


BMJ Open Geographic variation in tobacco use in India: a population-based multilevel cross-sectional study

Ankur Singh ,^{1,2} Monika Arora,³ Rebecca Bentley,^{1,2} Matthew J Spittal,⁴ Loc G Do,⁵ Nathan Grills,^{6,7} Dallas R English^{2,8}

To cite: Singh A, Arora M, Bentley R, *et al.* Geographic variation in tobacco use in India: a population-based multilevel cross-sectional study. *BMJ Open* 2020;**10**:e033178. doi:10.1136/bmjopen-2019-033178

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-033178>).

Received 24 July 2019

Revised 22 April 2020

Accepted 13 May 2020

ABSTRACT

Objective This study aims to quantify the extent to which people's use of tobacco products varies by local areas (city ward and village) across India and the variation in this clustering by tobacco products.

Design Cross-sectional study.

Setting and participants Data on 73 954 adults across 2547 city wards and villages were available for analysis from 30 states and 2 union territories in India.

Primary and secondary outcome measures We included as primary outcomes self-reported any tobacco use, current cigarette smoking, current bidi smoking, current smokeless tobacco use and a derived variable for dual use describing respondents who engaged in both smoking and smokeless tobacco use.

Results The median risk of an individual using tobacco was 1.64 times greater if a person hypothetically moved from an area of low to high risk of tobacco use (95% CI: 1.60 to 1.69). Area-level partitioning of variation differed by tobacco product used. Median ORs ranged from 1.77 for smokeless tobacco use to 2.53 for dual use.

Conclusions Tobacco use is highly clustered geographically in India. To be effective in India, policy interventions should be directed to influence specific local contextual factors on adult tobacco use. Where people live in India influences their use of tobacco, and this association may be greater than has been observed in other settings. Tailoring tobacco control policies for local areas in India may, therefore, provide substantial public health benefits.

INTRODUCTION

Four-fifths of the world's current smokers reside in low/middle-income countries (LMICs), creating enormous societal and public health challenges.¹ The number of deaths from tobacco-related causes and loss of productivity are rapidly increasing in these often resource-poor settings.^{2,3}

The latest Indian Global Adult Tobacco Survey (GATS 2016-17) found that nearly 30% of all Indian adults use tobacco.⁴ Additionally, the widespread use of smokeless tobacco presents a complex challenge for health systems and tobacco control because of its strong relationship with oral cancerous

Strengths and limitations of this study

- First time application of multilevel analysis to quantify variations in tobacco use among local areas in a low- and middle-income country.
- Multiple measures were estimated (intra-class correlation coefficient, median OR and discriminatory accuracy) that corroborated the importance of local areas in determining tobacco use.
- Large and nationally representative data on tobacco use were analysed.
- Individual-level policy and economic variables were excluded to avoid the atomistic fallacy.
- Lacking area-level variables restricted analysis of their role in area variations in tobacco use.

and precancerous lesions.⁵ Despite a nationwide smokeless tobacco ban implemented in 2013–2014, 20% of all tobacco users are smokeless tobacco users.⁴ Added to this, the burden of tobacco use in India is disproportionately high among socially disadvantaged people.^{6–8}

There is consistent evidence that local social and policy contexts shape patterns of tobacco use.⁹ Multilevel studies (that simultaneously examine individual-level and group-level determinants of health) from The Netherlands, Australia, South Africa, Mexico, Scotland, India, the USA and the UK suggest evidence of an association between area-level contextual factor (such as social disadvantage and local policy environments) and smoking at the individual level.^{10–22} For example, a study of Indian high school students from Mumbai reported the density of tobacco vendors around schools was associated with increased tobacco use by students.²³

Notably, the majority of multilevel studies on tobacco use to date investigate associations between specific area-level exposures and tobacco use (the specific contextual effect). Such models are used simply as an extension of single-level regression models



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Ankur Singh;
ankur.singh@unimelb.edu.au

enabling them to handle group-level variables as exposures and covariates. Variation in tobacco use across contexts (general contextual effects) can also be examined using multilevel models. Yet, this aspect of multilevel analysis has been underused in research to date.^{24 25} Using this approach, it is possible to describe the extent of geographic inequalities in tobacco use drawing attention to underlying contextual drivers unaddressed through individually directed interventions.^{26–29} This is important information. Tobacco control interventions targeting specific area-level exposures will only be effective if areas share significant interindividual variation in tobacco use.^{24 25}

To redress this important gap in evidence, this study aims to quantify the extent to which people's use of tobacco products varies by local areas (city ward and village) across India and the variation in this clustering by tobacco product.

METHODS

Study population

Data on tobacco use in India was obtained from the GATS-2 conducted in 2016 and 2017. GATS-2 is a multi-country household tobacco prevalence survey designed to support the implementation of tobacco control within study countries.⁴ Participants eligible for the survey were non-institutionalised individuals aged 15 years and older. The survey applied a multistage sampling procedure with different sampling hierarchies for urban and rural areas. For urban areas, city wards were the primary sampling unit from which census enumeration blocks, and then households, were selected. In rural areas, the primary sampling units were villages, from which households were selected. A total of 73954 adults across 2547 city wards and villages were available for analysis from 30 states and 2 union territories in India. The response rate was 93%.⁴

Data collection

GATS-2 collected data using household and individual questionnaires developed in English and translated into 19 regional languages. The interviewer-administered questionnaires collected data on demographic characteristics, tobacco smoking, smokeless tobacco use, secondhand smoke, socioeconomic position, media and knowledge, attitude and perceptions related to tobacco use. More details on sampling procedures and methods of data collection are published elsewhere.^{4 30 31}

Outcomes

We included as primary outcomes self-reported any tobacco use, current cigarette smoking, current bidi smoking, current smokeless tobacco use and a derived variable for dual use describing respondents who engaged in both smoking and smokeless tobacco use. Participants were asked 'On average, how many of the following products do you currently smoke each day?'.^{4 30 31} We categorised those who reported smoking one or more manufactured/

rolled tobacco in paper/leaf as current cigarette smokers. Similarly, we identified those who reported smoking one or more bidi as current bidi smokers. Regarding smokeless tobacco use, participants were asked 'Do you currently use smokeless tobacco on a daily basis, less than daily or not at all?'.^{4 30 31} We recorded those answering 'daily' or 'less than daily' as yes for current smokeless tobacco use. Those identified to be both current smokers (cigarette or bidi) and current smokeless tobacco users were identified as dual users. Therefore, we created five binary variables including any tobacco use, current cigarette smokers, current bidi smokers, current smokeless tobacco users and dual users.

Geographic level of aggregation (local areas)

Individuals from urban areas were clustered within city wards and those in rural areas were clustered within villages. In urban areas, city wards are the units for local government operations in India, responsible for essential community services including healthcare, education, housing, transport and so on.³² In rural areas, villages make up the boundary for local panchayat (traditional local self-governance).³²

Covariates

To account for compositional differences in populations within area-level clusters, we included individual-level demographic characteristics: age (as a continuous variable), sex and socioeconomic position (education: no formal education/less than primary/primary/secondary or more; occupation: unemployed/labourer/housewife, retired, student/self-employed/private/government; and household-level wealth: quintiles, 1=lowest, 5=highest) as covariates in the multilevel regression models. These variables were selected based on a previous study.⁷

Statistical analysis

We performed statistical analyses using Stata V.15.0 (Statacorp). We used survey commands to account for the complex survey design and to perform the weighted descriptive analysis. We plotted the prevalence and 95% CIs for any tobacco use and for different types of tobacco use to visually examine their variation by local areas. We fitted multilevel logistic regression models with random intercepts for local areas and fixed slopes with individuals nested in city wards or villages, respectively. Multilevel models operationalise studying population-level variations in health outcomes by examining the extent of clustering in health outcomes that exists at the group or contextual level.^{24 33–37} Using intraclass correlation coefficients (ICCs) and median ORs (MORs), we decomposed the variance in tobacco use at the city ward or village level. The ICC is expressed as a percentage and is interpreted in these analyses as the share of interindividual variation in health outcomes that exists at the group level. For example, an ICC of 8% at the village level means that of all the individual-level variation in tobacco use among rural areas, 8% is attributed to the

village level. The higher the individual correlation in health outcomes within a context, the more relevant is the context for understanding individual differences in the health outcome.²⁵ We estimated the MOR, which describes the area-level variance as an OR, as the median value of the distribution of ORs obtained when two individuals with the same covariate values are picked from two different areas, comparing the one from the higher prevalence area with the one from the area with lower prevalence.^{24 36 38} In the absence of any area-level variation, the MOR is equal to 1. We estimated both MORs and ICCs for binary outcomes as the partition of variance between different levels does not have the same intuitive interpretation as a linear model.^{24 38} We estimated ICCs and MORs from intercept-only models to examine the presence of clustering and heterogeneity between areas in the outcomes of tobacco use.

We also applied an alternate method to examine the relevance of area-level contexts for tobacco use by comparing discriminatory accuracies obtained from fitted single-level and multilevel logistic regression models. The area under the receiver operating characteristic curve was constructed by plotting the true-positive fraction (sensitivity) against the false-positive fraction (1-specificity). It measures the ability of the model to classify individuals with and without the outcome and takes a value between 0.5 and 1.0 where 1.0 is perfect discrimination and at 0.5 covariates have no predictive power.²⁴

We did this in three stages. First, we fitted a single-level logistic regression model with tobacco use as the outcome and included individual-level covariates (age, sex, education, household wealth and occupation) (model A). The ability of this model to classify tobacco use was quantified using the area under curve (AUC). Next, we fitted a multi-level logistic regression model (model B) for tobacco use that included the same individual-level covariates. In addition to quantifying the change in the AUC from Model A, MORs and ICCs were estimated from model B to examine the general contextual effect of areas. Finally, we added area of residence and states in model C as area-level covariates to examine any changes in AUC, MOR and ICCs.

We assessed goodness of fit by estimating the changes in the deviance information criterion (DIC). All models were fitted separately for each type of tobacco use (cigarette smoking, bidi smoking, smokeless tobacco use and dual use) to determine any differences in variations in tobacco use according to different types of tobacco use. We performed a sensitivity analysis to examine clustering in tobacco use in city wards and villages within states by fitting three-level hierarchical models: individual nested within city wards and villages nested within states.

Patient and public involvement

No patients or public were involved in this study.

Table 1 Descriptive characteristics of the sample (n=73 954)

| Variable | Categories | Percentage |
|-------------------|------------------------------------------|------------|
| Age (years) | 15–30 | 41.7 |
| | 31–45 | 29.7 |
| | 46–60 | 17.7 |
| | 61–75 | 8.9 |
| | ≥76 | 1.8 |
| Sex | Male | 51.1 |
| | Female | 48.9 |
| Wealth | Poorer | 23.4 |
| | Poor | 36.5 |
| | Middle | 15.0 |
| | Rich | 12.2 |
| | Richer | 12.9 |
| Education | No formal education | 26.4 |
| | Less than primary | 9.2 |
| | Primary | 28.2 |
| | Secondary or more | 36.2 |
| Occupation | Unemployed | 4.3 |
| | Labour | 21.2 |
| | Housewife/retired/student | 44.1 |
| | Self | 19.4 |
| | Private | 8.3 |
| | Government job | 2.7 |
| Area of residence | Urban | 34.5 |
| | Rural | 65.5 |
| Tobacco use | Non-user | 72.2 |
| | Cigarette smoking | 1.3 |
| | Bidi smoking | 4.6 |
| | Smokeless tobacco use | 18.6 |
| | Dual use (smokeless tobacco use+smoking) | 2.8 |
| | Bidi+cigarette | 0.5 |

Weighted percentages (using survey weights).

RESULTS

We analysed data for 73 954 individuals (99.9%) out of 74 037 survey participants. We did not analyse data on 83 participants due to missing covariates data. **Table 1** shows the descriptive characteristics of the sample according to residence status. Twenty-eight per cent of adults used tobacco products. The prevalence of smokeless tobacco use was 18.6% (**table 1**). Plots for prevalence and 95% CI for any tobacco use and different types of tobacco use by local areas showed substantial variations (online supplementary appendix).

Intercept-only models (null models with no covariate adjustment) estimated 22% (95% CI: 20% to 24%) of any

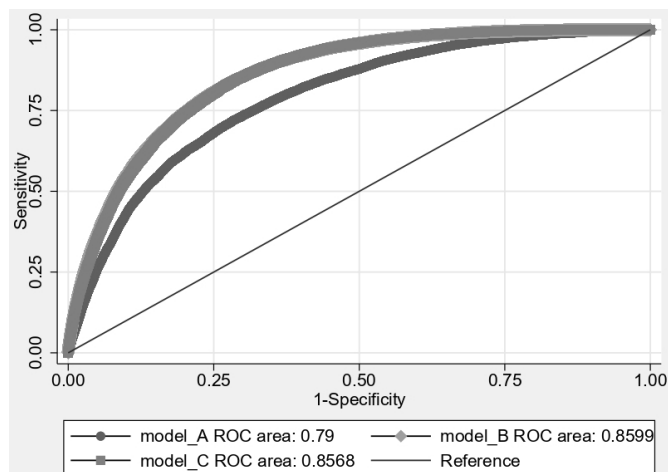


Figure 1 Area under the receiver operating characteristic (ROC) curve for tobacco use plotted separately for single and multilevel logistic regression models.

tobacco use was clustered at the city ward and village level. Cigarette smoking was clustered at 31%, bidi smoking at 28%, dual use at 40% and smokeless tobacco at 36%, respectively (estimates not reported in the tables). For each outcome, the AUC increased when multilevel logistic regression models were fitted. The AUC increased to 0.86 in model B as compared with 0.79 in a single-level logistic regression model (model A) implying the presence of a general contextual effect and the ability to better classify individuals according to tobacco use (figure 1). Changes

in AUC were highest for smokeless tobacco use 11% compared with 2% for cigarette smoking (tables 2 and 3).

After including all individual-level covariates, the proportion of variance attributable to the areas remained at 21% (95% CI: 20% to 22%) (table 2). Correspondingly, the MOR was 2.43 (95% CI: 2.35 to 2.52). These results suggest that the median odds of tobacco use are more than double for two individuals with the same covariates when comparing the one from city ward or village with high tobacco use to the other from a city ward or village with low tobacco use. Including the area of residence and state in model C substantially reduced the estimates of the proportion of variance attributable to areas and the respective MORs. The proportion of variance for any tobacco use reduced from 21% to 7.6% and the corresponding MOR from 2.42 to 1.64. Sensitivity analysis confirmed our findings of high clustering in any tobacco use within city wards or villages from the same state (online supplementary appendix, pp. 21).

The decrease in DIC values between the single-level models and multilevel models including covariates suggested better model fit (table 2).

Among the different types of tobacco use, the highest ICC (22%; 95% CI: 19% to 26%) and MOR (2.53; 95% CI: 2.32 to 2.74) were for dual use and the lowest for SLT use (ICC: 10%; 95% CI: 9% to 11% and MOR: 1.77; 95% CI: 1.71 to 1.83) (table 3). Similar to any tobacco use, substantial reductions in estimates of ICC and MOR were observed on inclusion of state and area of residence in model C compared with model B.

Table 2 Multilevel logistic regression models for any tobacco use among Indian adults (n=73 954 individuals nested in 2547 city wards and villages)

| | Model A Estimate 95% CI | Model B Estimate 95% CI | Model C Estimate 95% CI |
|------------|-------------------------------|-------------------------------|-------------------------------|
| AUC | 0.79 | 0.86 | 0.86 |
| AUC change | | 0.07 | 0 |
| Variance | | 0.87 (0.80 to 0.94) | 0.27 (0.24 to 0.30) |
| ICC | | 21% (20 to 22) | 8% (7 to 9) |
| MOR | | 2.43 (2.35 to 2.52) | 1.64 (1.60 to 1.69) |
| DIC | 71 171.7 | 66 619.6 | 64 702.3 |
| DIC change | | -4552.1 | -1917.3 |

Model A: single-level logistic regression model (covariates included: age, sex, area of residence, education, occupation and wealth). Model B: multilevel logistic regression model (covariates included: age, sex, education, occupation and wealth). Model C: multilevel logistic regression model (covariates included: age, sex, education, occupation, wealth, and area of residence and states). Reference group: no tobacco use. AUC, area under curve; DIC, deviance information criterion; ICC, intraclass correlation coefficient; MOR, median OR.

DISCUSSION

We found substantial variation in tobacco use across local areas in India. Individual-level social and demographic characteristics were not able to explain the high area-level variations in tobacco. Including states and areas of residence explained substantial area-level variation in tobacco use. However, the remaining variation in tobacco use was still high, indicating the importance of local areas. The degree of area-level variation in tobacco use differed according to the types of tobacco products. Dual use (smoking and smokeless) had the highest geographic clustering.

Strengths and limitations

This study had several strengths and limitations. To the best of our knowledge, this is the first study from LMICs that has studied variation in tobacco use at the local area level using a nationwide representative data.⁴ By using different measures (ICC, MOR and AUC), we not only inform the extent of variation but we comprehensively examine the degree of clustering, the heterogeneity in outcomes among areas as well as the ability of local areas to classify individuals according to tobacco use.^{24 33–36} This study also has limitations. We did not incorporate policy and economic variables related to tobacco use available in the GATS 2016–2017 in our analysis because the policy

Table 3 Multilevel logistic regression models for different types of tobacco use among Indian adults (n=73 954 individuals nested in 2547 city wards and villages)

| | Model A Estimate 95% CI | Model B Estimate 95% CI | Model C Estimate 95% CI |
|-------------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Cigarettes (n=54 648) | | | |
| Variance | | 1.44 (1.19 to 1.70) | 0.53 (0.38 to 0.68) |
| ICC (%) | | 30 (27 to 34) | 14 (11 to 18) |
| MOR | | 3.14 (2.83 to 3.46) | 2.00 (1.81 to 2.20) |
| DIC | 10 630.4 | 10 175.0 | 9480.5 |
| DIC change | | -455.4 | -694.5 |
| AUC | 0.68 | 0.70 | 0.69 |
| AUC change | | 0.02 | -0.01 |
| Bidi (n=56 814) | | | |
| Variance | | 1.53 (1.33 to 1.72) | 0.65 (0.53 to 0.76) |
| ICC (%) | | 32 (29 to 35) | 16 (14, 19) |
| MOR | | 3.25 (3.01 to 3.49) | 2.15 (2.01 to 2.30) |
| DIC | 18 822.5 | 17 680.8 | 16 765.4 |
| DIC change | | -1141.7 | -915.4 |
| AUC | 0.89 | 0.95 | 0.94 |
| AUC change | | 0.06 | -0.01 |
| Smokeless Tobacco (n=66 089) | | | |
| Variance | | 1.46 (1.34 to 1.59) | 0.36 (0.31 to 0.40) |
| ICC (%) | | 31 (29 to 33) | 10 (9 to 11) |
| MOR | | 3.17 (3.01 to 3.32) | 1.77 (1.71 to 1.83) |
| DIC | 56 207.3 | 51 179.1 | 48 915.1 |
| DIC change | | -5028.1 | -2264.0 |
| AUC | 0.76 | 0.87 | 0.86 |
| AUC change | | 0.11 | -0.01 |
| Dual use (n=55 522) | | | |
| Variance | | 2.41 (2.09 to 2.72) | 0.95 (0.78 to 1.12) |
| ICC (%) | | 42 (39 to 45) | 22 (19 to 26) |
| MOR | | 4.39 (3.96 to 4.82) | 2.53 (2.32 to 2.74) |
| DIC | 14 335.7 | 12 989.8 | 12 045.9 |
| DIC change | | -1345.9 | -943.9 |

Continued

Table 3 Continued

| | Model A Estimate 95% CI | Model B Estimate 95% CI | Model C Estimate 95% CI |
|------------|-------------------------------|-------------------------------|-------------------------------|
| AUC | 0.88 | 0.96 | 0.95 |
| AUC change | | 0.08 | -0.01 |

Model A: single-level logistic regression model (covariates included: age, sex, area of residence, education, occupation and wealth). Model B: multilevel logistic regression model (covariates included: age, sex, education, occupation and wealth). Model C: multilevel logistic regression model (covariates included: age, sex, education, occupation, wealth, area of residence and states). Reference group: no tobacco use. AUC, area under curve; DIC, deviance information criterion; ICC, intraclass correlation coefficient; MOR, median OR.

and economic variables were the respondent's perceptions rather than objective measures of availability and implementation of policies in local areas and because these data were only gathered from smokers. The non-ecological nature of these variables could lead to falsely attributing individual-level measures to area levels (the atomistic fallacy).³⁷

Discussion in context of current evidence

Our findings of high variations in tobacco use among local areas is new. A multilevel study on societal determinants of tobacco use from Scotland found no evidence of clustering in tobacco use at the area level.¹⁸ Other multilevel studies have not presented measures of variance, which limits comparisons.¹⁰⁻²² Our findings indicate much higher clustering of tobacco use at the area level than has previously been reported, suggesting that local area contexts and contextual determinants are highly relevant in India. Such variations, we speculate in the absence of data and available literature,¹⁰⁻²² may be due to differences in the availability and implementation of tobacco control policies, social environment (deprivation, area-level mean income, area-level income inequality and social capital) and shared cultural and social norms regarding tobacco use among people within an area.

Tobacco-specific variations in the values of ICC and MOR highlight potential differences in the relevance of contexts by type of tobacco product used. Evidence from other studies suggests that although wealthier and more educated individuals have higher odds of cigarette smoking than their disadvantaged counterparts, disadvantaged individuals have higher odds of bidi smoking and smokeless tobacco use.^{6, 7} Our study highlights the presence of both individual and geographic socioeconomic inequalities in tobacco use by products. For example, we observed a higher effect of individual social and demographic characteristics in smokeless tobacco use when compared with cigarette smoking and bidi smoking for contextual effects (change in ICC from 36% in the null model to 10% in the adjusted model). In addition, the proportion of variation for all types of tobacco use was



markedly explained by adding states into the model. This emphasises the role of cultural and regional diversity within India in determining tobacco use.³⁹ Both ICCs from the three-level hierarchical models and ORs estimated from regression models confirmed the pivotal role played by states in geographic inequities in tobacco use in India (see the online supplementary appendix).

Research and policy implications

Given the role of contexts in shaping individual health behaviours, this study builds a framework for operationalising contextual thinking in tobacco control activities, particularly in LMICs where social norms and cultural aspects may differ from high-income countries. High general contextual effects of local areas for tobacco use necessitates a thorough examination of factors at the area level that may be causally associated with individual tobacco use as well as those which can explain the high variations in tobacco use among local areas. This may only be possible if either data on individual-level tobacco use are linked with small area characteristics or if future population-based surveys collect both area-level and individual-level data relevant to tobacco use. Given the findings from our study, in future GATS surveys should consider the opportunities to comprehensively study both individual-level and area-level determinants of tobacco use within India and in other LMICs. First, it would be helpful if wards and villages were identifiable in future versions of GATS so that researchers and policy-makers can link in area-level covariates (social, policy, economic and physical environment) to examine their effects on tobacco use. Second, it would be useful if the administrative levels at which tobacco-related policies are implemented were recorded, allowing the examination of variation in tobacco use across multiple levels of geographical hierarchy. This would further help policy-makers compare clusters from an intervention perspective. Finally, identification of city wards and villages would also allow linking data to relevant area-level social, demographic, economic and policy variables increasing the ability to simultaneously examine area-level and individual-level determinants of tobacco use. Furthermore, current findings build the platform for more robust population-based studies that collectively examine area-level and individual-level determinants of tobacco use in India and other LMICs.

This study has several policy implications. Our findings confirm that context plays an important role in determining the use of tobacco. India's Cigarettes and other Tobacco Products Act (COTPA) is a national law, which is in line with the WHO's Framework Convention on Tobacco Control. States at the subnational level are responsible for implementing various tobacco control policy measures under COTPA. Comparison of GATS-2 and GATS-1, and household surveys, has highlighted changes in the prevalence of tobacco use due to differential implementation of these measures.^{40 41} States are also allowed to develop context-specific information, education and communication resources to match the local needs.^{41 42} Therefore,

health promotion and tobacco control interventions must be designed for contexts and applied contextually rather than being individually oriented.^{9 43} There is the potential to enhance National Tobacco Control Program's (NTCP) implementation at city ward, village and block level as well.⁴¹ NTCP is rolled out in 612 districts across 36 states/union territories in India and has a three-tier structure: National-, State- and District Tobacco Control Cell. District Tobacco Control Cells are established to train key stakeholders; information, education and communication activities; school programmes; monitor tobacco control laws; strengthen cessation facilities and coordinate tobacco control activities with Panchayati Raj (traditional local self-governance).⁴² High local-area variations in tobacco use reported in our study imply extending this structure more locally to city wards and villages to maximise public health benefits. Finally, our use of the multi-level approach in this study advances a 'proportionate universalism' approach suggesting tobacco control interventions applied nationally should be scaled according to local area-level disadvantage to reduce geographic inequalities.

CONCLUSIONS

Where people live in India influences their use of tobacco, and this association may be greater than has been observed in other settings. Tailoring tobacco control policies for local areas in India may, therefore, provide substantial public health benefits.

Author affiliations

¹Centre for Health Equity, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia

²Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia

³Health Promotion Division, Public Health Foundation of India, Gurugram, Haryana, India

⁴Centre for Mental Health, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia

⁵Australian Research Centre for Population Oral Health (ARCPOH), Adelaide Dental School, The University of Adelaide, Adelaide, South Australia, Australia

⁶Nossal Institute for Global Health, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia

⁷Australia India Institute, The University of Melbourne, Melbourne, Victoria, Australia

⁸Cancer Council Victoria, Melbourne, Victoria, Australia

Twitter Ankur Singh @drankursingh99

Contributors AS: conceptualised the study, acquisition of data, analysed the data, interpreted the results, led the manuscript preparation and the submission process. MA: interpreted local policy implications of the results and drafts of the manuscript. RB and NG: gave critical inputs on multiple draft of the manuscript and interpreted the results. MS and LD: developed analytical framework, interpreted the results and revised the manuscript. DRE: reviewed early and advanced drafts of the manuscript, developed analysis plan and interpreted the results.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Matthew Spittal is a recipient of an Australian Research Council Future Fellowship (project number: FT180100075) funded by the Australian Government. Rebecca Bentley was supported by an Australian Research Council Future Fellowship (FT150100131) funded by the Australian Government.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. De-identified data is available from GTSS (<https://www.cdc.gov/tobacco/global/gtss/index.htm>).

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Ankur Singh <http://orcid.org/0000-0003-1336-6493>

REFERENCES

- 1 WHO. *Tobacco: fact sheet*. Geneva: World Health Organization, 2018. <https://www.who.int/en/news-room/fact-sheets/detail/tobacco>
- 2 Hipple Walters B, Petrea I, Lando H. Tobacco control in low- and middle-income countries: changing the present to help the future. *J Smok Cessat* 2018;13:187–8.
- 3 Gilmore AB, Fooks G, Drope J, *et al*. Exposing and addressing tobacco industry conduct in low-income and middle-income countries. *Lancet* 2015;385:1029–43.
- 4 Tata Institute of Social Sciences (TISS) MaMoHaFW, Government of India. *Global adult tobacco survey GATS 2 India 2016-17*. New Delhi: Tata Institute of Social Sciences (TISS), Mumbai and Ministry of Health and Family Welfare, Government of India, 2018.
- 5 Gupta PC, Arora M, Sinha D, *et al*. *Smokeless tobacco and public health in India*. Ministry of Health & Family Welfare. New Delhi: Government of India, 2016.
- 6 Corsi DJ, Subramanian SV. Divergent socio-economic gradients in smoking by type of tobacco use in India. *Int J Tuberc Lung Dis* 2014;18:122–4.
- 7 Singh A, Arora M, English DR, *et al*. Socioeconomic gradients in different types of tobacco use in India: evidence from global adult tobacco survey 2009-10. *Biomed Res Int* 2015;2015:837804.
- 8 Bhan N, Srivastava S, Agrawal S, *et al*. Are socioeconomic disparities in tobacco consumption increasing in India? A repeated cross-sectional multilevel analysis. *BMJ Open* 2012;2:001348.
- 9 Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav* 1995;Spec No:80–94.
- 10 Rachele JN, Wood L, Nathan A, *et al*. Neighbourhood disadvantage and smoking: examining the role of neighbourhood-level psychosocial characteristics. *Health Place* 2016;40:98–105.
- 11 Ellaway A, Macintyre S. Are perceived neighbourhood problems associated with the likelihood of smoking? *J Epidemiol Community Health* 2009;63:78–80.
- 12 Morris T, Manley D, Van Ham M. Context or composition: how does neighbourhood deprivation impact upon adolescent smoking behaviour? *PLoS One* 2018;13:e0192566.
- 13 Mayne SL, Auchincloss AH, Moore KA, *et al*. Cross-Sectional and longitudinal associations of neighbourhood social environment and smoking behaviour: the multiethnic study of atherosclerosis. *J Epidemiol Community Health* 2017;71:396–403.
- 14 Chuang Y-C, Cubbin C, Ahn D, *et al*. Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. *J Epidemiol Community Health* 2005;59:568–73.
- 15 Fleischer NL, Lozano P, Arillo Santillán E, *et al*. The impact of neighbourhood violence and social cohesion on smoking behaviours among a cohort of smokers in Mexico. *J Epidemiol Community Health* 2015;69:1083–90.
- 16 Turrell G, Hewitt BA, Miller SA. The influence of neighbourhood disadvantage on smoking cessation and its contribution to inequalities in smoking status. *Drug Alcohol Rev* 2012;31:645–52.
- 17 van Lenthe FJ, Mackenbach JP. Neighbourhood and individual socioeconomic inequalities in smoking: the role of physical neighbourhood stressors. *J Epidemiol Community Health* 2006;60:699–705.
- 18 Shareck M, Ellaway A. Neighbourhood crime and smoking: the role of objective and perceived crime measures. *BMC Public Health* 2011;11:930.
- 19 Fleischer NL, Thrasher JF, Sáenz de Miera Juárez B, *et al*. Neighbourhood deprivation and smoking and quit behaviour among smokers in Mexico: findings from the ITC Mexico survey. *Tob Control* 2015;24 Suppl 3:iii56–63.
- 20 Lau YK, Tam J, Fleischer NL, *et al*. Neighbourhood deprivation, smoking, and race in South Africa: a cross-sectional analysis. *Prev Med Rep* 2018;11:202–8.
- 21 Timmermans EJ, Veldhuizen EM, Snijder MB, *et al*. Neighbourhood safety and smoking in population subgroups: the HELIUS study. *Prev Med* 2018;112:111–8.
- 22 Reijneveld SA. Neighbourhood socioeconomic context and self reported health and smoking: a secondary analysis of data on seven cities. *J Epidemiol Community Health* 2002;56:935–42.
- 23 Mistry R, Pednekar M, Pimple S, *et al*. Banning tobacco sales and advertisements near educational institutions may reduce students' tobacco use risk: evidence from Mumbai, India. *Tob Control* 2015;24:e100–7.
- 24 Merlo J, Wagner P, Ghith N, *et al*. An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health. *PLoS One* 2016;11:e0153778.
- 25 Merlo J, Wagner P, Austin PC, *et al*. General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: a conceptual tutorial. *SSM Popul Health* 2018;5:33–7.
- 26 Diez Roux AV. The study of group-level factors in epidemiology: rethinking variables, study designs, and analytical approaches. *Epidemiol Rev* 2004;26:104–11.
- 27 Merlo J. Multilevel analytical approaches in social epidemiology: measures of health variation compared with traditional measures of association. *J Epidemiol Community Health* 2003;57:550–2.
- 28 Merlo J, Chaix B, Yang M, *et al*. A brief conceptual tutorial on multilevel analysis in social epidemiology: interpreting neighbourhood differences and the effect of neighbourhood characteristics on individual health. *J Epidemiol Community Health* 2005;59:1022–9.
- 29 Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D, *et al*. Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain. *Soc Sci Med* 2012;75:1477–87.
- 30 Palipudi KM, Morton J, Hsia J, *et al*. Methodology of the Global Adult Tobacco Survey - 2008-2010. *Glob Health Promot* 2016;23:3–23.
- 31 GATS. *Global adult tobacco survey (GATS) India report 2009-2010*. India: Ministry of Health and Family Welfare, 2010.
- 32 Office of the Registrar General & Census Commissioner. Administrative divisions new Delhi, India: Ministry of home Affairs, government of India, 2018. Available: http://censusindia.gov.in/Census_And_You/Administrative_division.aspx [Accessed 30 Nov 2018].
- 33 Subramanian SV. The relevance of multilevel statistical methods for identifying causal neighborhood effects. *Soc Sci Med* 2004;58:1961–7.
- 34 Subramanian SV, Jones K, Kaddour A, *et al*. Revisiting Robinson: the perils of individualistic and ecologic fallacy. *Int J Epidemiol* 2009;38:342–60.
- 35 Merlo J, Chaix B, Yang M, *et al*. A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. *J Epidemiol Community Health* 2005;59:443–9.
- 36 Merlo J, Ohlsson H, Lynch KF, *et al*. Individual and collective bodies: using measures of variance and association in contextual epidemiology. *J Epidemiol Community Health* 2009;63:1043–8.
- 37 Singh A, Harford J, Peres MA. Investigating societal determinants of oral health-Oportunities and challenges in multilevel studies. *Community Dent Oral Epidemiol* 2018;46:317–27.
- 38 Merlo J, Chaix B, Ohlsson H, *et al*. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health* 2006;60:290–7.
- 39 Sarkar BK, Reddy KS. Priorities for tobacco control research in India. *Addiction* 2012;107:2066–8.
- 40 Ahluwalia IB, Arrazola RA, Zhao L, *et al*. Tobacco Use and Tobacco-Related Behaviors - 11 Countries, 2008-2017. *MMWR Morb Mortal Wkly Rep* 2019;68:928–33.
- 41 Nazar GP, Chang KC-M, Srivastava S, *et al*. Impact of India's national tobacco control programme on bidi and cigarette consumption: a difference-in-differences analysis. *Tob Control* 2020;29:103–10.
- 42 NHM. National Tobacco Control Programme (NTCP): National Health Mission, Ministry of Health and Family Welfare, Government of India; 2019, 2020. Available: <https://nhm.gov.in/index1.php?lang=1&level=2&sublinkid=1052&lid=607> [Accessed 6 Mar 2020].
- 43 Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav* 2010;51 Suppl:S28–40.