Scientific Research Publishing

Open Journal of Epidemiology, 2019, 9, 173-189
http://www.scirp.org/journal/ojepi
ISSN Online: 2165-7467
ISSN Print: 2165-7459

A Systematic Review of Observational Studies, Demonstrating Smoking among School Going Adolescents

Shafquat Rozi[1*], Gillian Lancaster[2], Sadia Mahmud[3], Zahid Ahmad Butt[4], Talat Roome[5], Nida Zahid[6]

[1]Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan
[2]Department of Mathematics and Statistics, Lancaster University, Lancashire, United Kingdom
[3]Visiting Faculty, Department of Medicine Consultancy, Department of Pediatrics & Child Health, Aga Khan University, Karachi, Pakistan
[4]School of Population and Public Health, University of British Columbia, Vancouver, Canada
[5]Department of Pathology, Dow University of Health Sciences, Karachi, Pakistan
[6]Department of Surgery, Aga Khan University, Karachi, Pakistan
Email: *shafquat.rozi@aku.edu

How to cite this paper: Rozi, S., Lancaster, G., Mahmud, S., Butt, Z.A., Roome, T. and Zahid, N. (2019) A Systematic Review of Observational Studies, Demonstrating Smoking among School Going Adolescents. *Open Journal of Epidemiology*, **9**, 173-189.
https://doi.org/10.4236/ojepi.2019.92015

Received: March 28, 2019
Accepted: May 27, 2019
Published: May 30, 2019

Abstract

Objective: To characterize the methods of design and analysis currently adopted in survey research of school-based observational studies for smoking, and to identify the common pitfalls made by researchers. **Methods:** The systematic review was conducted in 2009 and consisted of observational studies in school settings published between January 2005 and January 2009. Smoking status was the main outcome of interest. Following Cochrane style, five steps were followed: setting selection criteria for studies and conducting a literature search; review of abstracts; review of complete articles; data extraction and quality assessment of included studies; and, finally, synthesis of studies. **Results:** Of the 292 abstracts retrieved, 45 (15.4%) articles were selected for the final review. Inconsistencies were found in the definition of smoking behaviour which impeded generalisability. Individual-level factors had importance, but environmental level factors were also important in studying the aetiology of smoking. Results showed that studies inappropriately reported sample size estimation and important confounding factors. Hierarchical linear modelling, random effects modelling and structural equation modelling were employed in comparatively few studies. **Conclusions:** There were concerns regarding data analysis of complex surveys. Fifty five percent of reviewed studies ignored environmental effects which may have produced unreliable inferences. Multi-level analysis assisted in understanding school-level effects.

Keywords

Adolescent, Smoking, Observational Studies, School-Based Survey

1. Introduction

Tobacco is one of the greatest public health issues of modern times. It consumes a considerable amount of resources of the healthcare system in Pakistan for both treatment and prevention [1]. Currently, about 5.4 million deaths occur every year due to tobacco use [1]. More than 8 million tobacco deaths are expected by 2030 if urgent action is not taken. The dilemma is that 80% of tobacco-related deaths within a few decades will occur in the developing countries [1]. This devastating shift to the developing countries is due to the fact that the global tobacco industry is targeting young adults [2]. Currently, 150 million adolescents are tobacco users worldwide. According to recent studies, school environment has an impact on outcomes of adolescent behaviour that include substance abuse and committing crime [3]. A study from Karachi, Pakistan, reported prevalence of tobacco use among school-going children at 16.1% [4] [5].

Studies from Pakistan have documented the association of betel quid, areca and tobacco use with head and neck cancers [6] [7]. Although tobacco kills far more people than human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), unfortunately funding for tobacco control is less than that is required for its prevention [8]. About 27 countries have imposed tax rates of more than 75% of the retail price of tobacco to control its use [9]. There is no data available to measure the expenditure for treating diseases caused by tobacco use in Pakistan.

Mostly, survey-based studies have been designed to discuss health and smoking issues which are carried out either in communities, clinics, hospitals or schools. To consider environmental/school-level factors affecting adolescent smoking, surveys often employ multi-stage cluster sampling that would incur a hierarchical structure of the population.

If hierarchical data is analysed using analytical techniques that will not consider clustering of data, this will produce misleading inferences and conclusions regarding the association of smoking with other predictors. A researcher may need to consider intra-class correlation (ICC) in the sample size calculation, if accounting for clustering in the structure of the study. About 55% studies have revealed that a multi-level modelling approach shows that the environmental factors have considerable contribution in variation of smoking prevalence between schools besides individual-level factors [10]. A study suggested that there was more similarity in children within school compared to between schools [10]. The current review study was planned to determine the methodological approaches in recent surveys, and to identify the common pitfalls in the methodology, especially design and analysis, in school-based observational studies for teenage smoking.

2. Material and Methods

The systematic review was conducted in 2009 and comprised of observational studies in school settings published between January 2005 and January 2009 with smoking status as the main outcome of interest. Following the Cochrane methodology [11], five steps were followed: setting selection criteria for studies and conducting a literature search; review of the abstracts; review of the complete articles; data extraction and quality assessment of included studies; and, finally, synthesis of studies. The focus was not on pooling estimates.

Eligibility criteria

In the first step, the eligibility criteria for the inclusion of the studies and the search strategy to be used for selection of the studies for literature review were defined.

Inclusion criteria

The search was limited to peer review published studies in English language. All the studies selected for this systematic review were observational; cohort retrospective, prospective, case-control, longitudinal and cross-sectional with population aged 11 - 18 years in school settings from both the developed and the developing world with smoking status as the main outcome of interest.

Exclusion Criteria

We excluded non-peer reviewed articles, letters to editors, conference proceedings and articles published in a language other than English. All studies done on adolescents smoking in hospitals or clinics or communities were also excluded from the review. If the main outcome of interest of the study was something other than smoking then the study was not included in the systematic review.

Search Strategy for Identification of Studies

A comprehensive literature search was performed in both general databases, such as PubMed free access database, Embase, and subject-specific databases such as: PSYCINFO. Moreover, to retrieve publications reporting observational studies for smoking among teenage children attending school, we performed a combined search strategy that included the following terms as both medical subject heading (MeSH) terms and text words: Adolescent in MeSH; Adolescences; Adolescents; Adolescents, Female; Adolescents, Male; Teenagers; Teens; Youth; 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8; Schools in MeSH; Primary Schools; Schools, Secondary; Secondary Schools; 10 or 11 or 12 or 13; Smoking in MeSH; Cigarette Smoking; Smoking Tobacco; Tobacco Smoking; Tobacco Smoking Pollution; 15 or 16 or 17 or 18 or 19; and 9 and 14 and 20.

Filters were also used for observational studies during the literature search defined by the SIGN group and University of York [12]. This filtered the extracts of the observational studies from the retrieved lists after the POL search which is the population (adolescents), outcome (smoking) and location (secondary schools). Moreover, in the second step articles that were fulfilling our eligibility criteria were extracted (**Figure 1**).

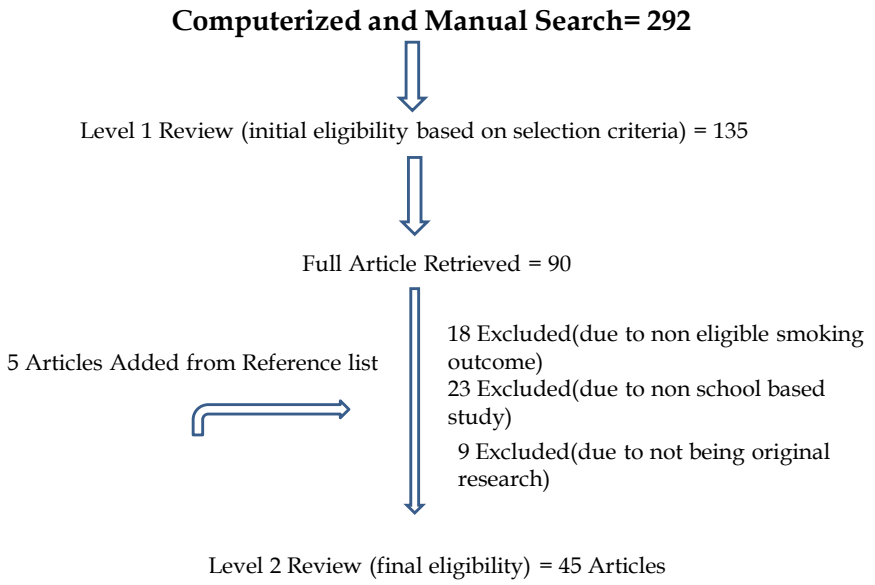

Figure 1. Search and selection process.

This resulted in selection of articles for second-level review of the eligibility of the retrieved full papers. One investigator carried out the second level review but uncertainty about inclusion of studies was resolved with the other reviewer. To validate the inclusion of studies, two reviewers independently reviewed the inclusion process on a 10% sample selected randomly by the principal investigator. Finally, 100% agreement was achieved on inclusion status.

In the 3rd step one investigator extracted data for the included studies on a pre-designed data-abstraction form. Most questions on the form were pre-coded. Validation of data-abstraction was performed by selecting a random sample of the included studies for the second reviewer to check (5 papers). In case of a disagreement, the relevant field was checked on all abstraction forms and a mutually agreed description was achieved. Data were primarily extracted to assess design and analysis, including whether there was a reported sample size calculation or if analysis had been conducted appropriately or not. We also assessed whether the investigators had provided details of sample size calculation and analysis considering clustering of data and intra-class correlation or not. Place of publication, date of publication, population characteristics, detailed description of outcomes, study design, sampling strategy and detail of individual and school-level factors relating to adolescents smoking were also extracted. The research team discussed any discrepancies.

Both internal and external validity were assessed during methodological quality assessment. The interpretation of the findings of a study depended on design, conduct and assessment (internal validity), as well as on populations and outcome measures (external validity). In the final step, the findings of the literature review were synthesized and the key points were summarised in tabular form.

Assessment of quality of study

The quality of each study was assessed independently. Any disagreements on

study quality were resolved through discussion with the other reviewer to confirm whether or not the study had achieved the quality dimension [13] [14]. The quality of reviewed studies based on internal and external validity was assessed which aimed at ascertaining whether the methods used in the design and conduct of the studies were likely to prevent systematic error. For this purpose, study design, method of analysis, sample size calculation, whether the researchers had discussed the limitations of their study or not, selection bias, consideration of potential confounders of the study (age, gender, socio-economic status), the quality of reporting, the generalizability of the results, conclusion, and recommendations based on study results were the factors assessed.

3. Results

Table 1 presents the methodological characteristics of the 45 studies, including: Setting, school type and grades, sample size, study design, and the sampling strategy. Of the 292 abstracts retrieved, 45 (15.4%) articles were selected for final review. About 32.4% of the published studies were conducted in the United States, followed by Canada (10.8%), China (8.6%), the United Kingdom (8.6%), and India (8.6%), while the remaining 32.8% were in other countries including Malaysia, Norway, Scotland, Lebanon, Taiwan, Australia, Sudan, Iceland, Brazil, New Zealand, Belgium, Greece, Netherland and Germany. The countries involved in joint projects were the USA, the UK, Greece, Romania, Denmark, Israel and other European countries.

The studies included were conducted in public and private schools, vocational schools, and technical education schools. However, most of the reviewed studies did not clearly state the proportion of public and private schools that they had included in their study. The studies under review included 77.7% cross-sectional and 22.3% prospective cohort or longitudinal follow-up studies. However, in some longitudinal studies (8.8%), the length of follow-up was not reported.

Twenty percent of studies comprised two-stage or multi-stage cluster sampling, 11.1% studies employed random sampling with stratification, 13.3% studies reported simple random sampling, 15.5% reported convenience sampling and 15.5% derived data from other projects for secondary analysis and did not clearly state their sampling strategy. However, 11.3% studies reported other sampling techniques such as systematic sampling, and single-stage cluster sampling, while the rest of the studies (13.3%) did not discuss sampling strategy.

The reviewed studies determined samples from schools in specific regional areas, whereas few studies obtained samples at national level. Some schools were selected using systematic random sampling. Mostly school-based studies conducted surveys on a particular school day. Those adolescents who were absent on the day of the interview or had dropped out of school were excluded, which may have biased the results and generalisability of the findings.

Sample sizes ranged from 384 to 91,778 school-children. About 20% studies employed single-stage or multi-stage cluster sampling. However, clustering was

Table 1. Summary of methodological aspects and analytical issues of studies reviewed.

Authors	Country & Setting	Study design	Sampling technique & sample size	Method of analysis
Arillo-santillan E *et al.* 2005 [2]	Mexico, Public schools from 98 - 99	Cross-sectional 1998-99	systematic random sampling 5825 boys 7468 girls	Multinominal model, Multivariate logistic regression stratified by school levels, Multilevel model
Bidstrup PE *et al.* 2008 [15]	Denmark, 118 Public schools 230 classes, 2 classes per school with 13 pupils per class	Cross-sectional 2004	Random sampling of schools (Data derived from Danish national institute of Public health in 2008) 1426 boys 1487 girls	Multilevel logistic regression
Bird Y*et al.* 2006 [9]	Juancz, Mexico, 6 Public & private schools. Data Collection year was not mentioned.	Cross-sectional study	Random sampling stratified by SES to select schools. All 6th grade students were selected 242 boys 264 girls	Multiple logistic regression by SES, school setting & gender
Bird Y *et al.* 2007 [16]	Ciudad Juarez, Mexico,6 Public & private schools in 2000	Cross-sectional study	Random sampling stratified by SES to select schools. All 6th grade students were selected 242 boys 264 girls	Chi-squared test, ANOVA and logistic regression
Bond L *et al.* 2007 [17]	Victoria, Australia, 26 Secondary schools. Year of data collection not mentioned.	Longitudinal study	Data derived from Gatehouse project 1124 boys 1276 girls	Prevalence estimates, & Logistic regression using robust" "information sandwich" estimates of standard errors to account for clustering within schools.
Carson NJ *et al.* 2005 [18]	Northern Virginia Philadelphia, USA, 5 Public high schools. Year of data collection not mentioned	Cross-sectional	Data derived from another Longitudinal study of behavioural predictors of smoking. 967 adolescents	SEM, EFA, CFA, model fit assessed
Chang F *et al.* 2006 [19]	Taipei, Taiwan, 16 Vocational high schools 39 classes from 2000-2002	Longitudinal study 3 years	Probability proportionate to size systematic random sampling to draw schools and classes. All 10th grade students selected 1695 boys and girls	GEE
Cleveland MJ *et al.* 2008 [20]	Pennsylvania, USA, 180 Public & private schools. from 2005 Pennsylvania Youth survey	Cross-sectional.	Random sampling stratified by 6 regions 91,778 boys and girls	Generalized linear mixed models, cumulative logit models (proportional odds model)
Conti DV *et al.* 2006 [21]	China, 62 Middle schools 83 Upper schools from Oct 2002-Dec 03	Longitudinal study	Multistage cluster sampling stratified by school type 5042 boys 5606 girls	Multilevel random coefficient modeling
Damianaki A *et al.* 2008 [22]	Semi urban area in Crete, Greece, 15 Public schools from 2004-05	Cross-sectional	Random sampling stratified by rural/urban status to select schools and students 386 boys 507 girls	Stepwise logistic regression, Mann-Whitney, t-test
Difranza JR *et al.* 2008 [23]	Urban and suburban region in central Massachusetts, USA, 2 Schools in 2007	Cross-sectional	Convenience sampling 1055 boys and girls	2-tailed Kendall's tau b test for correlations
Ditre JW *et al.* 2008 [24]	Florida, USA, 3 High schools in 1999	Cross-sectional	Convenience sampling 363 boys 394 girls	Mann-Whitney test, Kruskal Wallis, ANCOVA

Continued

Droomers M *et al.* 2005 [25]	New Zealand, Number and type of schools not stated, April 1972, and March 1973	Longitudinal study	Data derived from Dunedin multidisciplinary health behavior project 481 boys 466 girls	Longitudinal Logistic GEE
Eengels RC *et al.* 2005 [26]	Netherlands, 11 Secondary schools in 2004	Cross-sectional and short term longitudinal follow up	Data derived from a large scale survey in Netherlands 929 boys 932 girls	Binary logistic regression
El-Roueiheb Z *et al.* 2008 [27]	Beirut, Lebanon, 3 Public, 10 private schools from 2003-04	Cross-sectional	Random sampling stratified by district within Beirut to select schools. All students were interviewed from schools 1089 boys 1328 girls	Multivariable regression analysis with missing data
Ellickson PL *et al.* 2008 [28]	California and Oregon, USA, 30 Secondary schools from 2000-2005	longitudinal study over 5 years,	Data derived from ALERT project 1960 adolescents for smoking analysis. 2000 for problem behaviour analysis	Multivariate regression
Ennett ST *et al.* 2008 [29]	North Carolina USA, Wave 1: 8 middle schools, 2 comprehensive k-8 schools, 3 alternative schools Wave 2: 6 high schools, Middle & high grade from 2002-2004	Longitudinal study	Convenience sampling Wave 1: 5220 boys and girls Wave 5: 5017	Three level hierarchical growth models, Bayesian information criterion (BIC), Random intercept and random effect model
Faeh D *et al.* 2006 [30]	Seychelles, Indian ocean, 10 public schools 2 private schools, from global youth tobacco survey	Cross-sectional	Two stage cluster sampling design to obtain students Schools: probability proportional to school size Student: equal probability sampling 620 boys 654 girls	Percentage or proportion and CI
Finkelstein DM *et al.* 2006 [31]	Cincinnati, Ohio, USA, 1public high& 1 junior high public schools, date of study not mentioned.	Longitudinal study	Not stated 494 boys 527 girls	Chi-squared test, ANOVA with Tukey's test, logistic regression analysis' and GEE
Glanz K *et al.* 2007 [32]	Hawaii, 5 states Oahu, the big island of Hawaii, Maui, Molokai and Kawi, 22 Public & private schools from 2000-2002	Cross-sectional	Convenience sampling 844 boys 851 girls	Chi.-square & t-test, GEE, PROC logistic regression
Grenard JL *et al.* 2006 [33]	China 7 cities, Hanbin, Shenyang, Wuhan, Chengdu, Kunming, & Hangzhou, Qingdao, 147 Middle schools & high schools, date of study not mentioned	Longitudinal study 1 year	Multistage cluster sampling 6991 boys 7387 girls	Chi-squared test t-test, paired t-test, Multilevel analysis & ICC, generalized linear mixed models
Grotvdt L *et al.* 2008 [34]	6 countries in Norway including Oslo, 2 Southern inland 3 Northern countries, Public and private schools from 2000-04	Cross-sectional	Convenience sampling 7762 boys 7768 girls	Generalized linear model with binominal distribution

Continued

Gutschoven K *et al.* 2005 [35]	Flanders, Belgium, 15 Schools. Data collected from Leuven Study on Media and Adolescent Health (SOMAH)	Cross-sectional	Random sampling of schools. All students were selected from 1st & 4th year (Data derived from the media and adolescent health project) 1380 boys 1166 girls	ANOVA with post hoc Tukey's test
Henderson M *et al.* 2008 [36]	Tayside and Lothian, Scotland, 24 Scottish schools. Data collected in 1998 and 1999 at the first follow-up of the SHARE trial.	Longitudinal study	Not stated 5092 boys and girls	Two level logistic regression model
Henewinkel R *et al.*2007 [37]	Schleswig Holstein, Germany, 27 Schools in Oct-Nov 2005	Cross-sectional	Random sampling of schools. 2861 boys 2717 girls	Chi square test, Lowess (locally weighted scatter plots), Logistic regression
Henrikson L *et al.* 2008 [38]	California USA, 135 High schools in 2005-06	Cross-sectional	Random sampling of schools 384 boys and girls	Multiple regression for complex sampling design
Hublet A *et al.* 2006 [39]	9 European countries, 2 UK counties and Canada, Number and types of schools not specified. Study from 1990-2002	Cross-sectional	Cluster sampling of schools & classes 36, 767 boys 38,978 girls	Logistic regression
Kokkevi A *et al.* 2006 [40]	5 European countries (Bulgaria, Croatia Greece, Romania, Slovenia) and UK. High schools in 1999	Cross-sectional	Two stage sampling stratified by school type 16445 boys and girls	Survey logistic regression
Kristjansson AL *et al.* 2009 [41]	Iceland, Secondary schools, in March 2006.	Cross sectional	9th-10th grade students of all Icelandic secondary schools 3612 boys 3620 girls	Pearson's correlation & logistic regression
Leatherdale ST *et al.* 2005 [42]	Ontario, Canada, 57 Elementary schools. Data collection time not mentioned.	Cross-sectional	Not stated 2758 8th grade boys and girls 4286 6th or 7th grade boys and girls	Multilevel logistic regression
Leatherdale ST *et al.* 2007 [43]	Ontario, Canada, 29 Secondary schools, in 2000-2001	Cross-sectional	Not stated 511 9 boys and girls grades 9 - 12	Multilevel logistic regression
Lee LK *et al.* 2005 [44]	Malaysia 7 districts, Sermban, Port Dickson, Kualapilah, Jempaltampin& Remban, 14 Schools, from June 2001-August 2002	Cross-sectional	Multistage stratified random sampling, 6 schools from urban, 8 schools from rural 2 classes per school 2088 boys 2411 girls	Chi-squared test & CI
Lovato CY *et al.* 2007 [45]	British Columbia, Monitabo, Newfoundland, Ontario and Quebec 81 Secondary schools. Data collection year not mentioned.	Cross-sectional	Random sampling of schools 22,318 boys and girls Grades 10 - 11	Linear regression analysis, t-test
Mathur C *et al.* 2008 [46]	Delhi (n = 16) and Chennai (n = 16), India, 32 Public & private school in 2004	Cross–sectional	Convenience sampling 11,642 boys & girls	Mixed effects regression model

Continued

Muilenburg JL *et al.* 2008 [47]	Southeastern USA 6 Public schools in 2006	Cross-sectional	Part of National youth tobacco survey 990 boys 1071 girls	Ordered logit regression
Murnghan DA *et al.* 2007 [48]	Prince Edward Island, Canada, 10 English high schools from 1999-2001	Cross-sectional	Not stated 3965 boys and girls	Multilevel logistic regression
Poulin CC *et al.* 2007 [49]	Atlantic province (Nova Scotia, New Brunswick, Prince Edward Island, Newfoundland & Labrador, Secondary schools in 2002	Cross-sectional	Single stage cluster sample stratified by grade and either health regional or school district 6403 boys 6368 girls	Two level random effect model, Sensitivity analysis was performed using multinomial logistic regression
Ridout F *et al.* 2008 [50]	Hampshire and Central London, 12 Secondary schools. Data collection year not mentioned.	Cross-sectional	Convenience sampling 912 boys 853 girls	Chi-squared, logistic regression adjusted for clustering
Rodriguze D *et al.* 2007 [51]	USA, 5 Public schools. Data collected in 2003.	Longitudinal study over 4 years	Not stated 496 boys 467 girls	SEM, CFA
Schnohr CW *et al.* 2009 [52]	Denmark, Sweden, Norway, Finland, and UK secondary schools 2001-2002	Cross-sectional	Data derived from a project (The health behaviour in school children) 20,399 boys and girls	Binary logistic regression analysis stratified by country
Sinha DN *et al.* 2007 [53]	India (North, Eastern, Western, North East), 179 schools in 2006	Cross-sectional	Two stage cluster sampling Schools and classes selected randomly. All students selected within classes 12086 boys and girls	Logistic regression for clustered data
Sinha DN *et al.* 2008 [54]	India, 6 regions: North, South, East, West, Central & North East, 2003: 818 schools 2006:179 schools,	Cross-sectional	Two stage cluster sampling Schools selected proportional to number of students, and classes selected randomly. All students interviewed from selected classes 2003: 68,077 boys and girls 2006: 12,086 boys and girls	Logistic regression for correlated data
Thrasher JF *et al.* 2008 [55]	Mexico,18 Public & private schools in 2006	Cross-sectional	Proportional sampling strategy to randomly select schools & classes 3874 boys and girls	Multilevel adjusted logistic regression model, SEM
Wen X *et al.* 2007 [56]	Huang pu and Guangzhou, China, 6 Secondary schools in 2004	Cross-sectional	Cluster sampling of schools 2021 boys 1850 girls	Logistic regression
Wen X *et al.* 2008 [57]	Guangzhou, China, 3 Public schools, 1 Factory school 2 General paid private schools in 2004	Cross-sectional	Multistage sampling with stratification by school type 3957 boys and girls and 2870 parents of these students	Chi-squared & Kruskal Wallis

SEM: Structural equation modeling, EFA: Exploratory factor analysis, CFA: Confirmatory factor analysis; ETS: Environmental tobacco smoke; TAR; Tobacco advertisement receptivity, DFT: Drive for thinness; GEE: Generalized estimating equations; ICC: Intra-class correlation, ASO: General education, TSO: Technical education, BSO: Vocational training.

not taken into account during sample size estimation by all the 45 studies. Such studies that did not take clustering into account in sample size calculation suffered from a considerable loss of power since the design effect was multiplied by the

sample size calculated under simple random sampling to account for clustering.

However, we made no judgment regarding sample size calculations for studies which did not mention clustering pattern.

The definitions of adolescents' tobacco use varied from study to study and, hence, it was not possible to aggregate them together. Smoking use among adolescents was categorised into five levels or stages 1) experimentation, 2) regular use, 3) occasional use, 4) frequent use, and 5) heavy use across studies. About 75% of reviewed studies did not discuss level or stage of tobacco use but they explained whether or not adolescents had ever smoked and/or were currently smoking.

While most of the studies applied simple definitions of smoking behaviour, it is imperative to recognise that some research studies also reported progression across stages of tobacco use.

Few studies (4.4%) were purely descriptive (*i.e.* prevalence, rates or means). Most of the studies reviewed employed traditional methods of analysis e.g., 8.8% performed a t-test, 28.8% multivariable logistic regression and linear regression, 11.1% analysis of variance and 17.7% a chi-squared test. Also, 55% studies correctly accounted for clustering and performed hierarchical linear modelling, random effects modelling, and generalized estimating equation (GEE). Although in the 55% of studies appropriate statistical techniques were used to answer the research questions, some studies employed longitudinal samples, but this was not reflected in their analysis, reflecting lack of competence in advanced statistical methods.

Different statistical software packages were used to analyse the data. Most studies used Statistical Package for the Social Science (SPSS) (28.8%) and Statistical Analysis System (SAS) (24.4%). Some studies reported using STATA (17.7%), Multilevel software (MLWIN) (8.8%), or Survey data analysis (SUDAAN) (4.4%) and 4.4% of studies reported the use of MPLUS for statistical analysis. However, 20% studies did not mention the type of software used by them for their analysis.

Approximately 73.3% of reviewed studies reported about missing information in their datasets. About 17.7% of studies imputed missing values or used a regression technique with imputation and about 28.8% of them excluded the missing information from their analysis. However, 33.2% studies did not discuss anything about the missing data in their studies. About 11.1% of studies had no missing information.

The median prevalence of smoking across studies was 13.1% interquartile range [IQR] = 11.4). The demographic variables included in the review studies were adolescent age and gender, family, socio-economic status and education, and family smoking. The risk factor most commonly associated with teenage smoking leading them to become habitual or regular smokers were senior students smoking in school, peer smoking, public schools, school suspension, retailer near the school, higher level of exposure to movies, pupil's relation and attitude to school, school size, social environment of the school, and quality of

school teachers. In schools where anti-smoking policies were well implemented, the ratio of adolescents smoking to those who did not was lower compared to those schools where no tobacco monitoring policies were present. Good grades, school connectedness, and knowledge about public policies were prominent school-level protective factors that restricted adolescents from smoking.

The assessment of potential confounders and risk factors is vital. In most of the studies, smoking status of family members was considered. Some studies reported the effect of individual (mother's and father's) smoking habits on adolescents' smoking, while a few observed the combined effect of parents' smoking on adolescent smoking, Furthermore, the effect of sibling smoking was also assessed in some studies. A few studies combined parent, siblings and other relatives smoking as family member smoking.

The association between parents or family members smoking and adolescent smoking suggested that parental smoking influenced the initiation of adolescent smoking. When family members of adolescent were using tobacco, they had easy access to use it and it made them believe that it was socially acceptable.

Adolescents whose friends were smokers were more likely to smoke compared to those with no peer smoking. The average age of smoking in reviewed studies was approximately 13.1 years. At this age usually peer pressure or peer relationships become stronger than family relationships, so adolescents were more likely to be influenced by the behaviour of friends.

It was found that only a few studies fulfilled the criteria of good quality study. We scored studies according to their assessed validity, thus, the more valid a study was the higher the score it received. The quality of the reviewed studies was below average in approximately one-third (37%) of studies, 48% were marked as average, and 15% as good according to the criteria of strobe [14] and consort guidelines for randomised controlled trials (RCTs) [13]. This reflects an inadequate understanding of the substantive issues underpinning hierarchical data in a survey research.

4. Discussion

Discrepancies were observed in the definition of smoking in the reviewed studies and the tools that were used to measure adolescent smoking were not standardised, which may hinder generalisability. A number of studies were designed to assess the social influences of adolescent smoking along with individual-level covariates. Although we found different study designs, analytical approach and covariates, our results indicate that environmental level factors are as crucial as individual factors for studying the aetiology of smoking.

In all the reviewed studies, the questionnaire was self-administered by students who not only reported their own smoking habits but also the smoking habits of their parents, family and friends who currently smoke cigarettes. However, it was difficult to report accurately the frequency and number of cigarettes used by their family and friends. Studies reported that parents, siblings and peers

were powerful influencers for adolescent smoking.

As in most of the studies, the interviews were conducted on one particular day, hence it was expected that a few students might be absent on that particular day thus not including them in the study may have produced biased results and may also have an effect on generalisability.

Causal inferences cannot be determined from cross-sectional studies. The predictors of smoking reported by cross-sectional studies were not markedly different from that reported by longitudinal studies.

Confounders and biases are major concerns in observational studies, which need to be addressed appropriately. However, most of the studies failed to measure important confounding factors. In most of the studies, clustering was not considered during sample size calculation which may have caused a reduction in the power. In conclusion, the results of this review highlight concerns about the analysis of complex survey data. Hierarchical linear modelling, random effects modelling and structural equation modelling were used in comparatively fewer studies. About 45% studies ignored the contextual/ environmental (e.g. retailer near the school, social environment of school and implementation of smoking policies at school) factors in their studies which may have produced misleading inferences.

The results of reviewed studies guided us in comprehending the effect of school-level factors and variances between schools. Moreover, it also provided us with more intuitive information on school-level covariates that can have an influence on the adolescents smoking behaviour.

One of the limitations of our study is that we only reviewed studies in English. This could have resulted in some studies being excluded from our review.

It is recommended that future studies should consider environmental/social settings from where the individuals are drawn at the design and analysis stage as well as taking clustering into account to come up with a leading inference to deal with public health issues.

5. Conclusion

There were different study designs, analytical approaches and covariates in the studies reviewed, but results indicate that environmental level factors are as crucial as individual factors for studying the aetiology of smoking.

Disclaimer

None.

Source of Funding

None.

Conflicts of Interest

None.

References

[1] WHO (2008) WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER Package.

[2] Arillo-Santillan, E., Lazcano-Ponce, E., Hernandez-Avila, M., Fernández, E., Allen, B., Valdes, R., *et al.* (2005) Associations between Individual and Contextual Factors and Smoking in 13,293 Mexican Students. *American Journal of Preventive Medicine*, **28**, 41-51. https://doi.org/10.1016/j.amepre.2004.09.002

[3] World Health Organization. Global Tobacco Epidemic. https://www.who.int/tobacco/global_report/2017/en/

[4] Rozi, S., Butt, Z.A. and Akhtar, S. (2007) Correlates of Cigarette Smoking among Male College Students in Karachi, Pakistan. *BMC Public Health*, **7**, 312. https://doi.org/10.1186/1471-2458-7-312

[5] Rozi, S., Akhtar, S., Ali, S. and Khan, J. (2005) Prevalence and Factors Associated with Current Smoking among High School Adolescents in Karachi, Pakistan. *The Southeast Asian Journal of Tropical Medicine and Public Health*, **36**, 498-504.

[6] Khawaja, M.R., Mazahir, S., Majeed, A., Malik, F., Merchant, K.A., Maqsood, M., *et al.* (2006) Chewing of Betel, Areca and Tobacco: Perceptions and Knowledge Regarding Their Role in Head and Neck Cancers in an Urban Squatter Settlement in Pakistan. *Asian Pacific Journal of Cancer Prevention*, **7**, 95-100. https://doi.org/10.1186/1747-597X-1-10

[7] Mazahir, S., Malik, R., Maqsood, M., Merchant, K.A., Malik, F., Majeed, A., *et al.* (2006) Socio-Demographic Correlates of Betel, Areca and Smokeless Tobacco Use as a High Risk Behavior for Head and Neck Cancers in a Squatter Settlement of Karachi, Pakistan. *Substance Abuse Treatment, Prevention, and Policy*, **1**, 10. https://doi.org/10.1186/1747-597X-1-10

[8] Mackay, J. and Eriksen, M. (2002) The Tobacco Atlas. World Health Organization. Myriad Editions Limited, Oxford.

[9] Bird, Y., Moraros, J., Olsen, L.K., Coronado, G.D. and Thompson, B. (2006) Adolescents' Smoking Behaviors, Beliefs on the Risks of Smoking, and Exposure to ETS in Juarez, Mexico. *American Journal of Health Behavior*, **30**, 435-446. https://doi.org/10.5993/AJHB.30.4.9

[10] Pokorny, S.B., Jason, L.A. and Schoeny, M. (2004) Current Smoking among Young Adolescents: Assessing School Based Contextual Norms. *Tobacco Control*, **13**, 301-307. https://doi.org/10.1136/tc.2003.005363

[11] Higgins, J. and Green, S. (2011) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0.

[12] Centre for Review and Dissemination University of York. https://www.crd.york.ac.uk/CRDWeb

[13] Schulz, K.F., Altman, D.G. and Moher, D. (2010) CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. *BMC Medicine*, **8**, 18. https://doi.org/10.1186/1741-7015-8-18

[14] Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., *et al.* (2014) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *International Journal of Surgery*, **12**, 1495-1499. https://doi.org/10.1016/j.ijsu.2014.07.013

[15] Bidstrup, P.E., Frederiksen, K., Siersma, V., Mortensen, E.L., Ross, L., Vinther-Larsen, M., *et al.* (2008) Social-Cognitive and School Factors in Lifetime Smoking

among Adolescents. *Cancer Epidemiology Biomarkers & Prevention*, **17**, 1862-1871. https://doi.org/10.1158/1055-9965.EPI-07-2773

[16] Bird, Y., Moraros, J., Olsen, L.K., Forster-Cox, S., Buckingham, R.W. and Staines-Orozco, H. (2007) Smoking Practices, Risk Perception of Smoking, and Environmental Tobacco Smoke Exposure among 6th-Grade Students in Ciudad Juarez, Mexico. *Nicotine & Tobacco Research*, **9**, 195-203. https://doi.org/10.1080/14622200601078533

[17] Bond, L., Butler, H., Thomas, L., Carlin, J., Glover, S., Bowes, G., *et al.* (2007) Social and School Connectedness in Early Secondary School as Predictors of Late Teenage Substance Use, Mental Health, and Academic Outcomes. *Journal of Adolescent Health*, **40**, 357.e9-18. https://doi.org/10.1016/j.jadohealth.2006.10.013

[18] Carson, N.J., Rodriguez, D. and Audrain-McGovern, J. (2005) Investigation of Mechanisms Linking Media Exposure to Smoking in High School Students. *Preventive Medicine*, **41**, 511-520. https://doi.org/10.1016/j.ypmed.2005.01.002

[19] Chang, F.C., Lee, C.M., Lai, H.R., Chiang, J.T., Lee, P.H. and Chen, W.J. (2006) Social Influences and Self-Efficacy as Predictors of Youth Smoking Initiation and Cessation: A 3-Year Longitudinal Study of Vocational High School Students in Taiwan. *Addiction*, **101**, 1645-1655. https://doi.org/10.1111/j.1360-0443.2006.01607.x

[20] Cleveland, M.J., Feinberg, M.E., Bontempo, D.E. and Greenberg, M.T. (2008) The Role of Risk and Protective Factors in Substance Use across Adolescence. *Journal of Adolescent Health*, **43**, 157-164. https://doi.org/10.1016/j.jadohealth.2008.01.015

[21] Sun, W., Andreeva, V.A., Unger, J.B., Conti, D.V., Chou, C.-P., Palmer, P.H., *et al.* (2006) Age-Related Smoking Progression among Adolescents in China. *Journal of Adolescent Health*, **39**, 686-693. https://doi.org/10.1016/j.jadohealth.2006.04.023

[22] Damianaki, A., Kaklamani, S., Tsirakis, S., Clarke, R., Tzanakis, N. and Makris, D. (2008) Risk Factors for Smoking among School Adolescents in Greece. *Child: Care, Health and Development*, **34**, 310-315. https://doi.org/10.1111/j.1365-2214.2008.00839.x

[23] DiFranza, J.R. and Ursprung, W.S.A. (2008) The Latency to the Onset of Nicotine Withdrawal: A Test of the Sensitization-Homeostasis Theory. *Addictive Behaviors*, **33**, 1148-1153. https://doi.org/10.1016/j.addbeh.2008.04.011

[24] Ditre, J.W., Coraggio, J.T. and Herzog, T.A. (2008) Associations between Parental Smoking Restrictions and Adolescent Smoking. *Nicotine & Tobacco Research*, **10**, 975-983. https://doi.org/10.1080/14622200802087549

[25] Droomers, M., Schrijvers, C.T., Casswell, S. and Mackenbach, J.P. (2005) Father's Occupational Group and Daily Smoking during Adolescence: Patterns and Predictors. *American Journal of Public Health*, **95**, 681-688. https://doi.org/10.2105/AJPH.2002.002774

[26] Engels, R.C., Hale III, W.W., Noom, M. and Vries, H.D. (2005) Self-Efficacy and Emotional Adjustment as Precursors of Smoking in Early Adolescence. *Substance Use & Misuse*, **40**, 1883-1893. https://doi.org/10.1080/10826080500259612

[27] El-Roueiheb, Z., Tamim, H., Kanj, M., Jabbour, S., Alayan, I. and Musharrafieh, U. (2008) Cigarette and Waterpipe Smoking among Lebanese Adolescents: A Cross-Sectional Study, 2003-2004. *Nicotine & Tobacco Research*, **10**, 309-314. https://doi.org/10.1080/14622200701825775

[28] Ellickson, P.L., Tucker, J.S. and Klein, D.J. (2008) Reducing Early Smokers' Risk for Future Smoking and Other Problem Behavior: Insights from a Five-Year Longitudinal Study. *Journal of Adolescent Health*, **43**, 394-400.

https://doi.org/10.1016/j.jadohealth.2008.03.004

[29] Ennett, S.T., Faris, R., Hipp, J., Foshee, V.A., Bauman, K.E., Hussong, A., *et al.* (2008) Peer Smoking, Other Peer Attributes, and Adolescent Cigarette Smoking: A Social Network Analysis. *Prevention Science*, **9**, 88-98. https://doi.org/10.1007/s11121-008-0087-8

[30] Faeh, D., Viswanathan, B., Chiolero, A., Warren, W. and Bovet, P. (2006) Clustering of Smoking, Alcohol Drinking and Cannabis Use in Adolescents in a Rapidly Developing Country. *BMC Public Health*, **6**, 169. https://doi.org/10.1186/1471-2458-6-169

[31] Finkelstein, D.M., Kubzansky, L.D. and Goodman, E. (2006) Social Status, Stress, and Adolescent Smoking. *Journal of Adolescent Health*, **39**, 678-685. https://doi.org/10.1016/j.jadohealth.2006.04.011

[32] Glanz, K., Mau, M., Steffen, A., Maskarinec, G. and Jacob Arriola, K. (2007) Tobacco Use among Native Hawaiian Middle School Students: Its Prevalence, Correlates and Implications. *Ethnicity and Health*, **12**, 227-244. https://doi.org/10.1080/13557850701234948

[33] Grenard, J.L., Guo, Q., Jasuja, G.K., Unger, J.B., Chou, C.-P., Gallaher, P.E., *et al.* (2006) Influences Affecting Adolescent Smoking Behavior in China. *Nicotine & Tobacco Research*, **8**, 245-255. https://doi.org/10.1080/14622200600576610

[34] Grotvedt, L., Stigum, H., Hovengen, R. and Graff-Iversen, S. (2008) Social Differences in Smoking and Snuff Use among Norwegian Adolescents: A Population Based Survey. *BMC Public Health*, **8**, 322. https://doi.org/10.1186/1471-2458-8-322

[35] Gutschoven, K. and Van den Bulck, J. (2005) Television Viewing and Age at Smoking Initiation: Does a Relationship Exist between Higher Levels of Television Viewing and Earlier Onset of Smoking? *Nicotine & Tobacco Research*, **7**, 381-385. https://doi.org/10.1080/14622200500125260

[36] Henderson, M., Ecob, R., Wight, D. and Abraham, C. (2008) What Explains Between-School Differences in Rates of Smoking? *BMC Public Health*, **8**, 218. https://doi.org/10.1186/1471-2458-8-218

[37] Hanewinkel, R. and Sargent, J.D. (2007) Exposure to Smoking in Popular Contemporary Movies and Youth Smoking in Germany. *American Journal of Preventive Medicine*, **32**, 466-473. https://doi.org/10.1016/j.amepre.2007.02.025

[38] Henriksen, L., Feighery, E.C., Schleicher, N.C., Cowling, D.W., Kline, R.S. and Fortmann, S.P. (2008) Is Adolescent Smoking Related to the Density and Proximity of Tobacco Outlets and Retail Cigarette Advertising near Schools? *Preventive Medicine*, **47**, 210-214. https://doi.org/10.1016/j.ypmed.2008.04.008

[39] Hublet, A., De Bacquer, D., Valimaa, R., Godeau, E., Schmid, H., Rahav, G., *et al.* (2006) Smoking Trends among Adolescents from 1990 to 2002 in Ten European Countries and Canada. *BMC Public Health*, **6**, 280. https://doi.org/10.1186/1471-2458-6-280

[40] Kokkevi, A., Richardson, C., Florescu, S., Kuzman, M. and Stergar, E. (2007) Psychosocial Correlates of Substance Use in Adolescence: A Cross-National Study in Six European Countries. *Drug and Alcohol Dependence*, **86**, 67-74. https://doi.org/10.1016/j.drugalcdep.2006.05.018

[41] Kristjansson, A.L., Sigfusdottir, I.D., Allegrante, J.P. and Helgason, A.R. (2009) Parental Divorce and Adolescent Cigarette Smoking and Alcohol Use: Assessing the Importance of Family Conflict. *Acta Paediatrica*, **98**, 537-542. https://doi.org/10.1111/j.1651-2227.2008.01133.x

[42] Leatherdale, S.T. and Manske, S. (2005) The Relationship between Student Smoking

in the School Environment and Smoking Onset in Elementary School Students. *Cancer Epidemiology Biomarkers & Prevention*, **14**, 1762-1765. https://doi.org/10.1158/1055-9965.EPI-05-0065

[43] Leatherdale, S.T. and Strath, J.M. (2007) Tobacco Retailer Density Surrounding Schools and Cigarette Access Behaviors among Underage Smoking Students. *Annals of Behavioral Medicine*, **33**, 105-111. https://doi.org/10.1207/s15324796abm3301_12

[44] Lee, L., Paul, C., Kam, C. and Jagmohni, K. (2005) Smoking among Secondary School Students in Negeri Sembilan, Malaysia. *Asia-Pacific Journal of Public Health*, **17**, 130-136. https://doi.org/10.1177/101053950501700212

[45] Lovato, C., Sabiston, C., Hadd, V., Nykiforuk, C. and Campbell, H. (2007) The Impact of School Smoking Policies and Student Perceptions of Enforcement on School Smoking Prevalence and Location of Smoking. *Health Education Research*, **22**, 782-793. https://doi.org/10.1093/her/cyl102

[46] Mathur, C., Stigler, M.H., Perry, C.L., Arora, M. and Reddy, K.S. (2008) Differences in Prevalence of Tobacco Use among Indian Urban Youth: The Role of Socioeconomic Status. *Nicotine & Tobacco Research*, **10**, 109-116. https://doi.org/10.1080/14622200701767779

[47] Muilenburg, J.L. and Legge, J.S. (2008) African American Adolescents and Menthol Cigarettes: Smoking Behavior among Secondary School Students. *Journal of Adolescent Health*, **43**, 570-575. https://doi.org/10.1016/j.jadohealth.2008.08.017

[48] Murnaghan, D.A., Sihvonen, M., Leatherdale, S.T. and Kekki, P. (2007) The Relationship between School-Based Smoking Policies and Prevention Programs on Smoking Behavior among Grade 12 Students in Prince Edward Island: A Multilevel Analysis. *Preventive Medicine*, **44**, 317-322. https://doi.org/10.1016/j.ypmed.2007.01.003

[49] Poulin, C.C. (2007) School Smoking Bans: Do They Help/Do They Harm? *Drug and Alcohol Review*, **26**, 615-624. https://doi.org/10.1080/09595230701613619

[50] Ridout, F., Charlton, A. and Hutchison, I. (2008) Health Risks Information Reaches Secondary School Smokers. *Health Education Research*, **23**, 1039-1048. https://doi.org/10.1093/her/cym085

[51] Rodriguez, D., Romer, D. and Audrain-McGovern, J. (2007) Beliefs about the Risks of Smoking Mediate the Relationship between Exposure to Smoking and Smoking. *Psychosomatic Medicine*, **69**, 106-113. https://doi.org/10.1097/PSY.0b013e31802e0f0e

[52] Schnohr, C.W., Kreiner, S., Rasmussen, M., Due, P. and Diderichsen, F. (2009) School-Related Mediators in Social Inequalities in Smoking: a Comparative Cross-Sectional Study of 20,399 Adolescents. *International Journal for Equity in Health*, **8**, 17. https://doi.org/10.1186/1475-9276-8-17

[53] Sinha, D.N., Gupta, P.C. and Gangadharan, P. (2007) Tobacco Use among Students and School Personnel in India. *Asian Pacific Journal of Cancer Prevention*, **8**, 417.

[54] Sinha, D.N., Gupta, P.C., Reddy, K.S., Prasad, V.M., Rahman, K., Warren, C.W., *et al.* (2008) Linking Global Youth Tobacco Survey 2003 and 2006 Data to Tobacco Control Policy in India. *Journal of School Health*, **78**, 368-373. https://doi.org/10.1111/j.1746-1561.2008.00316.x

[55] Thrasher, J.F., Jackson, C., Arillo-Santillán, E. and Sargent, J.D. (2008) Exposure to Smoking Imagery in Popular Films and Adolescent Smoking in Mexico. *American Journal of Preventive Medicine*, **35**, 95-102. https://doi.org/10.1016/j.amepre.2008.03.036

[56] Wen, X., Chen, W., Muscat, J.E., Qian, Z., Lu, C., Zhang, C., *et al.* (2007) Modifiable Family and School Environmental Factors Associated with Smoking Status among Adolescents in Guangzhou, China. *Preventive Medicine*, **45**, 189-197. https://doi.org/10.1016/j.ypmed.2007.02.009

[57] Wen, X., Chen, W., Qian, Z., Muscat, J.E., Lu, C. and Ling, W. (2008) Differences in Students' Smoking-Related Knowledge, Attitudes, and Behaviors among Public, Factory, and Private Secondary Schools in Guangzhou, China. *Journal of School Health*, **78**, 46-53. https://doi.org/10.1111/j.1746-1561.2007.00265.x