Letter to Editor

"Multiple" and "Multivariate": Two Commonly Misused Statistical Terms in Research Methodology

Dear Editor

Regression modeling is a powerful tool for examining associations between dependent and independent variables. Two frequently misunderstood terms in this context are "multiple" and "multivariate" regression. In multiple regression, a single dependent variable is predicted by one or more independent variables.[1] In multivariate regression, multiple dependent variables are predicted by one or more independent variables.[1] The choice between these models could significantly impact research conclusions. Using multiple regression where a multivariate model is appropriate may obscure important interactions among dependent variables, leading to incomplete or misleading interpretations.[2]

Difference between Univariate, Multiple, and Multivariate Regression

Univariate Regression: Univariate regression involves one dependent and one independent variable. It models the relationship between an independent variable and a dependent variable.[3] This method is common in fields such as economics (e.g., predicting sales from advertising spend), biology (e.g., modeling the effect of temperature on plant growth), and social sciences (e.g., analyzing the impact of study hours on exam scores).

Example 1: A simple linear regression model assessed whether systolic blood pressure increases with age among diabetic patients.

Dependent variable: Systolic blood pressure (mmHg)

Independent variable: Age (years)

Multiple Regression: Multiple regression involves one dependent variable and several independent variables. It analyzes the collective effect of multiple predictors on a single outcome. This technique is widely used in finance (e.g., predicting stock prices using various economic indicators), healthcare (e.g., assessing the effect of various treatments on patient recovery), and marketing (e.g., evaluating the impact of different promotional strategies on sales).

Example 2: A multiple linear regression model was applied to determine predictors of serum cholesterol levels among adults.

Dependent variable: Serum cholesterol (mg/dL)

Independent variables: Age (years), body mass index (BMI, Kg/m²), physical activity level (low/moderate/high)

Multivariate Regression: Multivariate regression involves several dependent variables and one or more independent variables.[4] It examines relationships between multiple dependent variables and the effects of the independent variables on them simultaneously. Multivariate regression is commonly used in fields such as psychology (e.g., studying factors affecting various mental health outcomes), environmental science (e.g., assessing pollution's impact on multiple health indicators), and economics (e.g., analyzing the effects of policy changes on economic measures).

Example 3: A multivariate multiple regression analysis evaluated the simultaneous effects of a new antihypertensive treatment on both systolic and diastolic blood pressures.

Dependent variables: Systolic and diastolic blood pressures (mmHg)

Independent variable: Treatment group (new drug vs. control)

In conclusion, distinguishing between "multiple" and "multivariate" regression is paramount for researchers. This distinction ensures the selection of appropriate analytical methods, thereby enhancing the integrity and credibility of research findings. Multiple regression models are specifically designed to predict a single dependent variable using multiple independent variables. This approach allows researchers to explore how various factors collectively influence one outcome. Multiple regression

predicts a single outcome from multiple predictors, while multivariate regression models multiple outcomes simultaneously.

Although selecting an appropriate regression model is crucial for obtaining valid and interpretable results, a primary challenge in research lies not only in choosing the “true” model, but in its proper application and interpretation. Misinterpretation of regression techniques can lead to misleading conclusions, such as ignoring relationships between variables. Therefore, developing a strong conceptual understanding of these models is essential for correct implementation.

Ultimately, differentiating between multiple and multivariate regression models strengthens the robustness of statistical analyses and enhances the validity and generalizability of conclusions. This awareness is vital for advancing scientific knowledge and informing evidence-based decisions in policymaking, healthcare, and other fields where accurate data interpretation has a significant real-world impact.

Authors’ Contribution

F.M.: Study concept, data gathering, and drafting; S.B.: Study concept, data gathering, and drafting. All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Declaration of AI

We use ChatGPT to parapharise and improve English.

Conflict of Interest: None declared.

Keywords ● Multiple regression model ● Multivariate regression model ● Univariate regression model ● Research methodology

Farzan Madadizadeh[1,2], PhD; Sajjad Bahariniya[3], PhD

[1] Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran;

[2]Center for Healthcare Data Modeling, Departments of Biostatistics and Epidemiology, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran;

[3]Department of Healthcare Services Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

Correspondence:
Sajjad Bahariniya, PhD;
School of Health Management and Information Sciences, Iran University of Medical Sciences, Valiasr St., Above Mirdamad, Rashid Yasemi St., Postal code: 89179-33431, Tehran, Iran
Email: sajjadbahari98@gmail.com
Received: 11 July 2025
Revised: 19 September 2025
Accepted: 11 October 2025

Please cite this article as: Madadizadeh F, Bahariniya S. “Multiple” and “Multivariate”: Two Commonly Misused Statistical Terms in Research Methodology. Iran J Med Sci. doi: 10.30476/ijms.2025.107689.4221.

References

1 Sun Y, Wang X, Zhang C, Zuo M. Multiple regression: Methodology and applications. Highlights in Science, Engineering and Technology. 2023;49:542-8. doi: 10.54097/hset.v49i.8611.

2 Westad F, Marini F. Variable selection and redundancy in multivariate regression models. Frontiers in Analytical Science. 2022;2:897605. doi: 10.3389/frans.2022.897605.
3 Hongyue W, Jing P, Bokai W, Xiang L, Julia ZZ, Kejia W, et al. Inconsistency between univariate and multiple logistic regressions. Shanghai archives of psychiatry. 2017;29:124.
4 Schober P, Vetter TR. Logistic Regression in Medical Research. Anesth Analg. 2021;132:365-6. doi: 10.1213/ANE.0000000000005247. PubMed PMID: 33449558; PubMed Central PMCID: PMC7785709.