Tryptophan Metabolism Profiling for Psychiatric Diagnosis Support

João Pedro Belo Pereira
joaobelopereira@ist.utl.pt

Eindhoven University of Technology, Eindhoven, Netherlands
Instituto Superior Técnico, Lisboa, Portugal

March 2016

Abstract

Mental disorders represent one of the most prevalent and debilitating classes of chronic diseases in the general population, affecting 38.2 % of the European population each year, with an estimated total cost of 798 billion euros in 2010. Several of them have been associated with the Tryptophan metabolism, which may provide the basis for physiological diagnosis. Driven by the necessity to develop a cheap, more quantitative method for diagnosis of psychiatric disorders, Tryptophan metabolism was explored, simulated and a novel-classification technique was developed. The first part of the project involved pattern recognition using *k-means* on blood measurements of Major Depressive Disorder(MDD) patients from two hospitals. This approach allowed to identify metabolic fingerprints for this disease and how those vary between groups of different age and sex, confirming the hypothesis that the neurotoxic Quinolinic Acid (Quin) and the neuroprotective Kynurenic acid (KynA) play an important role in the development of disease. Moreover, it was found that the ratio between 5-Hydroxyindole acetic acid (5-HIAA) and 5-Hydroxy-Tryptophan (5-htrp), Picolinic Acid (Pic) and 3-Hydroxyanthranilic Acid (3HAA) could all have an influence in mental health, and so does 3-Hydroxy-Kynurenine (3-HK) in a complex context-based fashion. The effect of medication on the Trp metabolism was also analysed using MDD blood measurements from another source and applying *k-Medoids* with the *Dynamic Time Warp* (DTW) distance to capture shape evolution. Some coherent patterns among the subjects were found, and it is proposed that Selective Serotonin Reuptake Inhibitors (SSRIs) have anti-inflammatory properties.

The second part of the project involved mapping real urine values to an existing Genome-Scale Metabolic Model (GSMM), which was tested and modified to be used in the classifier. During this process, a user-interface tool was developed to facilitate the modification, testing and analysis of the model and its simulations.

The classifier technique takes advantage of the available biological information offered by the GSMM, as well as training sample expansion, to improve the classification performance in comparison to other classifiers (KNN, Decision Trees and Discriminant Analysis). Moreover, it enables visual metabolism comparison and thus provides important details to aid both medical diagnosis and therapeutic intervention, and possibly drug design.

Keywords: Tryptophan, Data-Mining, GSMM, FBA, Classifier

1. Introduction

Mental disorders represent a common general population disability, characterized by an early age-of-onset (AOO) and higher measures of impairment severity than physical disorders [6].

Moreover, not everyone who exhibits psychiatric disorder symptoms seeks professional help [12]. For this reason, it is difficult to establish a realistic estimate of the disease prevalence, and those patients could live under impairing conditions.

Regarding the diagnosis of this class of diseases, the current practice relies on diagnostic manuals such as the DSM 5 and ICD-10 which rely on symptomatic classification rather than aetiologic processes.

However, there is strong evidence suggesting that the metabolism of the amino acid Tryptophan is disrupted in the majority of patients with psychiatric disorders[13, 16]. Another important characteristic of Tryptophan is that most of its metabolites are measurable in urine, which besides being biological "waste" and thus easy to obtain, the analysis of its components can be low-cost [2] . Combining human Genome Scale Metabolic

Models (GSMM), which contain comprehensive information on the known metabolic processes, with individual metabolic profile analysis, could provide a method to objectively and quantitatively diagnose psychiatric disorders, in an automated and systematic fashion.

2. Background

2.1. Tryptophan metabolism

Tryptophan (Trp) is one of the eight essential aminoacids, and in the brain, it is a precursor of the neurotransmitter serotonin (5-HT). This metabolite has been shown multiple times to be associated with a variety of psychiatric disorders, particularly depression and anxiety[10], as well as Atention Deficit-Hyperactivity Disorder (ADHD) [3]. In the methoxyindoles pathway, Trp is degraded by Tryptophan-hydroxylase (TPH) into 5-hydroxytryptophan (5htrp), which can be converted into serotonin. Serotonin is degraded by monoamine oxidase (MAO) into 5-hydroxyindoleacetic acid (5HIAA), whose extracellular levels represent an index of the levels of newly synthesized 5-HT[4].

Tryptophan can also be metabolised into kynurenine by tryptophan 2,3-dioxygenase (TDO) in the liver (99 % of body tryptophan), or by indoleamine 2,3-dioxygenase (IDO) in the brain[9].

Kynurenine can then either be converted into 3-hydroxykynurenine (3HK) by kynurenine-2-monooxygenase (KMO), Anthranalinic Acid(AA), or Kynurenic Acid (KynA). 3HK can be further degradaded into Xanthurenic acid (Xan) or to 3-hydroxyanthranilic acid (3HAA) by kynureninase. 3HAA can then be converted into Quinolinate (Quin).

Quin is a N-methyl-D-aspartate receptor (NMDA-R) agonist. For this reason, the accumulation of Quin leads to NMDA-R overstimulation, which in turn could lead to excitotoxicity[1]. KynA, on the other hand, is a non-selective NMDA-R antagonist and therefore it is protective against the excitotoxicity of Quin[15, 18].

The ratio between KynA and Quin is defined as the neuroprotective ratio, because it is a relative measure of the extent to which KynA protects the neurons against the toxic effects of Quin in elevated concentrations.

3HK and 3HAA, the intermediate products between kynurenine and Quin, were speculated to intensify the neurotoxicity of Quin because their oxidation gives rise to reactive oxygen species(ROS) which in turn cause oxidative stress[11]. Besides the metabolites described above, Picolinic Acid (Pic) can also be formed in the kynurenine pathway, which is also an NMDA agonist [14].

During inflammation, the activity of IDO and KMO are increased by pro-inflammatory cytokines, leading to a decrease in the availability of Tryptophan for serotonin production and an increase of Quin and Xan production.

Interestingly, there is a study suggesting that the otherwise cellular and mitochondrial damaging metabolites 3HK and 3HAA, under inflammation conditions actually exhibit anti-inflammatory, anti-oxidative and neuroprotective activities [8]. The levels of 3-HK were found to be reduced in patients with ADHD, which was hypothesized to impair neuronal pruning[3].

2.2. Data Mining

Data Mining is an increasingly popular field, which dedicates to the discovery of patterns in data.

2.2.1. k-means

The *k-means* algorithm is a commonly used data-mining tool which groups data into k clusters with k cluster representatives, which are the cluster points' means. The algorithm aims to minimize a cost function composed of the distances of all records $X = (x_1, x_2, ..., x_n)$ to their respective cluster set $S = (S_1, S_2, ..., S_n)$ mean:

$$\min_{S} \sum_{i=1}^{k} \sum_{k \in S} \|x - \mu_i\| \tag{1}$$

Where μ_i is the mean of points in S_i. The steps of the algorithm are the following:

1. Start with random cluster centres
2. Attribute each sample to the nearest cluster
3. Update the cluster centres by calculating the mean of the new cluster set
4. Repeat until the clusters no longer change

2.2.2. Silhouette Index

The Silhouette is a cluster quality measure based on the average intra and inter-cluster points.

For each point in the data, the Silhouette width is calculated by:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{2}$$

Where $a(i)$ is the average within-cluster distance and $b(i)$ the minimum of the average inter-cluster distance. In order to calculate the quality of each cluster, one can then compute the mean of the silhouette widths for the points in that cluster:

$$S_k = \frac{1}{n_k} \sum_{i \in C_k} s(i) \tag{3}$$

Where C_k is the kth cluster. The global cluster quality (the global silhouette index) is the mean of the mean silhouettes of all clusters:

$$\mathcal{C}_k = \frac{1}{k}\sum_{k=1}^{K} S_k \tag{4}$$

where K is the total number of clusters.

2.2.3. K-Medoids

The K-Medoids also clusters data into k groups, using points from the data itself as representatives. The algorithm selects k points from the data as representatives, and attributes each point to the closest representative. Then it swaps the current representative by other point and if the cost of doing so increases the swap is undone, otherwise is kept. This process is repeated until there is no further change.

2.2.4. Dynamic Time Warp

Dynamic Time Warping (DTW) is a technique designed to capture optimal alignment between two time-sequences, by warping the sequences non-linearly so that they match. If i and j are ith and jth elements of two time-dependent signals X and Y , then the DTW distance is:

$$DTW(i,j) = dis(x_i, y_j) + min \begin{cases} DTW(i, j-1) \\ DTW(i-1, j) \\ DTW(i-1, j-1) \end{cases} \tag{5}$$

2.3. Metabolic Modeling

In order to study the metabolism of an organism, a model can be developed to predict functionality, discover disease mechanisms or design drugs.
If, for a given organism, there is sufficient genetic, metabolomic and enzymatic information available, then a genome-scale metabolic model (GSMM) can be built by converting the metabolic network into a mathematical format.

2.3.1. Constraint-based Analysis: FBA

Constraint-Based methods impose cellular limitations on biological networks to simulate flux through the various reactions in a metabolic model. One such technique is the Flux Balance Analysis, which calculates the internal metabolic fluxes at steady-state by solving a linear optimization problem, subject to constraints on the allowed fluxes for a given objective function.
For a given metabolic network, an ordinary differential equation (ODE) can be attributed to each metabolite to describe its rate of change:

$$\frac{dX}{dt} = -v_1 + v_2 + ... + v_n \tag{6}$$

Where v_1 to v_n are the reactions' fluxes in which metabolite X participates. The set of all metabolites' equations and reactions can be described in a matrix S, called the Stoichiometric matrix.
At steady-state, the rate of production is equal to the rate of consumption, and therefore the set of equations generated by $S \cdot v$ cancel each other, and the system reduces to a set of linear equations $S \cdot v = 0$, where v is the array that contains each reaction $v = (v_1, v_2, ..., v_n)$.
However, this system is under-determined. To further restrict the solution space, the flux through each reaction can be restricted as follows:

$$v = v_m \quad or \quad l_b < v < u_b \tag{7}$$

Then an objective function Z to maximize is defined as follows:

$$Z = c^T \cdot v \tag{8}$$

Where c is a vector of zeros except at the positions of the reactions to be defined as objective functions.
The final step is to use linear optimization to calculate the set of fluxes that maximize the given objective function, subject to the constraints of the lower and upper reactions' boundaries. The problem to solve can be expressed as:

$$\begin{cases} \text{maximize} & c^T \cdot v \\ \text{subject to} & Sv = 0 \\ & l_b \leq v \leq u_b \end{cases} \tag{9}$$

2.3.2. Mapping Real-Values to the Model

Consider a compartment that is empty at time t_o, for instance, the bladder. After a certain time dT, the compartment is filled with various metabolites with different concentrations, which will be denoted as X_m. The average flux over that time can be calculated as:

$$\frac{\int_{t_o}^{t_o+dT} \frac{dX}{dt} dt}{\int_{t_0}^{t_o dT} dt} \overset{t_0=0}{=} \frac{\int_0^{dT} dX}{\int_0^{dT} dt} = \frac{X(0) - X(dT)}{dT} \tag{10}$$

and since the compartment was empty then $X(0) = 0$. This means for the mth metabolite with concentration X_m and the nth metabolite with concentration X_n:

$$\begin{gathered} \frac{\frac{\int_{t_o}^{t_o+dT} \frac{dX_m}{dt} dt}{\int_{t_o}^{t_o+dT} dt}}{\frac{\int_{t_o}^{t_o+dT} \frac{dX_n}{dt} dt}{\int_{t_o}^{t_o+dT} dt}} = \frac{\int_{t_o}^{t_o+dT} \frac{dX_m}{dt} dt}{\int_{t_o}^{t_o+dT} \frac{dX_n}{dt} dt} \overset{t_o=0}{=} \\ \frac{X_m(dT) - X_m(0)}{X_n(dT) - X_n(0)} = \frac{X_m(dT)}{X_n(dT)} \end{gathered} \tag{11}$$

and thus the average flux ratio between the reactions that transport or produce the mth and nth

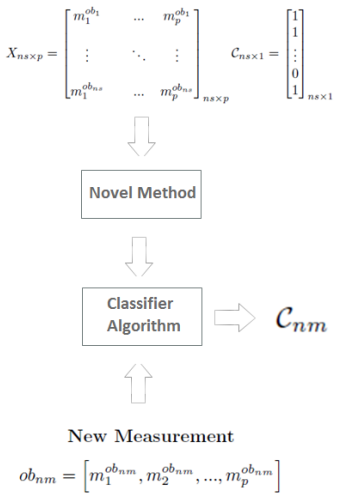

Figure 1: Goal of the classifier

metabolites into/in this compartment, is exactly the ratio between the final metabolite concentrations.
Thus, one way to map values into a GSMM is to use relative extracellular metabolite measurements to constrain the bounds of the excretion reactions for those metabolites.

2.4. Classification

Classification refers to the process of attributing a label to a given measurement (test set). An algorithm that implements classification is called a classifier.
In this project, a new classification algorithm was developed. The main idea is to use the metabolic network provided by the GSMM, to compare the metabolism of several already classified subjects with a new subject, to determine if it is a patient or not. This allows not only to objectify mental disorders as a disease with a physiological diagnostic, but also to provide a quantitative measure of the disease.

2.4.1. FBA$_1$ Classifier

Consider a sample $X_{ns\times p}$ with ns measurements for p excreted metabolites' concentrations, and a class array $\mathcal{C}_{ns\times 1}$ which contains the label of each measurement, for instance patient or healthy control.

If each observation j in the normalized sample $X'_{ns\times p}$, is allowed a positive and negative deviation Δ from its values, and those are imposed as the GSMM reactions' boundaries, an FBA solution can be calculated to simulate a variation of that observation metabolism. The hypothesis is that a new measurement might be experiencing the same metabolic mechanisms as some of the observations in the sample, but due to personal characteristics or time-related fluctuations, exhibit slight differences in the metabolic concentration pattern.
For that purpose, a matrix with all deviation combinations W_c is generated, and from this, a matrix with combinations of metabolite deviations W_Δ is generated:

$$W_c = \begin{bmatrix} 0 & 0 & \dots & 0 & \delta \\ 0 & 0 & \dots & \delta & 0 \\ & \vdots & \ddots & \vdots & \\ 0 & \delta & \dots & \delta & \delta \\ \delta & \delta & \dots & \delta & \delta \end{bmatrix}_{[(nsteps+1)^p-1]}, \quad \delta \in [0,1] \tag{12}$$

$$W_\Delta = W_c \cdot \Delta \tag{13}$$

Where $nsteps$ is the number of increments δ allowed in the combination matrix W_c.
Then, each array in W_{dev} is added or subtracted to the observation values, and those are re-normalized:

$$B_{ob_j}(u) = \frac{ob_j \pm W_\Delta(u,:)\cdot ob_j}{\sum_{i=1}^{p} ob_j(i) \pm W_\Delta(u,i)\cdot ob_j(i)} \cdot 100 \tag{14}$$

Each solution variation is calculated by performing FBA on the model constrained by each array in $B_{ob_j}(u)$, yielding $N = ns[2(nsteps+1)^p - 1]$ solutions. When classifying a new measurement, its values are used to restrict the model and an FBA solution is generated, to which each of the training solutions is compared.
Of the whole-model FBA solutions, a sub-tree of reactions is selected, and the fluxes of the T reactions in the tree will be denoted as solution tree: $\mathcal{S}_t$.
A reaction that carries flux through one or more downstream reactions, will be defined as a parent reaction of that/those reaction(s). Therefore, the flux value of a parent reaction can be defined as:

$$\mathcal{S}_t(\Pi_k) = \Pi_k^t = \sum_{i=1}^{chi} \mathcal{S}_t(\Gamma_k(i)) \tag{15}$$

Where chi is the number of children of the parent reaction Π_k, and Γ_k represents the children reactions from 1 to chi.
If the collection of parents of reaction R_j, that is, the direct parent of that reaction, its parent and so on, is denoted as:

$$\prod(R_j) = \{\Pi_1(R_j), \Pi_2(R_j), ..., \Pi_{np}(R_j)\} = \{\Pi_1(R_j), \Pi_1(\Pi_1(R_j)), ..., \Pi_1(\Pi_{np-1}(R_j))\} \tag{16}$$

Then a distance between solution trees which accounts for the network relations between metabolites can be established, and thus the distance between a new measurement and all other solutions is

calculated as follows:

$$\varepsilon(c) = \sum_{i=1}^{p} \frac{diff_c(i)}{sum_c(i)} + \sum_{par=1}^{npi} \frac{diff_p(i)}{sum_p(i)} \cdot \frac{sum_c(i)}{sum_p(i)}, \quad c \in [1, N] \tag{17}$$

with:

$$\begin{aligned} sum_c(i) &= |R_{nm}(i) + R_c(i)| \\ diff_c(i) &= |R_{nm}(i) - R_c(i)| \\ sum_p(i) &= \left|\Pi^t_{par}[R_{nm}(i)] + \Pi^t_{par}[R_c(i)]\right| \\ diff_p(i) &= \left|\Pi^t_{par}[R_{nm}(i)] - \Pi^t_{par}[R_c(i)]\right| \end{aligned} \tag{18}$$

2.4.2. FBA$_2$ Classifier

What differs in the second method is the inclusion of combinations of metabolites to draw flux from, instead of "penalizing" all other metabolites. Therefore, the combination matrix W_c is composed of arrays with all combinations with both positive and negative variations, yielding a total of $R = (2 \cdot nsteps+1)^p - 2[(nsteps+1)^p - 1] - 1$ arrays, where the combinations without both increase and decrease of flux are excluded. The combination matrix W_c can be divided into a positive and negative matrix:

$$W_c = W_{c+} + W_{c-} \tag{19}$$

Which also applies to the deviation matrix:

$$W_\Delta = W_c \times \Delta = (W_{c^+} + W_{c^-}) \times \Delta = W_{\Delta^+} + W_{\Delta^-} \tag{20}$$

Since the amount of flux drawn from certain metabolites might not be enough to cover the flux increase necessary for the others, those solutions need to adjusted.
Conceptually, this translates into including a scaling step prior the model restriction.
In order to calculate the total amount of flux difference for each combination:

$$X'_{ns \times p} \cdot W^\top_{\Delta^-} = \xi^-_{(ns \times N-1)} \tag{21}$$

$$X'_{ns \times p} \cdot W^\top_{\Delta^+} = \xi^+_{(ns \times N-1)} \tag{22}$$

Each row in ξ^+ and ξ^- contains the flux increase and decrease, respectively, for each variation of the observation in the same row in $X'_{ns \times p}$.
Consider the ith observation in X', ob_i and the kth combination in W_{Δ^+} and W_{Δ^-}. The flux variation ratio is calculated as:

$$\Xi = \frac{|\xi^-(i,k)|}{\xi^+(i,k)} = \frac{\sum_{m=1}^{p} |ob_i(m) \times W_{\Delta^-}(k,m)|}{\sum_{m=1}^{p} ob_i(m) \times W_{\Delta^+}(k,m)} \tag{23}$$

The flux adjustment can then be achieved by scaling the value attributed to the jth modified metabolite from:

$$ob'_i(j) = ob_i(j) + ob_i(j) \times W_{\Delta^\pm}(k,j) \tag{24}$$

to:

$$\begin{cases} \textbf{if } \Xi > 1 \textbf{ and } W_{\Delta^-}(k,j) < 0, \\ \quad ob'_i(j) = ob_i(j)\left[1 - W_{\Delta^-}(k,j) \times \Xi^{-1}\right] \\ \textbf{if } \Xi < 1 \textbf{ and } W_{\Delta^+}(k,j) > 0, \\ \quad ob'_i(j) = ob_i(j)\left[1 + W_{\Delta^+}(k,j) \times \Xi\right] \end{cases} \tag{25}$$

The rest of the concept is similar to the previous one.

3. Materials and Methods

Three datasets were explored in the development of this project, which were used for four different purposes.

3.1. MDD Data

Data from the hospital of Muenster was clustered to extract Tryptophan metabolic patterns from the groups with the most similar characteristics. The data included blood measurements of 12 Tryptophan metabolites from 156 control subjects and 249 patients with Major Depressive Disorder (MDD). There were male and female subjects aged between 18 and 65 years.
The data was pre-processed by removing all observations with at least 5 missing values. The remaining missing values were substituted by the nearest neighbour value, and the outliers were excluded based on the distance to the nearest neighbour for each observation. The data was then normalized for each parameter and subdivided by sex and age ($< 30, 30 <= \text{age} < 50, >= 50$) to account for the metabolic differences between those subgroups. Then each subgroup was clustered based on the 12 metabolites using 400 cycles of k-means with k ranging from 2 to 10, and selecting the best cluster based on the Silhouette Index
Some of the patterns found were inspected by selecting the subjects with those characteristics, and checking the class content.

3.2. Medication Data

Data from a US hospital was clustered to study how medication interacts with the Trp metabolism. It consisted of blood measurements of 6 Tryptophan metabolites from 90 MDD patients, both male and female, aged between 19 and 67 years. There were two subsequent measurements of those metabolites, on the 8th and 12th week after starting Selective Serotonin Reuptake Inhibitors (SSRIs) administration.
Only the records with no missing values were kept, a total of 16 subjects. In order to capture how the metabolic pattern evolves over the 12 weeks, the data was clustered running 300 cycles of k-medoids using Dynamic Time Warping (DTW), with time normalization to cluster based on shape evolution. Due to the use of DTW, subjects with a delayed

response to medication is still considered. The solution with the best Silhouette value was kept.

3.3. ADHD Urine data

In order to use extracellular metabolite concentrations to constrain the GSMM with real values and train the classifier, urine measurements of all Tryptophan end-products from 44 ADHD patients and 60 healthy controls were used.
Due to confidentiality issues there will be no further subject description.
There is no missing data. Instead, there are values that are below the detection threshold (b.d.t.).
The values are used were measured in $\mu mol/g$ of creatinine.
The data was processed by removing subjects with over 5 values b.d.t. and excluding the metabolites with b.d.t. values for more than 40 subjects, yielding 8 metabolites and 96 subjects (43 patients).
Each subject entry was normalized to the sum of its metabolites' values, and the values for each metabolite were used as boundaries of the respective metabolite extracellular transport reaction, to create several subject specific metabolic models. The starting model was the GSSM recon version 2.03 [5] (available in the website http://vmh.uni.lu), which was curated prior any analysis. Then, it was tested and modified until performing as intended for the classifier. Since manually investigating solutions for such a vast network can be a very time-consuming task, a User-Interface tool (*traceMets*), was implemented in $MATLAB^{®}$ to facilitate model modification, solution analysis and visualization.
This dataset was then used to train the FBA classifier, as detailed in section 2.4.1. The accuracy was compared with three other classifiers : Decision Trees, KNN-classifier and Discriminant Analysis.
To train the classifiers, a varying number (28-70) of randomly selected measurements was used, and the rest was used as test set. 80 cycles were ran and the accuracy averaged. The first method was compared to the three other classifiers, and then the second method was compared to the first.
The accuracy was calculated using 5 performance indexes:

$$\begin{aligned} Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \\ Sensitivity &= \frac{TP}{TP+FN} \\ Specificity &= \frac{TN}{FP+TN} \\ PPV &= \frac{TP}{TP+FP} \\ NPV &= \frac{TN}{FN+TN} \end{aligned} \tag{26}$$

where TP, TN, FP, FN are the number of true positive, true negative, false positive and false negative classifications.
The usefulness of using solution variations was also accessed for both FBA classifiers, running 500 classification cycles with a number of randomly selected subjects ranging from 40 to 60, and the cases where the closest measurement proposed by the classifier differed from the nearest neighbour were analysed.

4. Results

4.1. MDD Clustering

This analysis resulted in 56 clusters, of which 23 were single-class populated and 13 included a single subject. Those with only one class subjects were explored in an attempt to correlate their metabolic pattern with MDD or health.
Even though a high or low neuroprotective ratio was observed in control or patient clusters respectively, the opposite was often observed which challenges the idea that this marker can determine the state of mental health. This could be because a very high value of this marker could also represent hyposty-mulation (which was a characteristic of some Patient groups), and a low value could be compensated by the combined effects of the other metabolites on neural stimulation. For instance, it was often observed in control groups with low neuroprotective ratio a high level of 3-HK, which is consistent with the hypothesis by *Krause et al.* that this metabolite has anti-inflammatory and neuroprotective properties in early stages of inflammation. This could represent an early marker of inflammation which could result in resuming normal metabolism or, if it gets out of control, disease. Another important pattern is a high 5htrp and low 5HIAA concentration in patient groups (94 % of all subjects with this pattern were patients). Since 5htrp and 5HIAA are the precursor and product of serotonin, respectively, this pattern could indicate that serotonin is converted to 5-Hydroxy-kynurenine instead of 5HIAA leading to low serotonergic transmission. All of the single-class groups that exhibited high 3HAA were composed only of patients (95% of all high 3HAA subjects were patients), which challenges *Krause* hypothesis that this metabolite has neuroprotective properties.
Another interesting observation: all the single class groups with a high concentration of Picolinic acid, regardless of the neuroprotective ratio, are patient groups(91% of all high Pic subjects were patients). This could be tied with the fact that Pic is also an NMDA agonist, and enhances the formation of radicals [7], possibly contributing to neurotoxicity.

4.2. Medication Clustering

Even though the data size used was considerably small with only 16 observations, some patterns were found to be common to most observations. Most

notably:

1. Trp rises on the 8th week after starting medication in most subjects.

2. Most observations exhibit a decrease in Kyn in the 8th week.

3. Quin decreases on the 8th week in most cluster centres, which along with 1) and 2) and the proposed anti-inflammatory action of SSRIs [17], could indicate reduced IDO activity as a medication mechanism of action.

4. On the 12th week its interesting how the level of 3HK is below the one prior medication for all groups (and observations as well), which could indicate end of inflammation.

4.3. Urine Mapping

As explained in section 2.3.2 and 3.3, ADHD urine measurements were normalized, and the values of each measurement were used to constrain the model excretion reactions. The model has to be able to produce an FBA solution with a flux starting from the top reaction (Tryptophan Intake) and from that reaction distribute the flux through the excretion reactions to match the boundary values, to be usable in the classifier. In the initial curated model there were several reversible Trp exchange reactions allowing more Trp to be used than the desired maximum (100%), which was identified using the *traceMets* tool. This was dealt with by restricting all exchange reactions except for one, which was also applied to the other metabolites. In fig 2 it is represented the final modified model .

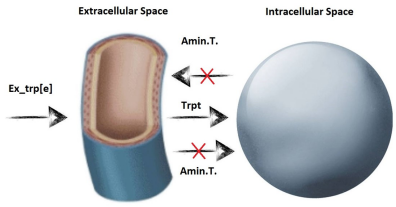

Figure 2: Scheme for the final version of the model. "Amin.T." stands for the aminoacid transporters present in the model, which were restricted.

This model was then used to produce the FBA solutions required by the classifier.

4.4. FBA_1 Accuracy

When performing the classification tests, the Decision Trees (DT), K-Nearest Neighbour (KNN) and Discriminant Analysis (DA) classifiers were tested using normalized and non-normalized values. However, since using non-normalized values yielded worse accuracy for all of them, those results will not be presented.
The average accuracy indexes is displayed in table 4.4, where Δ is the combinations deviation value used in the FBA_1 classifier .

$\boldsymbol{FBA_1}$ **Δ :**	**0.2**	**0.1**	**0.01**
Accuracy	0.5513	0.5408	0.544
Specificity	0.5677	0.5565	0.5584
NPV	0.6089	0.5987	0.6036
Sensitivity	0.5419	0.5329	0.5375
PPV	0.5025	0.4925	0.4936
Classifier :	**KNN**	**DT**	**DA**
Accuracy	0.5297	0.4984	0.47
Specificity	0.5446	0.5529	0.59
NPV	0.5905	0.5558	0.5297
Sensitivity	0.523	0.4405	0.3436
PPV	0.4788	0.4353	0.3949

Table 1: Comparison of FBA_1 with deviation values $\Delta = 0.2, 0.2, 0.01$ and other classifiers performance using mean accuracy, specificity and sensitivity values

NPV exhibited the highest values for all classifiers of all other indexes, which means that on average, the test measurements classified as being control have the highest probability of actually being so. This is important in case it is preferable to be "over-cautious" than to leave a condition untreated. A low PPV is particularly undesirable if the treatment consequences or cost outweigh the possibility of living with a given condition.
Comparing the FBA_1 classifiers with the others, it outperforms them in all indexes except for the DA's specificity. However, this high value is clearly due to the classifier's bias towards control, apparent from its high Specificity combined with the lowest NPV.
Comparing the different deviations of FBA_1, it is consistent for all indexes that $\Delta = 0.2 > 0.01 > 0.1$, which is rather interesting as the increase of the deviation value should improve or degrade the results. Whether using solution variations improves or degrades the results is therefore unclear using accuracy alone.
The next test involved analysing the cases when the chosen closest solution variation was not the same as the nearest neighbour, which were divided in:

- Improvement, the nearest neighbour would misclassify but the solution change prevented it

- Positive Neutral, when the proposed closest variation correctly classified but the nearest neighbour would also have

- Negative Neutral, when there was a misclassification but the nearest neighbour would also misclassify

- Degradation, when the nearest neighbour would correctly classify but the change prevented it

The results of this test are displayed in table 2.

Table 2: Solution variations' on FBA_1 quality. Cases Number is the number of cycles in which the solution variation technique had an influence, I- Improvement, PN- Positive Neutral, NN- Negative Neutral, D - Degradation

$\boldsymbol{FBA_1\ \Delta}$:	**0.01**	**0.1**	**0.2**	**0.3**
Cases Number	150	2661	6846	10394
I %	12.67	27.32	26.31	25.25
PN %	25.33	27.17	26.76	26.78
NN %	24	15.26	22.25	23.6
D %	38	30.25	24.69	24.37

As expected, the larger the deviation value the higher the number of classifications that differ from the nearest neighbour class.
Except for $\Delta = 0.2, 0.3$, the percentage of degradation is higher than that of improvement. This could explain the slightly increased accuracy for $\Delta = 0.2$. Even so, the difference between Improvement and Degradation is quite reduced and there is no guarantee this difference is not due to favourable random training/test set combination.
Therefore, the difference in classification accuracy between the FBA_1 classifiers and the others considered is most likely due to the developed tree distance (eq. 17).

4.5. $\mathbf{FBA_2}$ Accuracy Accuracy

The FBA_2 accuracy was compared with FBA_1 using the same test design.
Only a deviation of 0.2 for FBA_1 was used for comparison as it scored the highest accuracy values .
The results can be found in table 4.5

$\boldsymbol{FBA(\Delta)}$:	**1(0.2)**	**2(0.025)**	**2(0.028)**
Accuracy	0.5513	0.5525	0.5518
Specificity	0.5677	0.5582	0.5555
NPV	0.6089	0.6156	0.6173
Sensitivity	0.5419	0.5566	0.5605
PPV	0.5025	0.5002	0.5001

Table 3: Comparison of FBA_2 and FBA_1 performance

In comparison to FBA_1, the FBA_2 performance for the various deviations exhibits a slightly higher NPV and Sensitivity. This could hint that the effect of the solution variations for FBA_2 is to classify more instances as belonging to the patient class, which might be desirable if the hypothesis that there exist control subjects with an ill metabolism is correct. Therefore, FBA_2 should be more efficient at identifying ill metabolism, whether it belongs to the control class or not.
A similar test was employed to access whether the solution variations improve or degrade the analysis in FBA_2, and the results were included in table 4.

Table 4: Solution variations' on FBA_2 quality

$\boldsymbol{FBA_2\ \Delta}$:	**0.01**	**0.025**	**0.035**
Cases Number	70	1599	2113
I %	40	32.33	25.13
PN %	37.14	29.58	29.2
NN %	22.86	19.32	25.04
D %	0	18.76	20.63

The results are superior to the ones achieved with FBA_1, with a significant difference between the Improvement and Degradation.
This difference percentage decreases with increasing deviation value, but the number of cases where the algorithm intervenes also increases as expected. The hypothesis behind the solution variations is that a given network can be closer to the testing measurement (referred to as the Nearest-Neighbour), but there could be one which would become closer with small metabolic variations, and thus this network has the potential of becoming more similar to the one to classify. Since the goal of this project is to support medical decision, instead of a binary classification, the result could offer the most probable class, as well as the closest solution variation to inform the physician if the subject should or not be accompanied. For instance, if the nearest neighbour is a control, it is of the physician's interest to know exactly just how far is the test measurement from a patient, and what would it take for it to have a metabolism characteristic of a patient. Moreover, a visual representation of the metabolism might significantly ease the medical decision.
In figure 3 is illustrated how the metabolic network of a subject could be visualized, which was generated using a measurement from the urine data and a visualization technique which was implemented in *MATLAB*. By visualizing the network of the

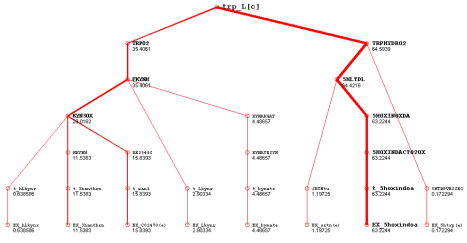

Figure 3: Network visualization of a randomly chosen control measurement

test measurement and the closest measurements in the training set, the physician could identify what would take for an apparently ill subject to approximate a healthier one (medical solution), or decide if a healthy subject is at risk (medical prevention).

5. Conclusions

The goal of this project was to provide methods to objectify and quantify psychiatric diseases. This would be achieved by shedding some insight on what are typical metabolic patterns in healthy and patient subjects, how medication impacts the metabolic network, and providing a diagnostic algorithm to aid medical decision.

Regarding the first, MDD was clustered and from the single-populated clusters the mechanisms that lead to health/disease were speculated and then investigated. The neuroprotective ratio was found not to be a good marker in all circumstances. It could be speculated that 3-HK actuates in an early stage on inflammation, apparent from the high 3hk, low neuroprotective ratio control groups. Of the subjects with high 3HAA, 95 % were patients. A high Pic could also be tied to disease, because 91 % of the subjects with this pattern were patients. Moreover, a high 5htrp low 5HIAA concentration was also commonly found in patients, which could indicate 5htrp is not being converted into serotonin. Even though some interesting patterns were found, a clear separation was not always achieved. The reason could be that in terms of metabolism there is not a binary class, and there could be subjects which have a deficient metabolism but still don't exhibit any physical/psychiatric manifestation, or are still in an early stage of inflammation, with which the body can still cope. Another reason could be that some of the included metabolites actually confound the analysis. Therefore, it would be interesting to start by analysing which are the metabolites that achieve the best Patient/HC separation, and then use those to search for patterns.

Data with measurements for three time points and medication information was clustered with DTW to capture shape evolution. From the results it is here proposed SSRI's work not only by affecting the serotonergic system but also by reducing inflammation, which was illustrated by an increase/decrease in Trp/Kyn in the 8th week (possibly due to decreased IDO activation) followed by a decrease of 3HK in the 12th week which could mark the end of inflammation. Along with the results from the MDD clustering , this outcome reinforces the hypothesis by *Krause et al.* that 3HK's function is to prevent cytokine-induced death during inflammation. The fact that those patterns were found in most observations suggests medication does indeed interact with the Trp metabolism. This effect should be further investigated though, as the data size was quite small, with only 16 instances.

In the metabolic model refinement to use in the classification, several reversible transport reactions were identified using the *traceMets* tool, which were allowing more Trp to be consumed than the desired (100%), and those were dealt with by imposing reaction irreversibility. This "trick" to solve the problem was acceptable, because this was required for the classifier. However, if one is to accurately model an individual's metabolism, for instance to predict the impact of medication, this simplistic approach wouldn't suffice, and therefore an alternative to deal with reversibility should be employed and the model tested against real quantitative values.

Mapping extracellular metabolite levels into the model is a currently active field of research. In this project, this was accomplished by normalizing urine measurements and restricting the excretion reaction boundaries to build subject specific models.

Regarding the classification tests performed, the developed method excelled the other methods in virtually all accuracy measures. Even though the obtained results are interesting, the accuracy for all the methods tested was quite low, and therefore in order to further access the utility of the present method, it should be tested on a larger size data set, produced by independent measurements.

More metabolites should be tested to improve the results, for instance Quin.

The proposed method of generating several metabolic alternatives, to account for small possible deviations in one's metabolism, did not constitute a major improvement for the first method. However, it did increase the global accuracy for the second method, with a better identification of Patient measurements and improved control accuracy for those classified as control. Besides, as demonstrated, the solution variations can aid the medical decision by visualizing how far the classified measurement is from a patient's characteristic metabolism.

This provides a valuable advantage over classifiers that operate like "black-boxes", such as neural networks, because the physician can easily grasp the process by which the classifier solution was generated, which is particularly relevant in case the solution is not the expected one.

In case the new measurement is classified as a Patient, another use for the solution variations would be to inspect the metabolism of the nearest control, and gain insight on what steps should be taken to make this Patient's metabolism closer

to a healthy one. In other words, it has the potential of providing not only the diagnostic, but also information for medical solution and drug design, by identifying which metabolites should be targeted.
Since the proposed classifier is the first of its kind, it was developed based on a small subset of metabolites and a small part of the whole Human Genome Scale Metabolic Model. However, the concept is applicable to any sub-network whose metabolites are measurable in urine. Thus, it would be interesting to inspect how the concept stands up in an improved and scaled up version, after which the question becomes: Is it possible to use all the available human metabolic information, and would that provide for a flawless method to identify metabolic deficiencies?

References

[1] D. A. Bender and G. M. McCreanor. Kynurenine hydroxylase: a potential rate-limiting enzyme in tryptophan metabolism. *Biochem Soc Trans 13, 441–443.*, 1985.

[2] S. Bouatra et al. The human urine metabolome. *PLoS ONE*, 8(9):1–28, 09 2013.

[3] B. Campbell et al. Kynurenines in CNS disease: regulation by inflammatory cytokines. *Front. Neurosci. 8:12*, 2014. doi: 10.3389/fnins.2014.00012.

[4] R. Dantzer et al. Cytokines, Stress, and Depression. *Springer Science and Business Media*, 23 November, 2007.

[5] T. et al. A community-driven global reconstruction of human metabolism. Nat Biotech, 2013.

[6] R. C. Kessler et al. The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *An update from the WHO World Mental Health (WMH) Surveys. Epidemiol Psichiatr Soc.; 18(1): 23–33*, 2011.

[7] K.Minakata et al. Effect on naturally occurring iron ion chelators on the formation of radicals in the reaction mixtures of rat liver microsomes with ADP, Fe^{3+} and NADPH. *J. Clin Biochem Nutr. 49(3):207–215*, 2011.

[8] D. Krause et al. The Tryptophan Metabolite 3–Hydroxyanthranilic Acid Plays Anti–Inflammatory and Neuroprotective Roles During Inflammation – Role of Hemeoxygenase–1. *The American Journal of Pathology, Vol. 179, No. 3 pages 1360–1372*, September 2011. doi: 10.1016/j.ajpath.2011.05.048.

[9] A. M. Myint. Kynurenine Pathway Metabolites Imbalances: The Link between Immune System Disturbances and Pathophysiology of Psychiatric Disorders. *Department of Psychiatry and Psychotherapy of the Faculty of Medicine of Otto-von-Guericke University of Magdeburg*, 2012.

[10] N. Nordquist and L. Oreland. Serotonin, genetic variability, behaviour, and psychiatric disorders - a review. *Upsala Journal of Medical Sciences 115(1): 2–10*, 2010.

[11] S. Okuda et al. 3-Hydroxykynurenine, an Endogenous Oxidative Stress Generator, Causes Neuronal Cell Death with Apoptotic Features and Region Selectivity. *J. Neurochem., 70, pp.299-307*, 1998.

[12] M. Oliver et al. Help–seeking behaviour in men and women with common mental health problems: cross–sectional study. *The British Journal of Psychiatry, 186(4) 297–301*, 2005.

[13] G. Oxenkrug. Tryptophan–Kynurenine Metabolism as a Common Mediator of Genetic and Environmental Impacts in Major Depressive Disorder: The Serotonin Hypothesis Revisited 40 Years Later. *The Israel journal of psychiatry and related sciences.;47(1):56–63.*, 2010.

[14] G. Oxenkrug. Serotonin – kynurenine hypothesis of depression: historical overview and recent developments. *Curr. Drug Targets 14(5): 514–521*, 2013.

[15] M. N. Perkins and T. W. Stone. An iontophoretic investigation of the actions of convulsant kynurenines and their interaction with the endogenous excitant quinolinic acid. *Brain Res 247, 184–187.*, 1982.

[16] D. Richard et al. L–Tryptophan: Basic Metabolic Functions, Behavioral Research and Therapeutic Indications. *International Journal of Tryptophan Research : IJTR.; 2:45–60*, 2009.

[17] R.J.Tynan et al. A comparative examination of the anti–inflammatory effects of SSRI and SNRI antidepressants on LPS stimulated microglia. *Brain, Behaviour and Immunity 26(3): 469–479*, 2012.

[18] M. B. Robinson et al. Structure-Function Relationships for Kynurenic Acid Analogues at Excitatory Pathways in the Rat Hippocampal Slice . *Brain Research, Volume 361 pages: 19–24*, 30 December 1985. doi:10.1016/0006-8993(85)91270-3.