Proceedings of the Nutrition Society (2024), **83 (OCE4)**, E335 doi:10.1017/S0029665124005731

Nutrition Society Congress 2024, 2–5 July 2024

Unsupervised machine learning to determine dietary protein distribution among New Zealand vegans across time periods of the day

Bi Xue Patricia Soh[1], Matthieu Vignes[2], Nick W Smith[1], Pamela R. von Hurst[3] and Warren C. McNabb[1]

[1]*Sustainable Nutrition Initiative, Riddet Institute, Massey University, Palmerston North, New Zealand*
[2]*School of Mathematical and Computational Sciences, Massey University, Palmerston North, New Zealand*
[3]*School of Sport Exercise and Nutrition, College of Health, Massey University, Auckland, New Zealand*

Excluding all animal-sourced foods may be associated with increased risks of nutrient deficiencies. As indispensable amino acids (IAAs) cannot be stored or endogenously produced, consistent protein consumption throughout the day is important to improve protein quality for optimal metabolic function[(1)]. Assessment of protein adequacy needs to be undertaken at the meal rather than daily intake level because food combinations within each meal can be complementary and influence the overall amino acid profile of the meal[(2)].

Outcomes of our previous review found that among plant-sourced foods, soy, legumes, nuts and seeds provide greater protein content and quality[(3)]. We hypothesise that variation in protein intake will exist both between vegan individuals and between observation days for the same individuals. Previous investigations of the relationship between meals and nutrient intake based on specified time windows for eating may have been subject to researcher bias in the definition of these windows. The main outcome of this study is to utilise time series clustering to determine the impact of dietary patterns on protein distribution, across the day.

Intake data was obtained using a four-day food diary from a cross-sectional survey of 193 New Zealand vegans (Ethical approval: HDEC 2022 EXP 12312). The inclusion criteria required participants to have followed an exclusive vegan diet for at least two years. A kernel density contour estimation was used to visualise protein distribution across eating occasions for all participants over four days. Dynamic Time Warping (DTW) was then used to align two temporal sequences (time series) to compute an output of distance[(4)] which was used for hierarchical clustering using the Ward.D2 method. An optimal cluster of 3 was identified using silhouette coefficient and domain knowledge.

Participants had a mean age of 39.4 years (SD = 12.3), with 90.1% having attained a tertiary-level education or higher. Overall, mean protein intake was 1.11 g/kg/d (SD = 0.39), with 8.29% of participants below the Estimated Average Requirements (EAR) and 24.3% of participants below the Recommended Dietary Allowance (RDA) for adults. The mean Acceptable Macronutrient Distribution Range (AMDR) for protein is 15.5% (SD = 4.16), with 96.9% of participants within the recommended AMDR range (10-35%). Peak protein consumption was observed at 1230 and 1900. Sequential colour scale representing density found higher distribution of data points representing protein intake of less than 10g per eating occasion. Time series similar in shape and amplitude were assigned to the same cluster. Preliminary findings identified three different protein intake profiles across the day.

A small percentage of participants has protein intake below the daily requirements for adults. More occasions with lower protein intake per eating moment was observed. This approach classifies dietary patterns objectively for analysis of daily protein and IAA intake.

References

1. Boye J *et al.* (2012) *British Journal of Nutrition*, **108**, 183–211.
2. Heerschop SN *et al.* (2023) *European Journal of Nutrition*, **62**, 2115–2128.
3. Soh BXP *et al.* (2023) *The Journal of Nutrition*, **154**, 300–313.
4. Giorgino T (2009) *Journal of Statistical Software*, **31**(7), 1–24.