Sociodemographic Factors and Stage of Cancer at Diagnosis: A Population-Based Study in South India

Aleyamma Mathew, PhD1; Preethi Sara George, PhD1; Kunnambath Ramadas, MD, PhD1; Beela Sarah Mathew, MD1; Aswin Kumar, MD1; Sivasevan Roshni, DMRT, DRB1; Krishnan Nair Lalithamma Jayakumar, MD2; and Christopher M. Booth, MD3

abstract

PURPOSE Lower socioeconomic status is associated with inferior cancer survival in high-income countries, but whether this applies to low- and middle-income countries is not well described. Here, we use a population-based cancer registry to explore the association between educational level and stage of cancer at diagnosis in South India.

METHODS We used the Trivandrum District population-based cancer registry to identify all cases of breast and cervical cancer (women) and oral cavity (OC) and lung cancer (men) who were diagnosed from 2012 to 2014. Educational status—classified as illiterate/primary school, middle school, or secondary school or higher—was the primary exposure of interest. Primary outcome was the proportion of patients with advanced stage disease at diagnosis defined as stage III and IV (breast, cervix, or OC) or regional/metastatic (lung).

RESULTS The study population included 4,547 patients with breast (n = 2,283), cervix (n = 481), OC (n = 797), and lung (n = 986) cancer. Educational status was 22%, 19%, and 26% for illiterate/primary, middle, and secondary school or higher, respectively. Educational status was missing for 33% of patients. The proportion of all patients with advanced stage disease was 37% (breast), 39% (cervix), 67% (OC), and 88% (lung). Patients with illiterate/primary school educational status were considerably more likely to have advanced breast cancer (50% v 39% v 36%; P < .001), cervix cancer (46% v 43% v 24%; P = .002), and OC cancer (77% v 76% v 59%; P < .001) compared with patients with higher educational levels. The proportion of patients with advanced lung cancer did not vary across educational levels (89% v 84% v 88%; P = .350).

CONCLUSION A substantial proportion of patients in South India have advanced cancer at the time of diagnosis. This is particularly true among those with the lowest levels of education. Future health awareness and preventive interventions must target less-educated communities to reduce delays in seeking medical care for cancer.
and cervical cancer in women and lung and oral cavity cancer in men, as they represent the highest burden of cancer in India.

**METHODS**

**Study Population**

The Population-Based Cancer Registry (PBCR) of Trivandrum District in South India, located in the Regional Cancer Centre (RCC), Trivandrum, was used to identify all incident cases of breast and cervical cancer in women and oral cavity and lung cancer in men who were diagnosed from 2012 to 2014. Trivandrum District has a population of 3.3 million—54% urban and 46% rural—and is located in the southern Indian state of Kerala (population 33 million). Kerala has the highest literacy rate (94% v 73% national rate), greatest life expectancy (age 74 years v 64 years), and lowest infant mortality rate in India.![Mathew et al](https://doi.org/10.1200/JCO.2019.37.14.0049)

**Data Sources**

The Trivandrum District PBCR is one of 27 cancer registries operating under the National Cancer Registry Program of India. The PBCR uses an active case-finding methodology that consists of visiting government/private hospitals and pathology laboratories. Data have been collected from more than 60 hospitals and seven pathology laboratories. Computerized information processing includes linkage of patient data obtained from various sources and review of duplicate/redundant records. Validity of the data is monitored by conducting data quality exercises periodically on abstraction of data from medical records and coding of the diagnosis. Microscopic confirmation, death certificate only, fatality ratio (%), and the proportion of unknown primary sites are used to assess the quality of the registry. Major sources for incidence data are the RCC (63% of cases) and the Government Medical College Hospital (24%...
Stage at diagnosis is routinely assigned by the treating oncologist for all patients seen at RCC and Government Medical College Hospital. On the basis of review of the case sheet, registry staff record clinical stage at time of diagnosis. If clinical stage is not available, pathologic stage is recorded. Staging systems vary by disease site—breast and oral cavity cancer are staged using the TNM classification system, cervical cancer is staged using the International Federation of Gynecology and Obstetrics system, and lung cancer is staged using the SEER clinical extent of disease system.

In cases of missing data, PBCR staff reviewed primary data sources to assign stage of disease and contacted patients by telephone to ascertain educational status. Among a random sample of 10% of patients, the agreement rate between oncologist-assigned stage and PBCR stage grouping was high: 92% for breast, 91% for cervix, 89% for oral cavity, and 88% for lung cancer. The primary end point of the study was to evaluate the association between stage of cancer at diagnosis and educational level. We also considered the following sociodemographic factors: age, marital status, religion, and urban/rural residence.

**Statistical Analysis**

We compared proportions between sociodemographic groups using the \( \chi^2 \) test. We used the Marascuilo procedure to compare multiple proportions. \( P \) values < .05 were deemed statistically significant. No adjustments were made for multiple comparisons.

**RESULTS**

**Characteristics of Study Population**

The study population included 4,547 patients with breast (\( n = 2,283 \)), cervix (\( n = 481 \)), lung (\( n = 986 \)), and oral cavity (\( n = 797 \)) cancer who were diagnosed in Trivandrum District from 2012 to 2014. Characteristics of the study population are listed in Table 1. The mean age was 58 years and 31% of cases were age 65 years or older. Eighty-three percent of patients were married. The majority (74%) of patients are Hindu and 26% of patients attended secondary school. Within the four cancer subgroups, patients with breast cancer were younger, more likely to live in urban areas, and more highly educated. A lower proportion of women with cervical cancer were married relative to those with the other cancers. Sixty-three percent (2,868 of 4,547) of patients were identified from RCC, 24% (1,090 of 4,547) from Government Medical College Hospital, and 13% (589 of 4,547) from other sources. Rates of data capture for stage and educational level were high for cases seen at RCC (range, 96% to 99% and 99% to 100%, respectively). Among patients seen at Government Medical College Hospital, the stage and educational capture rates were 62% to 92% and 7% to 12%, respectively, and corresponding rates from other sources were 35% to 71% and 12% to 19%.
Stage of Cancer at Diagnosis

As listed in Table 2, only 8% to 14% of patients presented with early-stage disease—that is, stage I/localized disease. Whereas 10% and 13% of patients with breast and cervical cancer presented with stage IV disease, the rate of metastatic disease at diagnosis was much higher for those with oral cavity (40%) and lung (46%) cancers. Overall, the proportion of patients with advanced disease—that is, stage III and IV or regional/metastatic—was 37% for breast, 39% cervix, 67% oral cavity, and 88% lung cancer.

Sociodemographic Factors and Stage of Cancer

Distribution of stage at diagnosis by educational level is shown in Table 3 and Figure 1. The proportion of patients with advanced disease—stage III and IV or regional/metastatic—is shown in Table 4. For breast, cervix, and oral cavity cancer, there is substantial association between stage at diagnosis and educational level. Among patients with breast cancer, the proportion of patients with advanced disease was 50%, 39%, 36% (P < .001) for illiterate/primary school, middle school, and secondary school or greater, respectively. Corresponding figures for cervix and oral cavity cancer are 46%, 43%, and 24% (P = .002) and 77%, 76%, and 59% (P < .001), respectively. The proportion of patients with lung cancer with advanced disease did not differ substantially across educational groups (89%, 84%, and 88%, respectively; P = .350).

Association between advanced stage of disease and age, marital status, religion, and rural/urban residence is shown in Table 5. Older men (age > 50 years) with oral cavity cancer were more likely than younger patients to have advanced disease (52% vs 42%; P = .035).

**DISCUSSION**

In this study, we used data from a PBCR in South India to evaluate whether there was an association between educational status and stage of cancer at diagnosis. Our study focused on the four most common cancers in India. Several important findings have emerged. First, our results demonstrate that the proportion of patients who are diagnosed with advanced stage disease in the general population of India is higher than in HICs. Second, our study demonstrates a strong association between educational status and stage of breast, cervical, and oral cavity cancer at diagnosis. This has important policy and public health implications. Finally, we did not observe any such association for patients with lung cancer.

It is worth comparing the distribution of stage at diagnosis in our study with existing literature. Eight percent of women in our study with breast cancer had stage I disease and 37% had stage III and IV disease, which is consistent with other reports from India; however, these figures are in stark contrast to data from HICs, where 30% to 45% of women have stage I disease and 8% to 22% stage III and IV disease at diagnosis. Likewise, we report that 14% and 39% of women with cervical cancer had stage I and stage III and IV disease, respectively—these data are consistent with that another study from India. However, corresponding figures in HICs are different; in Canada, 34% of women have stage I disease and 19% of women have stage III and IV disease at diagnosis. These differences in stage of cancer at diagnosis between India and HICs likely reflect differences in cancer awareness/education, availability of screening programs, and access to cancer care.

Stage distributions of oral cavity cancer and lung cancer in our study are less different than those reported in HICs. In our cohort, 10% and 67% of men with oral cavity were diagnosed with stage I or stage III and IV cancer, respectively. Whereas the proportion of patients with stage I disease in our study is lower than that in the United States (10% vs 29%), the proportion of patients with stage III and IV cancer is nearly the same.
**FIG 1.** Distribution of stage at diagnosis by educational level among all incident cases of breast, cervical, lung, and oral cavity cancer diagnosed from 2012 to 2014 in Trivandrum District, South India. (A) Breast cancer. (B) Cervical cancer. (C) Oral cavity cancer. (D) Lung cancer. (* Only female breast cancer. † Only male lung and oral cavity cancer.)
disease is similar (67% vs 67%).\(^9\) Distribution of stage at diagnosis for lung cancer observed in our cohort—13% localized and 80% regional/metastatic—is also similar to reports from Canada and the United States, where 65% and 66% of patients have stage III and IV disease, respectively.\(^11,12\)

The primary objective of our study was to evaluate the extent to which stage of cancer at diagnosis was associated with educational status, which was chosen as a surrogate for SES. Parallel work in HICs commonly use neighborhood median household income from census data as a measure for SES.\(^5,13,14\) Comparable census data in South India were lacking and we had concerns regarding the validity of self-reported household income in the Indian context. For these reasons, we elected to use educational level, commonly used in epidemiologic studies as a surrogate for SES.\(^5\)

Public health studies in India have demonstrated that educational level in India is correlated with other measures of SES, including occupation, housing, and social status.\(^16\) Educational level is a strong determinant of future employment and income. Education is also associated with cognitive function and health literacy, which can affect health-seeking behaviors.\(^17\)

Our study results demonstrate a substantial association between low educational status and advanced stage cancer at diagnosis for patients with breast, cervical, and oral cavity cancers. This association was not observed in patients with lung cancer. These results have face validity when one considers the mechanism by which educational status may be related to stage of disease. Patients with greater education who develop early signs of breast cancer (ie a breast lump), cervical cancer (ie, abnormal vaginal bleeding), and oral cancer (ie, a mouth lesion) may have greater awareness and access to health care, thus seeking earlier diagnosis.

| Education Level | Illiterate/Primary | Middle | ≥ Secondary | P       |
|-----------------|--------------------|--------|-------------|---------|
| Breast          | 50 (172/347)       | 39 (148/380) | 36 (269/749) | < .001  |
| Cervix          | 46 (84/184)        | 43 (50/115)  | 24 (21/88)   | .022    |
| Oral cavity     | 77 (199/259)       | 76 (148/195) | 59 (106/180) | < .001  |
| Lung            | 89 (196/221)       | 84 (136/162) | 88 (149/169) | .350    |

NOTE. Data are given as % (No./No.) unless otherwise noted.

*Defined as stage III or IV for breast, cervix, oral cavity, and SEER Clinical Extent of Disease regional/metastatic lung.

**TABLE 5.** Proportion of Patients With Breast, Cervical, Lung, and Oral Cavity Cancer Diagnosed From 2012 to 2014 in Trivandrum District, South India, With Regional/Metastatic Disease at the Time of Diagnosis

| Factor            | Breast (n = 2,283)† | Cervix Uteri (n = 481) | Oral Cavity (n = 797)‡ | Lung (n = 986)‡ |
|-------------------|---------------------|------------------------|------------------------|----------------|
| Age, years        |                     |                        |                        |                |
| ≤ 50              | 346/818 (42)        | 50/95 (53)             | 67/158 (42)            | 76/95 (79)     |
| > 50              | 607/1,465 (43)      | 188/386 (49)           | 331/639 (52)           | 708/959 (79)   |
| Marital status    |                     |                        |                        |                |
| Not married       | 25/61 (41)          | 1/2 (50)               | 16/31 (50)             | 22/249 (92)    |
| Married           | 773/1,805 (43)      | 166/343 (48)           | 365/721 (51)           | 734/912 (81)   |
| Religion          |                     |                        |                        |                |
| Hindu             | 693/1,678 (41)      | 182/372 (49)           | 275/565 (49)           | 585/728 (80)   |
| Muslim            | 116/264 (44)        | 19/42 (45)             | 37/81 (46)             | 121/154 (79)   |
| Christian         | 135/314 (43)        | 37/66 (56)             | 85/147 (58)            | 69/97 (79)     |
| Residence         |                     |                        |                        |                |
| Rural             | 588/1,345 (44)      | .021                   | 159/321 (50)           | 539/681 (79)   |
| Urban             | 365/938 (39)        | .021                   | 79/160 (49)            | 245/305 (80)   |
| Education         |                     |                        |                        |                |
| Illiterate/primary school | 172/347 (50) | < .001 | 84/184 (46) | .002 | 199/259 (77) | .001 | 196/221 (89) | .350 |
| Middle school     | 148/380 (39)        | 50/115 (44)            | 148/195 (76)           | 136/162 (84)   |
| ≥ Secondary school | 269/749 (36)       | 21/88 (48)             | 106/180 (59)           | 149/169 (88)   |

*Defined as stage III and IV disease for breast, cervix, oral cavity cancer, and SEER Clinical Extent of Disease regional/metastatic for lung cancer.
†Only female breast cancer.
‡Only male lung and oral cavity cancer.
medical treatment and receiving a diagnosis with earlier disease than patients from lower educational backgrounds. However, for patients with lung cancer, by the time symptoms manifest, the disease is often already advanced, in which case early treatment-seeking behavior would be expected to have less effect on stage at diagnosis.

In their overview of the existing literature, Woods et al conclude that stage of disease at diagnosis and access to optimal treatment explain a portion of the disparity in the survival of patients with cancer. Whereas much of the literature from the United States has found an association between SES and stage of disease at diagnosis, several large studies conducted in the United Kingdom and Canada have failed to confirm this observation. It remains unknown whether the results in Canada and the United Kingdom are different from the findings in the United States because of the availability of universal health insurance in the former but not the latter. We are not aware of any studies in LMICs exploring the association between stage, SES, and survival. Results presented in this study represent the first step in doing so. Future work should explore survival differences across socioeconomic groups and the relative extent to which stage and treatment delivery explain any observed survival differences.

Several hospital-based studies in India have explored the association between educational level and stage of cancer at diagnosis. Sathwara et al evaluated the association between sociodemographic factors and late-stage disease among 1,210 women with breast cancer who were treated at Tata Memorial Hospital (TMH) in 2008. Fifty-four percent of all women had stage III and IV disease at diagnosis among 765 women who were treated at TMH in 2007. Fifty-two percent of all women had stage III and IV disease at diagnosis. This rate was higher among illiterate patients than those reported in our study (37% and 39%), which may themselves reflect regional differences in SES between the Indian states of Maharashtra and Kerala.

Two studies have used the Hospital-Based Cancer Registry at the RCC in Trivandrum, Kerala, to explore this issue. Among 522 patients with breast cancer seen at RCC in 2006, Ali et al found that advanced disease stage was more common among women who were not married and who had lower educational levels. Kaku et al reported similar findings in their study of 349 women with cervical cancer seen at RCC in 2006. In their systematic review of barriers to breast cancer care in LMICs, Sharma et al identified eight studies (sample size range, 66 to 903) that demonstrated an association between SES and delayed presentation of breast cancer. The authors concluded that there was strong evidence that lower income and education contributed to barriers in breast cancer care. The existing literature is limited by the fact that, with one exception, each of the studies in the meta-analysis were institution based and therefore potentially limited by referral bias and selection bias. Given that lower SES is likely a major barrier to accessing cancer health services in LMICs, a population-based study, rather than a hospital-based study, is the optimal design. This is a notable strength of our current study.

Our study does have methodological limitations that warrant comment. As shown in Appendix Table A1, stage and educational level were more likely to be missing among those patients who were treated at the Government Medical College Hospital or other smaller institutions compared with the RCC in Trivandrum. This may limit the generalizability of our study results. Moreover, our study cannot explain why patients with lower educational status are more likely to have advanced cancer. It is likely that this reflects a delay in seeking medical attention, but future qualitative work is needed to explore this more fully to use this knowledge to improve the care of patients from impoverished backgrounds.

Cancer has emerged as a major cause of morbidity and mortality in LMICs where patients are more likely to be diagnosed with advanced cancer than in HICs. Optimal development and implementation of cancer control systems requires an understanding of existing patterns of disease, treatment, and outcomes in LMICs. Health services research can make use of existing data in cancer registries to provide unique insights into health system performance in LMICs. Whereas the association between SES, stage of cancer, and survival has been well described in HICs, there are limited reports from LMICs. Data from this study suggest that future health awareness, preventive, and treatment interventions in Kerala will need to specifically target less-educated communities in an effort to minimize delays in seeking medical care for cancer.

**AFFILIATIONS**

1. Regional Cancer Centre, Trivandrum, India
2. Medical College, Trivandrum, India
3. Queen’s University, Kingston, Ontario, Canada

**CORRESPONDING AUTHOR**

Aleyamma Mathew, PhD, Department of Epidemiology, Regional Cancer Centre, Trivandrum 695 011, Kerala, India; e-mail: aleyammamathewrcc@gmail.com.
AUTHOR CONTRIBUTIONS

Conception and design: Aleyamma Mathew, Aswin Kumar, Sivasevan Roshni, Christopher M. Booth

Financial support: Christopher M. Booth

Administrative support: Aleyamma Mathew, Preethi Sara George, Sivasevan Roshni, Christopher M. Booth

 Provision of study material or patients: Aleyamma Mathew, Kunnambath Ramadas, Beela Sarah Mathew, Aswin Kumar, Sivasevan Roshni, Krishnan Nair Lalithamma Jayakumar, Christopher M. Booth

 Collection and assembly of data: Aleyamma Mathew, Preethi Sara George, Sivasevan Roshni, Krishnan Nair Lalithamma Jayakumar, Christopher M. Booth

 Data analysis and interpretation: Aleyamma Mathew, Preethi Sara George, Kunnambath Ramadas, Beela Sarah Mathew, Aswin Kumar, Sivasevan Roshni, Christopher M. Booth

REFERENCES

1. Kogevinas M, Pearce M, Susser M, et al: (eds): Socioeconomic Differences in Cancer Survival: A Review of the Evidence. Lyon, France, IARC Scientific Publications, 1997, pp 177-206

2. Woods L, Rachet B, Coleman MP: Origins of socio-economic inequalities in cancer survival: A review. Ann Oncol 17:5-19, 2006

3. Boyd C, Zhang-Salmons JY, Groome PA, et al: Associations between community income and cancer survival in Ontario, Canada, and the United States. J Clin Oncol 17:2244-2255, 1999

4. Mackillop WJ, Zhang-Salmons J, Boyd CJ, et al: Associations between community income and cancer incidence in Canada and the United States. Cancer 89:901-912, 2000

5. Booth CM, Li G, Zhang-Salmons J, et al: The impact of socioeconomic status on stage of cancer at diagnosis and survival: A population-based study in Ontario, Canada. Cancer 116:4160-4167, 2010

6. Office of the Registrar General and Census Commissioner of India: 2011 census data. http://www.censusindia.gov.in

7. Yeole BB, Kumar AV, Kurkure A, et al: Population-based survival from cancers of breast, cervix and ovary in women in Mumbai, India. Asian Pac J Cancer Prev 5:308-315, 2004

8. Walters S, Maringe C, Butler J, et al: Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: A population-based study. Br J Cancer 108:1195-1208, 2013

9. Voruganti T, Moineddin R, Jembere N, et al: Comparing stage of diagnosis of cervical cancer at presentation in immigrant women and long-term residents of Ontario: A retrospective cohort study. CMAJ Open 4:E424-E430, 2016

10. National Cancer Institute: Cancer Stats Facts: Oral cavity and pharynx cancer. https://seer.cancer.gov/statfacts/html/oralcav.html

11. Morgenstern D, Ng SH, Gao F, et al: Trends in stage distribution for patients with non-small cell lung cancer: A National Cancer Data Base survey. J Thorac Oncol 5:29-33, 2010

12. Canadian Partnership Against Cancer: Distribution of cases by stage at diagnosis for non-small cell lung cancer. http://www.systemperformance.ca/disease-sites/lung/stage-distribution/

13. Byers TE, Wolf HJ, Bauer KR, et al: The impact of socioeconomic status on cancer incidence and stage at diagnosis: Selected findings from the Surveillance, Epidemiology, and End Results: National Longitudinal Mortality Study. Cancer Causes Control 20:417-435, 2009

14. Thomson CS, Hole DJ, Twelves CJ, et al: Prognostic factors in women with breast cancer: Distribution by socioeconomic status and effect on differences in survival. J Epidemiol Community Health 55:308-315, 2001

15. Brewster DH, Thomson CS, Hole DJ, et al: Relation between socioeconomic status and tumour stage in patients with breast, colorectal, ovarian, and lung cancer: Results from four national, population-based studies. BMJ 322:830-831, 2001

16. Sathwara JA, Balasubramaniam G, Bobbey SC, et al: Sociodemographic factors and late-stage diagnosis of breast cancer in India: A hospital-based study. Indian J Med Paediatr Oncol 38:277-281, 2017

17. Jain A, Ganesh B, Bobbey SC, et al: Sociodemographic and clinical profile of cervical cancer patients visiting in a tertiary care hospital in India. Indian J Med Paediatr Oncol 38:291-295, 2017

18. Ali R, Mathew A, Rajan B: Effects of socio-economic and demographic factors in delayed reporting and late-stage presentation among patients with breast cancer in a major cancer hospital in South India. Asian Pac J Cancer Prev 9:703-707, 2008

19. Kaku M, Mathew A, Rajan B: Impact of socio-economic factors in delayed reporting and late-stage presentation among patients with cervix cancer in a major cancer hospital in South India. Asian Pac J Cancer Prev 9:589-594, 2008

20. Sharma K, Costas A, Shulman LN, et al: A systematic review of barriers to breast cancer care in developing countries resulting in delayed patient presentation. J Oncol 2012:121873, 2012

21. Pineros M, Sánchez R, Cendales R, et al: Patient delay among Colombian women with breast cancer. Salud Publica Mex 51:372-380, 2009
27. Sullivan R, Badwe RA, Rath GK, et al: Cancer research in India: National priorities, global results. Lancet Oncol 15:e213-e222, 2014
28. Sankaranarayanan R, Boffetta P: Research on cancer prevention, detection and management in low- and medium-income countries. Ann Oncol 21:1935-1943, 2010
29. Coleman MP: Cancer survival: Global surveillance will stimulate health policy and improve equity. Lancet 383:564-573, 2014
## Table A1. Data Capture Rate For Incident Cases in Trivandrum PBCR From 2012 to 2014

| Variable | Breast | Cervix Uteri | Lung | Oral Cavity |
|----------|--------|--------------|------|-------------|
| Total cases in PBCR, No. | 2,283 | 481 | 986 | 797 |
| Cases from RCC, No. | 1,381 | 380 | 490 | 617 |
| No. (%) with stage | 1,353 (98) | 365 (96) | 481 (98) | 608 (99) |
| No. (%) with education level | 1,373 (100) | 377 (99) | 485 (99) | 611 (99) |
| Cases from Medical College, No. | 530 | 57 | 381 | 122 |
| No. (%) with stage | 409 (77) | 38 (67) | 350 (92) | 77 (63) |
| No. (%) with education level | 61 (12) | 4 (7) | 48 (13) | 12 (10) |
| Cases from other sources, No. | 372 | 44 | 115 | 58 |
| No. (%) with stage | 130 (35) | 20 (46) | 82 (71) | 20 (35) |
| No. (%) with education level | 42 (12) | 6 (14) | 19 (17) | 11 (19) |

NOTE. Stage was determined by TNM for breast and oral cavity, International Federation of Gynecology and Obstetrics for cervix, SEER Clinical Extent of Disease for lung cancer.

Abbreviations: PBCR, population-based cancer registry; RCC, Regional Cancer Centre.

*Breast cancer is female only.

†Lung and oral cavity cancer are male only.