

# Multinomial logistic regression modelling of cardiologists' awareness of the impact of air pollution on cardiovascular disease in Vietnam and the Philippines.

Nguyen, Q.T. , Naguib, R. , Papathomas, N. , Shaker, M.H. , Culaba, A.B. , Wickramasinghe, N. and Ton, T.N.

**Author post-print (accepted) deposited in CURVE June 2013**

**Original citation & hyperlink:**

Nguyen, Q.T. , Naguib, R. , Papathomas, N. , Shaker, M.H. , Culaba, A.B. , Wickramasinghe, N. and Ton, T.N. (2012) Multinomial logistic regression modelling of cardiologists' awareness of the impact of air pollution on cardiovascular disease in Vietnam and the Philippines. International Journal of Biomedical Engineering and Technology, volume 10 (1): 38-54 <http://dx.doi.org/10.1504/IJBET.2012.049324>

**Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.**

**This document is the author's post-print version of the journal article, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.**

**CURVE is the Institutional Repository for Coventry University**  
<http://curve.coventry.ac.uk/open>

---

## Multinomial logistic regression modelling of cardiologists' awareness of the impact of air pollution on cardiovascular disease in Vietnam and the Philippines

---

Q.T. Nguyen and R.N.G. Naguib

Biomedical Computing and Engineering Technologies (BIOCORE)  
Applied Research Group,  
Health Design Technology Institute and Faculty of Engineering and  
Computing, Coventry University, UK

Author:  
Please  
supply  
e-mail  
id and  
zipcode for  
remaining  
authors.

M. Papathomas

School of Mathematics and Statistics,  
University of St Andrews, UK

M.H. Shaker

Ecology and Environment Inc,  
Buffalo, USA

A.B. Culaba

Centre for Engineering and Sustainable Development Research,  
De La Salle University, Manila, Philippines

N. Wickramasinghe\*

Epworth Chair Health Information Management,  
School of Business Information Technology and Logistics,  
RMIT University,  
Melbourne, Australia  
E-mail: nilmini.work@gmail.com

\*Corresponding author

T.N. Ton

Environmental Health Unit,  
World Health Organization,  
Western Pacific Regional Office,  
Hanoi, Vietnam.

**Abstract:** The Air Pollution (AP) situation in Vietnam and the Philippines has deteriorated in recent years. During this time period, mortality rates from Cardiovascular Diseases (CVDs) have remained in the top ten of all causes

Copyright © 200x Inderscience Enterprises Ltd.

of deaths. The aim of this paper is to understand and model the awareness of cardiologists of the potential impact of AP on CVD in both nations. A full-scale survey covering 321 cardiologists was conducted in Vietnam (27 hospitals across the country) and the Philippines (members of the Philippine Heart Association with a subset from Davao Province). The paper reports on extensive results obtained through descriptive analysis and logistic regression modelling.

**Keywords:** particular matter; AP; air pollution; CVD; cardiovascular disease; pathophysiological pathway; developing countries.

**Reference** to this paper should be made as follows: Nguyen, Q.T., Naguib, R.N.G., Papatomas, M., Shaker, M.H., Culaba, A.B., Wickramasinghe, N. and Ton, T.N. (xxxx) 'Multinomial logistic regression modelling of cardiologists' awareness of the impact of Air Pollution on cardiovascular disease in Vietnam and the Philippines', *Int. J. Biomedical Engineering and Technology*, Vol. x, No. x, pp.xxx-xxx.

**Biographical notes:** Quynh T. Nguyen (PhD) obtained both her Master's Degree (with Distinction) in Business Decision Management and Doctorate in Computing Science from Coventry University, UK, in 2007 and 2011, respectively. Her research interests include healthcare informatics, information retrieval, e-health, electronic healthcare records systems and environmental health in developing countries.

Raouf N.G. Naguib (PhD) is a Professor of Biomedical Computing and Head of BIOCORE. Prior to this appointment, he was a Lecturer at Newcastle University, UK. He has published over 300 journal and conference papers and reports in many aspects of biomedical and digital signal processing, image processing, AI and evolutionary computation in cancer research. He was awarded the Fulbright Cancer Fellowship in 1995–1996 when he carried out research at the University of Hawaii in Mānoa, on the applications of artificial neural networks in breast cancer diagnosis and prognosis. He is a member of several national and international research committees and boards. He is also an Adjunct Research Professor at the University of Carleton, Ottawa, Canada, and an Honorary Professor at De La Salle University, Manila, Philippines.

Michail Papatomas (PhD) is a Lecturer in the School of Mathematics and Statistics at the University of St Andrews in the UK. Prior to this appointment, he was a Biostatistician at Coventry University's Centre of Excellence in Mathematics and Statistics Support. He was closely involved with a number of studies and statistical analyses undertaken within BIOCORE in relation to Environmental Health research. He is also an Honorary Research Associate in the School of Public Health at Imperial College, London, UK.

Mohyi H. Shaker (MD) is a Cardiologist whose 20 years professional experience spans across interventional cardiology, internal medicine, emergency medical support and intensive care, health systems and administration, and medical facility management. He has a wealth of experience in environmental health assessment and economics. He was the lead physician involved in the implementation and conduct of monitoring and assessment studies for the Environmental Damage/Public Health Claims in Saudi Arabia to identify the treatment costs for disease conditions that were epidemiologically associated with the public health consequences of the 1990–1991 Iraqi invasion of Kuwait. On several occasions in 2004, he attended the UN Compensation Commission hearings in Geneva, Switzerland, concerning Saudi Arabia's public health claim.

Alvin B. Culaba (PhD) is a Professor of Mechanical Engineering and Executive Vice-President for External Relations and Internationalisation at De La Salle University, Manila, Philippines. His current research interest focuses on the comprehensive analysis of low-carbon renewable energy systems using Life Cycle Assessment (LCA) and Multi-Criterion Decision Analysis (MCDA) techniques. In particular, he currently heads a multidisciplinary research project on the processing of algal biomass for biodiesel production. His other notable recent work includes optimisation of water recycling in industrial parks, identification of the best alternative technologies for light urban vehicles, modelling the dynamics of nascent energy supply chains and, in collaboration with Prof. Naguib, the evaluation of air pollution effects on health.

Nilmini Wickramasinghe (PhD) researches and teaches in several areas within Information Systems, including IT for competitive advantage, knowledge management, e-commerce and m-commerce and organisational impacts of technology. In addition, she specialises in the impacts of technologies on the healthcare industry and various aspects of medical informatics. She is well published in all these areas, having written books, numerous refereed scholarly papers and encyclopaedia entries. She regularly presents her work throughout North America as well as in Europe and Australasia. Currently, she is the Epworth Chair Health Information Management and a Professor at RMIT University's School of Business IT and Logistics.

Tuan Nghia Ton (MSES) earned his Master's Degree in Environmental Science from Indiana University in Bloomington, Indiana, USA, under a Fulbright Scholarship programme in 1995–1997. Since joining the WHO Representative Office in Vietnam in 2006, he has been a national professional officer in charge of environmental health covering water and sanitation, air pollution, climate change, chemical safety and healthcare waste management. One of his main recent interests is to successfully implement a Water Safety Plan in the urban water supply sector in Vietnam. He currently also works closely with the Vietnamese National Institute for Occupational and Environmental Health (NIOEH) in conducting research on health impacts of indoor air pollutants and setting up standards on indoor air quality.

---

## 1 Introduction

Cardiovascular Diseases (CVDs) are responsible for the largest number of deaths in the world, especially in low and middle-income countries (WHO, 2011). For over five decades, to prevent CVD mortality and morbidity, a spectrum of studies were instigated and revealed many potential causes for CVDs, such as cigarette smoking, alcohol, physical inactivity, diabetes, obesity, genetic predisposition and gender (Ayres, 2006). Scientists and professionals have additionally focused on the identification of novel Risk Factors (RFs), for example AP, in pursuance of the prediction of future development of CVDs, such as ischaemic heart disease and atherosclerosis (Bhatnagar, 2011). Thus, the consideration of the potential association between AP and CVD has recently been exposed and studied in developed countries, such as the USA and UK. According to Tofler and Muller (2006), occurrences of cardiac events can be posed by AP after days' or even hours' exposure, especially to small particles, also known as Particulate Matter (PM). In the 21st

century, the main sources of AP not only in the Western countries, but also in developing nations, are motor vehicle emissions and industrial pollutants, which usually comprise many complex constituents. Hence, the primary ambient air pollutants, such as ozone, sulphur dioxide and PM, have been the centre of focus to investigate their impact upon cardiovascular health.

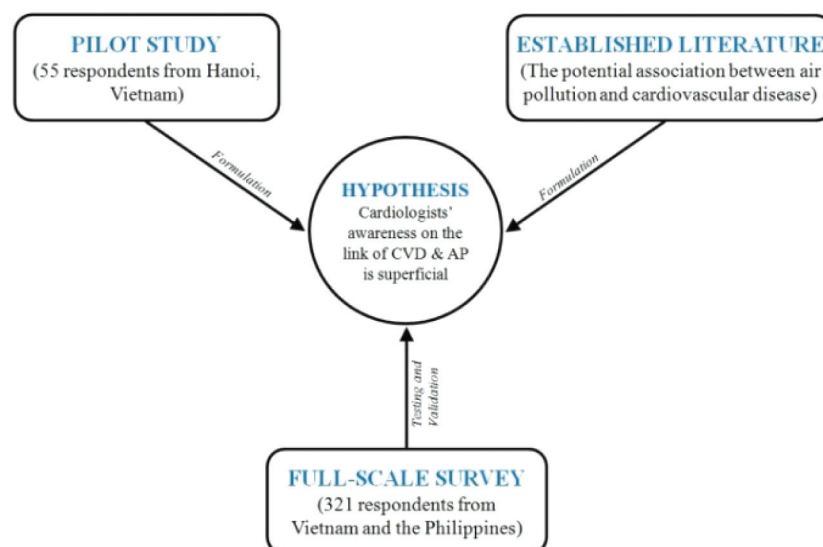
Scientifically, a wealth of epidemiologic, cohort and cross-sectional studies have affirmed that airborne pollutants have deleterious effects on mortality and morbidity of CVD in relation to long-term and short-term exposures (Bhatnagar, 2006). Developed countries, where the standards of the environment in ambient and indoor areas are in line with the Air Quality Standards from the World Health Organization (WHO), have widely acknowledged, and are fully aware of this issue. However, in many developing countries, such as Vietnam and the Philippines, which suffer from high concentrations of ambient PM and where the mortality from CVD is in the top ten of all causes of deaths, AP is a significant problem (Nguyen et al., 2011a). Su et al. (2011) reviewed all studies on AP impact on cardiovascular health in Asia that were published in peer-reviewed English journals since the 1980s and no paper on this topic was found from any country of the South East Asian region. It is, therefore, important to investigate the awareness of the populations in these countries about the association between environmental pollution and CVD to estimate the increase in risk and consequently contribute to strengthening effective preventive interventions. However, owing to the characteristic nature of AP, it is very difficult for lay people to differentiate between this and other more obvious RFs, such as (passive) smoking (Samet and Spengler, 1991). Therefore, this study is aimed at targeting cardiologists as an important group of healthcare professionals who combat CVD on a routine basis to ascertain and evaluate their level of awareness regarding the association between AP and CVD from a professional standpoint.

The study is divided as follows: Section 2 briefly explains the methodology of the study; Section 3 summarises the descriptive results of the survey; Section 4 gives information relating to the coding of the variables in SPSS®17 and the adopted statistical modelling approach; Section 5 discusses the effectiveness of multinomial Logistic Regression modelling; Section 6 reports on the results of the multivariate Logistic Regression analysis and the multicollinearity issue; finally, in Section 7, conclusions are given.

## 2 Methodology

A study of this nature necessitates the use of a formal approach to corroborate the hypothesis under investigation. Triangulation was applied in this study and comprises: the analysis and results of a pilot study, a comprehensive environmental and clinical literature review, a full-scale study conducted in Vietnam and the Philippines (Figure 1). The combination of the above-mentioned approaches leads to the formulation and support of a hypothesis relating to the awareness of South East Asian cardiologists of the potential links between AP and CVD.

Following the model for qualitative–quantitative triangulation proposed by Newman and Benz (1998), the pilot study, which was performed as a qualitative study, gives an account of the survey initially conducted at two specialised cardiac hospitals in Hanoi, Vietnam, with a sample size of 55 cardiologists. The data was analysed and indicated that the awareness of the impact of AP on CVD by the Vietnamese cardiologists is rather superficial and not commensurate with the depth of the subject, especially from the pathophysiological pathway understanding perspective. Could this be an isolated case in Vietnam or could it be

**Figure 1** Triangulation methodology as applied to the study (see online version for colours)

symptomatic of a wider problem that spans across South East Asia? This has become the hypothesis for a full-scale study that eventually covered many other provinces in Vietnam and also covered the Philippines.

The content of the survey endeavoured to cover a number of points, such as the metadata pertaining to the physicians, the current mortality rates of CVD in both countries, the changes in patterns of the diseases and patient profiles in the last 10 years, some CVD RFs of concern, the potential association between AP and CVD and some basic environmental knowledge of those cardiologists. The questionnaire was divided into either multiple-choice or open-ended questions with a total of 22 items. For Vietnam, the questionnaire was translated into the Vietnamese language to avoid any confusion that may have arisen and then translated back into English. Some of the questions resulted in one or two missing cases, and such missing cases were accounted for during the analysis. The analysis was classified into five main parts:

- Basic information about the interviewees
- Cardiologists' opinions with regard to CVD mortality
- Cardiologists' views about the current patterns of CVD and common RFs in either country
- Cardiologists' awareness of the impact of AP and of the effects of various emissions on CVD
- Suggestions to combat AP in the country.

Through an extensive review of the established literature, it is evident that, in developed countries, physicians, in general, and cardiologists, in particular, are fully aware of the possible impacts of AP on cardiac events and are fully versant in the mechanisms that may induce CVD as a result of ambient AP. However, as was clear from our initial investigation, this is not the case in many developing countries in South East Asia, which are renowned for their notoriously

polluted ambient air environment and where CVDs are the leading cause of mortality. To confirm the hypothesis of the pilot study, a full-scale study has thus been implemented covering, in total, 321 subjects from Vietnam and the Philippines. Comprehensive qualitative and quantitative analyses have been undertaken for investigating and testing this hypothesis.

### 3 Statistical descriptive summary

The complete full-scale study included 321 valid responses: 179 cardiologists from Vietnam and 142 from the Philippines.

#### 3.2 Demographic data

Cardiologists' experiences ranged from 1 to 43 years (mean: 9 years; median: 7 years). The mean experience (in years) of the Vietnamese cardiologists is slightly lower than that for their Philippine counterparts, with the longest experience being 43 and 35 years in the Philippines and Vietnam, respectively. The majority of cardiologists (76.7%) in the survey practise at tertiary or provincial hospitals/organisations, whereas the rest of the physicians practise at either secondary or rural health centres.

In the initial pilot study, one drawback was that there were not many different sub-specialisms, this being a common characteristic of cardiology departments in Vietnam; therefore, a related classification was not made. Subsequently, this issue was addressed in the ensuing full-scale study, which included a variety of sub-specialisms from the Philippines, as well as some additional ones from Vietnam. In the end, three types of cardiology were classified, including general, intervention and 'other' cardiology (such as cardiac anaesthesiology). This division was based on the actual number of sub-specialists and their equitable distribution.

#### 3.2 Preliminary Descriptive Analysis

When asked about their affirmation that CVD is the leading cause of mortality in the entire South East Asian region, 97.8% of the respondents was in agreement. Coincidentally, a similarly high percentage of cardiologists also agreed with a noticeable tendency for current additional demands for hospitalisation, medication and surgical interventions. This result also concurs with the actual situations in both countries according to the investigations outlined in the annual reports from the Ministry of Health (MOH), Vietnam, and Department of Health (DOH), the Philippines (MOH, 2009; DOH, 2010).

The majority (95.6%) of respondents noted that the population being afflicted by the disease has been getting increasingly younger over the last 10 years. Moreover, there is a noticeable widespread variation in the spectrum of patients, regardless of their socio-economic status or geographical location (whether urban or rural areas), which was observed by the cardiologists.

#### 3.3 Detailed descriptive results of the cardiologists' opinions about the association between air pollution and CVD

Analysis of the survey revealed that 64% of all cardiologists identified that many of their patients suffered from CVD without being subjected to any of the obvious RFs, such as alcohol, cigarette smoking, high blood pressure, obesity and diabetes. This

means that the majority of physicians have problems diagnosing patients who do not present with any of the RFs that commonly lead to the disease. There is, therefore, a need to clinically and scientifically explore any unusual RFs that may be linked to CVD. Furthermore, 56% of the sample tended to agree with the fact that AP is an RF to CVD. Out of the total of those who positively agreed that AP was an RF, 50.3% ( $n=90$ ) confidently proposed one or more mechanisms that involved a clear clinical pathway for cardiovascular events.

Approximately seven cardiologists suggested that AP, combined with other RFs or more general factors, such as heat or noise, would lead to CVD. Such affirmation is rather general and ambiguous; it does not provide a clinically valid explanation of the pathophysiological pathway of AP in relation to CVD. Therefore, seven responses were qualified as invalid mechanisms.

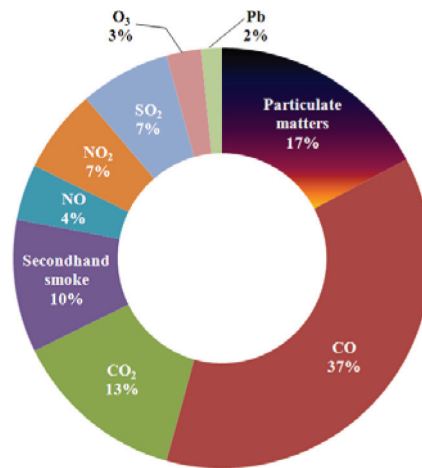
With regard to the proposed mechanisms not matching the established literature, stress and anxiety resulting from either long- or short-term exposure to ambient AP were proposed by 19 cardiologists. However, no scientifically supported evidence was proposed that they may also be a precursor to CVD. It can thus be deduced that the observations of those cardiologists who articulated a stress mechanism in relation to CVD are not based on established RFs, but that such stress mechanism is an indirect symptom that leads cardiologists to diagnose the disease. Therefore, this stress justification may not be strictly considered as a mechanism. One Philippine physician proposed that exposure to AP can promote oxidative stress and lead to the development of immature red blood cells. The former part of the mechanism complies with the literature, while the latter part (development of immature red blood cells) was not confirmed through previous studies. Hence, this mechanism once again qualified as not matching the literature.

For the mechanism(s) that matched the literature, only one physician mentioned that exposure to AP might result in congenital heart disease. Through the literature analysis, Nguyen (2011b) provided such evidence resulting from a few studies on gene-environment interaction in high-risk groups, especially in relation to AP. In Vietnam, there is a prevalence of children with congenital heart disease who live in the vicinity of industrial zones. The largest proportion of cardiologists (33 cardiologists) believes that CVD can develop from 'respiratory diseases' such as COPD. Hypertension was one of the cardiac events that was mentioned as a pathway; however, such pathway is not found in the established literature. An open question for research is to investigate whether hypertension is a consequence of respiratory disease owing to exposure to AP. Moreover, a Philippine cardiologist suggested cigarette smoking as a mechanism, which is an extremely broad reference. It can, therefore, be classified under the respiratory disease mechanism, although this may still not correctly reflect what he actually meant. The same situation is true of the passive smoking mechanism.

In relation to likely sources of pollution, only 133 out of 204 cardiologists could propose either one or more such sources, and these are described in Figure 2. Some other air pollutants were also mentioned in the survey, such as metal dusts, ammonia ( $\text{NH}_3$ ) and other mixtures, however the authors will not report on the details of those pollutants as their frequency of occurrence in the responses was extremely limited.

With regard to physicians' opinions about CVD exacerbation states, the proportional positive agreement towards the element of exacerbation of CVD conditions owing to AP, which was approximately 64% (204 respondents), was slightly higher than for AP as an RF (mentioned earlier). This figure also reflects the fact that, when asked, out of every three



**Figure 2** Number of occurrences of listed air pollutants by the cardiologists (see online version for colours)

physicians in either country, there would be one cardiologist considering no impact at all of AP on CVD exacerbation. This is an alarmingly high proportion, which reflects the lack of knowledge and awareness of the associated link between AP and the disease. In addition, those who agreed with the AP exacerbation effects were asked to list at least one individual harmful airborne pollutant.

This gives a clear picture that the deeper the knowledge of the physicians was examined, the more the respondents showed a low appreciation of the link between AP and CVD as reflected through the fewer number of positive responses the authors obtained. In fact, with each drilling-down step undertaken, there was an approximate decrease of one-third of the sample in the overall valid responses.

Focusing on the effects of PM, and as part of the questionnaire, the authors suggested three main mechanisms, which were mentioned by Brook et al. (2004) and recorded three types of responses: agreement, indecision and disagreement. Those mechanisms are:

- Inhaled particles accumulating in the lungs may cause systemic inflammation via oxidative stress, which mediates endothelial dysfunction and atherosclerosis (M1).
- Inhaled particles accumulating in the lungs may cause systemic inflammation, which may also increase blood coagulability (M2).
- Short-term exposure to air particles can be linked to Myocardial Infarction and Arrhythmia (M3).

The percentages of positive agreements decreased from the first through to the third mechanism, above, and these were 65, 58% and 31%, respectively. This, in a broad sense, matches with other studies on the PM pathways in relation to CVD and its cardiac events. For the first mechanism (M1), a spectrum of studies have shown that exposure to PM can enhance the progression and instability of underlying atherosclerosis. For the second mechanism (M2), PM inhalation can also enhance arterial thrombosis and

coagulation; however, some studies are still in contention in this respect. With regard to the third mechanism (M3), there has also been a great deal of debate and contentious studies undertaken. Studies, such as the one conducted by Pope et al. (2006), suggested that AP, especially fine particles, can trigger myocardial infarction. Brook et al. (2004) also proposed and provided evidence, in vitro, that AP can generate arrhythmia. However, there is no clinical evidence to support this mechanism in humans; hence, it remains controversial. The survey results also reflect this fact since 45% of the cardiologists were indeed undecided in relation to this mechanism.

With regard to the different possible sources of AP, the majority of physicians agreed that emissions from vehicles and industrial plants are the main contributors to CVD (70 and 72%, respectively). These figures are higher than the more general figure of 56% of cardiologists who agree that AP is an RF to CVD. This uncovers an element of misunderstanding, or shortage in knowledge, by the cardiologists that vehicles and industrial plants emissions are an important source of AP. According to the most recent report released by the Clean Air Initiative, Asia (CAI-Asia), emissions from motor vehicles in the Philippines and Vietnam are the major source of AP, followed by industrial plants (CAI-Asia, 2006).

Beyond their cardiac expertise, it is also important to uncover the general environmental knowledge of the respondents. After all, this is one of the core concerns of this study, and it is important to ascertain that physicians are aware and knowledgeable of the enormous measures developed by their respective authorities to combat ambient AP. To investigate this issue, a question probed the physicians' awareness of their respective governments' efforts in combating environmental pollution. It further encouraged cardiologists to raise additional suggestions that could practically lead to the development of effective environmental regulations and legislations. Statistically, more than half of the total numbers of physicians (55%) were not aware of any governmental efforts to battle AP issues.

#### 4 Precoding preparation

To investigate factors affecting the cardiologists' awareness of the potential association between AP and CVD, and based on the results of the above-mentioned section, a hypothesis can be established that the awareness of the cardiologists about the potential link between AP and CVD is associated with their experiences, their sub-specialisms, the type of organisations where they practise, their opinions about the pathophysiological pathway of AP in relation to the disease and their basic environmental knowledge. As a result of the nature of the data obtained from the questionnaire – mostly multinomial variables – multinomial and binary Logistic Regression models are suitable tools to use to predict and highlight the awareness of cardiologists about the potential link between AP and CVD, and to expose the relative relationships and strengths of the various RFs to CVD.

For the best model that fits the data to be achieved, all potential alternative Logistic Regression models, which comprise all possible main effects (9 predictors) and all possible two-way interactions, were included in the test. The reason why the authors selected two-way interactions is because there may be an effect on a Dependent Variable (DV) from two separate but interacting Independent Variables (IDVs). Table 1 explains the coding scheme adopted for the explanatory (or independent) variables, IDVs, and for DVs in SPSS®17.

**Table 1** Dependent and independent variables

---

*Dependent variables coding – DV*

APRF: whether a cardiologist considers AP as a risk factor to CVD or not. One of the 3 possible answers may be selected:  
*0 = No; 1 = Yes; 2 = Undecided*

APRFG: the same context as APRF, however the ‘No’ and ‘Undecided’ categories are merged:  
*0 = No or Undecided; 1 = Yes*

*Dependent variables coding – IDV*

Pathophysiological Mechanism 1 (PM1): Inhaled particles accumulating in the lungs may cause systemic inflammation via oxidative stress, which mediates endothelial dysfunction and atherosclerosis?

Pathophysiological Mechanism 2 (PM2): Inhaled particles accumulating in the lungs may cause systemic inflammation, which may also increase blood coagulability?

Pathophysiological Mechanism 3 (PM3): Short-term exposure to air particles can be linked to Myocardial Infarction and Arrhythmia?

Vehicle Emissions: Emissions from vehicles can be a significant contributor to CVD?

Industrial Plant Emissions: Emissions from industrial plants can be a significant contributor to CVD?

Those above IDVs have the same 3 possible answer of which 1 is selected:  
*0 = No; 1 = Yes; 2 = Undecided*

Grouping of cardiologists’ experience (EXPG): The length of experience of cardiologists taking part in the survey  
*0 = 1–3 years; 1 = 4–7 years; 2 = 8–11 years;  
 3 = 12 years and above*

Awareness of governmental measures against AP: cardiologists were asked of their knowledge of any governmental effort to combat air pollution. The answer can be:  
*0 = No; 1 = Yes*

Type of hospitals: Type of organisation where cardiologists are practising. It can be:  
*0 = Secondary or district hospital; 1 = Tertiary or provincial hospital; 99 = other hospital type*

Cardiologists’ sub-specialisms: The type of cardiologist which may be:  
*1 = General cardiology; 2 = Interventional cardiology;  
 0 = Other cardiology specialism*

---

The Logistic Regression model tests the significance of each variable and determines whether a variable should be removed from the model if its effect is of minimal significance. Stepwise methods, specifically Backward Elimination (using the Likelihood Ratio (LR)) in SPSS®17, are applied to automatically delete indicators of statistical insignificance in the model and achieve model parsimony. The default significance level for removal in SPSS®17 is set at 0.1.

Whilst Table 1 presents 9 IDVs, correlation analysis, nevertheless, yields low pairwise correlation between APRF/APRFG (as DV) with governmental measures, type of hospitals and cardiologists’ sub-specialisms. These three IDVs were consistently being omitted by the Logistic Regression test models, thus providing evidence that they were not of significance to the analysis.

Moreover, the pairwise correlation test shows that emissions from vehicles and industrial plants are highly correlated ( $r=0.87$ ); hence, Logistic Regression Backward LR also omits one of those two variables from the models, but will consider each one alternately. Therefore, for ongoing testing, all models will be examined with five main-effect predictors and two-way interactions.

## 5 The effectiveness of multinomial logistic regression modelling

Ordinarily, the Multinomial Logistic Regression model would be applied to investigate the effects of the explanatory variables to forecast a probability of the agreements since, originally, APRF has three values: 0 = 'No', 1 = 'Yes', 2 = 'Undecided'. One of the critically important elements of the modelling is to explore whether the model adequately fits the data of the survey. According to the literature, the Pearson Chi-square ( $\chi^2$ ) test is a reliable one to obtain judgement on the goodness-of-fit of a model (King et al., 2010). The null hypothesis ( $H_0$ ) of this test states that there is no significant difference between a theoretical frequency distribution and a frequency distribution of observed data for which each observation may fall into one of the several classes. Therefore, to infer that our model fits the data well, it is required to not reject  $H_0$ . This corresponds to the value of  $\chi^2$ , with an attached Degree of Freedom (df), falling (at least) below its value at the 95th percentile.

As an example, the Multinomial Logistic Regression with 5 IDVs for the above-mentioned model has a Pearson  $\chi^2$  distribution with 90 dfs, i.e.,  $\chi^2(90)$ , of 235.34. This is much larger than the value at the 95th percentile of the distribution, i.e.,  $[\chi^2(90)]_{95th}$  of 113.14. Thus, the  $H_0$  is rejected implying that this Multinomial Logistic Regression does not fit the data. The authors attempted to evaluate all other possible Multinomial Logistic Regression models, which combine the DV and all IDVs (main effects and two-way interactions) through SPSS®17 and the R Foundation for Statistical Computing software. However, the null hypothesis was always rejected owing to Pearson's  $\chi^2$  values being larger than those of the 95th percentile of the  $\chi^2$  distribution and its degrees of freedom. Therefore, this is not an effective means of modelling in terms of predicting the awareness of cardiologists about the potential link between AP and CVD. To attain a best-fitting model, merging the categories of the DV into APRFG, which is now a binary variable, is the only option available.

## 6 Multivariate logistic regression modelling

Statisticians are traditionally more in favour of using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in selecting the 'best' model. Both AIC and BIC provide a means for comparing between models and are not in themselves a test of the model in terms of goodness-of-fit hypothesis testing. The 'best' model would be obtained from the smallest values of AIC and BIC. The pseudo- $R^2$  measurement of fitness measures how effectively the model describes the response variable. For SPSS®17, in Logistic Regression, Nagelkerke's  $R^2$  is used to measure the goodness-of-fit for a given model. This measurement provides a simple and clear interpretation, takes values between 0 and 1, and becomes larger as the model 'fits better', in particular when adding more predictors. However, Nagelkerke's  $R^2$  does not bring to light the prediction errors associated with the model. Classification accuracy, which is also part of the Logistic Regression output, fulfils this requirement. The accuracy rate compares predicted group membership based on the logistic model with the actual known group membership, which is the value for the DV. Obviously, the higher the accuracy rate the better the Logistic Regression model is.

A total of 9 Logistic Regression models, including main effects and two-way interactions, were considered to predict the probability of the awareness of cardiologists about the potential association between AP and CVD with the DV; all of the models satisfy the null hypothesis of a good fit through the Pearson  $\chi^2$  test (Nguyen, 2011). Of these, two models,

**Table 2** Best-fitting models

<i>Model 1: Main Effects – 5 IDVs: EXPG, PM1, PM2, PM3 and Vehicle Emissions</i>	
Pearson distribution	$\chi^2 = 65.651$ : satisfied
( $\chi^2 = 96.22$ , df = 75)	
Akaike's Information (AIC)	160.278
Bayesian Information (BIC)	205.535
Nagelkerke $R^2$	0.517
Accuracy rate (%)	78.5%
<i>Model 2: Main Effects – 5 IDVs: EXPG, PM1, PM2, PM3 and Industrial Plant Emissions</i>	
Pearson distribution	$\chi^2 = 63.01$ : satisfied
( $\chi^2 = 96.22$ , df = 72)	
Akaike's Information (AIC)	154.2
Bayesian Information (BIC)	199.46
Nagelkerke $R^2$	0.492
Accuracy rate (%)	77.6%

which included only main effects, were selected owing to their lowest values of AIC and BIC and highest values of accuracy rate (%) and Nagelkerke  $R^2$  (Table 2).

Model 2 would be the best model owing to the lowest values of AIC and BIC. However, since vehicle emissions have long been established in the literature as a main source of AP – even more so than industrial plant emissions – then Model 1 will be the one to adopt.

In SPSS®17, Logistic Regression computes dummy coding predictors owing to the multinomial explanatory variables. The 0 value is chosen to be the reference group, i.e., for all of the mechanisms PM1–PM3, we use: 1 for Yes, 2 for Undecided and 0 as a reference group. In the case of experience, the EXPG groupings assume the following coding: 1 for 4–7 years, 2 for 8–11 years, 3 for 12 years and over and 0 for 1–3 years (as reference group). Figure 3 illustrates the outcomes of the Logistic Regression modelling for Model 1, together with regression coefficients (B) and their Standard Errors (SEs). Specifically, Logistic Regression provides the maximum likelihood estimates of the model parameters together with estimates of their SE. Then, a 95% confidence interval for the odds ratio associated with a unit rise/decline in a predictor is adjusted for the other covariates in the model (this is displayed in the last two columns of Figure 3 with the Wald upper and lower confidence intervals).

*Discussion:* Statistically, the odds (Exp (B)) of a physician who agrees with the harmful impact of emission from vehicles, compared with the odds of those who disagree, are such that the physician is approximately 10 times (95% CI: 2.01–54.18) more likely to agree positively that AP is a risk factor to CVD. Similarly, a physician who is in the 4–7 years experience group is three times more likely to agree that AP is an RF to CVD than those in the shortest experience group. For the cardiologists with the longest experience, the odds ratio shows that they are four times more likely to agree with this fact. With regard to the third mechanism, the odds of a cardiologist who does not express an opinion in this respect, compared with the odds of those who do not agree, show that he/she is 0.4 times (95% CI: 0.18–0.88) more likely to agree about AP as being an RF. In a similar way, odds of agreeing that inhaled particles accumulating in the lungs may cause systemic inflammation via oxidative stress, which might mediate endothelial dysfunction and atherosclerosis, show that the respective cardiologists are over seven times (95% CI: 1.1–47) more likely to agree that AP is an RF, compared with the odds of those who disagree; similarly, comparison

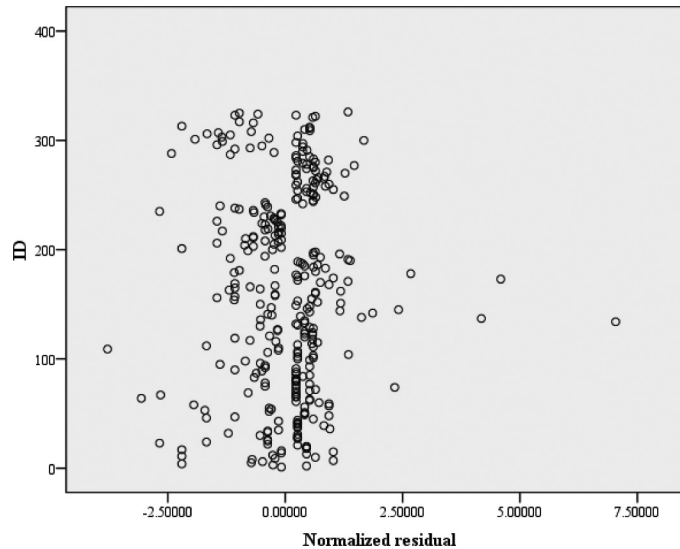
**Figure 3** Logistic regression of model 1, which characterises the positive agreement between AP as RF to CVD

	<i>B</i>	<i>S.E</i>	<i>Sig.</i>	<i>Exp(B)</i>	<i>95% CI for Exp(B)</i>	
					<i>Lower</i>	<i>Upper</i>
<i>Mechanism 1</i>			.078			
Mechanism 1 (Yes)	1.992	.947	.035	7.333	1.145	46.943
Mechanism 1 (Undecided)	1.557	.975	.110	4.747	.703	32.056
<i>Mechanism 2</i>			.090			
Mechanism 2 (Yes)	1.533	.708	.030	4.632	1.155	18.569
Mechanism 2 (Undecided)	1.245	.740	.093	3.474	.814	14.823
<i>Mechanism 3</i>			.000			
Mechanism 3 (Yes)	.694	.484	.152	2.001	.774	5.172
Mechanism 3 (Undecided)	-.934	.411	.023	.393	.176	.879
<i>Vehicles</i>			.000			
Vehicles(Yes)	2.346	.840	.005	10.440	2.012	54.178
Vehicles(Undecided)	.068	.892	.939	1.071	.186	6.153
<i>Experience group</i>			.004			
Experience group (4-7 years)	1.082	.433	.012	2.952	1.263	6.899
Experience group (8-11 years)	.189	.440	.668	1.208	.509	2.864
Experience group (from 12 years)	1.354	.432	.002	3.874	1.660	9.040
<b>Constant</b>	<b>-4.985</b>	<b>1.199</b>	<b>.000</b>	<b>.007</b>		

between the odds of agreeing that inhaled particles accumulating in the lungs may cause systemic inflammation, which may also increase blood coagulability and the odds of those who disagree show that the former group is over four times (95% CI: 1.2–18.6) more likely to consider AP as an RF.

*Multicollinearity:* In Logistic Regression, multicollinearity can be detected by the presence of very large SEs for the B coefficients (Sheskin, 2007). All the values of the SE from Figure 3 are close to or below 1.0. Cross-checking multicollinearity with SPSS®17, all Tolerance values are greater than 0.1 and the Variance Inflation Factor values are undoubtedly below 10. Therefore, there is no multicollinearity problem in Model 1.

*Outliers diagnosis:* Although the fit of Model 1 was completed in the previous section, the identification of outliers and their influence – necessary for checking the adequacy of the fitted model – were not addressed. Clearly, there are many methods to detect outliers that badly fit the Logistic Regression model. The simplest way to achieve this is through the output of Logistic Regression for SPSS®17, which provides the casewise diagnostics that enable to identify any cases of outliers outside two standard deviations (Weinberg and Abramowitz, 2002). On the basis of the scatter plot in Figure 4, all of those extreme values can be rechecked

**Figure 4** Outliers scatter plot for model 1

and cases that badly fit the model can be taken out. The Logistic Regression model can be tested once again after removing potential outliers to ascertain whether the model is a better fit. In our analysis, Model 1 detected 10 cases with standard deviations exceeding 2.0. However, the outliers should typically be in the range of  $[-3; 3]$ . Hence, five cases were considered as outliers and discarded.

After discarding those five outliers and testing Model 1 again, Logistic Regression gives similar results to the original Model 1; this is now a better model compared with the original owing to the smaller values of AIC and BIC values, which are 145.98 and 191.05, respectively.

## 7 Conclusions and limitations of the study

### 7.1 Conclusions

Logistic Regression modelling results suggest that a gap in cardiologists' knowledge exists when expressing the link between the various AP pathophysiological mechanisms and CVD. This gap was shown to exist in Model 1 where the physicians clearly exhibit more awareness of the harmful impacts from vehicular emissions. Moreover, the experience of the physicians also had a slight impact on their knowledge in relation to this topic. Similarly, cardiologists who were undecided that inhaled particles accumulating in the lungs may cause systematic inflammation via oxidative stress, which mediates endothelial dysfunction and atherosclerosis (M1) are approximately five times more likely to agree that AP is an RF to CVD (95% CI 0.7–32.06). As this confidence interval value includes 1.0, the difference between those who did not express their opinion and those who did not agree with this mechanism is not statistically significant at the 0.05 level (Lang and Secic, 2006). Nevertheless, the aforementioned

confidence levels in relation to that mechanism still provide some evidence for a degree of association with the physicians' decision that AP is an RF. In Model 1, applied separately to the two countries, mechanism 2 is not statistically associated with cardiologist's agreements that AP is an RF to CVD for the Philippine cardiologists only.

In contrast to prior expectations, the working environments and the sub-specialisms are not statistically significant in relation to the cardiologists knowing the potential link between AP and CVD. The outliers of Model 1 point to more cardiologists who are practising at secondary or rural organisations than in tertiary or provincial hospitals. This indicates that physicians who practise at secondary or rural hospitals may be confused about the AP impact on CVD more than physicians who work at tertiary or provincial hospitals.

In conclusion, it appears that the awareness of the impact of AP on CVD, in general, by Vietnamese and Philippine cardiologists is rather superficial and not commensurate with the depth of the subject, especially from the pathophysiological pathway understanding perspective. Besides, the obviously worsening mortality patterns of CVD in both countries and their associated RFs are not given enough attention from the health research perspective and from the level of disease management.

According to the WHO (2002), '50% death and disability from CVD can be reduced by a combination of simple effective national efforts and individual actions to reduce major cardiovascular risk'. On the basis of accumulated scientific evidence from epidemiological studies and in vivo research from developed countries on the potential association between AP and CVD, the South East Asian region cannot deny this existing association. Besides, the composition and relative contribution of ambient AP in the region are characteristically different from those in developed nations; hence, there is a pressing need to establish studies on the link in greater detail. Obviously, a spectrum and variety of cardiovascular risk prevention tools, strategies, programmes and frameworks have been formulated and put into action in different nations, especially in developed countries. Yet, guidelines or regulatory frameworks that include AP risk have not been established or embedded in any of the ones currently adopted by the respective governments or local environmental authorities (Nguyen et al., 2011b). Moreover, not only in Vietnam and the Philippines, but also in the whole region, the current environmental health legislations and regulations, as well as the preventive interventions for the disease, were exhaustively evaluated in terms of their effectiveness and efficiency. Given the critical emerging issues about the AP situation and the rapid escalation in morbidity and mortality rates from CVD in South East Asia, especially in Vietnam and the Philippines, it is vitally necessary to raise the awareness of all stakeholders about the link of the disease to AP.

## 7.2 Limitations

In terms of statistical analysis, one limitation of this study as indicated by Logistic Regression modelling is that the size of the sample should have been increased, with a greater number of responses, to support more comprehensive analyses. Second, the types of organisations/hospitals should have been more proportionate (e.g., through the inclusion of secondary hospitals) and types of sub-specialisms (such as paediatric or academic cardiology) also should have been included in the study to affirm whether the type of organisation or area of cardiology practised impinged on the cardiologists' knowledge in relation to AP impacts. Those can be indeed construed as (minor) weaknesses of the current analysis and best taken into consideration in any future such investigation.



The area of environmental health research is very wide. Even when narrowed down to consider the effects of AP alone on human health, it still remains a very broad area. It was thus necessary to limit the scope of the study to cardiologists to evaluate their appreciation of the effects of AP when diagnosing or treating CVD. By the same token, a study designed to cover an entire geographical region, such as South East Asia, is extremely intricate to handle. Not only are logistic or financial limitations a problem, but also the characteristics of many countries in the region are different, as highlighted through the two case countries presented in this study. This imposes an inevitable time limitation on the scope of the activities carried out for the duration of the study.

In addition, a further limitation of implementing a potentially regional framework is the possible inability to delineate gaps in environmental cardiovascular health knowledge applicable to each country in the region. This may be due to the disparity of the AP sources and their characteristics; healthcare cultures, intervention programmes and policies; or the involvement of a multitude of different stakeholders for each area/nation.

### 7.3 *Future work*

One direction for future work in this area of research might be the inclusion of other developing countries in the region, such as Cambodia, Laos and Myanmar within a similar study to provide a much wider scope of results, which, in theory, ought to lead to a more accurate assessment of clinical cardiology diagnostic and curative habits in relation to AP impacts. Such an exercise will also inevitably lead to a refinement in any regulatory framework suggested to respective governments.

The population afflicted by CVD was not categorised when the survey was conducted. Thus, no information about gender, age, educational level, employment, income, etc., was sought by the authors in the course of the survey. Even though some parts of the study touched upon a few perceptible (by the cardiologists) changes in CVD patterns and, in other parts, on congenital heart disease, it would be of benefit to consider the various different cardiovascular events that may incur as a result of AP exposure, based on age (especially for the newborn and the elderly), gender (especially pregnant women) or socio-economic predisposition.

Finally, this research considered the effects of AP on CVD. Many similar studies could be initiated that focus on other chronic diseases such as cancer (with its raft of different types) or diabetes. It would be of interest to uncover the findings for such other diseases and to develop corresponding national regulatory frameworks that cater for a multitude of ailments caused by AP.

## References

- Ayres, J.G. (2006) *Cardiovascular Disease and Air Pollution: A Report by the Committee on the Medical Effects of Air Pollutants*, [online] Available from: <<http://www.advisorybodies.doh.gov.uk/comeap/expertreview.htm>> [Accessed 14th September 2011].
- Brook, R.D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., Luepker, R., Mittleman, M., Samet, J., Smit, Jr, S.C. and Tager, I. (2004) 'Air pollution and cardiovascular disease: a statement for healthcare professionals from expert panel on population and prevention science of the American Heart Association', *Circulation*, Vol. 109, No. 21, pp.2655–2671.
- Bhatnagar, A. (2006) 'Environmental cardiology: studying mechanistic links between pollution and heart disease', *Circulation Research*, Vol. 99, No. 7, pp.692–705.

- Bhatnagar, A. (Ed.) (2011) *Environmental Cardiology: Pollution and Heart Disease*, Royal Society of Chemistry, Cambridge, UK.
- CAI-ASIA (2010) *Knowledge Management on Air Quality: Case Studies Report*, [online] Available from <<http://cleanairinitiative.org/portal/KMonCaseStudies>> [Accessed 15th September 2011].  
AUTHOR PLEASE CHECK IF THE HIGHLIGHTED YEAR OF PUBLICATION IS '2009' OR '2010'.
- Department of Health (2010) *Health Indicator Mortality and Morbidity*, [online] Available from <[http://www.doh.gov.ph/kp/statistics/leading\\_mortality.html](http://www.doh.gov.ph/kp/statistics/leading_mortality.html)> [Accessed 15th September 2011].
- King, B.M., Rosopa, P.J. and Minium, E.W. (2010) *Statistical Reasoning in the Behavioural Sciences*, 6th Ed., John Wiley & Sons, US.
- Lang, T.A. and Secic, M. (2006) *How to Report Statistics in Medicine: Annotated Guidelines for Authors*, 2nd Ed., American College of Physicians, US.
- MOH (Ministry of Health Vietnam) (2009) *Vietnam Health Statistics Yearbook 2008*, Hanoi.
- Newman, I. and Benz, C.R. (1998) *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum*, Southern Illinois University, USA.
- Nguyen, Q.T. (2011) *Assessment of Cardiologists' Awareness of the Impact of Air Pollution on Cardiovascular Disease*, PhD Thesis, Coventry University, UK.
- Nguyen, Q.T., Naguib, R.N.G., Wickramasingle, N.S. and Shaker, M.H. (2011a) 'Does ambient air particulate matter have a direct impact on cardiovascular disease incidence in the Philippines?', *International Journal of Biomedical Engineering and Technology*, Vol. 5, No. 1, pp.44–60.
- Nguyen, Q.T., Naguib, R.N.G., Papatomas, M., Shaker, M.H. and Culaba A.B. (2011b) 'Effects of air pollution on cardiovascular disease in the Philippines: results of a survey assessing cardiologists' awareness' in *Proc. Int. Conf. on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management*, Manila, Philippines, pp.70–78.
- Pope, C.A. III., Muhlestein, J.B., May, H.T., Renlund, D.G., Anderson, J.L. and Horne, B.D. (2006) 'Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution', *Circulation*, Vol. 114, No. 23, pp.2443–2448.
- Samet, J.M. and Spengler, J.D. (1991) *Indoor Air Pollution: A Health Perspective*, Johns Hopkins University Press, USA.
- Sheskin, D. (2007) *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, USA.
- Su, T-C., Chen, S-Y. and Chan, C-C. (2011) 'Progress of ambient air pollution and cardiovascular disease research in Asia', *Progress in Cardiovascular Diseases*, Vol. 53, No. 5, pp.369–378.
- Tofler, G.H. and Muller, J.E. (2006) 'Triggering of acute cardiovascular disease and potential preventive strategies', *Circulation*, Vol. 114, No. 17, pp.1863–1872.
- Weinberg, S. and Abramowitz, S.K. (2002) *Data Analysis for the Behavioural Sciences Using SPSS*, Cambridge Press, UK.
- WHO (2002) *Integrated Management of Cardiovascular Risk: Report of a WHO Meeting, Geneva*, [online] Available from <<http://site.ebrary.com/lib/coventry/docDetail.action?docID=10040290&p00=regulatory%20framework%20preventive%20cardiovascular%20disease>> [Accessed 13th August 2011].
- WHO (2011) *Cardiovascular Diseases (CVDs)*, [online] Available from <http://www.who.int/mediacentre/factsheets/fs317/en/index.html> [Accessed 23rd November 2011].