Graphical models of the lipidome for health status and nutritional biomarker discovery

Keywords: Biostatistics, Network Modelling, Mixture (Compositional) Modelling, Plate Notation, Machine Learning, Nutritional Epidemiology

Supervisors:
Dr James Smith (School of Food Science & Nutrition) and
Dr Jianhua Wu (Dental Translational Clinical Research Unit)

Collaborators:
Dr Lisa Marshall (School of Food Science & Nutrition) and
Dr Michael Zulyniak (School of Food Science & Nutrition)

In this work, we explore how combinations of fats, typically observed in lipidomics data, can be used as novel *biomarkers* in health and disease and for nutritional status **[1, 2]**. Fats are the building blocks of lipids. Singlet, duplet and triplet combinations of fats form individual lipids that contribute to pools. The mathematical question to be answered is how to formulate a mixture model that best represents the explicit contributions of lipid subpopulations within observed lipid pools. Only the building blocks (fats) and their final lipid pools are observable. Unfortunately, the subpopulations of the duplet and triplet fat combinations are unobservable and must be estimated. One approach is to consider the mixture model as a *composition model* **[3]** where the fractional contributions of the fats are measured from one analytical chemistry technique (FAME) and the fractional contributions of the lipid pools are measured from another technique (MS). *Plate descriptions* are a useful notation for describing mixture models and the hope is that the hierarchical description of the fats within the lipids within the lipid pools could be described as a multivariate Gaussian compositional mixture model **[3]**.

This project aims to develop a useful working definition for the *lipidome* **[4,5]** currently loosely defined as an individual's typical lipid **[6,7]** profile. This work will use models of directed acyclic graphs (exposure-outcome causal graphs) to illustrate how fat components combine into lipids and how the lipids, sharing equivalent configurations, form the lipid pools. *Exposure-outcome causal graphical models* provide a useful formalism for describing the combinations of the components. Graphical models illustrate different assumptions about population data; depending on their detail, the models can be used as *sufficient-component cause models* that illustrate specific hypotheses about mechanisms of action or *potential-outcome models* and *structural-equations models (SEMs)* that provide a basis for quantitative analysis of effects **[8]**. Using graphical modelling to predict reliable diagnostic or status biomarkers based on the latent (hidden) distribution structure of fat and lipid combinations would be game-changing.

This 10-week project will use prepared signal data of lipidomics profiles. The first objective is to build basic (toy) models in R (or an equivalent programming environment) that can be scaled up to demonstrate pair-wise combinations of fat mixtures forming the lipids in the lipid pools that reflect the lipidomics data.

This research project is cutting-edge and the hope is that student will be interested in presenting their work at a conference and publishing.

Criteria: Applicants should be interested in programming and in statistical modelling, although training will be provided.

References:

1. Combs, G. F., Jr., Trumbo, P. R., McKinley, M. C., Milner, J., Studenski, S., Kimura, T., et al. (2013). Biomarkers in nutrition: new frontiers in research and application. *Annals of the New York Academy of Sciences*, **1278(1)**, 1–10. doi:10.1111/nyas.12069

2. Hodge, A. M., Simpson, J. A., Gibson, R. A., Sinclair, A. J., Makrides, M., O'Dea, K., et al. (2007). Plasma phospholipid fatty acid composition as a biomarker of habitual dietary fat intake in an ethnically diverse cohort. *Nutrition, Metabolism and Cardiovascular Diseases*, **17(6)**, 415–426. doi:10.1016/j.numecd.2006.04.005

3. https://en.wikipedia.org/wiki/Mixture_model

4. https://en.wikipedia.org/wiki/Lipidome

5. Quehenberger, O. & Dennis, E. A. (2011). The Human Plasma Lipidome, *New England Journal of Medicine,* **365(10)**,1812-1823. doi: 10.1056/NEJMra1104901

6. Calder, P.C. (2015). Functional Roles of Fatty Acids and Their Effects on Human Health. *Journal of Parenteral and Enteral Nutrition*, **39(1)** Supplementary 18S-32S doi: 10.1177/0148607115595980

7. Rustan, A. C. & Drevon, C. A. (2001). *Fatty Acids: Structures and Properties* (pp. 1–7). Chichester: John Wiley & Sons, Ltd. doi: 10.1038/npg.els.0003894

8. Greenland, S., & Brumback, B. (2002). An overview of relations among causal modelling methods. *International Journal of Epidemiology*, **31(5)**, 1030–1037.