cancer surveillance or time trends. Among the 188 articles utilizing PBCR data, 70 articles (37%) reported estimates of cancer disaggregated by at least one sociodemographic variable (Table 4); education, marital status and religion were the most commonly used variables, and case control and cohort studies were most likely to present estimates disaggregated by sociodemographic variables. However, very few studies that used PBCR data were explicitly designed to examine cancer inequities. Supplementary Table S3 lists all the articles that presented disaggregated estimates by article category.

[Table 4]

Challenges of cancer registration

The challenges registry staff faced while case finding and collecting data included securing the cooperation of data sources, and abstracting information from primarily paper-based medical records in hospitals with vastly different record keep systems (Table 5). Most registries cited a lack of funding as the primary challenge of operating the registry which affected paying salaries, and retaining staff. In registries with direct patient contact, staff described the challenge of responding to bad patient experiences in efforts to seek treatment, including long journeys and lines, poor medical care, or the loss of a family member. Most registries faced major challenges with cancer stigma, however, this was more pronounced in the new registries (Sangrur and Chandigarh) where staff explained: *"there is low awareness of cancer. We can't say 'cancer' directly, we ask: 'have you had treatment?' and 'what did the doctor say?'"* In contrast, staff in Barshi, discussed the changes in cancer stigma since the 1980s:

> *"10-20 years ago, when we went to the house to talk about the symptoms…they would look at us with shock as if to say 'what are you saying?'…. now, they recognize us and know what we are going to ask. There is a lot more awareness and now we don't have problems in the villages".*

[Table 5]

Recommendations to improve cancer registration and implications for new PBCRs

Finally, reflections and recommendations offered by respondents to improve cancer registration have implications for the design, and data collection of new PBCRs, including ways to improve the collection of sociodemographic data (Table 6). Of note was the emphasis respondents placed on the importance of building relationships between PBCR staff and the institutions which served as data sources, village health workers and local government. Persevering in efforts to get data was another key recommendation as was safely sharing data across registries, including hospital-based registries. Respondents advocated for the development of new population-based registries in India, particularly in rural areas:

> *"In the last 35 years this is the only rural registry...we need more rural registries, because our data cannot be used for all of rural India...75% of the population in India is rural and most registries are urban." (Barshi)*

Discussion

We show that both urban and rural registries in India have developed a range of approaches to respond to the challenging environment for population-based cancer registration, and in many cases, are engaging in cancer screening, education, and stigma reduction alongside registration.[26, 27] Although some of the challenges registries face are shared by other registries globally[4, 7, 11, 28] – for example, in negotiating access to data for case finding, ensuring reliable data, and gaining adequate resources – the major challenges of registration we report are inseparable from the structural and systems level lacunae in the health system in India. Poor and unaffordable access to cancer care and diagnosis, weak death registration, and sparse health information systems[8-11] shape which places and populations are included in cancer data, and are implicated in cancer inequities in India,[4, 10] and globally.[29] However, our findings show that PBCRs in India use additional methods of case finding and cancer registration in response to the challenging social, cultural, and financial realities of cancer care seeking and caregiving. These approaches allow data on cancer incidence to be collected in contexts where patients often reach the hospital at later cancer stages, struggle financially, migrate for treatment, or seek multiple medical opinions. Such approaches to cancer registration could be instructive for efforts to initiate or expand cancer registration in other contexts in both high and low-income countries.

The diverse cancer burden seen across six PBCRs in four states is reflective of the heterogeneity in both cancer incidence and the leading sites of cancer across India.[9, 15] India's context where each state is the size of some of the world's most populous countries underscores the importance of state-level cancer registry development with a focus on rural areas. In the words of one respondent: "*the state government needs to do something and focus on high [incidence] cancers. Usually they just blame it on tobacco consumption",* which underscores the importance of the government delivering its commitments to expanding cancer care infrastructure and improving quality of care. We highlight the need to improve the dissemination and use of PBCR data to local government to design cancer screening and control programs at the local and state level:

such efforts could better realize the sentiment expressed by most respondents that "*cancer registration is for cancer control...it is not an end in itself*". India's national program for non-communicable diseases (NCDs) which launched in 2010 and includes cancer screening, prevention and treatment, develops tertiary cancer care facilities, and health systems strengthening for NCDs[8] could offer an expansion of registration alongside cancer treatment facilities. We found there was strong agreement among respondents that expanding cancer registration without links to the expansion of cancer treatment facilities was unethical. This argument was rooted in concerns about health equity stemming from offering diagnoses to patients without ensuring access to treatment.

A dual commitment to cancer registration and treatment is also relevant for other LMICs developing national cancer control plans,[30] especially given that in 2017, only 26% of low-income countries reported pathology services in the public sector, and 30% reported treatment services.[31] As cancer care expands in LMICs, our findings contribute evidence on the importance of ensuring hospital-based and population-based cancer registries are a key component of NCD and cancer control plans, and that these registries are able to document and analyze social inequities in cancer. Making data on the sociodemographic distribution of cancer visible will be essential in both preventing and addressing social inequities across the cancer continuum, and in bringing attention to the burden of cancer among populations living in poverty and in rural areas.[32]

Prior research has documented the critical role of disease surveillance in the development and implementation of health policy, including the role of cancer registries in cancer control, research, and monitoring.[28, 33] In line with this, many respondents articulated the belief that one respondent described: "registries *are not only a tool for data collection they are for action...surveillance means data collection for action*". Our study specifically examines sociodemographic data and findings show that, although a wealth of data currently exists in PBCRs, these data could be better utilized for research, for advocacy, for guiding and assessing cancer prevention, treatment and control, and for measuring social inequities in cancer. First, efforts to ensure data quality and completeness of the socidemographic data that are currently being collected by PBCRs are essential in enabling the use of sociodemographic data. Secondly, in addition to improve data quality, improvements in the utilization of existing PBCR data could include conducting analyses of cancer incidence disaggregated by geography, education, religion, language or using existing resources - like the US National Cancer Institute's health disparities calculator[34] - to estimate the magnitude of inequities.

Efforts to ensure the routine, and high quality, measurement of key sociodemographic variables that are comparable across PBCRs in India will enable comparable cross-sectional and longitudinal analyses of social inequalities in cancer incidence in India. Epidemiologic research from other countries that utilizes comparable data from multiple PBCRs reveals large social inequalities in cancer incidence,[35-37] mortality, and survival.[38-40] However, many of these

analyses rely on linking PBCR data with sociodemographic information and measures of deprivation that are not collected by PBCRs. We show that many PBCRs in India are collecting these data at the individual level with enormous potential to improve and use these data for equity analyses.

However, as our findings from the review of the literature indicated, limited peer-reviewed research has utilized PBCR data to explicitly examine social inequities in cancer, with variables such as caste and ethnicity or SC/SCT never considered, and variables like occupation and income very occasionally considered – even though several PBCRs reported they were currently collecting these data. These results are in line with a 2014 study on cancer research in India which showed that cancer genetics and medical oncology (chemotherapy) are the dominant forms of cancer research,[41] underscoring the need for further research on the role of poverty and inequality in shaping population distributions of cancer. Our findings show that while the uses of PBCR data may thus far be limited, many PBCRs are collecting sociodemographic data. Further research on the completeness and missingness of sociodemographic data in PBCRs will be important in improving the quality of these data. Analyses which use PBCR data to estimate social inequities in cancer outcomes, or which document efforts to improve registration in rural areas, including how cancer registries interact with the health system, are warranted, especially given the lack of research on cancer registration in LMICs.[7, 42] Our findings also underscore the value of engaging PBCR staff to understand the realities of registration practices and further examine the institutional efforts to address the challenges of registration we document.[43]

In summary, the expansion of cancer registration in India offers opportunities to both include rural populations in registries, integrate equity analyses of PBCR data into routine monitoring and reporting, and allow new registries to benefit from the lessons and good practices in existing registries. Given that many PBCRs in India receive technical assistance from NCRP, TMC and IARC, there are opportunities to share best practices between cancer registries as cancer registration expands within India.

To conclude, data collected by PBCRs drives cancer epidemiology research, and can be used to develop and target prevention and control programs in India. Efforts by IARC, NCRP, PBCRs, policy makers, and state-level governments to improve the quality of sociodemographic data in PBCRs could make visible how poverty and inequality contribute to unequal cancer incidence and are a crucial step in allowing cancer prevention and control programs to both monitor and address social inequalities in cancer.

Acknowledgements

Anbumani, Dr. Evelyn, Ms. Shravani Koyande, Dr. Aleyamma Mathew, Mr. NS Panse, Dr JS Thakur, and Dr. Eric Zomawia for their invaluable support with the study. Thank you to the Harvard India Center, particularly Dr. Ananya Awasthi and Dr. Vish Viswanath for their helpful suggestions in developing the study and for their assistance throughout the research process.

Author contributions

- Amiya Bhatia conceptualized paper, conducted analysis, and drafted the manuscript

- Cesar Victora contributed to conceptualizing the paper and analyses, reviewed results, and provided comments on all drafts

- Jason Beckfield contributed to conceptualizing the paper and analyses, reviewed results, and provided comments on all drafts

- Atul Budukh contributed to conceptualizing the paper, data collection strategy, and analyses, reviewed results, and provided comments on all drafts

- Nancy Krieger contributed to conceptualizing the paper and analyses, reviewed results, contributed to drafting of manuscript, and provided comments on all drafts

Ethics statement

Ethics approval was obtained from the Harvard T.H. Chan School of Public Health and the Tata Memorial Center Mumbai, Institutional Review Boards.

Conflict of interest

No conflicts to declare

Data accessibility

The data that supports the findings of the literature review are available in the supplementary material of this article. The qualitative quotations that support the findings from this study are also available in the supplementary material. Data that supports the findings of this study are available from the corresponding author upon reasonable request.

References

1. International Agency for Research on Cancer. Global Initiative for Cancer Registry Development (GICR), vol. 2018: International Agency for Research on Cancer (IARC) http://gicr.iarc.fr/en/, 2018.

2. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. The Lancet Oncology 2012;13:790-801.

3. Collingridge D. Three countries—half of the global cancer burden. The Lancet Oncology 2014;15.

4. Valsecchi MG, Steliarova-Foucher E. Cancer registration in developing countries: luxury or necessity? The Lancet Oncology 2008;9:159-67.

5. Bray F, Znaor A, Cueva P, Korir A, Swaminathan R, Ullrich A, Wang SA, Parkin DM, Planning and developing population-based cancer registration in low- and middle-income settings International Agency for Research on Cancer, 2015.

6. Sankaranarayanan R. Cancer prevention and care in India: an unfinished agenda. Lancet Oncol 2014;15:554-5.

7. Behera P, Patro BK. Population Based Cancer Registry of India - the Challenges and Opportunities. Asian Pac J Cancer Prev 2018;19:2885-89.

8. Menabde N. A health-system response to cancer in India. The Lancet Oncology 2014;15:485-87.

9. Mallath MK, Taylor DG, Badwe RA, Rath GK, Shanta V, Pramesh CS, Digumarti R, Sebastian P, Borthakur BB, Kalwar A, Kapoor S, Kumar S, et al. The growing burden of cancer in India: epidemiology and social context. Lancet Oncol 2014;15:e205-12.

10. Pramesh CS, Badwe RA, Borthakur BB, Chandra M, Raj EH, Kannan T, Kalwar A, Kapoor S, Malhotra H, Nayak S, Rath GK, Sagar TG, et al. Delivery of affordable and equitable cancer care in India. Lancet Oncol 2014;15:e223-33.

11. Goss PE, Strasser-Weippl K, Lee-Bychkovsky BL, Fan L, Li J, Chavarri-Guerra Y, Liedke PE, Pramesh CS, Badovinac-Crnjevic T, Sheikine Y, Chen Z, Qiao YL, et al. Challenges to effective cancer control in China, India, and Russia. Lancet Oncol 2014;15:489-538.

12. Chatterjee S, Chattopadhyay A, Senapati SN, Samanta DR, Elliott L, Loomis D, Mery L, Panigrahi P. Cancer Registration in India - Current Scenario and Future Perspectives. Asian Pac J Cancer Prev 2016;17:3687-96.

13. Bray F CM, Mery L, Piñeros M, Znaor A, Zanetti R and Ferlay J (editors), Cancer Incidence in Five Continents, Vol. XI (electronic version). International Agency for Research on Cancer http://ci5.iarc.fr, 2017.

14. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.

15. Dhillon PK, Mathur P, Nandakumar A, Fitzmaurice C, Kumar GA, Mehrotra R, Shukla DK, Rath GK, Gupta PC, Swaminathan R, Thakur JS, Dey S, et al. The burden of cancers and their variations across the states of India: the Global Burden of Disease Study 1990–2016. The Lancet Oncology 2018;19:1289-306.

16. Global Burden of Disease Study 2016 (GBD 2016) Cancer Incidence, Mortality, Years of Life Lost, Years Lived with Disability, and Disability-Adjusted Life Years 1990-2016, vol. 2018: Institute for Health Metrics and Evaluation (IHME), 2016.

17. Prasad JB, Dhar M. Projections of burden of cancers: A new approach for measuring incidence cases for India and its states - Till 2025. Journal of Cancer Policy 2018;16:57-62.

18. Beaglehole R, Bonita R, Magnusson R. Global cancer prevention: An important pathway to global health and development. Public Health 2011;125:821-31.

19. Ginsburg O, Bray F, Coleman MP, Vanderpuye V, Eniu A, Kotha SR, Sarker M, Huong TT, Allemani C, Dvaladze A, Gralow J, Yeates K, et al. The global burden of women's cancers: a grand challenge in global health. Lancet (London, England) 2017;389:847-60.

20. Krieger N. Follow the North Star: Why Space, Place, and Power Matter for Geospatial Approaches to Cancer Control and Health Equity. Cancer Epidemiology Biomarkers & Prevention 2017;26:476-79.

21. Commission on Social Determinants of Health, Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. World Health Organization, 2008.

22. Krieger N, Emmons KM, Williams D. Defining, Investigating, and Addressing Cancer Inequities: Critical Issuesed. New York, NY, 2009.

23. Vaccarella S, Lortet-Tieulent J, Saracci R, Conway D, Straif K, Wild C. Reducing Social Inequalities in Cancer: Evidence and Priorities for Research IARC Scientific Publication No. 168 Lyon, France: International Agency for Research on Cancer, 2019.

24. Whitehead M. The Concepts and Principles of Equity and Health. International Journal of Health Services 1992;22:429-45.

25. National Cancer Registry Programme, Three-Year Report of Population Based Cancer Registries 2012-2014. NCDIR-NCRP, 2016.

26. Jayant K, Nene BM, Badwe RA, Panse NS, Thorat RV, Khan FY. Rural cancer registry at Barshi, Maharashtra and its impact on cancer control. Natl Med J India 2010;23:274-7.

27. Bashar MA, Thakur JS, Budukh A, Kapoor R. Quality assessment of four new population based cancer registries (PBCRS) in Chandigarh & Punjab. European Journal of Cancer 2016;54:S54-S54.

28. Parkin DM. The evolution of the population-based cancer registry. Nature Reviews Cancer 2006;6:603.

29. Magrath I, Bey P, Shad A, Sutcliffe S. Cancer funding in developing countries: the next health-care crisis? The Lancet 2010;376:1827.

30. Romero Y, Trapani D, Johnson S, Tittenbrun Z, Given L, Hohman K, Stevens L, Torode JS, Boniol M, Ilbawi AM. National cancer control plans: a global analysis. Lancet Oncol 2018;19:e546-e55.

31. World Health Organization. Cancer Fact Sheet, vol. 2018: World Health Organization 2018.

32. Budukh A, Dikshit R, Thakur J, Kadam P, Dora T, Sancheti S, Chaudhary D, Badwe R. Linkage of cancer registration with cancer treatment in predominantly rural district: A model form Sangrur district, Punjab state, India. International Journal of Noncommunicable Diseases 2018;3:56-59.

33. Armstrong B. The role of the cancer registry in cancer control. Cancer Causes & Control 1992;3:569-79.

34. Breen N, Scott S, Percy-Laurry A, Lewis D, Glasgow R. Health disparities calculator: a methodologically rigorous tool for analyzing inequalities in population health. Am J Public Health 2014;104:1589-91.

35. Tweed EJ, Allardice GM, McLoone P, Morrison DS. Socio-economic inequalities in the incidence of four common cancers: a population-based registry study. Public Health 2018;154:1-10.

36. Anthes LE, Hajizadeh M. Socioeconomic inequalities in pancreatic cancer incidence in Canada: evidence from Cancer Registry data. Journal of Public Health 2020.

37. Bryere J, Dejardin O, Launay L, Colonna M, Grosclaude P, Launoy G, French Network of Cancer R. Socioeconomic status and site-specific cancer incidence, a Bayesian approach in a French Cancer Registries Network study. Eur J Cancer Prev 2018;27:391-98.

38. Fantin R, Santamaria-Ulloa C, Barboza-Solis C. Social inequalities in cancer survival: A population-based study using the Costa Rican Cancer Registry. Cancer Epidemiol 2020;65:101695.

39. Ellis L, Coleman MP, Rachet B. How many deaths would be avoidable if socioeconomic inequalities in cancer survival in England were eliminated? A national population-based study, 1996-2006. Eur J Cancer 2012;48:270-8.

40. Singh GK, Jemal A. Socioeconomic and Racial/Ethnic Disparities in Cancer Mortality, Incidence, and Survival in the United States, 1950-2014: Over Six Decades of Changing Patterns and Widening Inequalities. J Environ Public Health 2017;2017:2819372.

41. Sullivan R, Badwe RA, Rath GK, Pramesh CS, Shanta V, Digumarti R, D'Cruz A, Sharma SC, Viswanath L, Shet A, Vijayakumar M, Lewison G, et al. Cancer research in India: national priorities, global results. Lancet Oncol 2014;15:e213-22.

42. Znaor A, Eser S, Anton-Culver H, Fadhil I, Ryzhov A, Silverman BG, Bendahou K, Demetriou A, Nimri O, Yakut C, Bray F. Cancer surveillance in northern Africa, and central and western Asia: challenges and strategies in support of developing cancer registries. The Lancet Oncology 2018;19:e85-e92.

43. Holland J. Who counts?: the power of participatory statisticsed. Warwickshire: Practical Action Publishing, 2013.

Table 1: GLOBOCAN, GBD and NCRP estimates of cancer incidence in India

	GLOBOCAN[9]	GBD[10,11]	NCRP[17]	CI5, Vol XI[8]
Year	2018	2016	2012-2014	
Number of PBCRs included	26	42	27	16
Population Data for denominator and age-standardization	UN population estimates	GBD reference population	Census 2011 with population growth estimates	Registries submit population denominators
Cancer cases				
Estimate	1.16 million	1.07 million	No national estimates. All estimates are specific to 27 PBCRs	No national estimates. All estimates are specific to 16 PBCRs
95% Confidence/ Uncertainty interval	(1.15 million - 1.17 million)	(1.04 million - 1.10 million)		
Incidence Rate (Combined)				
Crude	85.5	81.2 (79.3 - 83.7)		
Age standardized	89.4	106.6 (104.2 - 109.7)		
Incidence Rate (Female)				
Crude	90.0	87.0 (84.4 - 91.8)		
Age standardized	90.0	106.1 (102.9 - 111.8)		
Incidence Rate (Male)				
Crude	81.3	75.9 (73.2 - 78.0)		
Age standardized	89.8	108.6 (105.0 - 111.8)		
Cancers with highest age-standardized incidence rate (ASR)				
	ASR (World Standard)	ASR (GBD Standard)		
Female	1. Breast (24.7) 2. Cervix uteri (14.7) 3. Ovary (5.5) 4. Lip, oral cavity (5.5)	1. Breast (21.6) 2. Cervical (13.5) 3. Lip and oral cavity (9.2) 4. Stomach (7.3)		
Male	1. Lip, oral cavity (13.9) 2. Lung (7.8) 3. Stomach (6.2) 4. Colorectum (5.8)	1. Lip and oral cavity (13.5) 2. Lung (10.7) 3. Other pharynx (9.35) 4. Prostate (9.0)		
Estimates of incidence disaggregated by sociodemographics	Age and Sex	Age and Sex	Age and Sex	Age and Sex

Table shows the most recent estimates of cancer incidence in India published by GLOBOCAN, Global Burden of Disease (GBD), the National Cancer Registration Program (NCRP), and Volume XI of Cancer Incidence in Five Continents (CI5). Supplementary Table S2 lists the registries included in each set of estimates.

GLOBOCAN estimates come from GLOBOCAN 2018: National incidence was estimated as follows:
1. Incidence rates for the North-Eastern region (Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura states) were computed using population weighted average of the rates from nine population-based cancer registries (29% coverage around 2011). These rates were applied to the corresponding 2018 population.
2. Remaining urban population: rates from seven cancer registries (Bangalore, Bhopal, Chennai, Mumbai, Delhi, Nagpur, Pune) (15% coverage of urban population) for the period 2003-2012 were projected to 2018 and applied to the 2018 urban population. These seven registries were supplemented by three cancer registries (Aurangabad, Kollam and Trivandrum) when projection cannot be performed (use of the most recent rates).
3. Remaining rural population: rates from six cancer registries (Ahmedabad, Ambilikkai, Barshi, Mansa, Sangrur and Wardha) (1% coverage of rural population) were applied to the 2018 rural population. The overall incidence estimate for India for 2018 is the sum of the three estimates.[8]

GBD estimates come from: Global Burden of Disease Study 2016. Cancer Incidence, Mortality, Years of Life Lost, Years Lived with Disability, and Disability-Adjusted Life Years 1990-2016.[10,11] The estimation of cancer incidence was driven by registry data from India. Estimates were produced initially for each state. For states with at least one population-based cancer registry, the incidence data were transformed to mortality by multiplying incidence data with an independently modelled urban or rural mortality-incidence (MI) ratio for the respective states. Authors computed the population-weighted mean of these state estimates as the estimate for India. Authors based age-standardized rates on the GBD global reference population. 95% uncertainty intervals (UIs) were based on 1000 runs of the models for each quantity of interest, with the mean considered as the point estimate and the 2·5th and 97·5th percentiles considered as the 95% UI.[10]

Table 2: Context, coverage, processes and cancer cases in six population-based cancer registries in India

	Mumbai, Maharashtra	Barshi, Maharashtra	Thiruvanthapuram, Kerala	Aizwal, Mizoram	Chandigarh, Punjab	Sangrur, Punjab
Registry Characteristics						
Year created	1963	1987	1994	2003	2013	2013
Location	Indian Cancer Society	Nargis Dutt Memorial Cancer Hospital	Regional Cancer Center	Civil Hospital	PGIMER, School of Public Health	Civil Hospital and Homi Bhabha Cancer Hospital
Catchment area	Greater Mumbai	Barshi	Thiruvanthapuram district	Mizoram state	Chandigarh (Union Territory)	Sangrur district
Population covered	25.2 million	1.5 million	3.3 million	1.1 million	1.1 million	1.7 million
% Rural	0%	100%	46.34%	47.89%	2.75%	69%
SqKm covered	603 sq.km	3713 sq.km	2129 sq.km	21,081 sq.km	114 sq.km	3625 sq.km
Population density (people per sq.km)	19,652	*Unknown*	1509	52	9258	457
Registry Cancer Outcomes in most recent report						
Most recent report	2 year report; 2013-14	3 year report; 2012-2014	3 year report; 2012-14	5 year report; 2010-14	1 year report; 2014	1 year report; 2014
Number of cases (Calendar years of data)	27623 (2013-14)	929 (2012-14)	15,694 (2012-14)	7,535 (2010-14)	840 (2014)	879 (2014)
Average number of cases per year	13,756	~309	~5231	~1552	870	879
AAR (Female)	119.7	60.4	132.5	164.3	102.6	57.6
AAR (Male)	111	53.9	122	210.5	96.9	47.8
Top 3 leading sites and % of cases (Female)	1. Breast (30.4%) 2. Reproductive organs (18.5%)* 3. Digestive organs (18.3%)	1. Cervix uteri (26.7%) 2. Breast (20.0%), 3. Ovary (4.6%)	1. Breast (28.8%) 2. Thyroid (10.4%) 3. Cervix Uteri (6.0%)	1. Lung (16.0%) 2. Cervix Uteri (15.8%) 3. Breast (12.5%)	1. Breast (32.1%) 2. Cervix Uteri (8.2%) 3. Ovary (7.7%)	1. Breast (23.8%) 2. Cervix Uteri (14.8%) 3. Oesophagus (9.9%)
Top 3 leading sites and % of cases (Male)	1. Lung (11.4%) 2. Mouth (10.1%) 3. Prostate (9.7%)	1. Mouth (8.6%) 2. Oesophagus (7.7%) 3. Liver (6.2%)	1. Lung (13.0%) 2. Oral Cavity (10.5%) 3. Prostate (7.2%)	1. Stomach (20.1%) 2. Esophagus (15.8%) 3. Lung (14.1%)	1. Prostate (10.9%) 2. Lung (10.4%) 3. Oesophagus (5.8%)	1. Oesophagus (9.7%) 2. Mouth (5.6%) 3. Tongue (5.6%)
Data and Reporting						
Approximate number of institutional data sources	50	56	75	45	33	29
Contact with patients	• All data comes from data sources	• Routine village and house visits	• Occasional phone calls and house visits	• Routine phone calls • Patient interviews in Civil Hospital • Occasional home visits	• Occasional phone calls and house visits	• Routine village and house visits • Verbal autopsy
Organizations data is sent directly to	NCRP IARC	NCRP IARC TMC	NCRP	NCRP	TMC	TMC
Ever sent data to CI5	Yes	Yes	Yes	Yes	N/A**	N/A**
Data included in GLOBOCAN 2018	Yes	Yes	Yes	Yes	No	Yes
Data publicly available	No	No	No	No	No	No
Funding and Technical Assistance						

	Mumbai		Barshi		Thiruvananthapuram		Aizwal		Chandigarh		Sangrur	
Funders	NCRP India Cancer Society		NCRP TMC		NCRP Cancer Society of Finland		NCRP		TMC		TMC	
Technical assistance	NCRP		NCRP TMC IARC		NCRP		NCRP		TMC PGIMER		TMC PGIMER	

*Reproductive organs includes ovary, cervix, corpus uteri
**N/A as registry created after most recent CI5 (2008-12)
Sources: Interviews, Registry Reports (see Supplementary Materials), GLOBOCAN 2018, CI5, India Census 2011.
Definitions: NCRP=National Cancer Registry Program; TMC=Tata Memorial Center; IARC=International Agency for Research on Cancer; PGIMER=Post Graduate Institute of Medical Education and Research, Chandigarh. Cases per year either estimated in the reports or calculated.

Table 3: Sociodemographic information collected and reported by six population-based cancer registries in India

Sociodemographic variables	Mumbai, Maharashtra		Barshi, Maharashtra		Thiruvanthapuram, Kerala		Aizwal, Mizoram		Chandigarh, Punjab		Sangrur, Punjab	
	Collected	Reported	Collected	Reported	Collected	Reported	Collected	Reported	Collected	Reported	Collected	Reported
Age	+	+	+	+	+	+	+	+	+	+	+	+
Sex	+	+	+	+	+	+	+	+	+	+	+	+
Geography	+	+	+	No	+	No	+	No	+	+	+	+
Income	No	No	No	No	No	No	No	No	+	No	+	No
Occupation	No	No	No	No	No	No	No	No	+	No	+	No
Education	+	+	+	No	+	No	+	No	+	No	+	No
Marital status	+	+	+	No	+	No	+	No	No	No	No	No
Mother tongue	+	+	+	No	+	No	+	No	+	No	+	No
Religion	+	+	+	No	+	+	+	No	+	No	+	No
Ethnicity, Scheduled Caste/Tribe	No	No	No	No	No	No	No	No	No	No	No	No
Caste	No	No	No	No	No	No	No	No	No	No	No	No

Table shows the sociodemographic variables collected and reported by six PBCRs.

Collected refers to the information collected from patients by the PBCR and is based on the data collection form used for incidence data. Reported refers to the information available in the most recent PBCR report (or from the NCRP report) and whether cancer incidence or case counts are disaggregated by the sociodemographic variables listed. Geography refers to incidence estimates disaggregated by village, district, urban or rural.

Table 4: Types of peer-reviewed research using data from population-based cancer registries in India presenting cancer inequities

Category	Definition	Total number of records		Records reporting cancer outcomes disaggregated by sociodemographic variables										
				Total*		By sociodemographic variable								
		n	%	n	%	Geography**	Income	Occupation	Education	Marital status	Mother tongue	Religion	Ethnicity, SC, ST	Caste
Cancer surveillance	Studies reporting the epidemiology of cancer, incidence, survival, DALYs, mortality	109	47%	45	41%	10	2	0	16	15	2	28	2	0
Cancer surveillance: Time Trends	Studies estimating trends in cancer incidence, or cumulative risk over time	50	22%	8	16%	5	0	0	0	0	0	3	0	0
Projections	Studies estimating cancer projections using PBCR data	9	4%	1	11%	1	0	0	0	0	0	0	0	0
Case-control, cohort, associational studies	Case-control or cohort studies using PBCR data, studies on the association between behaviors and cancer outcomes	20	9%	16	80%	3	7	8	12	5	4	13	0	0
Studies about PBCRs in India	Studies on data quality, methods for cancer registration, studies on the state of registration in India	38	17%											
Cost	Studies on the cost of cancer registration in India, or on the economic cost of cancer	4	2%											
	Total	**230**	**100%**	**70**										

Table shows the results of a literature search for articles using PBCR data, or articles about PBCRs in India. For articles utilizing PBCR data, the table shows which articles reported cancer outcomes disaggregated by sociodemographic variables.

A total of 1290 records were screened after duplicates were removed. 230 records were identified to fit the inclusion criteria. The full text and tables of 188 articles which utilized PBCR data were searched for estimates of cancer outcomes disaggregated by sociodemographic variables.

*Total n and % of articles reporting disaggregated estimates of cancer is based on a full text search. If the full text was not available, we relied on the abstract (n=7).

**Geography refers to articles which examined different by urban/rural location, wards or panchayats. Studies which used data from PBCRs in multiple places were not coded as 'yes' for geography.

Table 5: Challenges in data collection faced by six population-based cancer registries in India

Theme	Illustrative examples
Cooperation of data sources	*"Not all [sources] cooperate the same way and there are differences in the facilities and staff. We need to phone patients to address the [data] gaps, but even getting their contact information can be hard." (Mizoram)* *"This is a big challenge for registries in metros: urban registries don't go house to house or do patient interviews so are very reliant on hospital data...we need to convince the hospitals...it's about personal relations, the staff develop a relationship with hospitals" (Mumbai)* *"Most hospitals have electronic data but don't give us a copy...we can see it on the screen and fill in our forms...this makes it easier to complete." (Mumbai)* *"It is very hard to collect data from the hospitals...they are congested, sometimes we don't get a seat and have to sit on the floor...we adjust" (Thiruvanthapuram)* *"Hospitals say 'why should we give you our patients' records?'...or they say they don't have time or call us early in the morning. Some hospitals say to come on a Sunday because the hospital is closed and they are off that day and can take out all the records and not interrupt their work to give us the records. We just go on Sunday...the main thing we have learned is not to stop until we have the record...not to get angry, to go the next day if they say so...they can say no once or twice, but will they say no ten times? We just have to keep trying." (Barshi)*
Abstraction of data in medical records	*"We need to go through all the cases at the general hospitals which takes longer. For example, we go through 200 forms to get 10 cases" (Mumbai)* *"100,000 patients report [to the hospital] and there are 100,000 files to go through to see if they are from Thiruvanthapuram" (Thiruvanthapuram)* *"We need to go to the medical records department and to pathology to get complete information. Pathology does not have the address but we use pathology [reports] to confirm cases." (Thiruvanthapuram)* *"Poor medical records are a challenge, some institutes are not keeping proper records, or institutes sometimes don't want to give all the records. But we can get records from the villages." (Barshi)* *"In the government medical college there is no online system, there is a huge register for histopathology and cytopathology where they have reported all the patients...we have to identify whether it is malignant or not...data are month-wise and there are 300-400 reports in the register...they are manual registers...our staff has to sit with the histopathology and cytopathology [registers] which are different and abstract the data into our forms...we have to cross verify the registration number to see if it is a Sangrur area. It is a very challenging tough job." (Sangrur)*
Address verification	*"There are challenges with the address. We need the proper address and often people will just put the city name or the name of a town. Let's say someone come from Serchhip [a district in Mizoram] to Aizawl. They will put Aizawl as they are staying with their family here, and we have to confirm this." (Mizoram)*
Cancer stigma	*"Everyone thinks cancer is fatal.... there have been cases where people's break down, faint or cry." (Mizoram)* *"People really do not want to go to hospital" (Barshi)* *"Suppose one patient gets bad treatment in the hospital...then we are [the ones] going house to house and we have to deal with this. That is the main problem. If doctors and nurses are badly behaved, we are suffering, and the data is hard to get. If you go right after the incident 10-15 days, it is worse." (Barshi)* *"They [patients] are very fearful...it's a very dreaded thing...sometimes even the relatives don't want it to be disclosed to the patient. They won't take the name of cancer...they ask us not tell the patient." (Sangrur)* *"They ask is to okay if they stay in a family together, eat together, does it spread through touch. And once there is a cancer diagnosis in a family, whether both the partners can have sexual intercourse." (Sangrur)*
Gathering information about deceased patients	*"We contact the family, but the family discards or burns the information as they don't want to be reminded [of the death]. It's hard to get information on the primary site and morphology" (Mizoram)* *"We need to go to the Panchayat [local government] to figure out the home address and then do a home visit. In these records, sometimes cause of death is listed as 'illness or old age'. There is only aggregated data available at the Taluka level [administrative unit], so we also need to go to Panchayat." (Barshi)* *"In the villages, when there is a death, they burn all the documents. People don't want forms lying around that say 'cancer' due to the stigma. So, we do a verbal autopsy and meet the relatives...without forms and medical records it is hard to confirm the primary site" (Sangrur)*

"Many people still leave the state for cancer. We share the data with other registries but it's hard to get full information. For example, sometimes there is no biopsy report and people [families and patients] don't always keep papers." (Mizoram)

Migration   *"The furthest source is 400km away in Mumbai. For earning purposes, a lot of people go to Mumbai. If a parent has cancer they might go to Mumbai where their son is." (Barshi)*

"We go through the files, there is a 4-5 page form with the death certificate. If there is any mention of cancer or tumor, we collect this information even if the person is not resident in Mumbai. People come to Mumbai to die" (Mumbai)

Table 6: Recommendations for new registries and to improve cancer registration

Theme	Recommendations and Implications for new population-based cancer registries
Sociodemographic data	• Collect individual level sociodemographic information based on patient interviews or medical records, and conduct quality checks on this information • Explore address geocoding
Rural registries	• Work closely with the village level health system and local government • Collect data on cancer cases from hospitals, labs, crematoriums, burial grounds, death registries • Link cancer registration to cancer screening • Train health workers to recognize cancer symptoms and refer patients • Train registration staff on how to conduct education about cancer and stigma reduction
Design of PBCR	• Start slowly with key variables and develop institutional practices to collect data and ensure quality • Develop a HBCR and/or PBCR alongside the development of cancer hospitals • Link the data systems of HBCRs with PBCRs • Conduct training for doctors and staff in key data sources on cancer registration • Build in checks of data quality • Include several questions to confirm address and residence
Approaches to collecting data	• Develop strong institutional and personal relationships with data sources • Persevere while getting data • Develop processes to share data with other cancer registries to capture migration
Data linking	• Consider integrated cancer and NCD registries • Link cancer registries to e-health or state level health management information systems