Conditional Multiple Logistic Regression Can Confirm the Relationship Between Hypertension and Income Groups in a Matched Case-Control Study

Sir,

Regarding the original article: Sabri S, Bener A, Eapen V, Azhar AA, Abdishakure A, Singh J. Correlation between hypertension and income distribution among United Arab Emirates population. Med J Malaysia. 2005; 60(4): 416-25, we would like to comment on the following points.

1. Matching

It was clearly stated in the materials and methods that this is a matched case-control study. Matching variables such as age, sex, nationality and educational level were mentioned only in abstract. How the matching was done is not mentioned anywhere. We would like to know whether individual or group matching was used. More confusingly, under the heading of "Selection of Community Control", it was stated that "... involved a random sample of 500 subjects in the community". Although demographic information of two study groups was not described, according to Table I, there were 218 and 204 males in cases and controls respectively, and 191 and 205 females in cases and controls respectively. In addition, there were 187 and 193 UAE nationals in cases and controls respectively, and 222 and 216 expatriates in cases and controls respectively. These are also controversial if the study is really done matching for sex and nationality as stated because we are expecting an equal number of cases and controls for each matching variable[1].

Matching controls to cases assures that both groups are comparable with respect to each of the matching variable[1,2]. In our opinion, it is very important to describe the detail procedure of matching, and background comparison of study groups should be the first table in the result section in this kind of study. The authors should present the comparisons of socio-demographic characteristics between cases and controls such as mean age, income, occupation, nationality etc.

It may be possible that the authors applied age 20-65 years as an inclusion criterion for both cases and controls and referred it as "matching". In this case, it should be called "age restriction" and cannot be called "matching"[2].

Furthermore, the aim of matching is to control for confounders and once we have matched controls to cases according to certain variable/s, we cannot study that characteristic anymore[1].

2. Sampling procedure

According to the section of "Selection of Hypertensive Subjects" (p 417), the authors mentioned that they selected hypertensive subjects as consecutive cases. On the other hand, it stated that the selection was by a simple random process in the same sentence. Why not include the name of probability sampling method either "simple random sampling or systematic sampling"? This applies to the selection of community controls as well.

The "multi-stage stratified cluster sampling" method was used to select the subjects. We presumed that the sampling involved two-stages of administration divisions and clinics; and then stratified according to urban and semi-urban areas. Cluster approaches are used when it is impractical to compile a complete list of study population. In a city, a set of city blocks could be selected for sampling and all those living in the city blocks should be included in the study[3]. In this study the authors should include all hypertensive patients in those selected clinics if they are really using the cluster method.

3. Misclassification of cases and controls

The selection of cases and controls is questionable since it was based on self-reported results. It was shown that more than half of hypertensive patients were undiagnosed in the community in Malaysia[4]. How is it in UAE? The validity of definitions of cases and controls

should be mentioned and discussed as it is the major variable of this study. Serious misclassification could have occurred in this study if there are many undiagnosed cases in the community.

4. Sample size calculation

The authors used the prevalence of hypertension in the sample size calculation. In our opinion, this sample size calculation was not appropriate as the study was not a prevalence study. The appropriate one is based on comparing two proportions[5] since this is a case-control study. In addition, the use of 1% level of significance was not appropriate as the authors presented the result with 95% confidence intervals, and set the p value cut off point as <0.05.

5. Questionnaire

In page 417, the procedure of translating questionnaire was described. However, it was not mentioned that from which standard or original questionnaire the researchers translated. If the questionnaire was developed by the authors why was it necessary to translate? Should it be developed in Arabic directly?

It was also mentioned that a pilot study was conducted for validation of the questionnaire. However, the authors did not state the result of the validation, whether it is satisfactory or not.

6. Appropriate statistical analysis

It is not clear why Mann-Whitney test was used for non-parametric distribution. What "non-parametric distribution" is referred to? Do the authors mean non-normal distribution? In our opinion, the authors should use the specific name of distribution[6].

If the study is truly a matched case-control design, it should be analyzed by using conditional logistic regression[7]. All analyses used in this study such as Student's t test, Mann-Whitney test, and chi-square test need the statistical assumption that the samples are two independent random samples[8]. Conditional logistic regression is a special type of analysis to apply in a matched case-control study.

Other variables such as occupation, family history of hypertension, obesity, physical activity, rural-urban, and comorbidity should be considered to include in the analysis as they could be potential confounders. The application of multiple conditional logistic regression will give opportunities to study these confounding variables while confirming the relationship between income and hypertension[2].

In the footnote of Table I, it was stated that OR and 95% CI were calculated using Mantel-Haenszel (MH) test. Were all odds ratios in the Table I, calculated using MH test? Only the totals should be applied by this MH test.

6. Others

The authors mentioned that the participants were weighed in light clothing without shoes. However, they did not mention whether the participants should remove all the pockets' content, watches, belts and jewelries. Is it the reason why 1kg of weight is subtracted for clothing?

There was an inconsistency in the statement, in page 418, in the last sentence of first paragraph that no data were obtained on alcohol consumption since this is a conservative Muslim country. However, there were results on the alcohol in the Tables II, III and IV.

The classification of the smoking behaviour is not clear. If someone has stopped smoking for three months, which category will he or she be included in? It fits neither ex-smoker nor current smoker according to the definitions given on page 418. Furthermore, the authors did not define other important variables such as expatriate, fat and salt consumptions. We are sure that almost all of us consume salt and fat to some extent in our daily diet. Quantity and frequency of the consumption should be taken into account as well.

The following statements mentioned in the discussion were not stated in the result section:
Paragraph 1: "Hypertension was more prevalent among females above 40 years old";
Paragraph 2: "Hypertension was more prevalent among expatriates than UAE nationals"; Paragraph 3: "Hypertension was found to be associated with poor health status ...", Paragraph 4: "...this significant increase was not observed for diastolic blood pressure".

Finally, we noticed that there was a mixture type of referencing style. Throughout the article, Vancouver style was used, however Harvard style was noted in page 417, second paragraph.

B Norsa'adah*, MComMed, L Naing**, MPH, MMedStats
*Unit of Biostatistics and Research Methodology, School of Medical Sciences, **School of Dental Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan

References

1. Gordis L. Epidemiology. WB Saunders Co: Philadelphia, 1996.
2. Schlesselman JJ. Case-Control Studies. Design, Conduct, Analysis. Oxford University Press: New York, 1982.
3. Timmreck TC. An Introduction to Epidemiology. Jones & Bartlett Publishers: Boston, 1994.
4. Ministry of Health Malaysia. Report of the Second National Health and Morbidity Survey. Public Health Institute: Kuala Lumpur, 1997.
5. Lemeshow S, Hosmer Jr DW, Klar J, Lwanga SK. Adequacy of Sample Size in Health Studies. John Wiley & Sons: Singapore, 1990.
6. Hastings NAJ, Peacock JB. Statistical Distributions: A Handbook for Students and Practitioners. John Wiley & Sons, Inc: New York, 1974.
7. Hosmer DW, Lemeshow S. Applied Logistic Regression. John Wiley & Sons, Inc: New York, 2000.
8. Daniel WW. Biostatistics: A Foundation for Analysis in the Health Sciences (7th Edition). John Wiley & Sons, Inc: New York, 1999.

(Note: The letter has been sent to the authors of the article being referred to for their response)