Sanquin

Laboratory of Proteomics, Sanquin Research, Amsterdam

Time series clustering of integrated clinical and proteomic data

Blood plasma is frequently used for biomarker studies and to monitor patient health and diseases status over time. Unbiased plasma proteomics facilitates this through the formation of new hypotheses about disease mechanisms and through the identification of new diagnostic targets. Using current discovery driven mass spectrometry-based proteomics we can detect > 400 proteins per plasma sample.

In this project we aim to apply this method on a cohort of neonates to investigate protein trajectories during development and the course of disease. Integration of such protein levels with clinical data is highly called for and represent an unmet clinical need for monitoring and predicting disease progression at an individual patient level. However, identifying proteins with similar trends over time and assigning these trends to specific patients of development or more importantly clinical observations is challenging.

In this project you will evaluate various algorithms that are suitable for time-series clustering of both clinical and plasma proteomic datasets. The most suitable algorithm will be applied to delineate proteins that associates with short-term and long-term clinical outcomes in preterm neonates.

University students with an interest in bioinformatics and an affinity for python and/or R who are looking for a dynamic and interesting internship are encouraged to contact Eva Smit: e.smit@sanquin.nl or Maartje van den Biggelaar: m.vandenbiggelaar@sanquin.nl