**Impact of the pre-examination phase on multicenter metabolomic studies**

Veronica Ghini, Peter M. Abuja, Ozren Polasek, Lukasz Kozera, Päivi Laiho, Gabriele Anton, Marie Zins, Janis Klovins, Andres Metspalu, H.-Erich Wichmann, Christian Gieger, Claudio Luchinat, Kurt Zatloukal\*, Paola Turano\*

Supplementary material

***Table S1****. List of the metabolites assigned and analyzed in plasma/serum samples of NSCLC patients. The Human Metabolome Database (HMDB) compound ID of each metabolite is reported.*

|  |  |
| --- | --- |
| **Metabolites** | **HMDB ID** |
|  |  |
| 3-methyl-2-oxovalerate | HMDB00491 |
| Acetate | HMDB00042 |
| L-alanine | HMDB00161 |
| Citrate | HMDB00094 |
| Creatine | HMDB00064 |
| Creatinine | HMDB00562 |
| Formate | HMDB00142 |
| D-glucose | HMDB00122 |
| L-glutamate | HMDB00148 |
| L-glutamine | HMDB00641 |
| L-histidine | HMDB00177 |
| L-isoleucine | HMDB00172 |
| Lactate | HMDB00190 |
| L-leucine | HMDB00687 |
| L-phenylalanine | HMDB00159 |
| Pyruvate | HMDB00243 |
| L-proline | HMDB00162 |
| L-tyrosine | HMDB00158 |
| L-valine | HMDB00883 |

***Table S2****. Discrimination accuracy values among samples collected from different biobanks from the BBMRI-LPC consortium using different supervised statistical techniques (Random Forest, OPLS-DA, PCA-CA, PLS-CA). The p-value of the permutation test is also reported.*

|  |
| --- |
| **RANDOM FOREST** |
| **Plasma** |  |  | **Serum** |  |
| Total Accuracy | **90**(p-value = 0.01) |  | Total Accuracy | **96.7**(p-value = 0.01) |
| Class Accuracy |  |  | Class Accuracy |  |
| BB1 | 82.8 |  | BB2 | 96.9 |
| BB2 | 97 |  | BB3 | 100 |
| BB3 | 89.7 |  | BB5 | 90.3 |
| BB5 | 76.7 |  | BB6 | 100 |
| BB6 | 100 |  | BB7 | 96.7 |
| BB7 | 100 |  |  |
| BB8 | 82.8 |  |
| **OPLS-DA** |
| **Plasma** |  |  | **Serum** |  |
| Total Accuracy | **92.5** (p-value = 0.01) |  | Total Accuracy | **98.5**(p-value = 0.01) |
| Class Accuracy |  |  | Class Accuracy |  |
| BB1 | 91.3 |  | BB2 | 100 |
| BB2 | 98.1 |  | BB3 | 100 |
| BB3 | 96.2 |  | BB5 | 95.0 |
| BB5 | 95.6 |  | BB6 | 100 |
| BB6 | 100 |  | BB7 | 94.9 |
| BB7 | 99.8 |  |  |
| BB8 | 92.5 |  |
| **PCA-CA** |
| **Plasma** |  |  | **Serum** |  |
| Total Accuracy | **94**(p-value = 0.01) |  | Total Accuracy | **99**(p-value = 0.01) |
| Class Accuracy |  |  | Class Accuracy |  |
| BB1 | 89.9 |  | BB2 | 97.8 |
| BB2 | 97 |  | BB3 | 99.3 |
| BB3 | 92.5 |  | BB5 | 96.9 |
| BB5 | 95 |  | BB6 | 100 |
| BB6 | 96.2 |  | BB7 | 99.5 |
| BB7 | 100 |  |  |
| BB8 | 92 |  |
| **PLS-CA** |
| **Plasma** |  |  | **Serum** |  |
| Total Accuracy | **95**(p-value = 0.01) |  | Total Accuracy | **98**(p-value = 0.01) |
| Class Accuracy |  |  | Class Accuracy |  |
| BB1 | 90.4 |  | BB2 | 98.2 |
| BB2 | 94.3 |  | BB3 | 100 |
| BB3 | 92.3 |  | BB5 | 94.2 |
| BB5 | 96.6 |  | BB6 | 100 |
| BB6 | 99.5 |  | BB7 | 95.3 |
| BB7 | 99.4 |  |  |
| BB8 | 93.8 |  |