CIR ESP. 2022;100(11):731–733

CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia

Methodological letter

Multivariate analysis in surgical studies

Análisis multivariante en investigación quirúrgica

Pere Rebasa Cladera

Unitat Cirurgia Esòfago-Gàstrica, Servei de Cirurgia General i de l'Aparell Digestiu, Hospital Universitari Parc Taulí, Sabadell, Professor Associat Medicina Universitat Autònoma de Barcelona, Barcelona, Spain

Why should we use regression models in surgical research? Isn't the Student's *t*-test, X^2 or an analysis of variance test sufficient to provide sufficient evidence of a relationship between an event and a consequence?

Of course, to try to obtain scientific truth, a univariate test such as those cited above may be useful, but it is unlikely to be sufficient, or to lead to robust cause-effect conclusions, which is what we as surgeons tend to look for in the vast majority of our research (do sutures fail more if the patient is hypoproteinaemic, is tumour recurrence more likely if I have transfused the patient?).

In general, biological variables are strongly related to each other. It is difficult to attribute to a single variable X the effect on Y. Most commonly, there are a myriad of other variables that can alter that relationship: can hypoproteinaemia in suture failure be somehow conditioned by the patient having a known liver disease? In biology, therefore, pure univariate analysis (relationship between X and Y without regard to any other considerations) can hardly be used except in the case of highly controlled experimental studies, which are usually only possible in a laboratory setting and with experimental animals.

Biological systems often follow chaotic patterns, where small changes in initial conditions lead to huge changes in outcomes. This is what we as surgeons face every day when performing an anastomosis: the uncertainty that by doing the same thing every time, our results may differ greatly from patient to patient. How, therefore, can I succeed in establishing powerful cause-effect relationships in a chaotic biological system? With multivariate analysis.

Types of multivariate analysis

Are all multivariate analyses the same, and is it as easy as having a grid of data in SPSS and hitting enter on the "multivariate analysis" tab? No and no. When are we going to use a simple regression model? Whenever we seek to test and quantify the relationship between variable X and variable Y: for example, what is the relationship between preoperative haemoglobin level and postoperative infectious complications. But in most clinical situations, other variables must be taken into account. This is where multiple regression[1] comes in. Here are some examples:

- Epidemiological studies of risk factors. Imagine that we have collected data on a potential risk factor for gastric cancer (e.g., HER2 gene mutation) in a cohort of patients at time t, and we follow these patients for five years to see how many of them have developed gastric cancer. Now, some of these patients were smokers, and others were *Helicobacter pylori* positive. The right question here is not *what is the relationship between the* HER2 *mutation and gastric cancer, but how much of a*

E-mail address: prebasa@tauli.cat

relationship between the HER2 mutation and gastric cancer is there that cannot be explained by the relationship between smoking, Helicobacter pylori and gastric cancer?

- Studies of prognostic factors. In these studies, we want to identify patient factors that help us to make a prediction regarding a particular disease or outcome. This could be a new prognostic factor, trying to answer the question: *does this new predictive factor improve the prognosis compared to the ones we use today?* Or with the idea of creating a new prognostic index, *how can we combine the values of several prognostic factors to create a predictive score?*
- Diagnostic studies. Very similar to the prognostic factor studies mentioned in the previous paragraph, in this case the aim is to diagnose a disease. Do the Alvarado criteria help us to diagnose acute appendicitis? These studies have a differential characteristic with respect to prognostic factor studies, and that is that in the end, what we are looking for is a yes/no answer in order to affirm or deny that the patient has the disease, and not only to estimate the probability that the patient has the disease.
- Multifactorial studies. Less common in our field, these studies seek to investigate several factors simultaneously by trying to take advantage of the use of available material (usually experimental animals). For example (and here there are countless possible examples), we can use three different doses of growth hormone and three different doses of *insulin-like growth* factor in nine animals to test bacterial translocation following trauma
- Effect modification. Finally, we may be interested in knowing the effect that one variable may have on another, or in a clear example, if the new chemotherapy is superior to the one we have for a population with oesophageal neoplasia, and if this effect is maintained taking into account multiple known covariates that may modify it (tobacco, alcohol).

In short, multiple regression models (linear, logistic, Cox) not only provide us with a framework to describe the effect of one variable on another (which univariate models can already do), but also allow us to describe how these effects depend on other variables, and how much they modify it.

P and multivariate studies. How to draw conclusions (inferences) in a multivariate study

For many years now, surgeons (and not only us, in fact) have relied too blindly on p-values to reinforce our conclusions. There are good articles in this journal[2] explaining how p works and how it is correctly interpreted, and even the journal Nature has found it necessary to go into the subject[3]. How can we draw conclusions from multivariate analysis? We must remember that the p, standard error and confidence intervals, perfectly calculated by our statistical confidence programme, depend entirely on the technique we use to calculate them, and not on the intrinsic value of the relationship between our data.

Without going into a deep analysis, there are two different techniques to obtain inferences in multiple regression[4], a "classical" one that any statistical programme has implemented, and a "modern" one available as an option in the best programmes. But both depend absolutely on two assumptions regarding the Y variable (in this analysis, the Y variable is the errors, in statistical terminology): on the one hand, the assumption of linearity, on the other hand, the assumption of independence (which includes homoscedasticity and normality of the Y variables). It is far beyond the scope of this paper to explain these concepts, but it should be known that they cannot be ignored, and that someone who knows perfectly well how to investigate them and how to control them should be responsible for the analysis of our data. The "modern" way of drawing inferences will eliminate the need to explore homoscedasticity and normality in the distribution of errors.

How do we know how valid our regression results are?[5,6] With the principle of least squares, with the goodness of fit. How? Weren't they those nice p's that appear in SPSS after you hit enter? Not exactly. With a $p < .05$, for every 20 comparisons to see if nail length is related to suture failure, one of them is going to say yes, regardless of whether or not they have a real cause and effect relationship. In a regression analysis to see if in addition to nail length, albumin, age and 10 other variables influence suture failure, hundreds of comparisons are going to be made and therefore the probability of one of them falling below the magic .05 is very high. We will have to be extremely careful when we publish our result saying that nail length has a $p = .028$ and is related to suture failure. That is not evidence of association. It is a misinterpretation of a result provided by a statistical programme that knows nothing about medicine.

What should we do when considering a multivariate study?7 Clinical significance test with confidence intervals, explain the methods used to select the independent variables, explain the specific methods used to generate the models, explain how, if at all, the interaction between variables has been calculated, describe whether we have actually obtained 10 events per independent variable, check that we have described the linearity of the residuals, describe whether we have checked for collinearity, describe how we have validated the model, describe the goodness-of-fit, the discrimination statistics, and provide full information on how we have coded the variables[7].

To summarise: in medicine, a multivariate study should always be chosen to look for cause-effect relationships between the variables we are interested in. Univariate can give us the first clues as to which variables may be relevant to take into account in the multivariate. We need an expert in multivariate analysis with us at our side for the analysis, and many checks will have to be made before publishing any results so that we are not publishing an anomalous result and not a cause-effect relationship.

REFERENCES

1. Boateng EY, Abaye DA. A review of the logistic regression model with emphasis on medical research. J Data Anal Inf Process. 2019;7:190–207.
2. Rebasa P. Entendiendo la ‘‘p<0.001’’. Cir Esp. 2003;73:361–5.
3. Amrhein V, Greenland S. McShane Blake. Retire statistical significance. Nature. 2019;567:306–7.

4. Anderson RP, Jin R, Grunkemeier GL. Understanding logistic regression analysis in clinical reports: An introduction. Ann Thorac Surg. 2003;75:753–7.
5. Hosmer DW, Lemeshow S. Applied logistic regression, 2nd edition. New York: John Wiley & sons; 2000.
6. Vach W. Regression models as a tool in medical research. Boca Ratón: CRC Press; 2013.
7. Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standarts for use and reporting with particular attention to one medical domain. J Clin Epidemiol. 2001;54:979–85.