Statistical Corner

Chung-Hong CHAN 陳仲康 and **Daniel Kwok-Keung NG** 吳國強
Department of Paediatrics, Kwong Wah Hospital, Hong Kong

Question 1:
What are the scales of measurement?

Answer :
Measurement in clinical research can be classified into 4 scales
a. Nominal/Categorical (e.g. Male, Female)
b. Ordinal (Mild, moderate, severe)
c. Interval (Temperature in C)
d. Ratio (Temperature in K)

Question 2:
Why dichotomising a continuous variables into a categorical variables?

Answer:
Most of the measurements in clinical research are continuous variables, i.e. intervals and ratio. Example included blood pressure in mmHg, BMI in Kg/M^2 and apnoea hypopnoea index. Clinically, it is very common to dichotomise continous variables into categorical variables based on a cut-off point. e.g. AHI larger than 1.5 as having obstructive sleep apnoea. In clinical research, dichotomising continous variables into categorical variables is also helpful for presentation of results. For example, we can present the proportion of children having obstructive sleep apnoea, i.e. prevalence based on the cut-off value of apnoea hypopnoea index.

Question 3:
What is the advantage of dichotomising a continuous variable into a categorical variables in statistical analysis?

Answer:
It is easier to analyse categorical variables rather than continuous variables. Categorical variables can be analysed by simple technique such as logistic regression, chi-square test and t-test. However, continuous variables are not always normally distributed. Analysis of continuous variables may require transformation or some relatively uncommon regression technique. (e.g. negative binomial regression for count/rate data)

Question 4:
What is the problem of dichotomising a continuous variables into a categorical variables in statistical analysis?

Answer:
Dichotomising a continous variable into categorical variable for the ease of analysis may lead to problems as listed below.
a. Massive waste of information and thus reducing statistical power.
b. Positive results from logistic regression of a dichotomised continous variables as the dependent variables are very likely to be false positive.
c. When regression is used to control the effect of confounding factors on a dichotomised continous variable, a substantial amount of confounding is likely to remain in effect.
d. The cut-off point of choice for dichotomisation is usually controversial. e.g. AHI >1.5 is not an universally accepted cut-off point for obstructive sleep apnoea in children.

In conclusion, one should always analyse continuous variables as continuous variables in statistical analysis. (Common sense) Transformation or alternative regression model may be used to cater for non-normal distribution of the continuous variable.

References
1. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.
2. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127-41.