JOURNAL OF CLINICAL RESEARCH BEST PRACTICES

Vol. 6, No. 3, March 2010 "Can You Handle the Truth?"

Making Sense of Biostatistics: Multivariate Regression Analysis

By Ronald E. Dechert

In the last two columns, we discussed two analytical methods — correlation and simple linear regression analysis — commonly employed to evaluate the strength of association between a dependent (Y) and independent variable (X). Although both techniques are useful when there is only a single independent variable, they are less useful in explaining complex associations involving multiple independent variables. For example, they can measure the association between the height of a parent and his or her child, but not between both parents and their child, especially in the presence of other variables, such as diet, age and weight. In this article, we will expand the conversation to multivariate regression analysis.

Multivariate regression analysis is a logical extension of simple linear regression. In multivariate regression analysis, we add additional independent variables, also known as covariates, to the right side of the formula, as shown in the following equation:

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3.....B_nX_n$$

Although the concept of multivariate regression analysis is fairly intuitive, you have to select the right covariates. One obvious approach is to first use univariate analysis to identify covariates that, in isolation, are statistically associated with the dependent variable. The role of multivariate analysis is then to measure the strength of each association in the presence of all the covariates.

Suppose Dr. Arlene Saturday is interested in a new drug (Phoebus) that, in preliminary studies, demonstrated the potential for decreasing the duration of mechanical ventilation in adult intensive care patients. Dr. Saturday tests Phoebus versus placebo in a randomized, double-blind, single-center study of adult ICU patients who were consented and randomized (1:1) in her clinical study. A total of 800 subjects were enrolled and randomized over the course of the clinical trial. At the conclusion of the trial, Dr. Saturday performed a univariate analysis to examine differences between the treatment (Phoebus) and control (placebo) groups. The results of her analysis are presented in Table 1:

Table 1. Univariate Analyses

Variable	Phoebus	Placebo	Significance*
Age	35 $\pm$ 10	34 $\pm$ 11	0.382
Inspired Oxygen (FiO2)	0.59 $\pm$ 0.21	0.54 $\pm$ 0.18	0.001
Mean airway pressure (MAP)	17 $\pm$ 7	14 $\pm$ 5	0.001
PEEP	10 $\pm$ 4	9 $\pm$ 4	0.031
Peak Inspiratory Pressure (PIP)	32 $\pm$ 9	29 $\pm$ 10	0.001
Ventilation Days (LOS)	12 $\pm$ 5	11 $\pm$ 4	0.001

* Independent Student's t-test, p=

In reviewing the univariate analyses, all of the covariates demonstrated statistical significance ($p <= 0.05$), with the exception of age. The drug vs. placebo association seems significant, but it may be an artifact of the other covariates. Dr. Saturday therefore wants to examine the effect that study group assignment (Phoebus versus Placebo) has on the primary outcome of interest (ventilation days) if the other covariates are controlled for in the model. In other words, if all the other covariates had a single value, what would be the effect of drug vs. placebo? To do this, Dr. Saturday performs a multivariate regression analysis using the following model:

$$\text{Ventilation Days (Y)} = \text{constant } (B0) + \text{FiO2 } (B1) + \text{MAP } (B2) + \text{PEEP } (B3) + \text{PIP } (B4) + \text{Group } (B5)$$

This model allows Dr. Saturday to determine if the difference in ventilation days is associated with the group assignment when controlling for the effects that the other covariates have on ventilation days. The results of her analysis are presented in Table 2.

The analysis demonstrates that there is still a statistically significant ($p<0.05$) association between ventilation days and group assignment, even when controlling for the other covariates identified in the univariate analysis. In addition, examining the coefficient for group assignment (Placebo=0, Phoebus=1), Dr. Saturday can report that, on average, use of Phoebus decreased ventilation days by 0.143 days when the other covariates were accounted for.

Table 2. Multivariate Analysis

Covariate	**Coefficient**	**Significance**
Group (0,1)	-0.143	0.001
PEEP	0.003	0.968
MAP	0.031	0.520
FiO2	0.071	0.228
PIP	-0.051	0.217

As a result, Dr. Saturday concludes that clinical use of Phoebus in adult ICU patients who require ventilation would be beneficial because Phoebus, on average, decreases the average duration of ventilation in that population.

Given the many ways in which clinical trial populations can vary, multivariate regression analysis is a very useful tool for determining the actual effect of a study treatment.

Author

Ronald E. Dechert, DPH, is Associate Director of the Mott Respiratory Care at the University of Michigan Medical Center. Contact him at 734.936.5237 or rdechert@umich.edu.