5. Perrone RD, Abebe KZ, Watnick TJ, et al. Primary results of the randomized trial of metformin administration in polycystic kidney disease (TAME PKD). *Kidney Int*. 2021;100:684–696.
6. Brosnahan GM, Wang W, Gitomer B, et al. Metformin therapy in autosomal dominant polycystic kidney disease: a feasibility study. *Am J Kidney Dis*. 2022;79:518–526.

Olivier Devuyst[1,2], Amy Earley[3] and Vicente E. Torres[4]
[1]*Institute of Physiology, University of Zurich, Zurich, Switzerland;* [2]*UCLouvain Division of Nephrology, UCLouvain Medical School, Brussels, Belgium;* [3]*Kidney Disease: Improving Global Outcomes (KDIGO), Brussels, Belgium; and* [4]*Divison of Nephrology and Hypertension, Mayo Clinic, Rochester, Minnesota, USA*

Correspondence: *Olivier Devuyst, Cliniques Universitaires Saint-Luc, UCLouvain Medical School, 10 Avenue Hippocrate, B-1200 Brussels, Belgium. E-mail: olivier.devuyst@saintluc.uclouvain.be or olivier.devuyst@uzh.ch*

Kidney International (2025) **108,** 150–151; https://doi.org/10.1016/j.kint.2025.02.012

Reevaluating principal component analysis: advocating for nonlinear and nonparametric methods in multi-omics data analysis

To the editor: In multi-omics research, accurate results depend on methods that capture complex, nonlinear interactions among high-dimensional data from genomic, transcriptomic, proteomic, and metabolomic sources. Traditional linear models, such as principal component analysis, are widely used for feature reduction and visualization; however, their reliance on linear and parametric assumptions may lead to misleading conclusions in nonlinear proteomic analyses. For instance, although principal component analysis successfully highlighted circadian gene expression patterns in kidney studies,[1] it can oversimplify multifaceted biological interactions that better align with nonlinear behavior.

To address these challenges, nonlinear and nonparametric robust statistical methods are recommended. Techniques like Spearman rank correlation with P values and Kendall τ with P values can uncover associations without assuming data normality, making them better suited for complex, heterogeneous data sets.[2,3] These methods assess the strength and direction of relationships using data ranks rather than raw values, thus capturing nuances that principal component analysis might miss.

Additionally, prior assessment of feature collinearity through measures like the variance inflation factor is essential to avoid distortion of analytical outcomes. Proper collinearity diagnostics help ensure that each variable contributes uniquely, enhancing the interpretability and reliability of the analysis.[4]

Ultimately, combining principal component analysis with nonlinear, nonparametric approaches provides a more comprehensive understanding of the intricate molecular networks underlying nonlinear health and disease, leading to more robust and accurate conclusions.

DISCLOSURE

The author declared no competing interests.

AUTHOR CONTRIBUTIONS

YT completed this research and wrote this article.

1. Preston R, Chrisp R, Dudek M, et al. The glomerular circadian clock temporally regulates basement membrane dynamics and the podocyte glucocorticoid response. *Kidney Int*. 2025;107:99–115.
2. Shen X, Wang C, Zhou X, et al. Nonlinear dynamics of multi-omics profiles during human aging. *Nat Aging*. 2024;4:1619–1634.
3. Pattaroni C, Begka C, Cardwell B, et al. Multi-omics integration reveals a nonlinear signature that precedes progression of lung fibrosis. *Clin Transl Immunol*. 2024;13:e1485.
4. Jacob J, Varadharajan R. Robust variance inflation factor: a promising approach for collinearity diagnostics in the presence of outliers. *Sankhya B*. 2024;86:845–871.

Yoshiyasu Takefuji[1]
[1]*Faculty of Data Science, Musashino University, Tokyo, Japan*

Correspondence: *Yoshiyasu Takefuji, Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan. E-mail: takefuji@keio.jp*

Kidney International (2025) **108,** 151; https://doi.org/10.1016/j.kint.2025.03.002

The authors reply: As the author highlights in their recent letter to the editor,[1] principal components analysis (PCA) is widely used for feature reduction and visualization of complex data sets. To achieve this, through a series of linear transformations, PCA reduces the dimensionality of data, while retaining most of the variability, to obtain new variables (or components) to explain variation within data. This was indeed the purpose of the PCAs employed in our study[2]: to demonstrate separation of the samples based on time of day. While we accept that such linear and parametric assumptions may lead to misleading conclusions in nonlinear omics' data sets, the conclusion that we have successfully demonstrated is to be expected, as the author kindly points out. As circadian rhythm is typically the major source of variance in circadian data sets, this approach is especially appropriate for the extraction of rhythmic patterns of gene and protein expression. That is, if we assume that the main source of variability is from the biological condition of interest (time of day), we expect to see samples clustering accordingly. As recently demonstrated, although alternative methods, such as linear discriminant analysis, may be better able to determine the contribution of individual rhythmic genes, PCA identifies