AUTOMATED DIAGNOSIS OF INBORN ERRORS OF METABOLISM USING HIGH-THROUGHPUT NMR SPECTROSCOPY

Ajit Singh¹, Russell Greiner¹, Victor Dorian² and Brent Lefebvre²
¹Department of Computing Science, University of Alberta  ²Chenomx Inc.

CCMA : CONCURRENT CHEMICAL MIXTURE ANALYSIS

Inborn errors of metabolism (IEMs) are an extremely rare class of genetic diseases.

The Versatile Automated Sampler Transport (VAST) rapidly injects samples into the NMR spectroscope.

High-throughput Nuclear Magnetic Resonance (NMR) spectroscopy.

Chenomx's CCMA software automatically identifies 75 small metabolites by their spectral signature.

Accurate identification is normally a time-consuming process for experts.

DxTOOL : DIAGNOSTIC TOOL

1-Methylhistidine 0.0687
Acetamide 0.0353
Adipic Acid 0.0000
Leucine 0.0449
Sucrose 0.0144

DxTool uses
- Bayesian networks
- Bayesian error bars for confidence of diagnosis.

Inputs
- Compound-concentration pairs
- Noisy concentrations
- Occasionally incomplete data

DxTool produces
- Probability of each IEM
- Confidence intervals
- Variance of the response
- Linear-normal assumption

This work was funded in part by Chenomx and NSERC CRCNS.
### BN2O (Binary Node 2-Layer Noisy-OR)
- One layer of disease states
- One layer of symptom States
- Encodes $P(\text{Symptom} | \text{Disease})$

### Design

### Error Bars

### Example

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>True (mmol/L)</th>
<th>False (mmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methylmalonic aciduria</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Propionic aciduria</td>
<td>0.2477</td>
<td>0.0000</td>
</tr>
<tr>
<td>Leukopenia</td>
<td>0.973</td>
<td>0.027</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
<td>0.637</td>
<td>0.363</td>
</tr>
<tr>
<td>3-D-Hydroxybutyrate</td>
<td>0.0000</td>
<td>0.2477</td>
</tr>
<tr>
<td>3-Furoic acid</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>5-Hydroxylysine</td>
<td>1.0444</td>
<td>0.0000</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.4411</td>
<td>0.0000</td>
</tr>
<tr>
<td>Methylmalonic acid</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- 3-D-Hydroxybutyrate: abnormal
- 3-Furoic acid: normal
- 5-Hydroxylysine: abnormal
- Alanine: normal
- Methylmalonic acid: normal

### Pros / Cons
- **Small Number of Parameters**
  - $O(n)$ instead of $O(2^n)$ per CPTable
- **Parameters have clear interpretation**
  - Use of medical literature and experts
- **Discretization to normal/abnormal works well for most symptoms**
- **Inference can be done in $O(2^{\text{positive findings}})$**
  - Quickscore Algorithm [1]
- **Not yet possible w/ Bayesian error bars**
- **Allows modelling of a patient with two similar IEMs**
  - (e.g. Both 3-Methylglutaconic aciduria and 3-Methylglutaric aciduria)

### Conclusions