



Brief review

Review and practical excursus on the comparison between traditional statics methods and Classification And Regression Tree (CART) in real-life data: Low protein diet compared to Mediterranean diet in patients with chronic kidney disease



Vincenzo Calabrese ^{a,*}, Daniela Metro ^{b,1}, Angela Alibrandi ^c, Domenico Maviglia ^d, Valeria Cernaro ^b, Veronica Maressa ^b, Elisa Longhitano ^b, Guido Gembillo ^b, Domenico Santoro ^b

^a Unit of Nephrology and Dialysis, Department of Medicine and Surgery, University of Enna “Kore”, Italy

^b Unit of Nephrology and Dialysis, Department of Clinical and Experimental Medicine, A.O.U. ‘G. Martino’, University of Messina, Messina, Italy

^c Department of Economics, Unit of Statistical and Mathematical Sciences, University of Messina, Italy

^d University of Messina, Messina, Italy

ARTICLE INFO

Article history:

Received 10 September 2024

Accepted 23 January 2025

Keywords:

CART

CKD

Diet

Machine learning

Real-life studies

Mediterranean diet

Low protein diet

Renal disease

Kidney impairment

ABSTRACT

The Classification and Regression Tree (CART) is a supervised learning approach useful used to segment the space of and the predictors/features space into smaller homogeneous regions that are, represented in a decision tree. It computes the selection of features automatically, in contrast to traditional statistic methods. In this review, we compared CART to traditional statistics in patients who did not attempt a diet to patients who followed a low-protein diet (LPD) or the Mediterranean diet in chronic kidney disease (CKD) patients, and we analyzed them using linear regression and CART methods. In our example, diet adherence proved to be the factor with the greatest impact on renal function decline, but CART failed to detect significant differences between LPD and Mediterranean diet. Similar results were found using traditional statics, but CART gave a model with the proportion of the explained outcome by the model (R^2) higher by about 20%, thus a stronger model. In addition, CART allows quick and easy identification of the variables affecting the outcome with a simple visual representation through the decision tree, increasing the interpretability of the results. In summary, no difference in the impact of variables has been detected with the two methods, but CART gave us a more detailed model with a faster and easier interpretation.

© 2025 Published by Elsevier España, S.L.U. on behalf of Sociedad Española de Nefrología.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: v.calabrese@outlook.it (V. Calabrese).

¹ These authors contributed equally to this work.

<http://dx.doi.org/10.1016/j.nefro.2025.04.008>

2013-2514/© 2025 Published by Elsevier España, S.L.U. on behalf of Sociedad Española de Nefrología. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Revisión y excursos práctico sobre la comparación entre los métodos estadísticos tradicionales y el Árbol de Clasificación y Regresión en la práctica clínica: dieta baja en proteínas comparada con la dieta mediterránea en pacientes con enfermedad renal crónica

RESUMEN

Palabras clave:

CART
CKD
Dieta
Machine learning
Práctica clínica
Dieta Mediterránea
Dieta baja en proteínas
Enfermedad renal
Insuficiencia renal crónica

El Árbol de Clasificación y Regresión (CART) es un enfoque de aprendizaje supervisado usado para segmentar el espacio de los predictores/características en regiones homogéneas más pequeñas que se representan en un árbol de decisión. Calcula la selección de las características automáticamente, a diferencia de los métodos estadísticos tradicionales. En esta revisión, comparamos CART con la estadística tradicional en pacientes que no intentaron una dieta con pacientes que siguieron una Low Protein Diet (LPD) o la dieta mediterránea en pacientes con enfermedad renal crónica (ERC), y los analizamos utilizando regresión lineal y métodos CART. En nuestro ejemplo, la adherencia a la dieta demostró ser el factor con el mayor impacto en el deterioro de la función renal, pero CART no detectó diferencias significativas entre la LPD y la dieta mediterránea. Se encontraron resultados similares utilizando la estadística tradicional, pero CART proporcionó un modelo con la proporción del resultado explicado por el modelo (R^2) superior en aproximadamente un 20%, lo que resulta en un modelo más robusto. Además, CART permite la identificación rápida y sencilla de las variables que afectan el resultado con una representación visual simple a través del árbol de decisión, facilita la interpretabilidad de los resultados. Resumiendo, no se han detectado diferencias en el impacto de las variables con los dos métodos, pero CART nos proporcionó un modelo más detallado con una interpretación más rápida y sencilla.

© 2025 Publicado por Elsevier España, S.L.U. en nombre de Sociedad Española de Nefrología. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

Introduction

Although randomized clinical trials (RCTs) are the gold standard for evaluating treatment efficacy, observational studies are fundamental to generating hypotheses, evaluating the treatment in real life, and generating audits. However, observational are subjected to many biases and the correct statistical methods should be applied to give realistic results. Indeed, each statistical analysis method has its own advantages and limitations. In this paper, we conduct a detailed study comparing the characteristics of Classification and Regression Trees (CART) with traditional statistical methods, focusing specifically on their application in a real-world evaluation of the impact of diet on renal function.

Diet in CKD

The reduction in glomerular filtration rate (GFR) after the age of 40 is a parapsychological event; hypertension, diabetes, cardiovascular disease and inadequate diets are the major risk factors for developing moderate to advanced degrees of chronic kidney disease (CKD).

Nutritional therapy can be useful in slowing the progression of CKD and delaying the need for a preventive kidney transplant or the start of renal replacement therapy, while continuously improving patients' quality of life.¹

Healthy dietary habits are crucial to counteract the progression of chronic diseases such as CKD and the risk factors associated with their development. A customized diet

based on patients' eating habits can promote adherence to nutritional therapy and thus improve the conservative management of CKD patients. Optimal nutrition is crucial for patients with impaired kidney function. It is an important modifiable lifestyle factor in the primary prevention of CKD progression.^{2,3}

Machine learning

Machine learning is a group of automatic performers to solve problems in most sciences useful to improve diagnostics accuracy establish a prognosis and monitor physiological data. They can be divided into supervised and unsupervised ML methods, based on whether the output is known or not, respectively.⁴

Unsupervised methods are often used to split into clusters or find subsamples with high likelihood, and they are often in genetic mapping.

Conversely, supervised ML are useful to compute the strength of the association between variables and outcome. Among them, we can find Classification and Regression Trees (CART), Random Forests (RF), and Neural Networks (NNs). In this paper, we developed the first of them.

Classification and regression tree (CART)

Classification and regression tree is a supervised learning approach useful to segment the predictors/feature space into smaller regions with more homogeneous outcomes. This set of segments is summarized in a decision tree. CART can be used for regression or classification, without methodological

Table 1 – Baseline feature split for diet.

Variable	Whole sample (n = 107)	Group A (n = 32)	Group B (n = 40)	Group C (n = 35)	p
Age (years)	72 (10)	75 (6)	68 (9)	74 (13)	0.01
Sex (M)	64 (60%)	23 (72%)	24 (60%)	17 (49%)	0.15
BMI (kg/mq)	28.0 [25.2–30.7]	29.6 [27.6–31.8]	28.7 [25.0–34.7]	26.4 [24.0–28.0]	<0.01
CrCl (ml/min)	34.0 (13.1)	39.1 (12)	35.2 (13)	28.0 (13.7)	<0.01
Serum urea (mg/dl)	73 [60–98]	75 [58–98]	73 [63–98]	73 [60–121]	0.75
Serum creatinine (mg/dl)	1.8 [1.5–2.3]	1.8 [1.4–2.2]	1.7 [1.4–2.3]	2.0 [1.6–2.7]	0.05
Serum potassium (mmol/L)	4.7 (0.6)	5.1 (0.4)	4.6 (0.4)	4.6 (0.7)	0.90
Hyperkalemia (n)	24 (22%)	4 (13%)	9 (23%)	11 (31%)	0.20
Serum phosphate (mg/dl)	3.7 (0.7)	3.7 (0.4)	3.8 (0.5)	3.8 (0.4)	<0.01
Hyperphosphatemia (n)	9 (8%)	0	4 (10%)	5 (14%)	0.09
WBC (cc/mmcc)	7500 [6200–8630]	6425 [5440–7622]	6675 [5275–7400]	9100 [7700–5800]	0.04

BMI = body mass index; CrCl = creatinine clearance; WBC = white blood cell. Variables are reported as mean (standard deviation) or median [interquartile range] or number (%), as appropriate concerning the nature and the distribution of the variables. Hyperkalemia is defined as serum potassium higher than 5.1 mmol/L and hyperphosphatemia is defined as a serum phosphate higher than 4.5 mg/dl.

differences, with the exploratory scope. It is not the gold standard to test hypothesis.⁵ Each region corresponds to a node in the tree, and two child nodes will occupy different regions of the parent node. These splits aim to reduce the residual sum of squared error (RSS), thus increasing the precision of our prediction.⁶ Similar to regression analysis, CART can be performed using only one dependent variable – comparable to the univariate regression model – and includes all the features that are comparable to multivariate regression analysis. However, conversely to traditional statistics, in exhaustive CART the selection of the variables is automatically computed.

One of the advantages of the tree-structured approach is to do an automatic stepwise variable selection and it can analyze different kinds of features, independently from their nature or distribution of them. Furthermore, CART is easy to interpret through its clear visual reading.⁷

Linear regression

The association between two variables, often a risk factor and an outcome, can be assessed by linear regression. It provides us with a coefficient which, when it is multiplied by the value of our risk factor, estimates the value of our outcome.⁸

The causal impact of a variable on the outcome can be influenced by other variables, the so-called “confounding variables”. According to Tripepi G. et al., confounders should be related to the outcome, should be associated with a risk factor and should not be involved in the pathophysiological pathway between risk factor and outcome. Confounders can increase, decrease or obscure the association between our risk factor and the outcome.⁹

Multivariate analysis is an analysis that includes in the model all confounding variables and variables related to allocation and outcome, which are not included in the pathophysiological pathway. The multivariate analysis makes it possible to adjust the relationship value determined in the univariate analysis for no more than 1 confounding variable per 10 events. However, the choice of confounding variables may not be wise.

Aim of the review

The aim of our study is to compare data about the effects of no diet, a low-protein diet and a Mediterranean diet on the progression of CKD in non-diabetic patients with traditional statistics and classification and regression trees (CART).

Practical example

To compare this method, we used a dataset of 105 non-randomized patients with CKD stage 2–5 on non-dialysis therapy. Patients were followed-up for one year. 32 patients were not subjected to any diet (group A), 40 patients were subjected to a low-protein diet (group B), and 35 patients were subjected to a Mediterranean diet (group C).

Allocation followed a real-life observational design, based on the dietitian’s recommendation, and the patient’s adherence. Anamnesis was collected at baseline, whereas laboratory data were collected at baseline and at annual follow-up.

Adherence to diet has been computed at the annual follow-up, corresponding to the end of our evaluation. Patients reported complete adherence to the diet in 82% (Mediterranean diet is 86% vs LPD 80%, $p=0.79$). For patients who did not follow a specific nutritional regimen was not possible to detect adherence, due to there being no diet to follow.

As reported in Table 1, the three groups differed in Age, BMI, basal creatinine clearance, serum phosphate, and white blood cells (WBC). These variables were selected as confounders and included in the multivariate regression model. Variables are reported as mean \pm standard deviation or median [interquartile range] or number (%), as appropriate concerning the nature and the distribution of the variables.

The univariate regression model showed a significant impact of the Mediterranean diet and LPD on creatinine clearance difference compared to patients not subjected to any diet (Table 2). These results were confirmed by multivariate analysis. Furthermore, also creatinine clearance at baseline had a significant association with our outcome and a trend of significance for BMI. This multivariate model had an R^2 of 0.43 (Table 2).

Table 2 – Univariate and multivariate linear regression analysis using the differences of creatinine clearance as dependent variable.

Variables	Univariate model			Multivariate model		
	β	95% CI	<i>p</i>	β	95% CI	<i>p</i>
Mediterranean diet (yes)	8.4	5.9/10.9	<0.001	8.3	4.3/12.4	<0.001
Low-protein diet (yes)	14.6	11.2/18.1	<0.001	14.8	11.1/18.4	<0.001
Age (years)				-0.02	-0.16/0.12	0.752
BMI (kg/mq)				0.23	-0.04/0.50	0.091
Baseline CrCl (ml/min)				-0.18	-0.29/-0.06	0.003
Serum phosphate (mg/dl)				-0.80	-3.20/1.60	0.511
WBC (10^3 cc/ml)				0.001	-0.001/0.002	0.326

BMI = body mass index; CrCl = creatinine clearance; WBC = white blood cell.

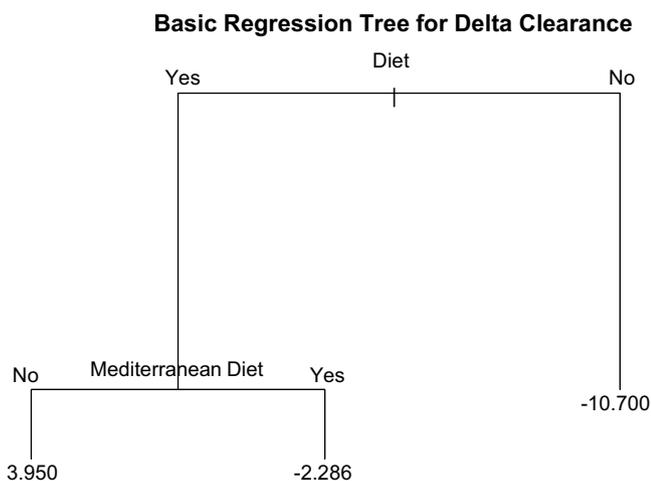


Fig. 1 – Basic classification and regression tree. Basic tree showed as a first node is based on the presence or absence of specific diet, focused as it was the variable with major impact on the variance of creatinine clearance.

CARTs were performed in the basic model and the exhaustive model, using the difference in creatinine clearance at follow-up as the continuous dependent variable, i.e., regression tree. The model was trained in a training test that included a randomly selected 70% of the sample. The remaining 30% was used for the test set. We used the complexity parameter, which better performed in terms of cross-validated accuracy in terms of repeated cross-validated mean square error.

Starting from a univariate basic CART model, we found that patients who did not follow a diet lost about 11 ml/min in one year, whereas differences in patients who followed a Mediterranean or low-protein diet had no clinical significance. According to their position on the first node, adherence or non-adherence to a particular diet was the most important variable.

Similar to the basic CART model (Fig. 1), the attempt or lack of a specific diet was the variable with the greatest influence in a multivariate CART model (Fig. 2). Moreover, no other split was performed based on type of diet in this last tree. Secondary splits were performed based on BMI and baseline CrCl value, which revealed a high impact of these characteristics on the course of clearance progression. The other nodes,

based on age and WBC, were lost in the pruned model, as its impact existed but was not very relevant.

Indeed, it is to highlight that the splits in the pruned CART were based on the same variable with significant impact in the multivariate regression model deleting variables not significant in the regression analysis. The RMSE was 11.87. Moreover, the R-square of the pruned model was 0.58 (Fig. 3), higher than the R-square of the multivariate regression model by about 20%.

Usefulness and limits of CART

Usefulness of CART

Observational studies based on real-life data are often subjected to many biases and missing information, and data do not always have a normal distribution. However, they can be used to do audits and surveys, which allow analyzing of the delivery of care and the prevalence of diseases at different levels,¹⁰ from a single ward to whole national data.^{11,12} Audits based on real-life data are often used in Intra-departmental assessments, and they need to be easily explainable and interpretable, and CART is in keeping with this assumption.¹³ Splitting the population in the subsample with high likelihood, allows us to detect the part of the population more related to our outcome. Indeed, CART deeply sums up the impact of variables on the outcome, and its plot is visually easier to understand.

Furthermore, this method provides often a model stronger than traditional statics, with a higher R^2 and lower error measures, and it does not require assumptions about the variable distributions or outcome types (continuous or dichotomous).

Limitation of CART

The first limitation of CART is given by the lack of conditional inference, thus has not concept of statistical significance. Indeed, it is based on the likelihood of the splits, and variables selection bias and overfitting risk are not solved. Since the algorithm only searches for the better split, the heuristic variable can be selected to deep the process, giving useless steps and it is not able to detect if some variable is or is not a “noise variable”.¹⁴

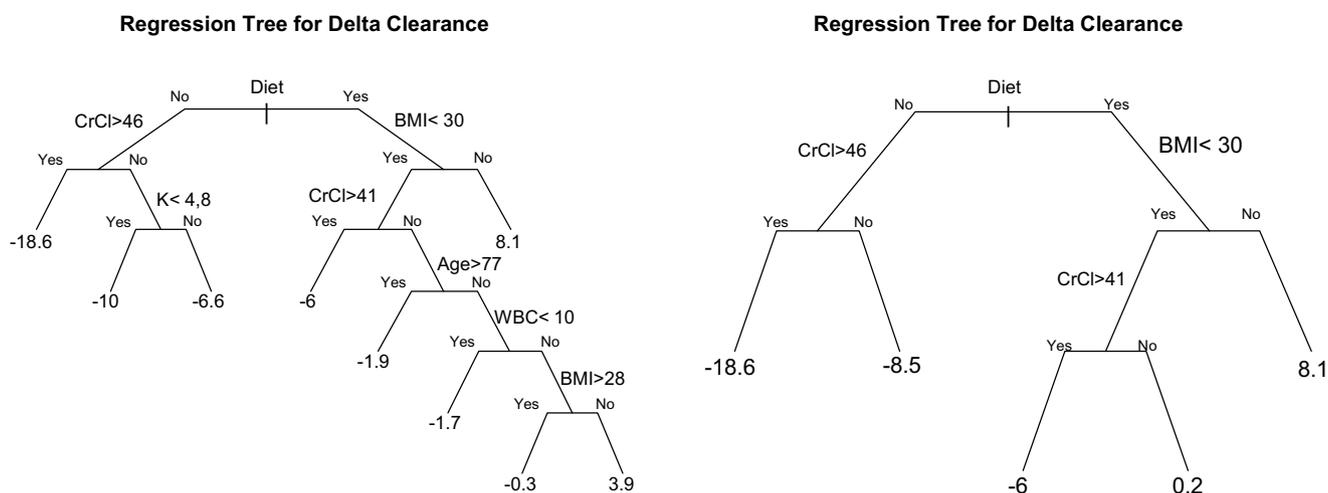


Fig. 2 – Basic exhaustive classification and regression tree (A) and pruned exhaustive classification and regression tree (B). In both tree models, the first node is based on the presence or absence of a particular diet, as this was the variable with the greatest impact on the variance of creatinine clearance. BMI = body mass index; CrCl = creatinine clearance; WBC = white blood cell count.

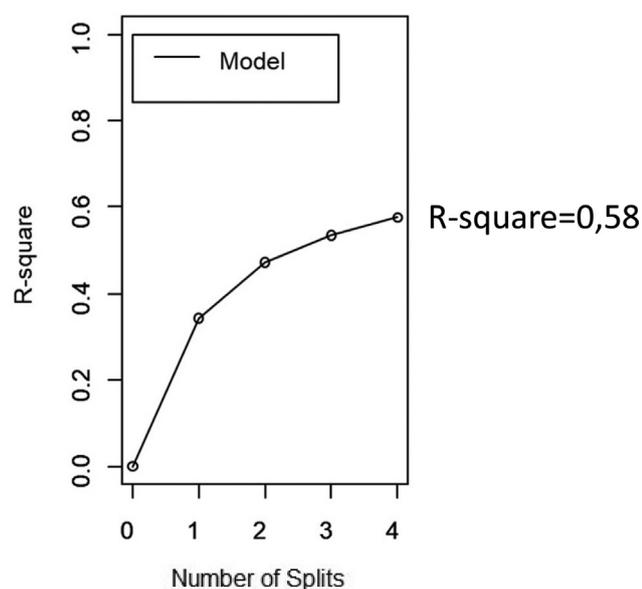


Fig. 3 – R-square representation of the pruned exhaustive classification and regression tree in each its split.

Discussion

Our analysis showed different trends in CKD progression in patients treated with a specific diet (low-protein or Mediterranean) and in patients without a specific diet. The analyses were performed using both traditional statistical analyses and machine learning methods. In both cases, the lack of a specific diet appeared to be the most important factor in the deterioration of creatinine clearance.

Although the results of the two methods are similar, CART probably allows easier detection of the variables involved in

the change in results. In addition, the comprehensibility of the decision tree is no less important, as is the visible different strength of the influence that each variable has on the outcome. The position of the variables in the tree, as well as the removal of some variables from the pruned tree, make it possible to understand well how much a variable affects the outcome.

Even though less considered, the difference in the R^2 between traditional methods and machine learning ought to be emphasized. Indeed, this indicates the proportion of the outcome that the model explains, and machine learning often allows for a higher R-square, despite of the easier choice of the selected variables.

Our results are in keeping with the literature. Indeed, is well established that adequate nutrition is a critical lifestyle factor in the primary prevention of CKD progression^{2,15} and avoids dysregulation of fluid, acid-base and electrolyte homeostasis.^{16–18}

Although the role of LPD is beyond doubt, patient adherence to this diet is crucial.¹⁹ On the other hand, the effect of LPD in the elderly is not as strong as in the young, as the geriatric population has poor adherence to dietary management.²⁰ Indeed, in patients refusing LPD with malnutrition as a consequence, the Mediterranean diet could be considered a valid choice.³ Although our analysis cannot confirm this, both the traditional analysis and CART were able to detect differences between patients who followed a diet and those who did not. In addition, CART did not detect split based on the Mediterranean diet or LPD. This focused on the high impact of having or not adhering to a specific alimentary regimen.

We need to highlight that CART automatically based one of the major splits on BMI, detecting 30 kg/m² as a splitting value. This evidenced the role of Obesity in CKD progression, well-known in the literature and detected by the ML algorithm.^{21–24}

Conclusion

In conclusion, considering as a major topic of our paper the comparison between statistical methods, we can conclude that supervised machine learning approaches such as CART provide a detailed summary of the impact of variables on the outcome that is consistent with traditional statistical methods. Furthermore, the model is visually easier to understand and better explained by using a cross-validated cut-off that allows the model to be more specific. Last but not least, the automatic stepwise selection of the variable used to split the groups leads to faster and more effortless procedures, giving us more strength models.

Support and financial disclosure

Authors declare any interest, financial support or relationship that may pose a conflict of interest.

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

- Yen CL, Fan PC, Kuo G, Chen CY, Cheng YL, Hsu HH, et al. Supplemented low-protein diet may delay the need for preemptive kidney transplantation: a nationwide population-based cohort study. *Nutrients*. 2021;13:3002, <http://dx.doi.org/10.3390/nu13093002>. Published 2021 Aug 28.
- Kelly JT, Su G, Zhang L, Qin X, Marshall S, González-Ortiz A, et al. Modifiable lifestyle factors for primary prevention of CKD: a systematic review and meta-analysis. *J Am Soc Nephrol*. 2021;32:239–53, <http://dx.doi.org/10.1681/ASN.2020030384>.
- Calabrese V, Gembillo G, Buda M, Cernaro V, Longhitano E, Pontoriero A, et al. Review and practical excursus of the propensity score: low protein diet compared to Mediterranean diet in patients with chronic kidney disease. *G Ital Nefrol*. 2023;40:2023. Published 2023 Dec 22.
- Obermeyer Z, Emanuel EJ. Predicting the future – big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216–9, <http://dx.doi.org/10.1056/NEJMp1606181>.
- James G, Witten D, Hastie T, Tibshirani R. Tree-based methods; 2013, http://dx.doi.org/10.1007/978-1-4614-7138-7_8.
- Hsich E, Gorodeski EZ, Blackstone EH, Ishwaran H, Lauer MS. Identifying important risk factors for survival in patient with systolic heart failure using random survival forests. *Circ Cardiovasc Qual Outcomes*. 2011;4:39–45, <http://dx.doi.org/10.1161/CIRCOUTCOMES.110.939371>.
- Henrard S, Speybroeck N, Hermans C. Classification and regression tree analysis vs. multivariable linear and logistic regression methods as statistical tools for studying haemophilia. *Haemophilia*. 2015;21:715–22, <http://dx.doi.org/10.1111/hae.12778>.
- Schober P, Vetter TR. Linear regression in medical research. *Anesth Analg*. 2021;132:108–9, <http://dx.doi.org/10.1213/ANE.0000000000005206>.
- Provenzano F, Versace MC, Tripepi R, Zoccali C, Tripepi G. Il confondimento negli studi epidemiologici [Confounding in epidemiology]. *G Ital Nefrol*. 2010;27:664–7.
- Calabrese V, Labbozzetta V, Lipari A, Labbozzetta V, Gembillo G, Cernaro V, et al. Comparison between Creatinine Clearance and eGFR_{crea}: a real-life experience. *G Ital Nefrol*. 2022;39:2022. PMID: 36073332.
- Sykes M, Thomson R, Kolehmainen N, Allan L, Finch T. Opportunities to enhance ward audit: a multi-site qualitative study. *BMC Health Serv Res*. 2021;21:226, <http://dx.doi.org/10.1186/s12913-021-06239-0>. PMID: 33712006; PMCID: PMC7971099.
- Spencer E, Walshe K. National quality improvement policies and strategies in European healthcare systems. *Qual Saf Health Care*. 2009;18 Suppl. 1:i22–7, <http://dx.doi.org/10.1136/qshc.2008.029355>. PMID: 19188457; PMCID: PMC2629881.
- Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA). 2018. p. 80–9.
- Kuhn M, Johnson K. Applied predictive modeling. New York: Springer; 2013, <http://dx.doi.org/10.1007/978-1-4614-6849-3>.
- Yen CL, Fan PC, Kuo G, Chen CY, Cheng YL, Hsu HH, et al. Supplemented low-protein diet may delay the need for preemptive kidney transplantation: a nationwide population-based cohort study. *Nutrients*. 2021;13:3002, <http://dx.doi.org/10.3390/nu13093002>. Published 2021 Aug 28.
- Calabrese V, Cernaro V, Battaglia V, Gembillo G, Longhitano E, Siligato R, et al. Correlation between hyperkalemia and the duration of several hospitalizations in patients with chronic kidney disease. *J Clin Med*. 2022;11:244, <http://dx.doi.org/10.3390/jcm11010244>. Published 2022 Jan 4.
- Sprague SM, Martin KJ, Coyne DW. Phosphate balance and CKD-mineral bone disease. *Kidney Int Rep*. 2021;6:2049–58, <http://dx.doi.org/10.1016/j.ekir.2021.05.012>. Published 2021 May 17.
- Angeloco LRN, Arces de Souza GC, Romão EA, Frassetto L, Chiarello PG. Association of dietary acid load with serum bicarbonate in chronic kidney disease (CKD) patients. *Eur J Clin Nutr*. 2020;74 Suppl. 1:69–75, <http://dx.doi.org/10.1038/s41430-020-0689-1>.
- Liu AYL, Low S, Yeoh E, Lim EK, Renaud CJ, Teoh STY, et al. A real-world study on SGLT2 inhibitors and diabetic kidney disease progression. *Clin Kidney J*. 2022;15:1403–14, <http://dx.doi.org/10.1093/ckj/sfac044>. Published 2022 Feb 16.
- Levine ME, Suarez JA, Brandhorst S, Balasubramanian P, Cheng CW, Madia F, et al. Low protein intake is associated with a major reduction in IGF-1, cancer, and overall mortality in the 65 and younger but not older population. *Cell Metab*. 2014;19:407–17, <http://dx.doi.org/10.1016/j.cmet.2014.02.006>.
- Jiang Z, Wang Y, Zhao X, Cui H, Han M, Ren X, et al. Obesity and chronic kidney disease. *Am J Physiol Endocrinol Metab*. 2023;324:E24–41, <http://dx.doi.org/10.1152/ajpendo.00179.2022>.
- Kjaergaard AD, Teumer A, Witte DR, Stanzick KJ, Winkler TW, Burgess S, et al. Obesity and kidney function: a two-sample Mendelian randomization study. *Clin Chem*. 2022;68:461–72, <http://dx.doi.org/10.1093/clinchem/hvab249>.
- Chintam K, Chang AR. Strategies to treat obesity in patients with CKD. *Am J Kidney Dis*. 2021;77:427–39, <http://dx.doi.org/10.1053/j.ajkd.2020.08.016>.
- Pérez-Torres A, Caverni-Muñoz A, González García E. Mediterranean diet and chronic kidney disease (CKD): a practical approach. *Nutrients*. 2022;15:97, <http://dx.doi.org/10.3390/nu15010097>. Published 2022 Dec 25.