Distinct metabolic phenotypes in adolescents with obesity identified by unsupervised learning: associations with insulin resistance and resting energy expenditure

A. Sonza, S. Tamini, A. Bondesan, D. Caroli, L. Abbruzzese, A. Sartorio

Frontiers in Endocrinology 17, 2026

Background: pediatric obesity is characterized by substantial metabolic heterogeneity, including adiposity, insulin resistance, inflammation, and energy metabolism. Traditional categorical classifications may inadequately capture this complexity.

Objective: to identify clinically meaningful metabolic phenotypes using unsupervised clustering and to compare the statistical performance and clinical interpretability of two- and three-cluster solutions.

Methods: in this cross-sectional study, 758 children and adolescents with obesity underwent anthropometric, biochemical, and indirect calorimetry assessments. Fat-free mass was estimated using validated bioimpedance-based equations, and fat-free mass index (FFMI) was standardized before clustering. K-means clustering (k = 2 and k = 3) was performed using FFMI z-score, BMI-SDS, HOMA-IR, triglycerides, and HDL-cholesterol. Cluster validity was assessed using the Silhouette, Elbow, Davies–Bouldin index, Calinski–Harabasz index, and Gap statistic methods. Validation employed variables not included in clustering, such as glucose-insulin dynamics, inflammatory and hepatic markers, blood pressure, respiratory quotient, and resting energy expenditure (REE). Group comparisons were performed using Kruskal–Wallis and Dunn's post-hoc tests with false discovery rate correction.

Results: both clustering solutions demonstrated significant separation across metabolic domains. The two-cluster model identified metabolically favourable and unfavourable phenotypes differing in adiposity, insulin resistance, inflammatory burden, hepatic markers, blood pressure, and REE. The three cluster solution revealed a more granular metabolic stratification with an intermediate phenotype characterized by partial metabolic impairment. Validation confirmed robust differences across physiological variables (all FDR adjusted $p < 0.05$). Although the two-cluster solution showed slightly superiorinternal validity, the three-cluster model provided greater clinical granularity and phenotypic resolution.

Conclusions: unsupervised clustering identified biologically coherent metabolic phenotypes in pediatric obesity. While the two-cluster solution provided statistically robust stratification, the three-cluster configuration captured a broader spectrum of metabolic heterogeneity and risk profiles, supporting the application of higher-resolution phenotyping approaches in pediatric endocrinology and metabolic risk assessment.