

# How to interpret and report the results from multivariable analyses

Neus Valveny and Stephen Gilliver

TFS Develop, Medical Writing Unit, Barcelona, Spain / Lund, Sweden

## Correspondence to:

Neus Valveny  
TFS Develop  
Consell de Cent, 334-336, 4th floor  
ES-08009 Barcelona, Spain  
+ 34 617 414 124  
neus.valveny@tfsacro.com

## Abstract

Multivariable analyses are some of the central statistical methods of clinical trials, and yet some medical writers may be unsure as to what they are and how best to interpret and report the results. In this article we provide an overview of multivariable analyses, introducing some of the core models biostatisticians use to analyse trial data. We focus on odds ratios, hazard ratios, and  $\beta$  coefficients as key parameters and provide guidance on important considerations when reporting them.

## What is a multivariable analysis?

*Univariate analyses* – analyses involving only a single variable – are descriptive by nature. They allow us to describe the distribution of a variable in a sample of  $n$  individuals or  $n$  tumour biopsies, for example. In univariate analyses we commonly use parameters such as the median, mean, and standard deviation to describe *quantitative* (or *continuous*) variables and frequencies and percentages to describe *categorical variables*. We can also estimate population parameters by calculating 95% *confidence intervals* (CIs) for the aforementioned summary statistics (median,

mean, percentage). With univariate analyses we can only answer “descriptive questions” in a single arm or cohort, such as “What is the rate of responders to drug X?” or “What is the mean survival time in patients treated with drug Y?”

But what about situations where we wish to analyse more than one variable at a time? The purpose of *bivariate* and *multivariable analyses* is to probe the relationships between two (bivariate) or more than two (multivariable) variables. These types of analyses allow us to test a previously defined hypothesis (e.g. the primary efficacy analysis of a confirmatory study) or to explore the existing relationships between the collected variables (e.g. between-arm analyses, subgroup analyses, exploratory analyses). With bivariate and multivariable analyses we can answer “analytical questions” in one or more cohorts, such as “What is the overall survival with drug X compared with drug Y?”, “What is the efficacy of drug Z, based on the reduction in cholesterol levels, compared with placebo?”, or “What is the relationship between response rate to drug X and the level of biomarker Y?”

In both bivariate and multivariable

analyses the participating variables can be classified into:

- *Dependent* (or *outcome* or *predicted*) variables and
- *Independent* (or *predictor* or *explanatory*) variables, which in some models can be further classified into factors and covariates (or confounding factors).

In a bivariate analysis (sometimes referred to as *univariate* – see Box 1 below) there is only one independent and one dependent variable.

In a multivariable analysis there are:

- **One** dependent variable and
- **Two or more** independent variables.

**BOX 1:** Bivariate analyses that analyse the relationship between one independent variable and one dependent variable are often referred to as “univariate” analyses to distinguish them from multivariable analyses, in which two or more independent variables are assessed in relation to a dependent outcome. In this context, the term “univariate” is correct and replaces the term “bivariate”.



Multivariable analyses should not be confused with *multivariate analyses*, which are used to assess the relationships of several predictors with two or more dependent variables or outcomes at the same time. In this article we will not review multivariate analyses. However, medical writers should be aware that the terms multivariate and multivariable are often used interchangeably. Do not be surprised to see multivariable analyses described as multivariate.

To correctly interpret a multivariable analysis it is highly recommendable to first look at the bivariate analyses between the variables that were involved in the multivariable modelling. They show you: 1. the raw relationships between the dependent and independent variables (which allow the unadjusted associations to be quantified) and 2. correlations or associations between independent variables (which, if present, may require changes to the model).

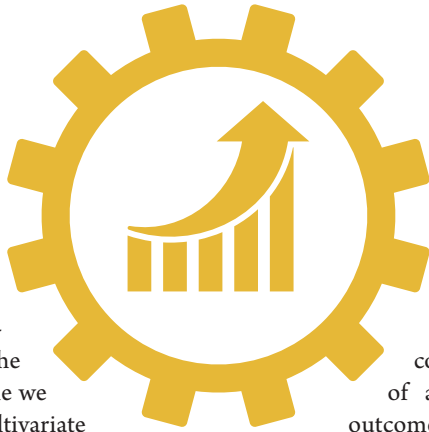
### Variables: Dependent vs independent / Quantitative vs categorical

It is very important to note that **both the dependent and independent variables can be either quantitative or categorical, and correct identification of these statistical properties is essential for the medical writer to correctly interpret and report the results.**

Common *quantitative outcomes* include cholesterol levels, blood pressure, and questionnaire scores and common *categorical outcomes* are survival (yes/no), response to treatment, and presence/absence of a specific event (e.g. cardiovascular event, relapse).

Common *quantitative predictors* include age, BMI, and baseline values for the outcomes. Common *categorical predictors* include treatment arm, gender, and baseline disease severity.

Note that categorical variables with only two categories are referred to as *dichotomous*.



The dependent variable is the one that is assessed with the study. Sometimes it is referred to as the *endpoint*. Usually this term is reserved for the combination of the outcome plus the timepoint(s) of assessment (e.g. if the outcome is “mortality”, the endpoint could be “mortality rate at 6 months”).

The independent variables define the subgroups of patients in which the outcome will be compared (e.g. treatment arms).

- If the independent variable is categorical (e.g. treatment arm, gender), the parameters of the multivariable models we will review in later sections – the odds ratio (OR), hazard ratio (HR), and beta coefficient ( $\beta$ ) – always estimate the effect on the outcome of **one or more categories versus a reference category** (e.g. placebo or female gender), which must be defined *a priori*.

- If the independent variable is quantitative (e.g. age), no subgroups are compared and the OR,  $\beta$ , and HR estimate the effect on the outcome of **each 1-unit increase in the independent variable** (e.g. “for each 1 mg/dl increase in baseline cholesterol”).

It is very common for continuous predictors to be transformed into categorical variables prior to the multivariable analysis using a previously defined cut-off point (from the literature). This is because the parameters of the models are much easier for physicians to interpret if they compare one category to another than if they inform about the risk associated with a 1-unit increase in the predictor. However, this leads to a loss of statistical power and to the risk of not finding significant results. If the model includes the original continuous predictor, the medical writer may facilitate interpretation of the results by reporting the risk associated with, for example, a 10-unit increase in the predictor.

In interpreting a multivariable analysis

we must also consider that some independent variables may be entered in the model because they are *confounding variables* (sometimes also denoted as *covariates*). Confounding variables are factors related to both the dependent and independent variables. Unless we adjust our multivariable analysis for confounding variables, we may end up with an inaccurate or incorrect representation of the true relationships between the dependent and independent variables. For example, in many clinical trials the baseline value for a quantitative outcome (e.g. baseline blood pressure in a hypertension trial) is a potential confounding variable if it is not fully balanced between the two treatment arms, despite randomisation of the patients, because it is also related to the outcome. For this reason, the primary efficacy analysis should always include the baseline value for the quantitative outcome as a covariate.

### When to apply a multivariable analysis

A multivariable analysis is needed in the following cases:

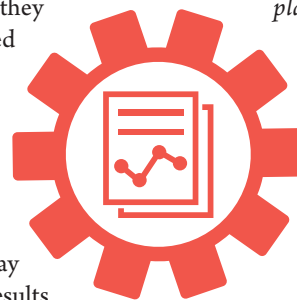
1. If there is **one** main independent variable of interest (the other independent variables being secondary factors):
  - a. To evaluate the relationship between the variable of interest and the outcome **after adjusting (or controlling) for other independent variables that may also be related to the outcome (confounding factors or covariates).**

*Examples:*

“Patients treated with drug A had significantly higher cholesterol levels at 6 months compared to patients treated with placebo, after adjustment for baseline cholesterol.”

“Higher biomarker X levels were significantly associated with a higher response rate, independently of/after adjusting for age and gender.” (Box 2 opposite)

2. If there are **two or more** main independent variables of interest:
  - a. To explore which of the independent variables are **independently** associated



with the outcome, i.e. they keep a significant p-value in the model despite the inclusion of other independent variables:

**exploratory models.**

These models are commonly used to look for “causal relationships”, although the results must always be interpreted with caution because associations may be due to confounding factors that were not accounted for.

**Example:**

“In patients with disease Z, male gender and higher blood pressure were independently associated with higher mortality.”

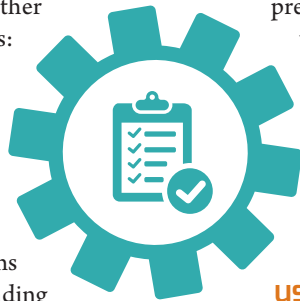
- b. To predict an outcome with independent variables that are known to be associated with the outcome:
- predictive models.**

These models are commonly used in oncology to establish prognostic factors that may be useful to select candidate patients for more aggressive therapies. They can also be used to predict response, compliance, and quality of life.

**Example:**

“In patients with disease Z, the independent factors predicting response to drug X were tumour stage at diagnosis and baseline beta-2-microglobulin level. The model with these two variables correctly predicted the response in 65% of patients”.

**BOX 2:** When describing associations between different variables, a common mistake is to not give the direction of the association, e.g. “Biomarker X levels were significantly associated with the response rate.” From this sentence, the reader cannot ascertain whether a higher response rate is associated with high or low biomarker X levels. Although, if not otherwise indicated, such an association would usually be interpreted as positive, a good medical writer should clearly indicate the direction of the association.



**Multivariable analyses commonly used in biomedical studies**

Please note that the words “independently associated” and “independent factor/predictor” imply that a multivariable model has been used and that the described relationship has been adjusted for at least one additional factor.

There are several different types of multivariable analysis. Three of the most commonly used analyses are *multiple logistic regression*, *multiple Cox regression*, and *multiple linear regression/multiple analysis of variance (ANOVA)/analysis of covariance (ANCOVA)* (Table 1 overleaf). It is important to note that multiple regression and *multivariate regression* are not the same thing. In multiple regression there is only one dependent variable; multivariate regression involves two or more main dependent variables and is less commonly used.

With multiple logistic regression the aim is to determine how one **dichotomous dependent variable** varies according to two or more independent (quantitative or categorical) variables. Multiple logistic regression might, for example, be used to test the relationships of weekly alcohol consumption at age 30 and gender (independent variables) with probability of developing liver cancer during a 10 year period (dependent variable). Liver cancer is a categorical variable with two categories at the end of the follow-up period: “cancer” and “no cancer”.

Multiple Cox regression is similar to multiple logistic regression but it explores the relationships between independent variables and a **time-to-event dependent variable** (dichotomous), e.g. time to death. If we wanted to determine whether a new treatment (independent variable) affects probability of disease progression (dependent variable) in patients with renal cell carcinomas of different clinical stages at baseline (second independent variable that may be considered a covariate), we could

potentially use multiple Cox regression.

Finally, multiple linear regression, multiple ANOVA, and ANCOVA are multivariable models in which the **dependent variable is continuous**, i.e. it can theoretically take any value in its given range. Despite being slightly different from each other, these models can be considered equivalent from a medical writer’s point of view. An example scenario would be to determine whether a new treatment (independent variable) reduces the score for disease index X (dependent variable) after adjusting for country and baseline disease index X score (independent variables considered as covariates).

These multivariable analyses will be discussed in further detail below. The aim is not to explain how to run the analyses, rather how to interpret and report the results they give. The focus will be on ORs, HRs, and  $\beta$  coefficients.

**Multiple logistic regression: What is an odds ratio?**

What is an OR? Let’s define two groups of subjects: a test group we are interested in and a reference group we wish to compare the test group to. The OR is the **ratio of two sets of odds**: the *odds* of an event occurring *in the test group* divided by the *odds* of the same event *in the reference group*. Note that odds are *not* the same as probability: the odds are the probability of an event (e.g. death) occurring *divided by* the probability of it not occurring. While probability ranges from 0 to 1, the odds may range from 0 to positive infinity.

Going back to our example above, how do weekly alcohol consumption and gender affect the odds of developing liver cancer? Here we can define two reference groups: one for weekly alcohol consumption and one for gender. The reference group for weekly alcohol consumption might be “0 units” and let’s say the one for gender is “female”. (If you’re wondering how a categorical independent variable such as gender may be entered into a mathematical model, this can be achieved by creating a *dummy variable* with a value of 0 or 1. In the present example, females may be given a value of 0 and males 1.)

	Multiple logistic regression	Multiple Cox regression	Multiple linear regression / Multiple ANOVA / ANCOVA
Dependent variable	<b>Dichotomous</b> (no information about timepoint)  <i>Example: Treatment response (yes/no)</i>	<b>Time to event</b> (dichotomous with information about timepoint)  <i>Example: Overall survival</i>	<b>Quantitative</b>  <i>Example: Blood pressure</i>
Independent variables	2 or more quantitative or categorical variables	2 or more quantitative or categorical variables	2 or more quantitative or categorical variables <sup>a</sup>
Equation <sup>b</sup>	$\text{logit}(p) = a + b_1x_1 + b_2x_2 \dots$	$\log(h_i(t)) = a + b_1x_1 + b_2x_2 \dots$	$y = a + b_1x_1 + b_2x_2 \dots$
Parameter	OR (= Exp(b))	HR (= Exp(b))	$\beta$ (= b)
Interpretation	Odds for: <ul style="list-style-type: none"> <li>• Category X vs reference category (if independent variable is categorical)</li> <li>• A 1-unit increase (if independent variable is quantitative)</li> </ul>	Instantaneous risk/hazard (hazard per unit time) for: <ul style="list-style-type: none"> <li>• Category X vs reference category (if independent variable is categorical)</li> <li>• A 1-unit increase (if independent variable is quantitative)</li> </ul>	Size of the effect on the outcome (in outcome units) for: <ul style="list-style-type: none"> <li>• Category X vs reference category (if independent variable is categorical)</li> <li>• A 1-unit increase (if independent variable is quantitative)</li> </ul>
Example of reporting	<i>"... odds of treatment failure were 3 times higher in men than in women"</i>	<i>"... risk of death was 3 times higher in men versus women"</i>	<i>"... systolic blood pressure was 3 mmHg higher in men than in women"</i>

<sup>a</sup> For ANOVA and ANCOVA at least 1 categorical variable is needed

<sup>b</sup>  $\text{logit}(p)$  is  $\log(p/1-p)$ , where  $p$  is the probability of the outcome;  $a$  denotes a constant,  $b_n$  denotes the coefficient for each independent variable,  $x_n$  denotes an independent variable,  $h_i(t)$  is the hazard to individual  $i$  at time  $t$ , and  $y$  denotes a dependent variable

**Table 1. Types of multivariable models commonly used in biomedical studies**

Say we obtain an OR for liver cancer of 1.68 for people who consume 40+ units of alcohol per week versus those who consume 0 units per week. This means that the odds of liver cancer are 1.68 times as high (or 68% higher) for those consuming 40+ units of alcohol per week than for teetotallers. Similarly, an OR of 1.22 for males versus females would mean that males have 22% higher odds of developing liver cancer compared to females.

ORs are typically presented with CIs. In general terms, the CI is a range of values within which the true value of a parameter **in the population** (not in the study sample) is expected to lie. A narrow CI indicates good precision in our OR estimate; a wider CI would indicate more uncertainty.

Narrower intervals are obtained with larger samples. For an OR, a CI that includes 1 (e.g. 0.9 to 2.5) prevents us from inferring a significant difference between groups.

If we adjust our multiple logistic regression model for confounder variables, then the ORs we obtain will be referred to as *adjusted ORs*. If in the present example we calculate a 95% CI of 1.25 to 2.13 for our OR of 1.68, we could describe the results of the multiple logistic regression thus:

*Compared to teetotallers, those who consumed 40+ units of alcohol per week at age 30 had higher odds of developing liver cancer (adjusted OR=1.68, 95% CI=1.25 to 2.13). Males had higher odds of liver cancer than females (adjusted OR=1.22, 95% CI=1.03 to 1.44).*

Note that we are not claiming that alcohol consumption *causes* liver cancer (although there is ample evidence to suggest this is the case). Rather, we are merely saying that excessive alcohol consumption *is associated with* liver cancer; it may or may not cause liver cancer.

*Risk* has a particular meaning in statistics, with *relative risk* (RR) implying a comparison of probabilities, not odds. In the above example, the odds of liver cancer were 1.22 times higher in males compared to females; we should not write that the “risk” of liver cancer was 1.22 times higher in males, because this would be inaccurate. Phrases that indicate or imply probability, such as “X times as likely to” and “a 50% higher probability of”, should also be

avoided when reporting ORs. Note that the OR gives a reasonable approximation for the RR when the event is rare, but not when the event is common.

**Multiple Cox regression:  
What is a hazard ratio?**

Multiple Cox regression is used to calculate HRs. An HR indicates the *instantaneous risk* or *hazard* (hazard per unit time, usually 1 day) of an event (e.g. death) in a test group relative to a reference group. Let's return to the example of the new treatment for renal cell carcinoma. The new treatment (test group) gives an HR for death of 0.5 versus the existing gold standard treatment (reference group). How do we interpret this?

In this example, the HR indicates the relative rates of death per day in the two treatment groups. The value of 0.5 indicates that the rate of death at any time during the follow-up period is twice as high with the gold standard treatment compared to the new treatment. A value of 1.0 would indicate no difference in rate of death between the two treatments.

Like ORs and  $\beta$  coefficients, HRs are typically presented with CIs. Assuming we adjust our multiple Cox regression model for several confounder variables, we could report the results in the present example as:

*Compared to the gold standard treatment, the new treatment was associated with a significantly lower rate of death (adjusted HR=0.5, 95% CI=0.25 to 0.75).*

In descriptions of survival analyses, RR cannot be used instead of HR, since the two terms are not synonymous. Though they can be interpreted in more or less the same way, HRs and RRs are calculated differently. Notably, RRs do not account for the timing of the events of interest. Don't write relative risk when you mean hazard ratio!

**Multiple linear regression /  
Multiple ANOVA / ANCOVA:  
What is the  $\beta$  coefficient?**

In multiple linear regression, multiple

ANOVA, and ANCOVA, the dependent variable is continuous. One such variable is height at age 18. What is its relationship with birth length and age at puberty onset (independent variables)?

In addressing this question by multiple linear regression we obtain one  $\beta$  coefficient for each quantitative independent variable and for each non-reference category of each categorical independent variable.

For continuous independent variables such as birth length the  $\beta$  coefficient indicates how a 1-unit change in the value of the independent variable would affect the value of the dependent variable if all other variables in the model were held constant, and the units for  $\beta$  are the units for the dependent variable divided by those for the independent variable. For categorical independent variables the units of the  $\beta$  coefficient are the same as those of the dependent variable. It is very important to understand this to correctly describe the results of the model.

If the  $\beta$  coefficient for birth length is positive (e.g. 1.2 cm/cm), then a higher birth length will be associated with a greater height at age 18. A negative  $\beta$  coefficient for age at puberty onset (e.g. -0.3 cm/year) indicates a negative association between age at puberty onset and height at age 18. The statistical significance of these results should be reported using the p-value associated with each  $\beta$  coefficient.

For categorical independent variables such as gender the  $\beta$  coefficient and the **corresponding** p-value will indicate whether the category is associated with greater height at age 18 compared to the reference category. A positive  $\beta$  coefficient for males relative to females with a p-value of <0.05 would indicate that males are likely to be taller than females at age 18.

$\beta$  coefficients are often presented with corresponding CIs; sometimes the CI is replaced by the standard error (SE) and the p-value. **If a CI does not include 0**, the association between the independent



variable and the dependent variable (after adjusting for covariates) is significant. Thus we could describe the results of the current analysis as:

*Higher birth length was associated with greater height at age 18 ( $\beta=1.2$  cm/cm, 95% CI=0.93 to 1.49). Age at puberty onset was inversely associated with height at age 18 ( $\beta=-0.3$  cm/year, 95% CI=-0.19 to -0.45).*

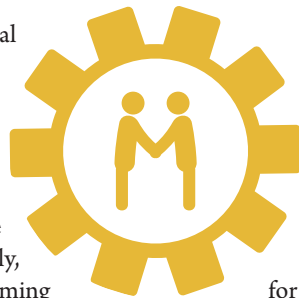
As a final remark regarding  $\beta$  coefficients, please be aware that they are sometimes also provided for multiple logistic regression and Cox regression models. In such cases,  $\beta$  is simply the natural logarithm (ln) of the OR (logistic regression) or HR (Cox regression).

**How to report the results from multivariable models**

Whatever the model used, good medical writing practice is to list all the factors that were taken into account in the multivariable analysis, including those that were discarded during the modelling process. Also, remember always to include the parameter that indicates the strength **and direction** of the association (i.e. the OR, HR, or  $\beta$  coefficient), preferably with the 95% CI and/or the p-value **for the variable** (different from the overall p-value for the model). If you are at all unsure as to the direction of a particular association, ask a statistician for clarification.

When reporting the parameter, the writing differs depending on the direction of the association. We round off our introduction to multivariable analyses with some illustrative examples:

- For ORs (logistic regression) and HRs (Cox regression), results are significant when the 95% CI does not include 1:
- A value <1 implies that the factor is negatively associated with (i.e. protects against) the outcome. The percentage decrease in the odds (OR) or risk (HR) is  $(1 - \text{OR or HR}) \times 100$ . Example: "Category X protected against mortality (adjusted OR=0.8, 95% CI=0.6 to 0.9 versus reference category)



Y)” or “Compared to Y, X was associated with a 20% reduction in the odds of death.”

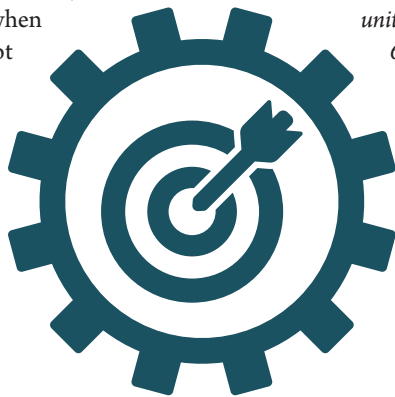
- A value >1 implies that the factor is positively associated with (i.e. increases the risk of) the outcome. The percentage increase in the odds (OR) or risk (HR) is (OR or HR - 1) × 100.

Example: “Category X was a risk factor for mortality (adjusted HR=1.5, 95% CI=1.1 to 1.9 versus reference category Y)” or “Compared to Y, X was associated with a 50% increase in the risk of death.”

When the percentage is ≥100, the “number of times” construction is often used:

Example: “Patients with X had a risk of death approximately three times higher compared to those with Y (adjusted HR=3.2, 95% CI=2.1 to 4.9).”

- For  $\beta$  coefficients (multiple regression, multiple ANOVA, ANCOVA), results are significant when the 95% CI does not include 0:



- A value <0 implies that the factor is negatively associated with the outcome.

Examples:

Quantitative factor: “A 1-unit increase in X was associated with a decrease of 3 mmHg in systolic blood pressure at 4 weeks ( $\beta=-3$ , 95% CI=-2.1 to -3.9).”

Categorical factor: “Compared to placebo, treatment with drug Z was associated with a decrease of 3 mmHg in systolic blood pressure at 4 weeks ( $\beta=-3$ , 95% CI=-2.1 to 3.9).”

- A value >0 implies that the factor is positively associated with the outcome.

Examples:

Quantitative factor: “A 1 mg/dl increase in X was associated with a 2-unit increase in quality of life score at 6 months ( $\beta=2$ , SE=0.3,  $p=0.025$ ).”

Categorical factor: “Compared to patients with mild disease at baseline, severe disease was associated with a 2-unit lower quality of life score at 6 months ( $\beta=-2$ , SE=0.3,  $p=0.025$ ).”

## Acknowledgements

We would like to thank Elena Stolyarova, a biostatistician at TFS, for her insightful comments on an early draft of this article. We also thank Adam Jacobs, Senior Principal Statistician at Premier Research, for providing constructive feedback and suggesting improvements.

## Conflicts of Interest and Disclaimers

The content of this article reflects the view of the authors, and not those of TFS. There are no conflicts of interest.

## Author information

**Neus Valveny** has been a medical writer for the last 15 years, preparing more than 40 manuscripts, 50 abstracts and posters, 20 study protocols, and 15 study reports. She has also worked as a statistician. She is now Associate Director of Medical Writing at TFS, a medium-sized international CRO founded in Sweden.

Previously a full-time science editor and part-time copy editor, **Stephen Gilliver** is now a medical writer at TFS. He is also Co-Editor of *Medical Writing*.

# Call for Companies

The 2nd Medical Writing Internship Forum will be held at our May 2017 Conference in Birmingham, UK.

Please contact [internship@emwa.org](mailto:internship@emwa.org) for more information.