

# Supplementary Figures for Revealing Subject-specific Temporal Patterns from Longitudinal Data

Christos Chatzis<sup>1,2\*</sup>, David Horner<sup>3</sup>, Rasmus Bro<sup>4</sup>, Ann-Marie Malby Schoos<sup>3,5,6,7</sup>, Morten A. Rasmussen<sup>3,4</sup>, and Evrim Acar<sup>1\*</sup>

<sup>1</sup>Department of Data Science and Knowledge Discovery, Simula Metropolitan Center for Digital Engineering, Oslo, Norway

<sup>2</sup>Faculty of Technology, Art and Design, Oslo Metropolitan University, Oslo, Norway

<sup>3</sup>COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark

<sup>4</sup>Department of Food Science, University of Copenhagen, Copenhagen, Denmark

<sup>5</sup>Department of Pediatrics, Copenhagen University Hospital - Næstved, Slagelse and Ringsted, Slagelse, Denmark

<sup>6</sup>Department of Pediatrics, Copenhagen University Hospital - Amager and Hvidovre, Hvidovre, Denmark

<sup>7</sup>Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

\*Corresponding authors: christos@simula.no (C.C.) and evrim@simula.no (E.A.).

## 1 Metabolomics

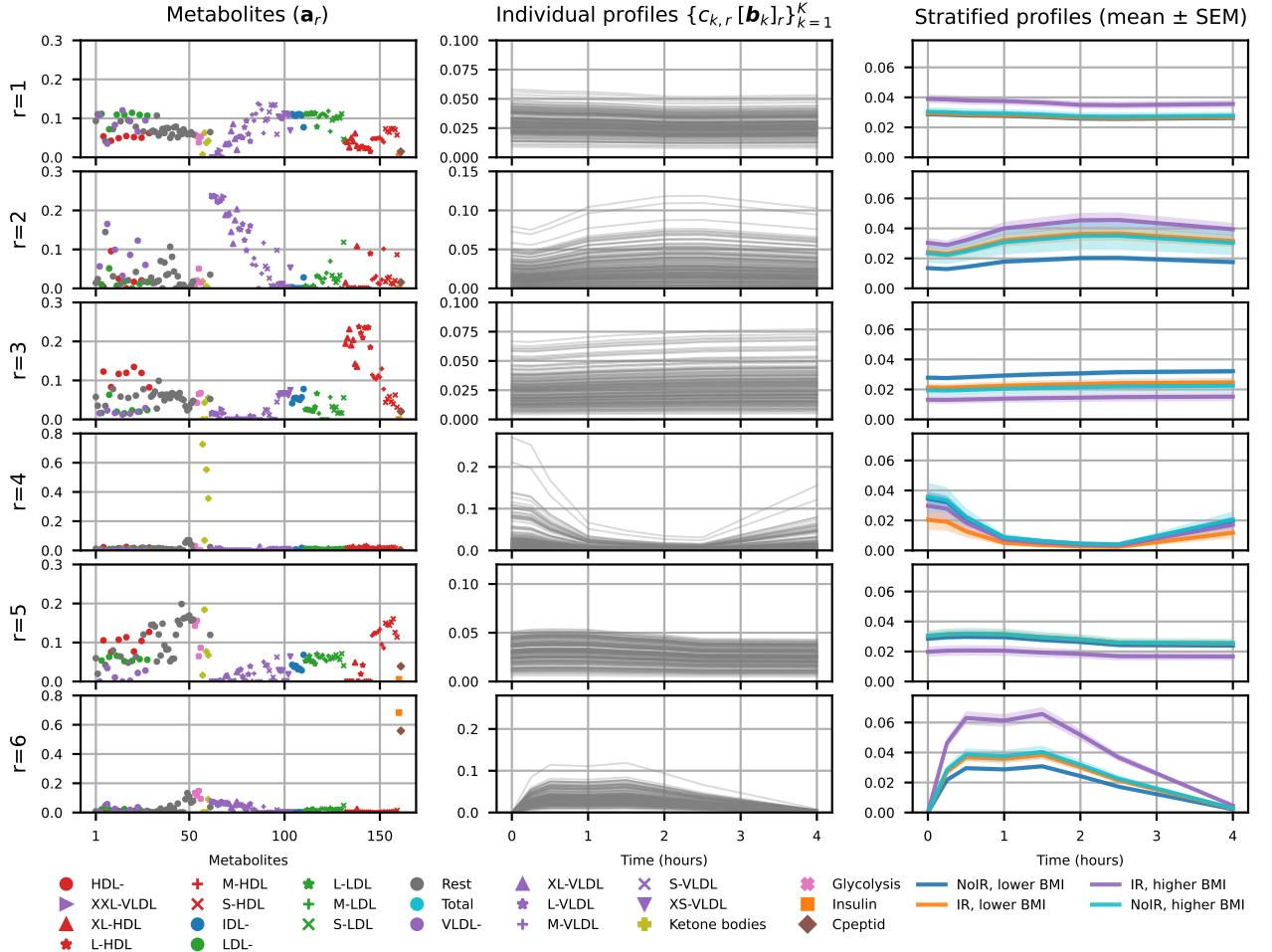


Figure S.1: Metabolomics. Components of a 6-component CP model (with nonnegativity constraints in all modes) of the metabolomics data from males.  $\mathbf{a}_r$  denotes the pattern in the metabolites mode, where metabolites are colored by lipoprotein classes. Different shapes are used for lipoprotein subclasses. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  for each component are shown in the middle column. The last column shows scaled subject-specific time profiles colored according to four BMI/IR groups.

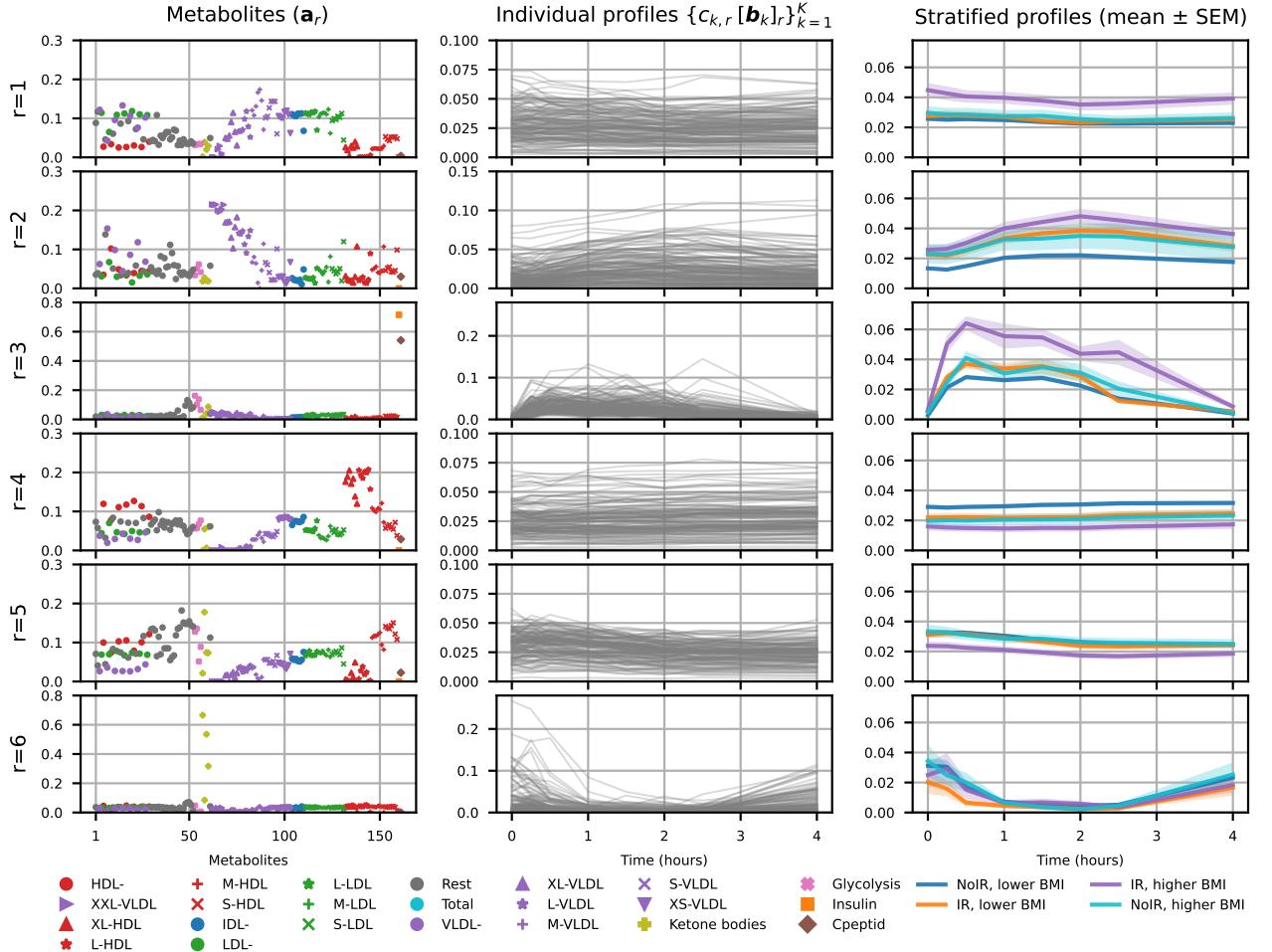


Figure S.2: Metabolomics. Components of a 6-component PARAFAC2 model (with nonnegativity constraints in all modes) of the metabolomics data from males.  $\mathbf{a}_r$  denotes the pattern in the metabolites mode, where metabolites are colored by lipoprotein classes. Different shapes are used for lipoprotein subclasses. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  for each component are shown in the middle column. The last column shows scaled subject-specific time profiles colored according to four BMI/IR groups.

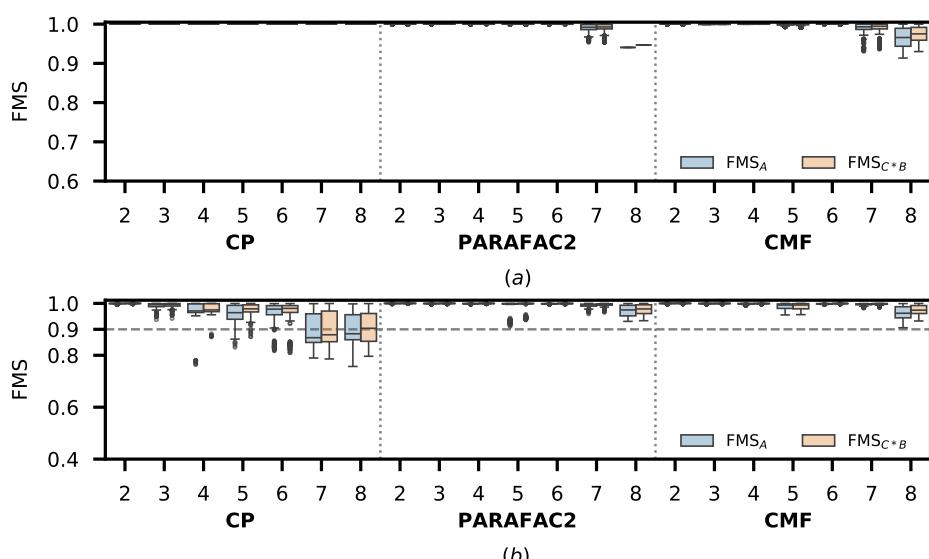


Figure S.3: (a) Reproducibility, and (b) Replicability of different models of the metabolomics data using different number of components ( $R$ ) for the subset of the female subjects.

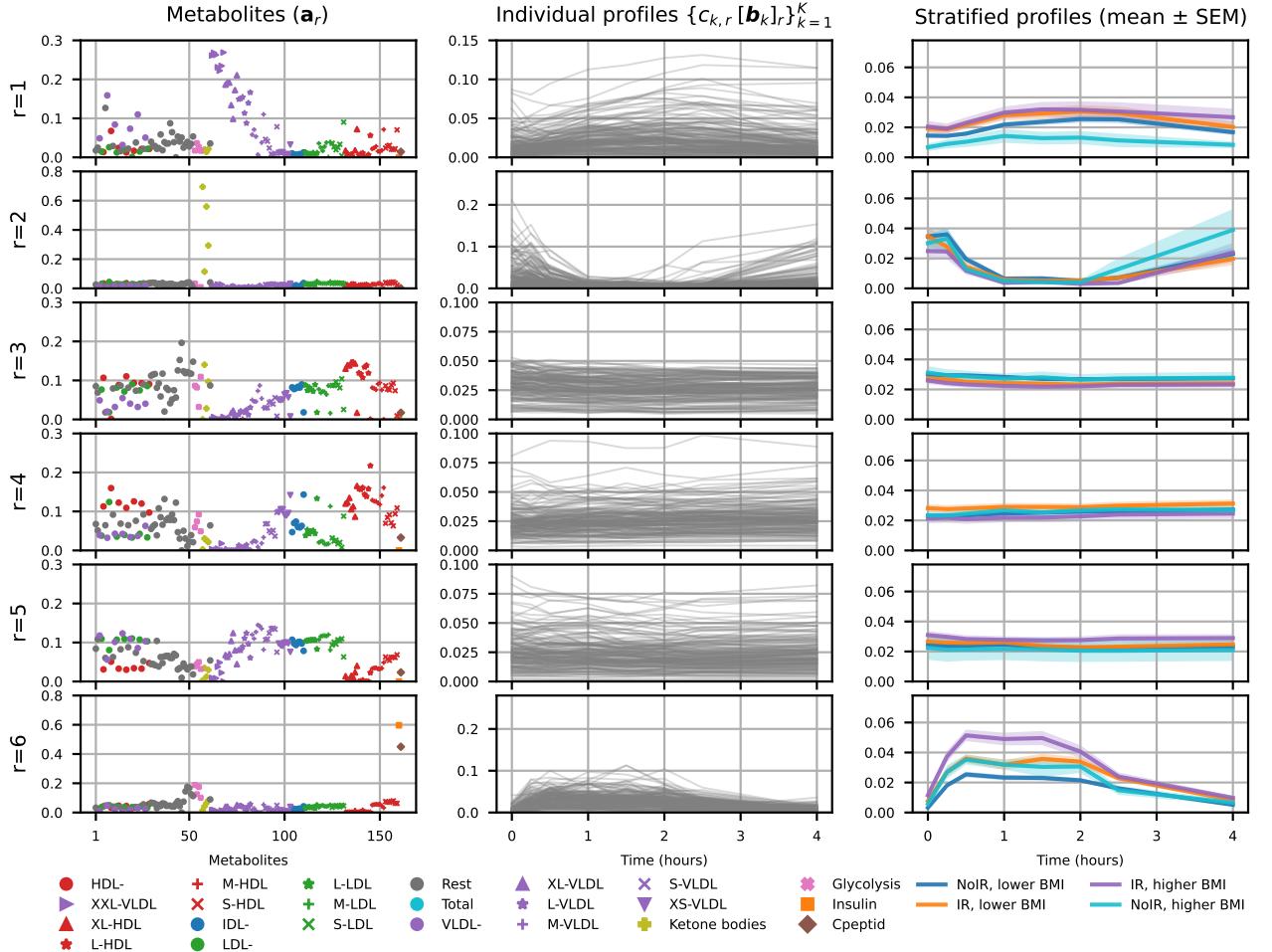


Figure S.4: Metabolomics. Components of a 6-component PARAFAC2 model (with nonnegativity constraints in all modes) of the metabolomics data from females.  $\mathbf{a}_r$  denotes the pattern in the metabolites mode, where metabolites are colored by lipoprotein classes. Different shapes are used for lipoprotein subclasses. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  for each component are shown in the middle column. The last column shows scaled subject-specific time profiles colored according to four BMI/IR groups.

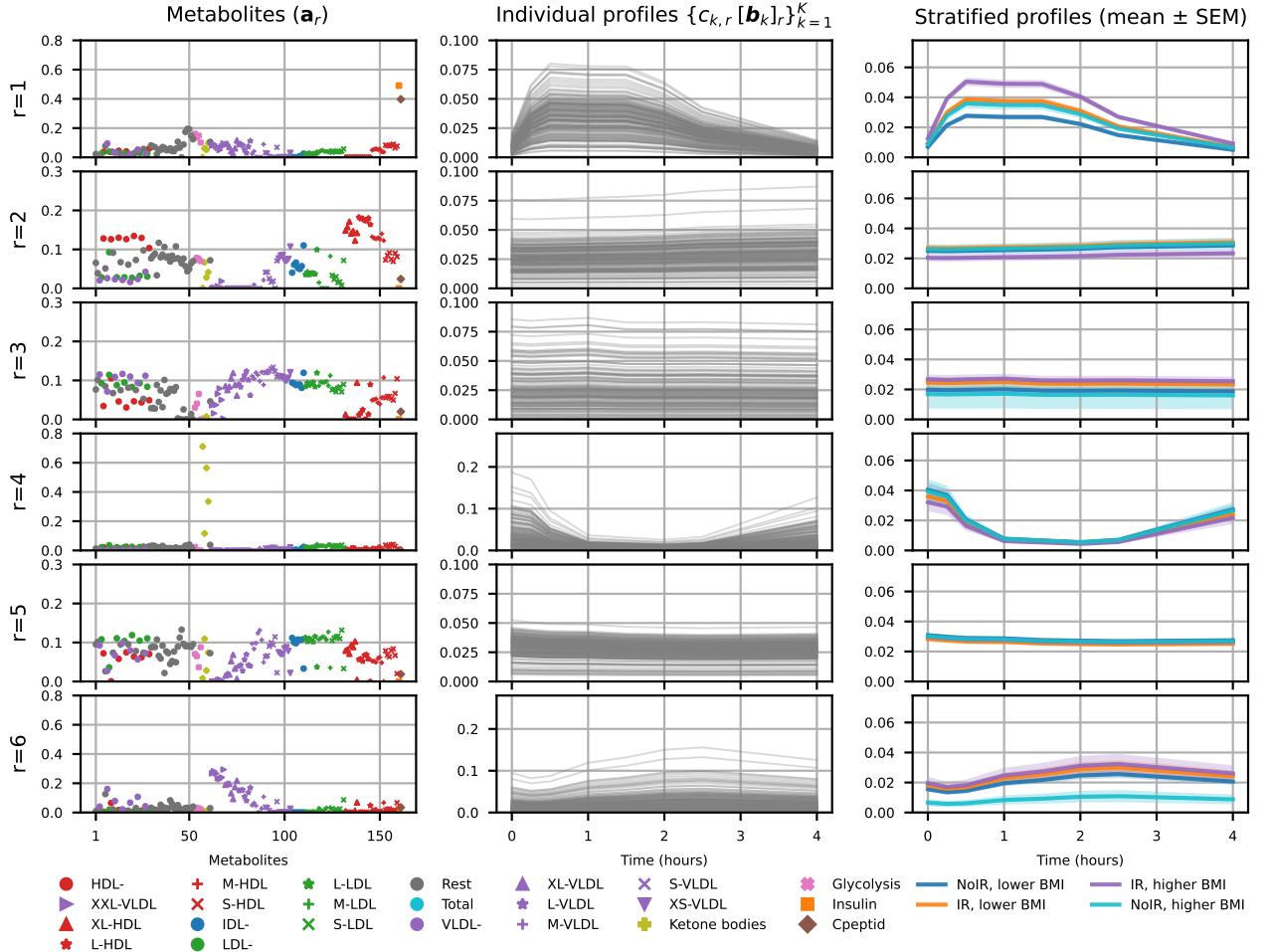


Figure S.5: Metabolomics. Components of a 6-component CP model (with nonnegativity constraints in all modes) of the metabolomics data from females.  $\mathbf{a}_r$  denotes the pattern in the metabolites mode, where metabolites are colored by lipoprotein classes. Different shapes are used for lipoprotein subclasses. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  for each component are shown in the middle column. The last column shows scaled subject-specific time profiles colored according to four BMI/IR groups.

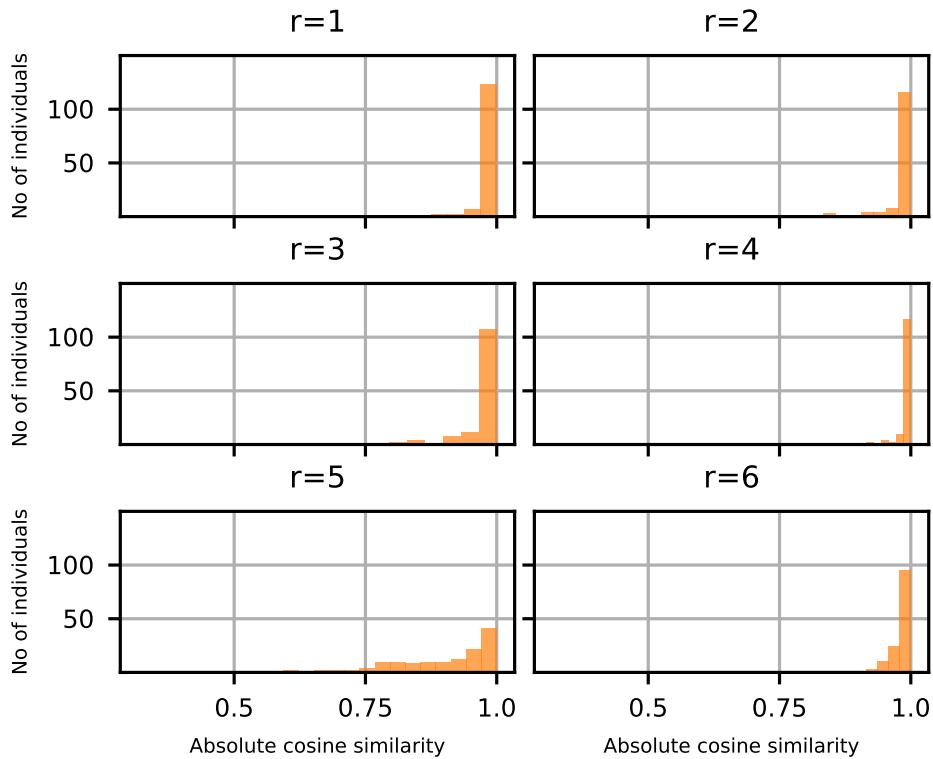


Figure S.6: Comparison of subject-specific time profiles ( $c_{k,r}[\mathbf{b}_k]_r$ ) captured by CMF and PARAFAC2 for each component for metabolomics, where each histogram contains 140 data points, one for each male subject.

## 2 Sensitization

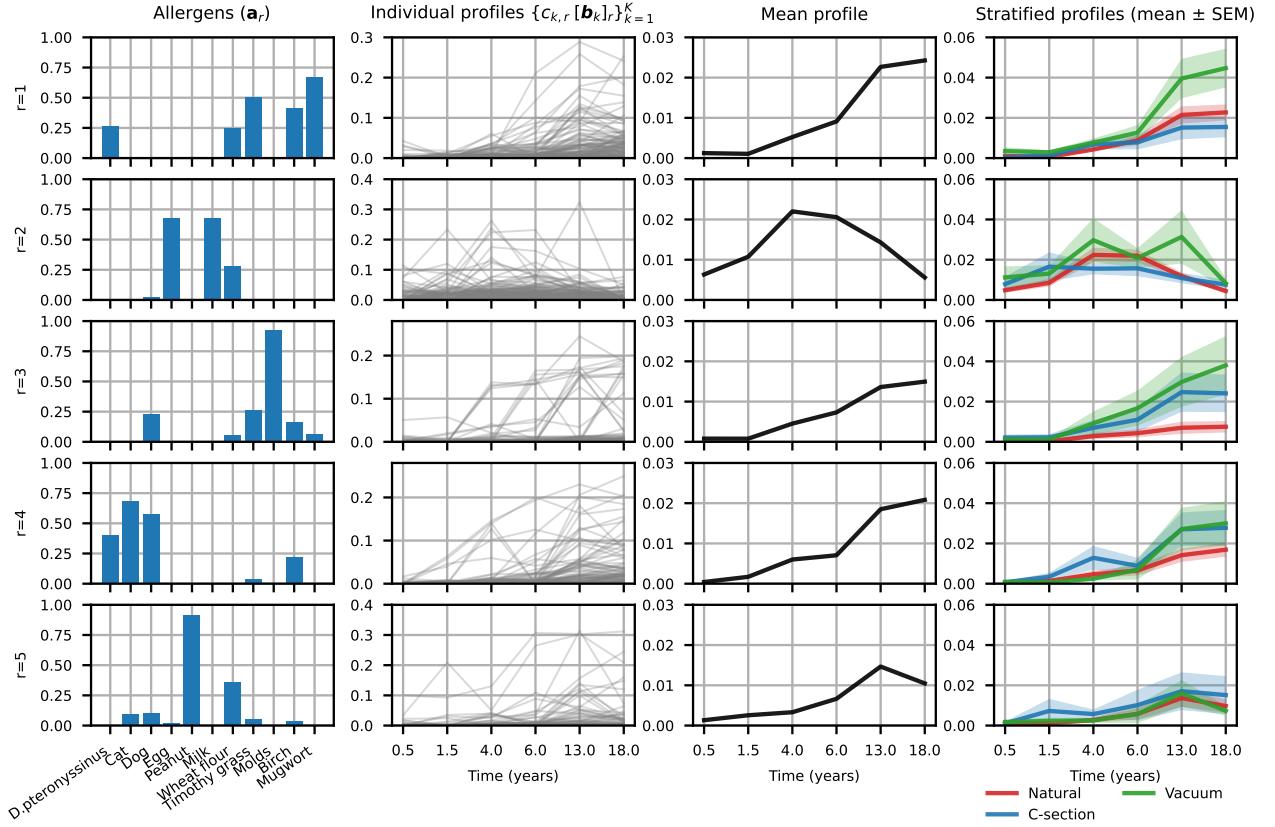


Figure S.7: Sensitization. Components of a 5-component PARAFAC2 model (with nonnegativity constraints in all modes) of the sensitization data.  $\mathbf{a}_r$  denotes the pattern in the allergens mode. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  are shown in the middle column. Mean of scaled subject-specific profiles are plotted in the third column while the last column shows mean (and standard error of mean) patterns of scaled subject-specific time profiles colored according to delivery/birth mode groups.

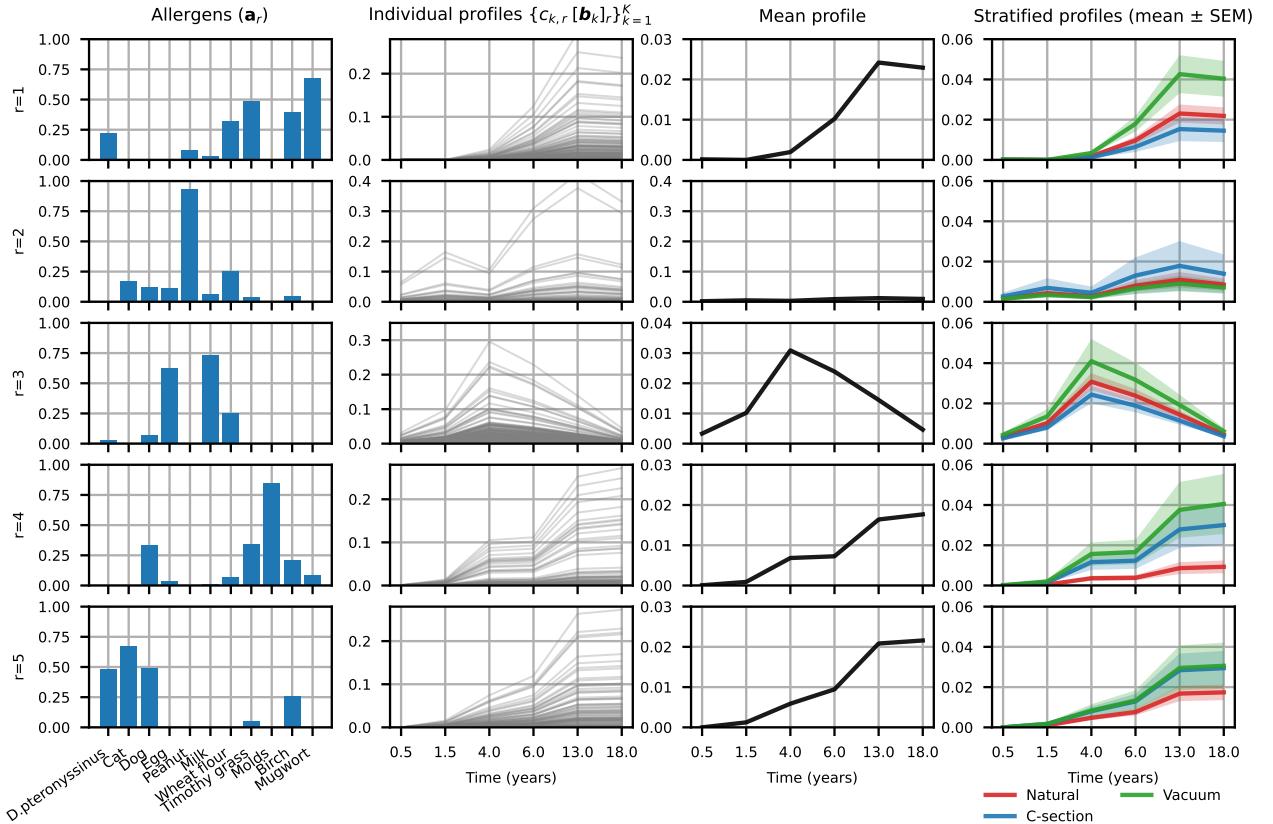


Figure S.8: Sensitization. Components of a 5-component CP model (with nonnegativity constraints in all modes) of the sensitization data.  $\mathbf{a}_r$  denotes the pattern in the allergens mode. Subject-specific time profiles scaled by the corresponding subject scores, i.e.,  $c_{k,r}[\mathbf{b}_k]_r$  are shown in the middle column. Mean of scaled subject-specific profiles are plotted in the third column while the last column shows mean (and standard error of mean) patterns of scaled subject-specific time profiles colored according to delivery/birth mode groups.

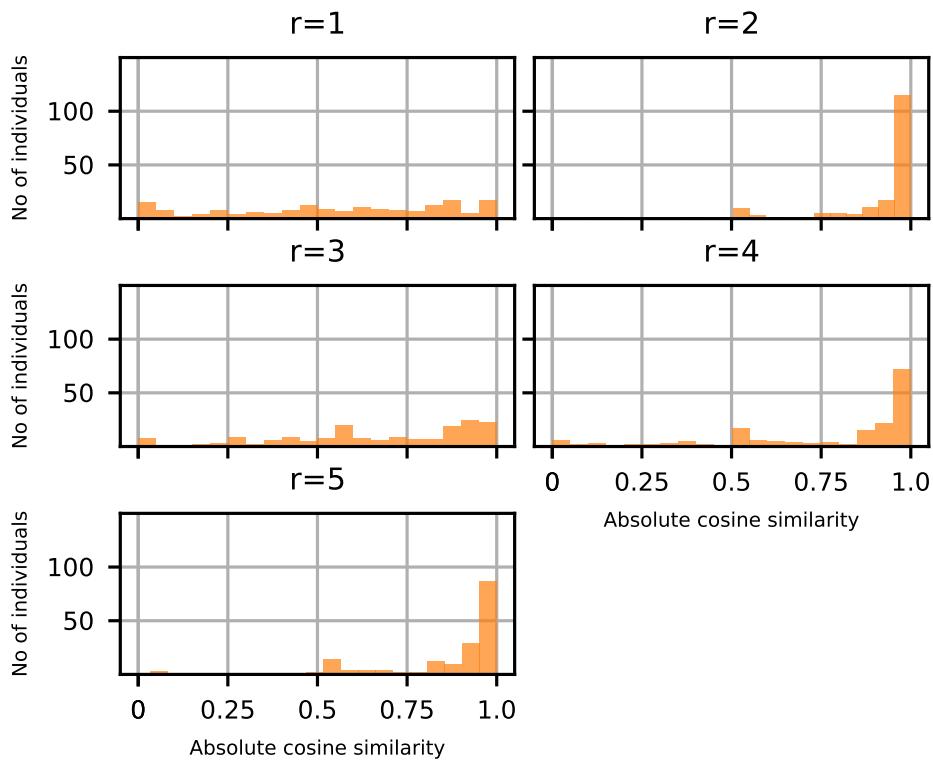


Figure S.9: Comparison of subject-specific time profiles ( $c_{k,r}[\mathbf{b}_k]_r$ ) captured by CMF and PARAFAC2 for each component for metabolomics, where each histogram contains 176 data points, one for each subject.