High quality randomised clinical trials (RCTs) and meta-analysis represent the highest levels of evidence, but in everyday clinical practice, observational studies are often exploited as a quick and easy way to understand the performance of clinical and interventional strategies. In this setting, multivariate analyses are exploited to drive useful and independent information, but due to potentially confounding messages, should be critically appraised and used in everyday clinical practice.

**Keywords**: Multivariate analysis, PCI, stent.

In recent years patients, physicians, and governments have been looking for the most accurate and economically sustainable combination of new drugs, diagnostics, and interventional technologies in a rapidly changing economic scenario, pursuing several options in this quest, including comparative effectiveness research.¹

Actually, from a scientific point of view, a growing bulk of new pharmacological and technological choices have been offered, especially in the cardiovascular field.²⁻⁹ According to widespread opinion, well-conducted randomised controlled trials provide the most valid estimates of the relative efficacy of competing healthcare interventions.¹⁰ A meta-analysis of randomized clinical trials (RCTs) that directly (head-to-head) compares two different interventions or drugs is thus considered the highest quality evidence. However, many interventions and drugs have not been directly compared in RCTs. This may relate to a large number of factors, ranging from the need for important resources,¹ fear of negative results for direct comparisons, and the underreporting of non-significant or negative data.¹¹ For example, placebo-controlled trials are often enough to obtain the regulatory approval of a new drug, once again limiting any ensuing direct comparisons.

On the other hand, especially in interventional cardiology, a high number of non-randomised studies are still performed in order to save...
Economical resources, to create hypotheses, especially for non-randomisable patients, or to shed light on the generalisability of results from existing randomised experiments.

In an attempt to exploit the broad potential of observational databases, various statistical models are currently employed. Several different multivariable approaches are available to control for systematic baseline differences naturally occurring between groups in the non-randomised setting. Even more, their striking importance lies on defining the impact of several independent variables on a single dependent variable, thus avoiding confounding effects coming from observed variables in non-randomised studies.

Nevertheless, multivariable analysis should be performed according to precise statistical issues, in order to offer understandable results and to offer a more prevalent impact on everyday practice.

**TIPS AND TRICKS TO PERFORMING AND INTERPRETING MULTIVARIATE ANALYSIS**

**In Theory**

The first important step for any researcher performing multivariate analysis, and for those reading articles, is to choose the most accurate model.

This choice should be performed according to a simple selection of parameters, that firstly, have differences or similarities in follow-up and secondly, a number of events for covariates.

One of the most historically exploited models is represented by binary logistic regression, which evaluates the independent predictive role of one or more independent variables of interest. Actually, to appraise the logit of the probability of an event (dependent variable) given one or more dependent variables, event probabilities are appraised as a function in order to appraise. This model performs accurately, especially for studies with a similar follow-up, not adjusting for time-variation, and independently from number of events for covariate.

On the contrary, Cox proportional hazard analysis also adjusts for differences in follow-up duration and censored data, by assessing the relationship of explanatory variables to survival time controlling for covariates and known confounders.

Last but not least, propensity score, which is defined as the conditional probability of receiving an exposure or treatment given a vector of measured covariates. Propensity could be exploited to perform a matching analysis (by obtaining two sample sizes of patients with a similar risk baseline profile) or may be incorporated into Cox multivariate models, and should be exploited for studies with a low ratio of events per covariate. For both of these models, some similar points should be accurately assessed.

The first choice of variables should be based on prior epidemiological evidence (i.e. an established association from prior well-conducted experimental or clinical studies) and strong associations (e.g. $p<0.10$ or $p<0.05$ at bivariate analysis) stemming from the specific dataset of interest.

Specifically, for propensity scores both the calibration and possible discrimination of the model should be evaluated. With calibration, the distance between the observed (treatment, yes or no) and the predicted outcome from the model (propensity score) are assessed through the Hosmer–Lemeshow goodness of fit test. On the contrary, with discrimination (through area under-the-curve), authors understand how the predicted probabilities, derived from the model, classify patients into their actual treatment group.

**In Practice**

In a recent clinical review of our group, we analysed all observational studies comparing bare metal and drug-eluting stents (DES), which demonstrated that independently from any impact factor, a better exploitation and methodological appraisal of multivariable analysis is needed in order to improve the clinical and research impact and reliability of non-randomised studies.

In all studies, a low number of events per variable was a common feature, potentially suggesting overfitted data and misleading associations. Another difficult finding was the lack of reporting and perhaps conducting of internal control, as it was frequently not possible to assess calibration or censoring appraisal. Moreover, any omission of the methodological assessment was not related to the quality rating of the journal in which the paper was published: we found no substantial differences among studies stratified according to the journal of publication’s impact factor, thus stressing the need for more careful attention from peer reviewers concerning studies reporting multivariable adjustments.
REFERENCES

15. Weitzen S, Lapane KL, Agostoni P, et al. Are propensity scores beneficial for percutaneous coronary intervention of bifurcation lesions? The authors demonstrated the increased risk of ST for DES, data that have never been confirmed in randomised evidence. As a result of the potentially dramatic clinical impact, the work caused a reduction of more than one-third of DES implantation, particularly in North America. This example stresses the crucial point of the limitations of multivariate analysis, even when accurately performed, because they could not account for non-recorded or evaluated features, thus leaving potentially fundamental clinical or interventional properties unanalysed.