PCA to reveal non-adherence and non-analytical outliers using global metabolomics and lipidomics in clinical settings

Barbora Pisklakova[1,2], Ales Kvasnicka[1,2], Sander Johannes Thorbjørnsen Guttorm[1,2], Anja Østeby Vassli[1,2], Helge Rootwelt[1,2], Katja Benedikte Prestø Elgstøen[1,2] ✉

[1]Section for Metabolomics and Lipidomics, Department of Medical Biochemistry, Division ofLaboratory Medicine, Oslo University Hospital, Oslo, Norway
[2]Core Facility for Global Metabolomics and Lipidomics, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway

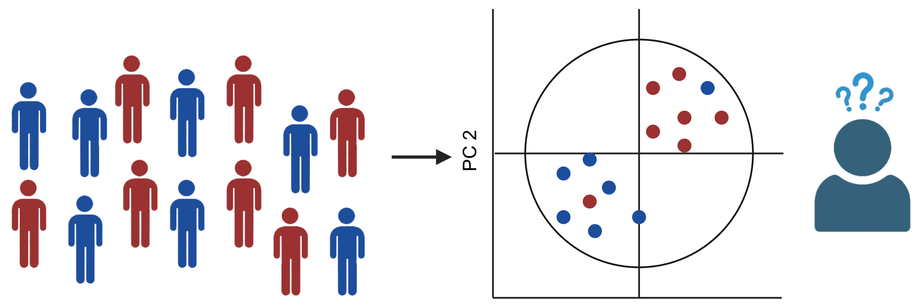
✉ kelgstoe@ous-hf.no
Website: https://www.ous-research.no/home/elgstoen/home

Section for MetabolOmics and Lipidomics
OSLO UNIVERSITY HOSPITAL

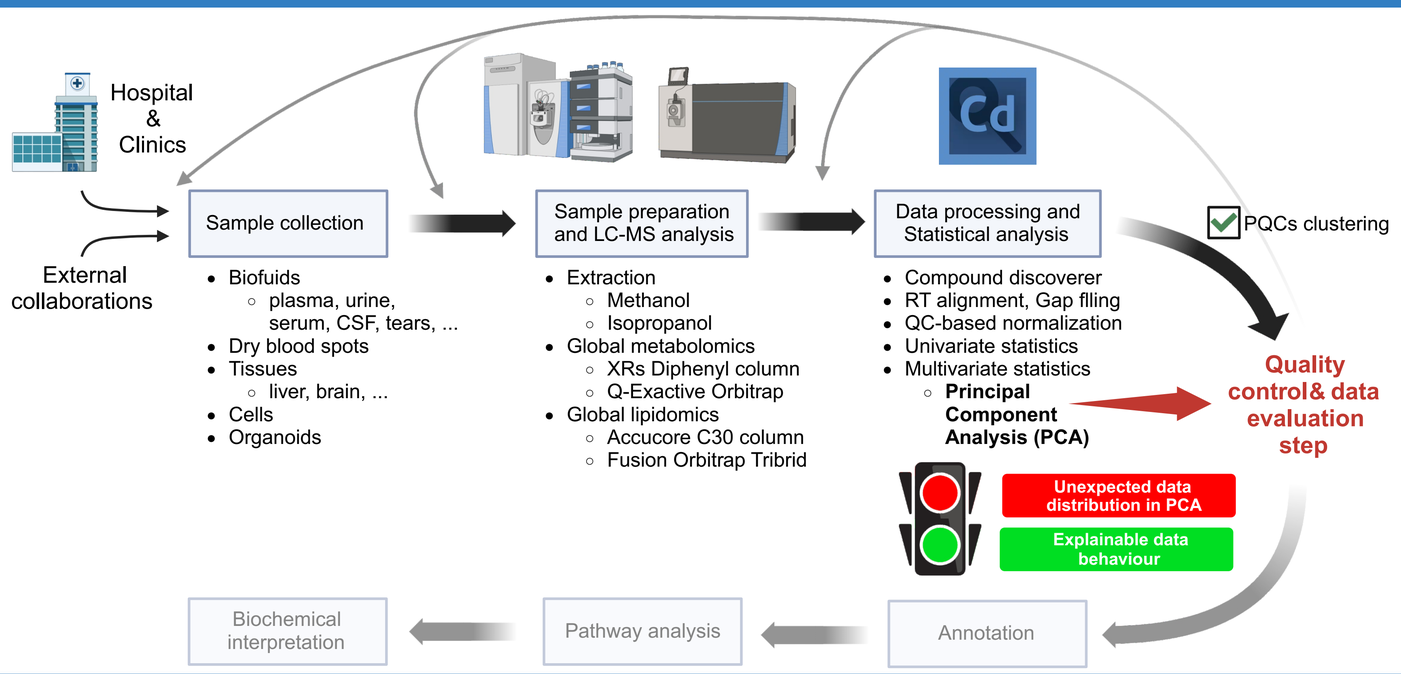
UNIVERSITETET I OSLO

INTRODUCTION

Non-adherence is common in clinical settings, leading to loss oftreatment response. In clinical practice this will often prompt institution ofnew and often inferior treatments with higher risk of complications. In aclinical research setting, the results will be biased and not based on realclinical effects but on the degree of participant adherence. Methods ofdetecting non-adherence are often insufficient, and since adherence is usuallyexpected and based on a signed agreement, non-adherence is usually not detected.We here aim to reveal the extent of non-adherence in clinical settings,and how the utilization of global metabolomics and lipidomics can detectnon-adherence at an early stage. Early detection provides an opportunity tocorrect and reinstate adherence and effective patient treatment, and to enableexclusion of study participants and samples that would otherwise contribute toloss of detection of significant findings and improper conclusions.



METHODS



Sample preparation: Serum or plasma metabolites and lipids were extracted by adding MeOH or IPA, respectively. Samples were then mixed before centrifugation (10 min, 4 °C, 21 100 RCF), supernatant aliquots were mixed to create the pooled quality control (PQC). Remaining supernatant was placed in HPLC vials for analysis. Dried blood spots were prepared as 3.2 mm punches with 100 µL of 80% MeOH with 0.1% formic acid and mixed for 45 min (at 45 °C, 700 rpm). Samples were transferred to an HPLC vial for analysis directly after extraction.

Global LC-MS Metabolomics and Lipidomics Analysis: Metabolites were separated by XRs Diphenyl column and analyzed using the Ultimate 3000 HPLC coupled to Q-Exactive Orbitrap MS (Thermo Scientific). Lipids were separated and analysed using Accucore C30 HPLC column and Vanquish Horizon UHPLC coupled to Fusion Orbitrap Tribrid.

Data processing and statistical analysis: Compound Discoverer (3.3, Thermo Scientific) was used for peak integration, alignment, SERRF QC correction, univariate and multivariate statistics and data presentation.

RESULTS

Our findings highlight the importance of evaluating metabolomics and lipidomics data from clinical studies using unsupervised multivariate statistical analysis, such as Principal Component Analysis (PCA). After evaluating PQC samples and ruling out analytical errors, outlier samples can sometimes be observed. Outlier samples or those clustering with a different experimental group or sample type may highlight potential errors. This poster presents clinical experience-based cases demonstrating how global metabolomics and lipidomics analyses can identify non-adherent patients, inconsistencies in sample types, and other anomalies using multivariate statistical tools. In the first case (**Fig. 1**), we identified an incorrect sample type (heparin plasma) that differed from the rest of the study samples (EDTA plasma). Hence, the detection of EDTA in plasma samples ensured the consistency of sample materials. Comparing EDTA peaks across samples provided further insights into proper sample collection practices. Another bright example shows the detected non-adherence to the dietary regimen of the study (**Fig. 2 and 3**) and deviations from the fasting schedule (**Fig. 5**). Lastly, we uncovered inaccurate clinical information for one patient, who was incorrectly reported as having an infection, unlike the other patients in the study (**Fig. 4**).

Case 1: Inconsistency in the sample type

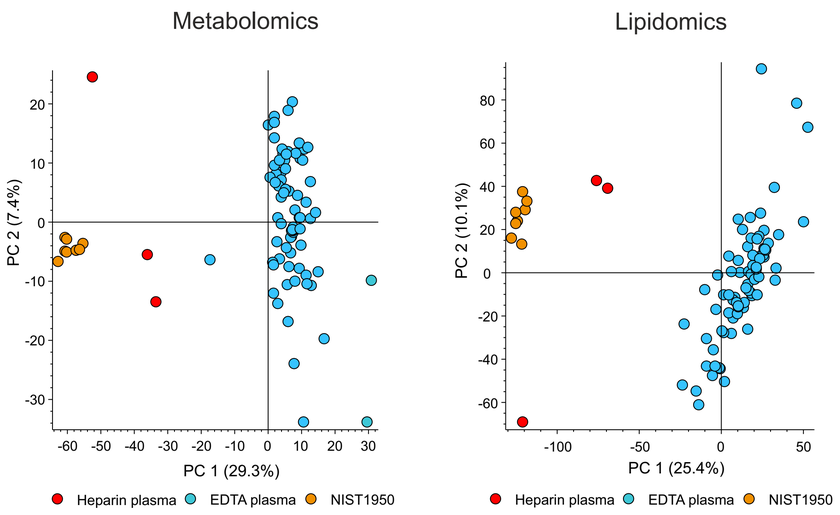


Fig. 1. PCA plot of metabolomic and lipidomic analysis of plasma samples and NIST1950 heparin plasma with noticeable clustering of three samples (in red). Further investigation showed that the clustering of these samples away from rest of the samples is not due to analytical error but to the fact that these samples are not EDTA plasma but heparin plasma and therefore cluster towards heparin NIST1950 plasma.

Case 2: Non-adherence with the prescribed diet

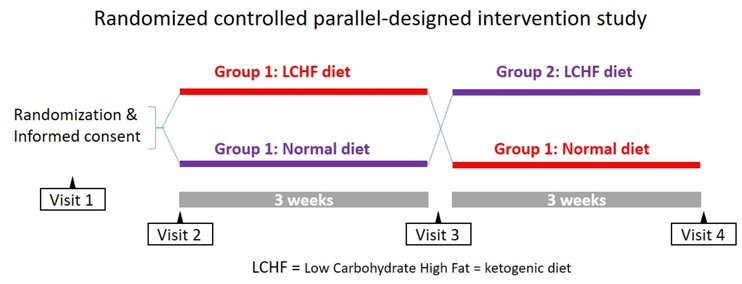


Fig. 2. Schematic overview of the study design for the metabolomics study of the ketogenic and normal diet. All subjects were sampled and divided into two groups according to the following diet: either normal diet with no restrictions or the low-carbohydrate high-fat diet. After three weeks samples were collected. After a week wash-out period, individuals then switched the diets and maintained it for three weeks. After this period the individuals were sampled and the effect of the diet on the metabolism was then studied.

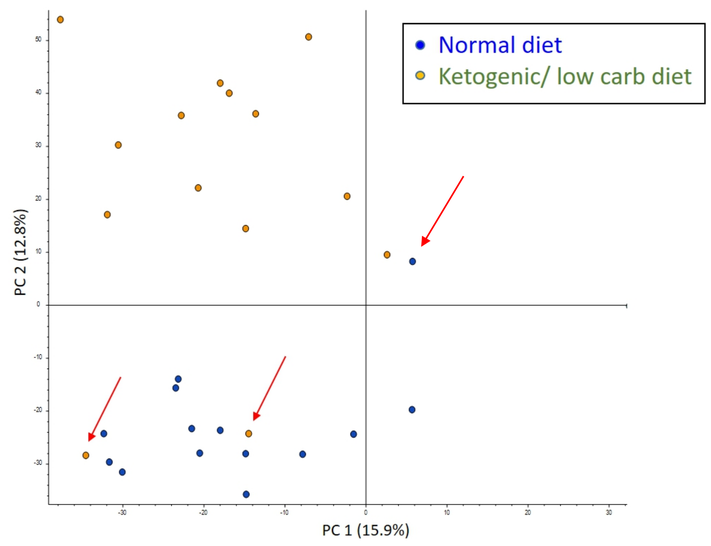


Fig. 3. PCA of metabolomic analysis of individuals on ketogenic or normal diet. Three samples show unexplained clustering together with the comparison group, later discovered to be due to dietary non-adherance.

Case 3: Incorrect clinical information about a patient

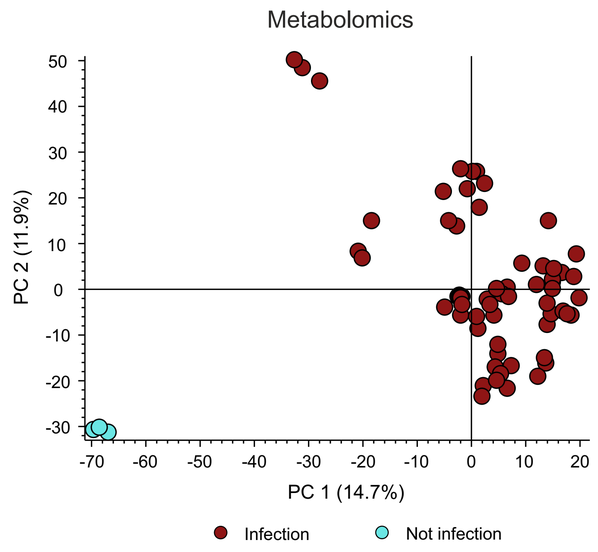


Fig. 4. PCA plot of metabolomic analysis of EDTA plasma samples from patients with an infection. Three samples from one patient (in blue) cluster disticntly apart from the remaining samples. After further investigation, it was revealed that this patient did not have an infection and was thus incorrectly assigned to this group.

Case 4: Non-adherence with the fasting schedule

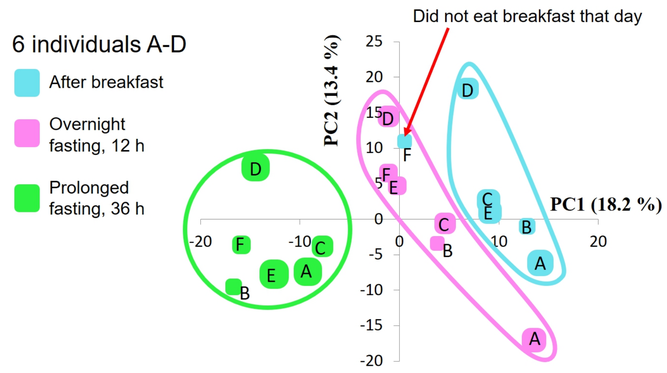


Fig. 5. PCA of metabolomic analysis of DBS samples from individuals with different time of fasting (devided into 3 groups). One sample from the group of regular eating is clustered with the group of overnight fasting. Further investigation revealed non-adherence of this individual who did not follow the diet. From Skogvold *et al.* 2021.[1]

CONCLUSION

Our findings demonstrate the crucial role of multivariate statistical analysis, such as PCA, in ensuring data quality and integrity in clinical metabolomics and lipidomics studies. By identifying sample inconsistencies, patient non-adherence, and misleading clinical information, these tools enhance the reliability of study outcomes, can elimintate bias and provide valuable insights for improving study design and execution.

References:
[1] J. Proteome Res. 2021, 20, 8, 4010–4021