**EMERGING THEMES
IN EPIDEMIOLOGY**

ANALYTIC PERSPECTIVE   **Open Access**

The individualistic fallacy, ecological studies and instrumental variables: a causal interpretation

Tom Loney[*] and Nico J Nagelkerke

Abstract

The validity of ecological studies in epidemiology for inferring causal relationships has been widely challenged as observed associations could be biased by the Ecological Fallacy. We reconsider the important design components of ecological studies, and discuss the conditions that may lead to spurious associations. Ecological associations are useful and valid when the ecological exposures can be interpreted as Instrumental Variables. A suitable example may be a time series analysis of environmental pollution (e.g. particulate matter with an aerodynamic diameter of <10 micrometres; PM_{10}) and health outcomes (e.g. hospital admissions for acute myocardial infarction) as environmental pollution levels are a cause of individual exposure levels and not just an aggregate measurement. Ecological exposures may also be employed in situations (perhaps rare) where individual exposures are known but their associations with health outcomes are confounded by unknown or unquantifiable factors. Ecological associations have a notorious reputation in epidemiology and individualistic associations are considered superior to ecological associations because of the "ecological fallacy". We have argued that this is incorrect in situations in which ecological or aggregate exposures can serve as an instrumental variable and associations between individual exposure and outcome are likely to be confounded by unmeasured variables.

Keywords: Aggregate studies, Bias, Ecological fallacy, Environmental health, Instrumental variables, Methodological individualism

Introduction

Ecological studies are epidemiological investigations in which either the units of analysis are populations or groups of people, as opposed to individuals, or exposures are only known at the population level while outcomes may be known at the individual level. Specifically, ecological variables are properties of groups, organisations, or places, whereas individual-level variables are properties of each person [1]. Generally, public and environmental health researchers utilise ecological study designs to explore potential causal associations between one or more exposures and a specific health outcome when alternative study designs (e.g. case–control, cohort, randomised controlled trial) are not possible or relevant. For example, an ecologic study is the most appropriate research design if we were interested in the effect of a macro-level governmental policy change, such as an inner-city traffic congestion charge to reduce carbon emissions

and improve air quality, on a particular aggregate health outcome (e.g. number of consultations for childhood respiratory disorders).

Ecological fallacy

An ecological study design may also be utilised when the underlying question regards individuals, such as when one is interested in the effects of air quality on health. In this context, ecological studies are potentially susceptible to the "ecological fallacy"; biases that may occur when an observed relationship between aggregated variables differs from the true, i.e. causal, association at an individual level [2]. Indeed, since Robinson in 1950 showed that correlations on an individual level can differ markedly from those on an aggregate, ecological level, epidemiological textbooks, papers and courses have warned us about the dangers of this ecological fallacy [3,4].

Although ecological studies are still widely carried out, this sceptical attitude towards ecological studies has led to the "individualistic primacy", i.e. the belief that associations on an individual level are intrinsically more truthful,

* Correspondence: tom.loney@uaeu.ac.ae
Institute of Public Health, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

i.e. better reflecting causal relationships, than those on an ecological level [5]. Public and environmental health practitioners in particular tend to assume low level validity of ecological studies. An additional reason for this scepticism may be that the implicit objective of mainstream epidemiological research is the understanding of disease aetiology, which is intrinsically an individualistic process. For example, asbestos exposure in a neighbourhood can only cause disease in individuals who were themselves exposed.

The supposed (ab)use of ecological associations has even given rise to some occasional mockery (spoof), for instance by showing an association between Nobel prizes awarded to a country, as a proxy for cognitive functioning, and per capita chocolate consumption [6]. This rather facetious example illustrates the potential weaknesses of ecological studies. First, of course, there are many obvious confounders that may explain the association between aggregate chocolate consumption and the number of Nobel prizes in a country, but also – more seriously, one can doubt whether aggregate consumption can be considered a good predictor or "cause" of individual chocolate consumption (by the Nobel laureates). Although originally formulated with nuance, the Simpson's "paradox" which has also become part of the canon of epidemiology, has strengthened this attitude [7]. Simpson's paradox (or the Yule-Simpson effect) is a relationship between two variables that is observed in different groups of data that reverses when these groups are combined or aggregated. That said, Rothman [8] posited that Simpson's paradox is not a 'true' paradox, rather the logical consequence of failing to recognise the presence of confounding variables. In the above example, it might be that academics tend to have the lowest chocolate consumption of all citizens, so that within a country the association is the inverse of that on a country level.

The concept of ecological fallacy, however, seems to encompass several, different, potential biases, which may explain why many authors seem to have difficulty precisely delineating this concept. One of these biases, as explained above, is classical confounding, i.e. that ecological, aggregated, units differ in more (unmeasured) aspects than the (aggregate) exposure of interest, and that these aspects may be related to the outcome. This ecological confounding may also be caused by confounding on an individual level (say by age or sex) when those confounders cannot be measured or observed and their distribution differs among ecological units. Another form of bias is model specification bias, which can occur when the relationship between individual exposure and outcome is non-linear and the mean outcome and the outcome of the mean exposure therefore differ [9]. The resulting difference between the expectation of an estimator from an ecological study and the individual-level parameter of interest is also known as ecological or cross-level bias. In

general, this latter form of bias is unlikely to introduce associations that do not exist, but rather leads to biased estimates in the correct direction of effects, which is often less troubling than the possibility of introducing spurious associations and thus qualitatively incorrect conclusions. Tests for the null hypothesis of no associations are still unbiased. The same is true for the apparent difference in association that can occur due to the process of aggregation itself. The correlation between individual income and health expenditure on an individual level and that of *average* income and *average* health expenditure on a municipal level is bound to differ, either magnified or diluted, even when the latter association is completely explained by the former. In that case however, conclusions would still hold qualitatively as the direction of the association will be the same. As such, it is qualitative error, i.e. reversion of the association direction due to aggregation, or the introduction of spurious effects, that is often of greatest concern in epidemiology and is the main issue we consider in this paper.

Individualistic fallacy

Although this negative attitude towards ecological studies has been criticised for putting too much emphasis on the individual as the unit for analysis [10], which even led to the term "individualistic fallacy", this has gained much less attention in the epidemiological literature. Some of this critique pointed at the intrinsic, conceptual difference between variables on an individual and group level [2]. For example, that in addition to personal poverty (say) the poverty or deprivation of the society or area in which one lives also matters. In addition, some variables only exist, or make sense, at an aggregated or societal level, such as whether one lives in a democracy or not. In other contexts, societal level variables also obviously impact relevantly on an individual's risk. For example, the risk of contracting a sexually transmitted infection is not only determined by an individual's risk behaviour but also by the population prevalence of the infection, which in turn is largely determined by "societal" factors such as the connectedness of sexual networks. While this criticism of the individual primacy is correct, we feel that this critique overlooks some important points. Here we argue that even when the emphasis is clearly on the individual, with the ultimate objective of understanding the aetiology of disease, ecological analyses can be more "truthful" than individualistic analyses and may avoid a large portion of confounding on the individual level.

Objectives

The aim of this paper is to extricate the confusion surrounding ecological studies by re-evaluating the important design components and conditions that may lead to

bias. We propose that ecological associations are useful and informative for individual associations when the ecological exposures can be interpreted as Instrumental Variables (IV).
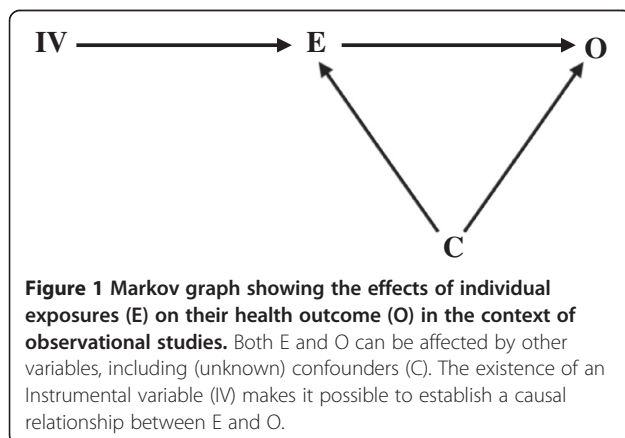
Instrumental variables

Public health epidemiology cannot lay claim to discovering the utility of IV in ecological studies as the earliest applications of "curve shifters" (variables with IV characteristics) can be traced to the econometrics literature in the 1930's [11,12]. Indeed, PG Wright introduced IV's for estimating the elasticities of supply and demand for flaxseed, the source of linseed oil [11,12]. For our re-assessment of the value of ecological associations, we point at the analogy between some ecological variables and IV and posit that ecological analyses are useful and meaningful when ecological variables can be interpreted as an IV *i.e.* a variable that is correlated with, or preferably a "cause of" the individual exposure, but whose association with the outcome of interest is not itself confounded, at least not by unmeasured variables (Figure 1). In addition, the instrumental variable should only affect the outcome via the individual exposure. That is, there should be no causal pathways linking IV and outcome that bypass the exposure. An epidemiological example where instrumental variables have shown to be useful is that of "Mendelian randomisation" (MR) [13]. For example, the blood pressures of individuals with alcohol aldehyde dehydrogenase deficiency mutations are compared to those without such mutations, as "on average" carriers of this mutation drink less than non-carriers [14]. Note that in this example, the use of MR also obviates the need to measure individual alcohol consumption, something that in addition to potential confounding by behavioural variables is fraught with methodological problems anyway [15]. However, MR has also been fruitfully used to evaluate a causal effect of high-density lipoprotein in blood on cardiovascular disease [16]. In addition, intention-to-treat analysis of a randomised controlled trial is also essentially based on instrumental

variables, as when compliance is <100% random assignment to treatment is not identical to actual treatment received, but is not influenced by the potentially confounding factors that influence individual treatment compliance.

Noteworthy is that intention-to-treat analysis does not estimate the average causal effect among compliers and is therefore subject to "dilution" bias i.e. a bias towards the null due to misclassified exposure. Treatment allocation is typically a good instrumental variable when by using blinding there are no causal links between the randomisation arm and outcome that bypass actual treatment. In the context of clinical trials, it has been established that monotonicity is an essential requirement of treatment assignment to be a good IV. Basically, in the context of ecological studies, monotonicity means that increases at the ecological level do not lead to decreases at the individual level. For example, suppose that in heavily air polluted areas people stay indoors so often in fact that they are less exposed to particulate matters than individuals living in areas with less air pollution then there is no monotonicity. Linearity is not required, although non-linear relationships can be challenging in the analysis [17,18]. In randomised controlled trial causal analysis terminology, this means that there are no "defiers" [19]. This idea can be simply extrapolated to other IVs. In the case of MR and alcohol consumption for example, monotonicity is plausible, as it is unlikely that anyone without the aldehyde dehydrogenase deficiency mutation would have had a higher weekly alcohol consumption in the counterfactual situation in which he/she carried the mutation. In the context of ecological associations, the IV idea can be easily translated to time series investigations when over the period of varying ecological exposure, such as air pollution or the introduction of specific legislation, the population remains the same. Monotonicity, however, could be violated if health warning about levels of environmental exposure would change an individual's behaviour.

An appealing feature of using an ecological IV is the potential to answer research questions that may not be possible using alternative observational research designs at the individual level. For this reason, ecological studies are also popular in studies where individual observations are often hard to obtain [16]. For example, the introduction of a legal obligation to wear bicycle helmets has been explored as an IV for a study on the impact of wearing (bicycle) helmets on risk of injury. On an individual level wearing helmets is probably confounded with factors such as "carefulness" and "risk taking". However, a legal obligation to wear helmets is unlikely to affect "carefulness" or other traffic behaviour but is an independent reason for wearing such helmets. Although the existence of defiers (people who stopped wearing

Figure 1 Markov graph showing the effects of individual exposures (E) on their health outcome (O) in the context of observational studies. Both E and O can be affected by other variables, including (unknown) confounders (C). The existence of an Instrumental variable (IV) makes it possible to establish a causal relationship between E and O.

helmets because of the obligation to wear one) cannot be completely ruled out they are probably rare.

Robinson [20], specifically, employed an ecological approach to explore the effectiveness of introducing a mandatory bicycle helmet law in Australia on reducing overall cyclist injury morbidity and mortality. Robinson [20] utilised population level data for injured cyclist hospital admissions and matched aggregate survey data on helmet wearing rates pre-law, first year of law and second year of law to argue that cycle helmets do not provide marked safety benefits for cyclists at the population level. Despite large increases in (mandatory) helmet use as a result of these helmet laws, the proportions of cyclists and pedestrians admitted to hospital for head injuries followed a similar, declining trend. As such, Robinson [20] attributed the observed reduction in cyclists with head injuries to the implementation of other major road safety initiatives at the same time such as the enforcement of speed limits, drink-driving laws, treatment of accident black spots, etc. Interestingly, the greatest effect of introducing the cycle helmet law appeared to discourage cycling as the pre- and post-law surveys showed reductions in numbers of child cyclists 15 and 2.2 times greater than the increase in numbers of children wearing helmets; an observation that would seem to invalidate the condition that an IV should only affect the outcome via the exposure of interest. As another example, alcohol consumption has many, often adverse, effects on health. However, on an individual level alcohol consumption is associated with many other behaviours, such as smoking and sexual risk taking, all of which may act as confounders when causally related to a health outcome of interest. Some interventions that affect alcohol consumption, such as the "natural experiment" of the 1920–1933 prohibition in the United States of America (USA), may offer opportunities to bypass this confounding and may thus act as instrumental variables. However, for this to be valid it is essential that the population could be considered (more or less) constant so that the concepts of complier and defier make sense. This latter condition would make long term time series (years, decades) questionable.

In other ecological contexts this criterion of monotonicity may not be so simple to apply; for example, when we compare regions A and B with different populations and different ecological exposure levels X. Direct comparison of an outcome measure y between A and B may not be appropriate as A and B may have different levels of other factors C that causally influence y, and are therefore confounders or effect modifiers. Another limitation of IVs is that they only can be expected to say something about the causal effect of individual exposures x_i on y if we assume that the relationship between x_i and y is monotonic. Thus, any exposure that has, say,

a U-shaped relationship with outcome should preferably not be the subject of ecological analysis.

Estimation

The existence of an ecological IV makes it possible to establish causality of the association between exposure and outcome, even when individual exposures are not observed but the IV for each individual is known. However, this is only a qualitative conclusion. In addition, it may also be desirable to estimate the strength of the relationship between individual exposures and outcome. When individual exposures are also measured this is straightforward using either two-stage least squares of residual insertion [21,22]. In other situations this can be challenging especially in the context of non-linear relationships.

Air pollution time series

One example in which an ecological variable would seem to have proper IV characteristics is in the case of time-varying air pollution, such as particulate matter with an aerodynamic diameter of <10 micrometres; PM_{10}. In many places such concentrations vary considerably, and often rapidly (within hours or days), over time. Although measured average levels may not accurately reflect individual exposures, it seems plausible that there exist close correlations and that therefore the ecological associations can be used to study the association between air pollution and (say) episodes of asthma exacerbation. Plausible confounders, for instance temperature, are measurable and can therefore be adjusted for. Also, since largely the same individuals are observed during different time periods, the concept of defier is easy to interpret, *viz.* someone who has lower individual exposure during times of high environmental levels. Of course, such studies would seem less appropriate for outcomes that are the result of chronic cumulative exposures e.g. cardiopulmonary- or cancer-related mortality such as in the American Cancer Society Study [23].

Geographical comparisons

More difficult is the ecological comparison between geographical areas, for example two areas with largely different air pollution levels (e.g. rural agricultural region versus an urban industrial zone). As the two populations are different, the concept of "defier" only has a counterfactual interpretation, *viz.* as someone, if he had happened to have resided in the area with the higher pollution levels would – personally, have lower exposure to air pollution (or *vice versa*). Such (counterfactual) individuals are probably rare if the areas are similar in other (relevant) aspects. However, this may not always be the case. Individuals with asthma, for example, may choose not to live in areas with high air pollution levels

and are thus defiers. There is no gold standard test to assess whether geographically defined environmental exposures have IV properties but this has to be assessed on the basis of substantive knowledge. Mean individual exposures may sometimes be plausible as IVs but sometimes not.

Poverty and health

A similar situation may exist in the widely discussed relationship between "wealth", "poverty", "social economic status", "employment status" and various (health) outcomes. On an individual level, there may be a strong relationship between unemployment, say, and (poor) health and mortality, suggesting that unemployment is causally related to adverse health outcomes. However, in addition to selection biases, this relationship may be confounded by many variables, which may make a causal interpretation questionable. The unemployed individuals may have lost their jobs due to poorer physical and/or mental health or may be less socially adaptable. An economic crisis or depression with rapid massive increases in unemployment may act as an IV, and if (specific) mortality or morbidity does not increase during such depressions then one may question causal interpretations of associations on an individualistic level. Again, since largely the same individuals are involved, it is (at least conceptually) easy to assess whether there are (significant numbers of) defiers, specifically people who economically benefit (or find employment) during times of depression. Although such individuals may exist, they are probably rare. Economic depressions may also potentially affect health through pathways not involving unemployment. For example, during depressions the (still) employed may avoid behaviours that would put their employment in jeopardy. Although these alternate causal pathways could potentially invalidate the value of depressions as an IV their importance would seem minor. Similarly, natural disasters or other "natural experiments" may similarly occasionally provide good instrumental variables. A case in point is the widely accepted association between tuberculosis and poverty [24-27]; while almost all cross-sectional studies show this association the rapid increase in poverty in the USA and Europe during the 1930's did not give rise to increases in tuberculosis mortality [28].

As a further example of the use of IV in intervention, rather than observational, studies, we consider the association between regional *sex* worker interventions and human immunodeficiency virus (HIV) prevalence in India. This is to illustrate situations in which ecological (geographical) levels of a variable may have suitable IV characteristics. In the four high HIV-prevalence states (Andhra Pradesh, Maharashtra, Karnataka, Tamil Nadu) of Southern India interventions targeting sex workers (e.g.

condom use, sexually transmitted disease treatment) are implemented on a district level. Effects of the interventions on district-level HIV prevalence and incidence are also measured on a district level by the district-specific HIV prevalence in young (<25 years) ante-natal care attendees. The idea behind these interventions is that prevention of transmission (by safer sex practices) between sex workers and their clients would reduce overall HIV transmission. These interventions received strong support from the Bill and Melinda Gates Foundation in the form of the "Avahan" project [29]. The existence of defiers, i.e. people with a lower risk of HIV infection in the absence of an intervention, is unlikely, thus monotonicity seems plausible. So in principle, intervention at a district level could be a good IV.

Conclusion

Researchers should be reminded that no perfect epidemiological study exists and all research designs have both strengths and weaknesses. Hence, the selection of an appropriate research design requires deep consideration of the design components, research question and ability to estimate or quantify important exposure variables related to the outcome of interest. While ecological associations have an infamous reputation in epidemiology and differences between individualistic and ecological associations are often attributed to the "ecological fallacy" we have argued that this is incorrect in situations in which ecological or aggregate exposures can serve as an instrumental variable and associations between individual exposure and outcome are likely to be confounded by unmeasured variables. Thus, instrumental variables, and by consequence some ecological studies, have the advantage of reducing bias. We believe that this paper delineates situations in which ecological associations may provide reliable causal interpretations and thus provides a yardstick for assessing whether the ecological fallacy can be expected. For assessment of whether a causal relationship exists, neither individual studies, nor ecological (IV) studies by themselves suffice. Even experimental studies, often inadequately mimicking "wild type" exposures, do not suffice. Deep conclusions often require synthesis of evidence of different kinds [30]. It is our opinion that proper ecological studies can contribute to this evidence.

References

1. Greenland S: Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol* 2001, **30**(6):1343–1350.
2. Schwartz S: The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *Am J Public Health* 1994, **84**(5):819–824.
3. Robinson WS: Ecological correlations and the behaviour of individuals. *Am Soc Rev* 1950, **15**(3):351–357.
4. Wakefield J: Ecologic studies revisited. *Annu Rev Public Health* 2008, **29**:75–90.
5. Dufault B, Klar N: The quality of modern cross-sectional ecologic studies: a bibliometric review. *Am J Epidemiol* 2011, **174**(10):1101–1107.
6. Messerli FH: Chocolate consumption, cognitive function, and Nobel laureates. *N Engl J Med* 2012, **367**(16):1562–1564.
7. Simpson EH: The interpretation of interaction in contingency tables. *J R Stat Soc Ser B* 1951, **13**:238–241.
8. Rothman KJ: A pictorial representation of confounding in epidemiologic studies. *J Chron Dis* 1975, **28**:101–108.
9. Sheppard L: Insights on bias and information in group-level studies. *Biostatistics* 2003, **4**(2):265–278.
10. Alker HA Jr: A typology of ecological fallacies. In *Quantitative Ecological Analysis*. Edited by Dogan M, Rokkan S. Cambridge: Massachusetts Institute of Technology; 1969:69–86.
11. Angrist JD, Krueger AB: Instrumental variables and the search for identification: from supply and demand to natural experiments. *J Econ Perspect* 2001, **15**(4):69–85.
12. Stock JH, Trebbi F: Who invented instrumental variable regression? *J Econ Perspec* 2003, **17**(3):177–194. 85.
13. Sheehan NA, Meng S, Didelez V: Mendelian randomisation: a tool for assessing causality in observational epidemiology. *Methods Mol Biol* 2011, **713**:153–166.
14. Chen L, Davey Smith G, Harbord RM, Lewis SJ: Alcohol intake and blood pressure: a systematic review implementing a Mendelian randomisation approach. *PLoS Med* 2008, **5**(3):e52.
15. Høyer G, Nilssen O, Brenn T, Schirmer H: The Svalbard study 1988–89: a unique setting for validation of self-reported alcohol consumption. *Addiction* 1995, **90**(4):539–544.
16. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Hólm H, Ding EL, Johnson T, Schunkert H, Samani NJ, Clarke R, Hopewell JC, Thompson JF, Li M, Thorleifsson G, Newton-Cheh C, Musunuru K, Pirruccello JP, Saleheen D, Chen L, Stewart A, Schillert A, Thorsteinsdottir U, Thorgeirsson G, Anand S, Engert JC, Morgan T, Spertus J, et al: Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet* 2012, **380**(9841):572–580.
17. Prentice RL, Sheppard L: Validity of international, time trend, and migrant studies of dietary factors and disease risk. *Prev Med* 1989, **18**(2):167–179.
18. Prentice RL, Sheppard L: On the reliability and precision of within- and between- population estimates of relative rate parameters. *Biometrics* 1995, **51**(3):853–863.
19. Freedman DA: Statistical models for causation: what inferential leverage do they provide? *Eval Rev* 2006, **30**(6):691–713.
20. Robinson DL: Head injuries and bicycle helmet laws. *Acc Anal Prev* 1996, **28**(4):463–475.
21. Nagelkerke N, Fidler V, Bernsen R, Borgdorff M: Estimating treatment effects in randomised clinical trials in the presence of non-compliance. *Stat Med* 2000, **19**(14):1849–1864.
22. Wooldridge JM: Chapter 15: Instrumental variables and two stage least squares. In *Introductory Econometrics: A Modern Approach*. 4th edition. Edited by Wooldridge JM. Mason: South-Western Cengage Learning; 2009:506–545.
23. Krewski D, Burnett R, Jerrett M, Pope CA, Rainham D, Calle E, Thurston G, Thun M: Mortality and long-term exposure to ambient air pollution: ongoing analyses based on the American Cancer Society cohort. *J Toxicol Environ Health A* 2005, **68**(13–14):1093–1109.
24. Dewan PK, Arguin PM, Kiryanova H, Kondroshova NV, Khorosheva TM, Laserson K, Kluge H, Jakubowiak W, Wells C, Kazionny B: Risk factors for death during tuberculosis treatment in Orel, Russia. *Int J Tuberc Lung Dis* 2004, **8**(5):598–602.
25. Hoa NB, Tiemersma EW, Sy DN, Nhung NV, Gebhard A, Borgdorff MW, Cobelens FG: Household expenditure and tuberculosis prevalence in Vietnam: prediction by a set of household indicators. *Int J Tuberc Lung Dis* 2011, **15**(1):32–37.
26. Jackson S, Sleigh AC, Wang GJ, Liu XL: Poverty and the economic effects of TB in rural China. *Int J Tuberc Lung Dis* 2006, **10**(10):1104–1110.
27. Pang PT, Leung CC, Lee SS: Neighbourhood risk factors for tuberculosis in Hong Kong. *Int J Tuberc Lung Dis* 2010, **14**(5):585–592.
28. Nagelkerke NJD: *Courtesans and Consumption*. Delft, Netherlands: Eburon Academic Publishers; 2012.
29. Ng M, Gakidou E, Levin-Rector A, Khera A, Murray CJ, Dandona L: Assessment of population-level effect of Avahan, an HIV-prevention initiative in India. *Lancet* 2011, **378**(9803):1643–1652.
30. Cox DR: Commentary: smoking and lung cancer: reflections on a pioneering paper. *Int J Epidemiol* 2009, **38**(5):1192–1193.