Raw metabolite integrals were normalized to the median value (median value = 1) of each metabolite for a given run day. Metabolite identification was performed using automated comparison to features (including retention time, *m/z*, preferred adducts, in-source fragments, and mass spectra) of ~8,000 purified chemical standards in the Metabolon reference library (Dehaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. Journal of Cheminformatics. 2010;2(1):1-12.). Metabolites that remained unknown, but are reproducibly captured, after comparison with standards were annotated with a unique and persistent "X-number" for potential future identification.

Abundances for a total of 1,546 grouped and aligned metabolites were reported by the Metabolon platform. Metabolites annotated as xenobiotics and with ≥ 75% missing values were removed, yielding 1,108 metabolites in the final dataset. Remaining missing values were imputed as the half-minimum intensity for a given metabolite. Metabolite intensities were log-transformed and *pareto*-scaled for use in regression models.