# Identifying Potential Biomarkers for Duchenne Muscular Dystrophy in Urinary Metabolomics Data.

Krabbendam, Jana

J.L. Krabbendam

# Identifying Potential Biomarkers for Duchenne Muscular Dystrophy in Urinary Metabolomics Data.

Bachelor thesis

6 July 2025

Thesis supervisors: dr. M. Signorelli
dr. P. Spitali


Leiden University
Mathematical Institute

# Abstract

Duchenne Muscular Dystrophy (DMD) is a serious muscular disease that ultimately causes premature death. To research DMD, biomarkers are used to track its progression or evaluate potential treatments. Obtaining biomarkers from skeletal muscle however, is a very invasive process and does not provide information on the overall status of the body. So, easy accessible alternatives that can also give overall information, such as urine, are researched. This thesis contributes to that by aiming to identify new, potential biomarkers in urine from DMD patients and mice.

With urine data from mice available, we used linear mixed models (LMMs) to model the expression linearly for each metabolite. With these modelled trajectories, we were able to test some hypotheses to compare different mouse groups. This resulted in five interesting metabolites showing significant difference in expression levels between wild type and *mdx* mice. From these five metabolites, three metabolites show a clear difference between healthy and DMD mice; two show this in their predicted intercept and one in its slope. This indicates a potential in using urine for Duchenne biomarkers, despite the available data having its limitations.

The available human data was more difficult to analyse, because of the missing values it has and the asymmetric distribution of the samples. This means that we were not able thoroughly analyse this data. However, we did do a descriptive analysis from which a conclusion could be drawn. We found the metabolite creatine to be a promising biomarker for DMD in human urine, as it shows a clear difference between healthy and DMD patients.

Overall, while not being able to couple the datasets together due to the missing labels in the mouse data, we did conclude that there is potential in urine for obtaining biomarkers for DMD.

# Contents

# 1 Introduction

Duchenne Muscular Dystrophy (DMD) is a severe, progressive muscular disease primarily affecting boys. It is a genetic disorder caused by a mutation on the DMD gene, which leads to muscle damage and eventually premature death. To monitor and research DMD, biomarkers are used as objective measurements of an individual's biological state. Of course, muscle biopsies can be used, but obtaining those can be very invasive and just provide local information. Thus, less invasive biomarkers are needed, obtained from for example biofluids, that can also be representative for the overall body condition. As blood already shows to have potential biomarkers for DMD, an even less invasive biofluid would be urine (Signorelli et al., 2021). This thesis is part of a pilot study aiming to identify new potential biomarkers for DMD patients and mice from urine data, as this might be a promising biofluid.

To reach this goal, we were given two datasets containing metabolomic data from urine samples, that is data containing measurements involving metabolites. One dataset consists of mouse samples, while the other one contains human samples. Furthermore, the mouse dataset is longitudinal, meaning that the multiple samples were obtained from the same mouse over a period of time, while the human data is cross-sectional and thus has measurements from one time point only.

To guide this thesis into reaching our goal, we formulated three research questions. The main research question is: "Are there metabolites whose expression level differ between healthy and dystrophic individuals?". And to support this question, the other questions are: "Are there metabolites that behave differently among the different mouse groups and if yes, how do they differ?" and "Are there metabolites that behave differently between the healthy and DMD patients?". We will answer these questions using linear mixed models (LMMs) and a descriptive analysis. By modelling the data with LMMs we can predict the expression of metabolites, which will enable us to test meaningful null hypotheses on the models and hopefully answer the research questions.

This thesis is organized in multiple chapters. Chapter 2 explains the background information of this thesis and introduces the available data and research questions. Next follows a chapter describing the data analysis pipeline and then a chapter that presents the findings from this analysis. Chapter 5 concludes the results found in the previous chapter, following by a chapter explaining the limitations of this research.

# 2 Context and problem definition

The purpose of this chapter is to provide the reader with enough knowledge to properly understand the background information and setup of this thesis. The first section explains the background of the topic, followed by a section introducing the available data. The last section presents the research questions that will be answered in this thesis.

## 2.1 Background

This section covers the context of the thesis. First, the muscular disease Duchenne will be explained, following an explanation of what biomarkers and metabolites are. These subsections together will also clarify the relevance.

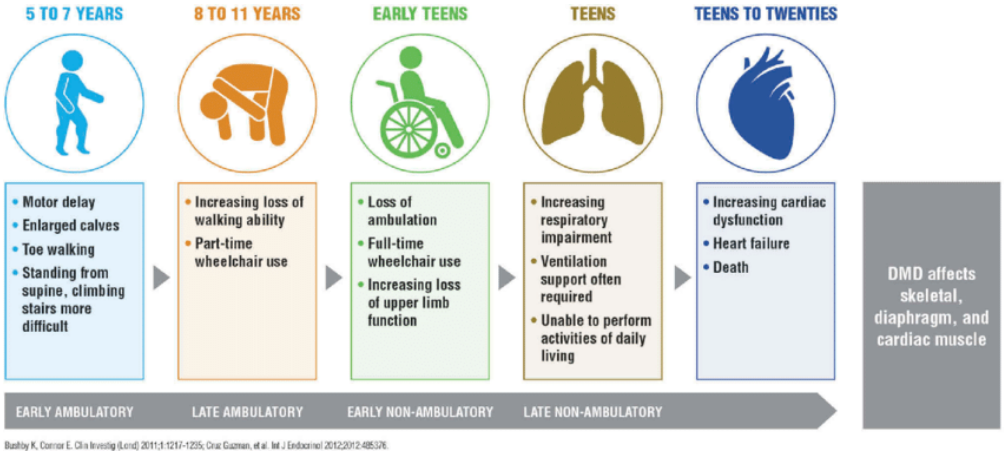



Figure 1: Illustrative representation of the disease progression of Duchenne Muscular Dystrophy. The first symptoms are expressed when patients are toddlers; they experience trouble in sitting, walking and talking. This progresses further to complete loss of ambulation in the early teens and eventually results in respiratory and cardiac failure. Patients typically pass away prematurely in their late twenties. Source: Asher et al. (2020).

### 2.1.1 Duchenne Muscular Dystrophy

Duchenne Muscular Dystrophy, commonly abbreviated DMD, is a muscular disorder that progressively causes severe loss in muscle function. DMD patients experience delays in early developmental milestones, such as sitting, walking and talking, which often leads to diagnosis at 3-5 years old. This progresses further to motor delays in the early ambulatory phase at 5-7 years old and complete loss of ambulation in the early teens, where patients become wheelchair dependent. The disorder ultimately causes respiratory and cardiac failure and thus results in premature death, typically before the patients thirtieth birthday (Figure 1) (Asher et al., 2020).

Duchenne is a genetic disorder that is caused by mutations in the *DMD* gene that encodes for the protein dystrophin. Dystrophin is important for maintaining healthy muscle fibres and without it, muscles are prone to damage (Figure 2). The mutations that cause Duchenne result in early truncation during protein translation, which leads ultimately to no production of dystrophin. Other mutations in the *DMD* gene that result in shorter, but functional, dystrophin cause Becker muscular dystrophy (BMD). BMD is a less severe form of Duchenne that has a slower progression (Duan et al., 2021).

DMD has X-linked recessive inheritance (Figure 3), which means that the DMD gene is inherited via the X chromosome. As males have only one X chromosome and females have two, Duchenne mostly occurs in males (less than 10 in 100.000 males) and is extremely rare in females (less than 1 in a million females). If females have the mutation in the DMD gene, they are often carriers of Duchenne as they have another X chromosome to encode for dystrophin. Being a carrier does not affect them, but causes the disease in their sons when they pass the mutated X chromosome on to them (Duan et al., 2021).

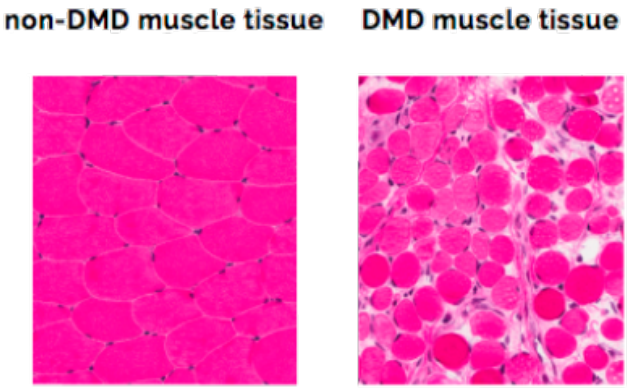



Figure 2: The difference in muscle tissue between healthy and sick patients. Due to the lack of dystrophin in DMD patients, muscle fibres progressively get replaced with fat and scar tissue, which is not as resilient and leads to decreased muscle function. Source: Duchenne Parent Project (2024)

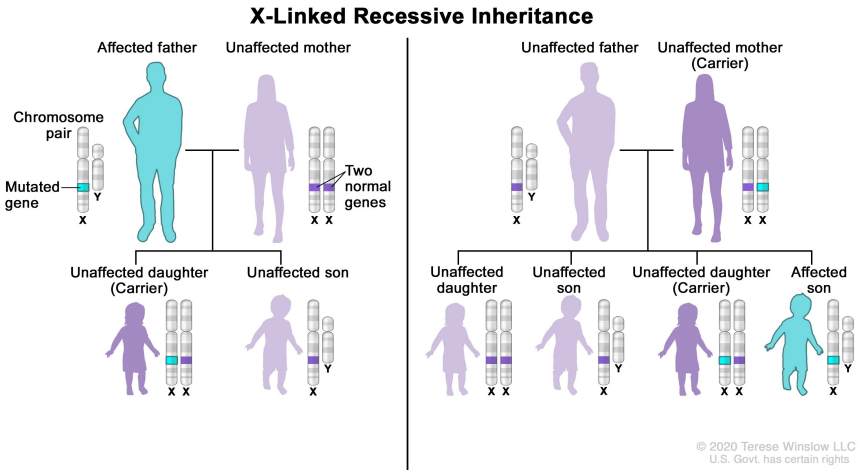



Figure 3: Illustrative representation of the inheritance of Duchenne Muscular Dystrophy. DMD is X-linked recessive, meaning that it is mostly present in males due to the fact that they have only one X chromosome. Source: National Cancer Institute (2024)

### 2.1.2 Biomarkers

Biomarkers are objective measures of all sorts of biological states, processes and responses and there are many different types of biomarkers (FDA-NIH Biomarker Working Group, 2016). For example, the body temperature is a simple diagnostic biomarker as an indicator for fever. Biomarkers can be used to give a diagnosis or track processes, such as diseases. These indicators thus help to develop clinical trials and offer objective assessments (Szigyarto and Spitali, 2018).

Historically, the most direct way to monitor Duchenne is analysing muscle biopsies, from skeletal muscle (Bucciolini Di Sagni et al., 1982; Turk et al., 2006). However, these biopsies are relatively difficult and invasive to obtain, especially since patients are so young, and additionally do not provide overall information of the body's condition. Therefore, less invasive alternatives using biofluids, such as blood or urine, are being researched (Spitali et al., 2018; Signorelli et al., 2021). A common biomarker for Duchenne nowadays is creatine kinase in blood, where elevated levels indicate DMD (Timonen et al., 2019). Also urine can be promising for identifying biomarkers as it is even less invasive to obtain than blood.

### 2.1.3 Metabolites

Metabolites are small molecules that are used or created during metabolic reactions. The metabolism is a collective term for all processes and reactions that are, among others, related to energy production of the human body (Chandel, 2021). As muscles need energy to function, these tissues play a role in the metabolism and are thus related to metabolites.

Urine is a bio fluid produced by the kidneys that filters out an individual's water-soluble waste, such as metabolites. Roughly 5500 metabolites are identified to be excreted through urine (Bouatra et al., 2013).

## 2.2 Available data

Two datasets were available for use in this thesis. Both contain urinary metabolite data, but are otherwise different (Table 1). Nevertheless, both data sets come from research experiments that contribute to DMD research and can therefore be relatively messy. Consequently, the datasets have missing values and require preprocessing and quality checks.

| | Data set 1 | Data set 2 |
|---|---|---|
| Samples | Mice | Humans |
| Type | Longitudinal | Cross-sectional |
| Method | LC-MS | NMR |

Table 1: The two available datasets with their differences. The first data set contains longitudinal data about mice measured using LC-MS. The second data set contains cross-sectional data about humans using NMR.

### 2.2.1 Mouse data

The first data set, containing experimental mouse information, is longitudinal. This means that for each mouse, the same measurements are repeated over time. The data is obtained using Liquid Chromatography-Mass Spectrometry (LC-MS) (Pitt, 2009).

The data contains metabolite measurements from four mouse groups across five time points. These four groups differ in terms of diagnosis, genetic background and number of functional

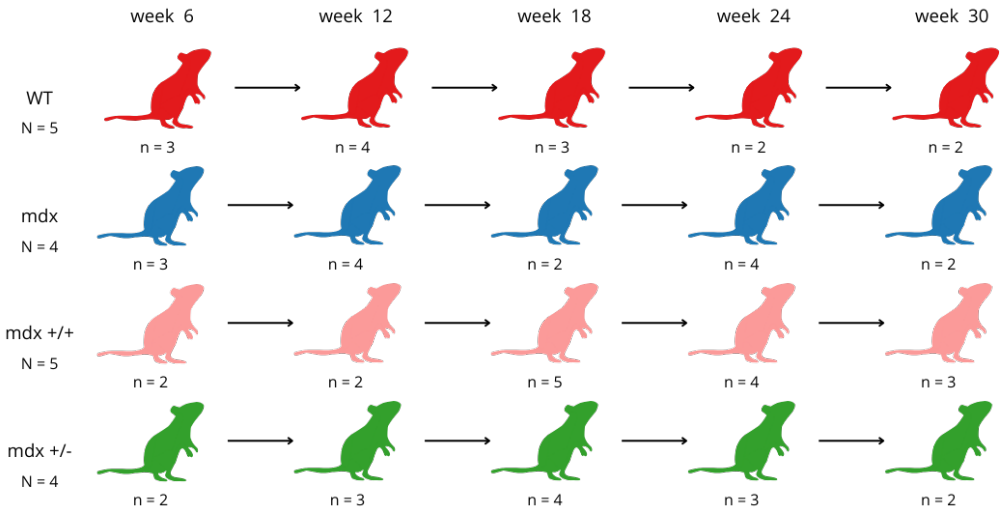



Figure 4: Visualization of the mouse data set and its sample sizes across the different groups and time points. It can be noted that the samples are not evenly distributed, i.e. the data is unbalanced.

utrophin alleles. Utrophin is a protein that is encoded for on a gene closely related to the *DMD* gene. This gene however, unlike the *DMD* gene, is not located on the X chromosome, but on another one for which two copies are available. In *mdx* mice, utrophin is upregulated as a result of the lack of dystrophin and the number of functional alleles has an effect on the diseases severity (Perkins and Davies, 2002).

| Mouse group | Diagnosis | Genetic background | Functional utrophin alleles |
|---|---|---|---|
| Wild type (WT) | Healthy | A | 2 |
| *mdx* | Sick | A | 2 |
| *mdx+/+* | Sick | B | 2 |
| *mdx+/-* | Sick | B | 1 |

Table 2: Overview of the differences between mouse groups. The wild-type mice are healthy and possess two functional alleles for utrophin. All other mice groups (*mdx*, *mdx+/+*, and *mdx+/-*) are sick. The *mdx* mice share the same genetic background with the WT mice and have two functional utrophin alleles. The *mdx+/+* and *mdx+/-* groups have a different genetic background, with two and one functional utrophin alleles, respectively.

In total, the data set has 59 different samples from 18 mice distributed across the different mouse groups and time points (Figure 4). It can be noted that the samples for each group are not evenly distributed across time points, i.e. the data is unbalanced, meaning that some measurements were failed to obtain. This is due to the experimental artifacts, such as lack of urination or early death of the mice. Furthermore, each sample has measurements for 456 metabolites. The metabolites are not labelled with a name, but are numbered.

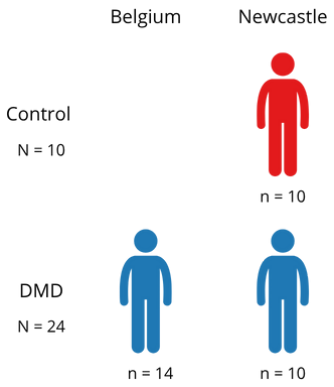



Figure 5: Visualization of human data set and its sample sizes across different groups. It can be noted that there is no control group for patients from Belgium.

### 2.2.2 Human data

The second data set contains data about humans, specifically Duchenne patients and age-matched unaffected boys. This data set is cross-sectional, meaning that the measurements are taken at one time point, unlike the mouse data. The human data is obtained using Nuclear Magnetic Resonance (NMR) Spectroscopy (Gerothanassis et al., 2002).

The data set has 34 different samples distributed across the different groups, control and DMD, and countries, Belgium and Newcastle (Figure 5). It is important to note that the data does not have a Belgian control group. This data set contains measurements for 54 metabolites, however some samples have some missing values.

## 2.3 Research questions

The goal of this thesis is to assess the feasibility of urine for potential, new biomarkers for Duchenne Muscular Dystrophy. Therefore, the main research question to be answered is:

RQ1) Are there metabolites whose expression level differ between healthy and dystrophic individuals?

To help answer this research question, two other questions will be answered.

RQ2) Are there metabolites that behave differently among the different mouse groups and if yes, how do they differ?

RQ3) Are there metabolites that behave differently between the healthy and DMD patients?

For RQ2, we plan to test hierarchically, meaning that we only test the difference between, e.g., *mdx* and *mdx+/+* mice if there is a difference between WT and *mdx* mice.

# 3 Modelling

This chapter describes how the analysis pipeline for this thesis is done and, with that, how the research questions will be answered. The first section describes data exploration, then a section about the normalization of the data follows. After this, the next section explains the statistical

models that are fitted on the data. The final section describes what hypotheses are tested on the models to ultimately answer the research questions.

The research questions described in Section 2.3 are used to find out whether there is potential in using urine for biomarkers to indicate DMD. The questions refer to the difference in expression levels of metabolites for different groups.

## 3.1 Exploration

In the analysis of metabolomic data, it is customary to perform some preliminary exploratory analysis before proceeding to the modelling phase. We do this to identify possible problems in the experimental data and outliers.

### 3.1.1 Mouse data

As described earlier in Section 2.2.1, the mouse data set has measurements from 18 different mice from 4 mouse groups on 5 different time points for 456 metabolites. As we do not know the name of the metabolites, it is not possible to comments on its biological features later on. We already noted that the data is unbalanced, which is possibly due to factors related to the fact that obtaining this data was not the main objective of the experiment or to the fact that obtaining urine was not always possible. From now on, the actual values about the metabolites obtained from the experiment are mentioned as measurements.

To give an initial look on the data, we create some plots presenting the measurements for the metabolites. To do this we first add a new column containing the total value of all metabolites per sample. These values are then used to make a boxplot and a trajectory plot.

Next, to identify patterns or outliers, we perform a Principal Component Analysis (PCA) (Greenacre et al., 2022). Using PCA, high-dimensional data can be reduced into less, new variables called principal components that contain, relatively, a lot of information. This way, we can visualize our data in two dimensions instead of needing over 450.

Finally, we analyse the contribution of individual metabolites to see if there are any outliers among the samples. We create two plots for this; one displaying the contribution for all metabolites per sample and one displaying the relative contribution to the total per metabolite.

### 3.1.2 Human data

The human data set has 34 samples from either the control or DMD group and either Belgium or Newcastle, as described in Section 2.2.2. We already noted that there is no Belgian control group available in the data. We did a similar exploration for the human data as we did for the mouse data.

Similarly to the mouse data, we create a boxplot to give an initial look on the human data. We also added a column with the total value for each sample to this dataset, which is used in creating this boxplot.

Then we perform the same PCA for the human data, as well as the metabolite contribution analysis.

### 3.1.3 Missing value analysis for human data

Unlike the mouse data, the human data has some missing values (NA's) among the measurements. We are interested in whether these missing values would have some sort of pattern that would indicate them not missing at random, but containing some relevant information. It could happen, for example, that the measuring equipment has a certain threshold that would make either or both extremely small and large values disappear. Or that missing values only appear in certain groups, which would obviously be a strong indicator for a potential biomarker. To study this, we do a missing value analysis on the human data set.

To analyse the first possible explanation for the missing values, regarding the possibility of an equipment threshold, we create a scatter plot showing the number of NA's against the median measurement value on a logarithmic scale per metabolite.

The second explanation for missing values, where values would not be missing at random, will be analysed by doing some statistical tests. We use the chi-square goodness-of-fit test to test if the missing values are distributed according to either group size or NA's per group. This test is a statistical test to assess whether a variable, in this case the number of missing values, aligns with an expected distribution. The null hypothesis to be tested here, would be that the variable indeed follows the expected distribution. The chi-square test statistic is then defined as follows,

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

where

- $k$ is the number of groups,
- $O_i$ is the observed frequency for group $i$,
- $E_i$ is the expected frequency for group $i$.

This test statistic follows the chi-squared distribution with $k-1$ degrees of freedom ($df$) under the null hypothesis. So, we can find the corresponding p-value by calculating,

$$p = P(\chi^2_{(\alpha,df)} > \chi^2 \mid H_0).$$

By comparing this p-value to our significance level $\alpha = 0,05$, we can either reject or accept the null hypothesis.

Here, we want to test two null hypotheses for each metabolite. The first null hypothesis says that the NA's are distributed according to the sample size per group. The expected frequency for group $i$ under this null hypothesis is:

$$E_{i,(1)} = \frac{N_i}{N},$$

where

- $N_i$ is the sample size for group $i$,
- $N$ is the total sample size in the study.

The second null hypothesis says that the NA's are distributed according to the total number of NA's per group. The expected frequency for group $i$ under this null hypothesis is:

$$E_{i,(2)} = \frac{\#\mathrm{NA}_i}{\#\mathrm{NA}},$$

where

- $\#\text{NA}_i$ is the total number of NA's for group $i$, so the number of NA's for group $i$ summed up for each metabolite,
- #NA is the total number of NA's in the study.

## 3.2 Normalization

Usually after exploring the data, a normalization is done to make data comparable and consistent. Especially for this urine data, which can be very unstable, it is important to scale the measurements. As we decided not to continue analysing the human data, for which the reasoning can be found in Section 4.1.2, we from now on only focus on the mouse data set.

We normalize the mouse data using Probabilistic Quotient Normalization (PQN) (Dieterle et al., 2006). Urine can be a very variable fluid across samples, as it depends highly on dilution, and thus the water intake of the individual, and the time of obtaining it (Bottin et al., 2016). This is why it is important to normalize it. PQN is a very robust and accurate method when normalizing such complex biofluids containing metabolites. PQN scales data based on a a most probable dilution factor. Unlike other widespread normalization methods, PQN does not assume that the total metabolite concentration is constant across samples, which makes it a more suitable method for urine.

## 3.3 Model specification

When dealing with longitudinal data, linear mixed models (LMMs) are a good option to use for analysing. Longitudinal data is usually correlated between measurements of the same sample and a LMM can handle this more properly. A linear mixed model can be seen as an extension of the well-known linear regression model that has more flexibility in terms of assumptions.

The general formula for a LMM estimating $y_{ij}$, the response variable of the i-th sample on the j-th time point, is as follows,

$$y_{ij} = x_{ij}^\top \beta + z_{ij}^\top u_i + \varepsilon_{ij}, \tag{1}$$

where

- $x_{ij}$ and $z_{ij}$ are covariate vectors, where usually the covariates for the random effects $z_{ij}$ are a subset of the covariates for the fixed effects $x_{ij}$,
- $\beta$ is a vector of fixed-effects parameters,
- $u_i$ is a vector of random-effects parameters that is distributed according $N(0, D)$, a normal distribution around zero,
- $\varepsilon_{ij}$ is a vector of independent errors that is also distributed according a normal distribution around zero.

It is assumed that both the errors and the random-effects parameters are independently distributed and that they are mutually independent as well.

When fitting a model, the goal is to find the optimal parameters that minimize the likelihood function. In simple equations, this can be done by hand, but for a LMM, this is often done computationally, for example, in R. We indeed used R with the default optimizer from the `lme4` package.

### 3.3.1 Linear mixed models for mouse data

Since the mouse data is longitudinal, we used a LMM on it. In this case, we wanted to fit a LMM for each metabolite to estimate its expression level; which would result in a total of 456 models.

However, as there are *only* 59 samples available, it can happen that while finding the optimal parameters, the process to find a fit does not fully succeed. It can fail to converge to a solution, returning a failed fit, or there is little to no variance in the data and the returned fit is singular. When this happens, we attempt to fit another, simplified version of the original model instead.

The model tried first on each metabolite, which is thus the most extensive one, estimates the expression level $y$ of the i-th mouse at time point j for metabolite r. The model is as follows,

$$y_{rij} = \beta_{r0} + \beta_{r1}a_i + \beta_{r2}b_i + \beta_{r3}c_i + \beta_{r4}t_{ij} + \beta_{r5}a_i t_{ij} + \beta_{r6}b_i t_{ij} + \beta_{r7}c_i t_{ij} + u_{ri0} + u_{ri1}t_{ij} + \varepsilon_{rij}, \quad (2)$$

where

- $\beta_{r0}, \beta_{r1}, \beta_{r2}, \beta_{r3}, \beta_{r4}, \beta_{r5}, \beta_{r6}$ and $\beta_{r7}$ are the fixed effect parameters,
- $a_i, b_i$ and $c_i$ are the binary covariates for the group of mouse $i$,
- $t_{ij}$ is the covariate for time $j$ for mouse $i$,
- $u_{ri0}, u_{ri1}$ are the random effect parameters for the $r$-th metabolite for mouse $i$ which are both independently distributed across mice according $N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{u0} & \sigma_{u01}\\ \sigma_{u01} & \sigma^2_{u1}\end{bmatrix}\right)$,
- $\epsilon_{rij}$ is the error for the $r$-th metabolite for mouse $i$ at time $j$ which is independently distributed across mice and mutually independent from the random effect parameters according $N(0, \sigma^2_\epsilon)$.

The covariates $a_i, b_i$ and $c_i$ are binary and indicate the group for mouse $i$. As they are binary, we can describe four groups with three covariates. If all covariates are 0, we fall into the fourth group, called the reference group, which is in this case the group with WT mice. Furthermore, $a_i$ refers to the group with *mdx* mice, $b_i$ to the group with *mdx+/+* mice and $c_i$ to the group with *mdx+/-* mice. The covariate for time is modelled continuous.

As said, R was used to fit the models to the data. For 190 metabolites, the most extensive model (2) is fitted successfully. For 227 metabolites, we fit a more simplified model that is the same as the more extensive model (2), but without the term for the random slope; $u_{ri1}t_{ij}$. The simplified model is as follows,

$$y_{rij} = \beta_{r0} + \beta_{r1}a_i + \beta_{r2}b_i + \beta_{r3}c_i + \beta_{r4}t_{ij} + \beta_{r5}a_i t_{ij} + \beta_{r6}b_i t_{ij} + \beta_{r7}c_i t_{ij} + u_{ri0} + \varepsilon_{rij}. \quad (3)$$

For the remaining 39 metabolites, no model is fitted and we leave them out of the further analysis.

For the first metabolite, we were able to fit the model with just the random intercept, so the simplified model. R fitted this model on the data points, resulting in four lines predicting the metabolites expression per group (Figure 6).

## 3.4 Hypothesis specification

To test hypotheses on a LMM, one option is using so-called F-tests. F-tests can be used when testing hypotheses related to just the fixed effect parameters.

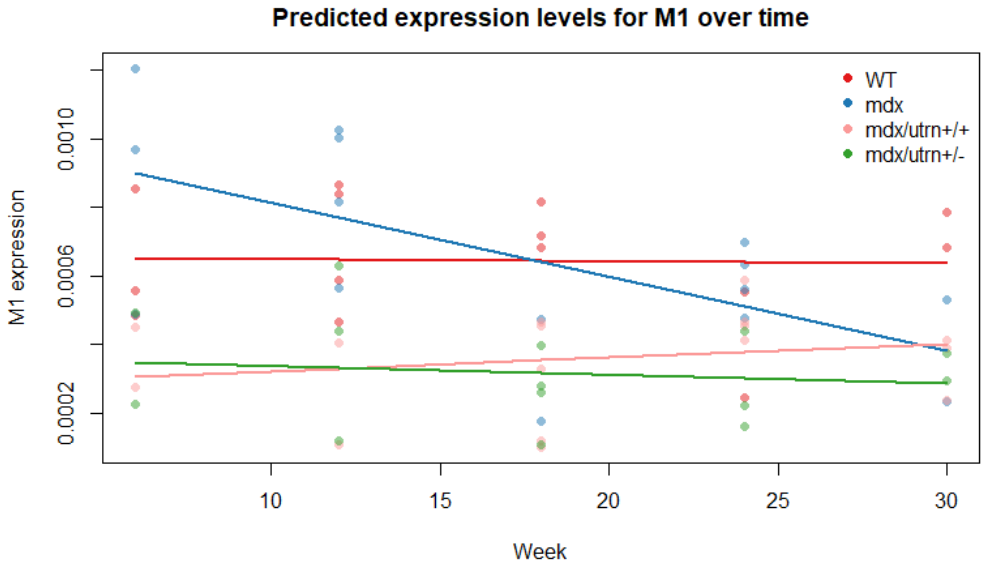



Figure 6: Predicted expression for metabolite 1. The model fitted on this data was the simplified model with just the random intercept.

The general notation for the null hypothesis of an F-test uses a matrix notation and is as follows,

$$H_0 : K\beta = k, \tag{4}$$

where

- $K$ is a contrast matrix specifying linear combinations of the fixed effect parameters to be tested,
- $\beta$ is the vector of fixed effect parameters,
- $k$ is a vector of constants specifying the null hypothesis values for those linear combinations.

We use Satterthwaite's method for approximating the degrees of freedom.

### 3.4.1 Null hypotheses for mouse data

To answer the research question regarding the different mouse groups, we want to test for a significant difference between these groups. To do that, it is sufficient to look at and compare the appropriate fixed effect parameters. To compare all four groups, there are six possible unique pairs to compare, listed below.

T1. WT and *mdx*

T2. WT and *mdx+/+*

T3. WT and *mdx+/-*

T4. *mdx* and *mdx+/+*

T5. *mdx* and *mdx+/-*

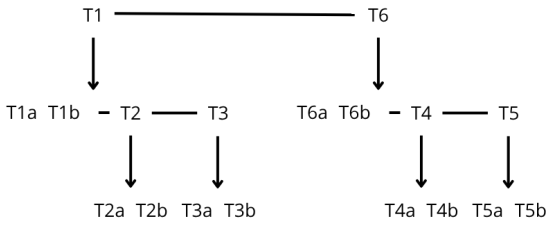



Figure 7: Schematic overview of the hierarchical structure of the hypothesis testing. The nodes correspond with testing the six pairs of groups and the notions of "a" and "b" refer to their sub-hypotheses.

T6. *mdx+/+* and *mdx+/-*

In our model, comparing between two groups means adjusting some of the values of the $\beta$'s such that the equations for both groups turn out the same. For example, when we want to compare the first pair, we set $\beta_{r1} = 0$ and $\beta_{r5} = 0$ as the null hypothesis. Using the notation for an F-test introduced earlier, we would have the following null hypothesis,

$$H_0 : \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \vdots \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5}$$

Similarly, we have five other null hypothesis for comparing the other five pairs, which can be found in full in the Appendix A.1.

For the metabolites that have a significant difference between two groups, we also want to find what that difference is. Specifically; does the intercept and/or the slope of the estimated expression differ? To do this, we can test the same $\beta$'s, but instead of together, we test them separately. The lower $\beta$ of the two tests for the intercept while the higher one tests for the slope. From now on, these separate tests are referred to as the sub-hypotheses. It can happen that while the test for any difference is significant, we can not find what the difference is exactly. This happens sometimes when there is too little data and the test for the sub-hypothesis becomes less powerful.

### 3.4.2 Hierarchical testing

Sometimes we are only interested in testing a pair if another pair is significant. This is why we do the hypothesis testing hierarchically. Using the numbering of the pairs we introduced earlier and referring to the sub-hypotheses using a and b, we created a tree that indicates in what order to test (Figure 7). The tree has two main branches, where the next comparison for a certain metabolite is done if the previous node has a significant result after applying a multiple testing correction. The first branch indicates that we first test hypothesis 1 and if that is significant we test hypotheses 1a and 1b and 2 and 3. Then, if 2 or 3 have a significant result, we test for their sub-hypotheses. The same goes for the second branch but instead of 1, 2 and 3 we test for hypotheses 6, 4 and 5.

### 3.4.3 Multiple testing correction

When doing multiple hypothesis tests simultaneously, some of the tests might give a false positive result. This kind of error is called a type I error and is more probable to happen the more tests

you do. With a significance level ($\alpha$) of $0,05$, there is a 5% chance of a false positive occurring. This means that for the mouse data, where we do over 400 tests simultaneously, there would be roughly 20 false positive results, which is a pretty high amount.

To avoid this, we apply a multiple testing correction. This adjusts all p-values to reduce the false discovery rate. We used the Benjamini-Hochberg procedure to do this (Li and Barber, 2018). The BH procedure compares each p-value to its calculated threshold and says its significant when it is smaller than the threshold.

Looking at the tree describing the order of the hypothesis testing, we apply the multiple testing correction in different ways. For hypotheses 1 and 6 we correct vertically, meaning that we correct for the same test over multiple metabolites. For hypotheses 1a and 1b, 2 and 3, 2a and 2b, 3a and 3b, 6a and 6b, 4 and 5, 4a and 4b and 5a and 5b we correct horizontally, meaning that for all these pairs we correct for the same metabolite over the two tests done simultaneously.

# 4 Results

This chapter presents the findings of the data analysis done for both datasets. It is organized in the different steps, similar as described in Chapter 3.

## 4.1 Exploration

The data exploration is done to possibly identify outliers or issues of the data. This section presents the results of the exploration, splitted in sections for the mouse and human data.

### 4.1.1 Mouse data

This section describes the exploration done for the mouse data, explained in Section 3.1.1. First, we created two initial explorative figures. The first plot is a boxplot comparing the four different mouse groups (Figure 8). What can be seen is the total value of all the measurements per sample. It can be noted that the average values do not really differ among the four groups, but there are some outliers in the *mdx+/-* group that are relatively large. Furthermore, a trajectory plot is made to explore the data (Figure 9). Again, the total values for each sample are displayed. Samples from the same mouse at different time point are connected by a trajectory. So, following a line from a data point, leads to the next measurements for the same mouse at the next time.

Then we performed a PCA and created a plot with the first two principal components (Figure 10). Each sample is given a score for both components and is coloured and shaped according its genotype and time point respectively. The first two principal components explain roughly 46% of the variance in the data, which means that they capture almost half of it. We can see some clustering of the genotypes, where the WT mice are clustered and *mdx* mice seem to share features across all groups.

Next, we analysed the overall metabolite contribution to check for unexpected measurements. We did this by creating two plots. The first plot shows a bar for each sample where the length and colour of each segment in the bar represents the proportional contribution of one metabolite to that sample (Figure 11). The metabolites are sorted according to their mean abundance, with the highest abundance on the right of the bar and the lowest on the left. Only the segments with a contribution higher than 3% are displayed to improve readability. It can be seen that all bars have similar order of colouring, meaning that there are no samples that have unexpectedly high or low metabolite measurements. Furthermore, there are no metabolites that are contributing

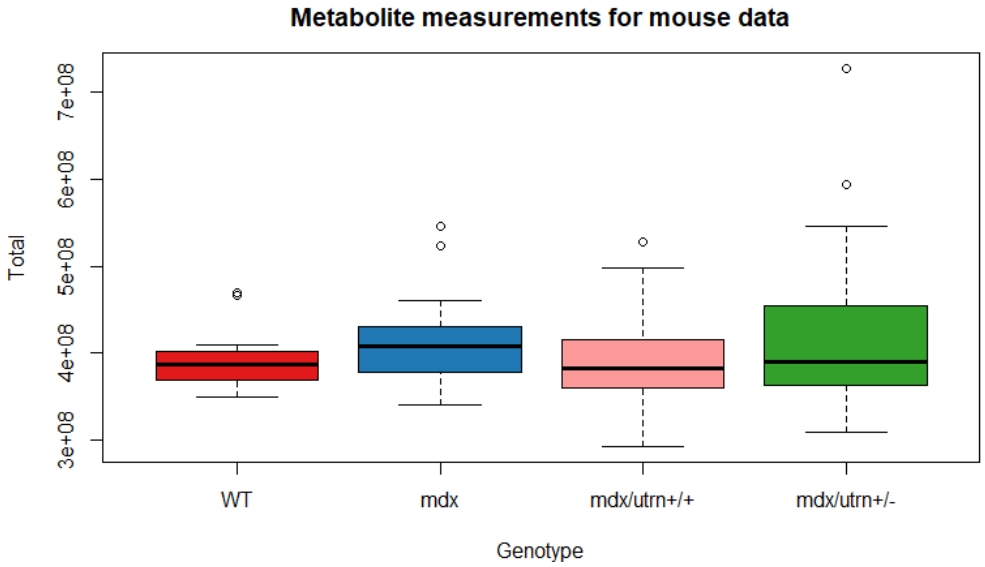



Figure 8: Boxplot of the mouse data. For each sample, the total value of all the metabolite measurements is calculated and this is what is displayed.

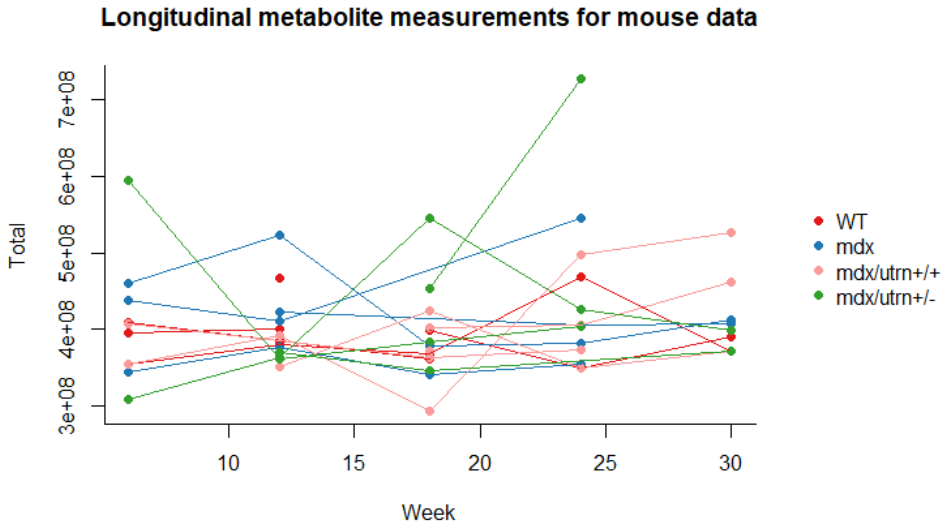



Figure 9: Trajectory plot of the mouse data. For each sample, the total value of all the metabolite measurements is calculated and this is what is displayed. Samples from the same mouse are connected by a trajectory.

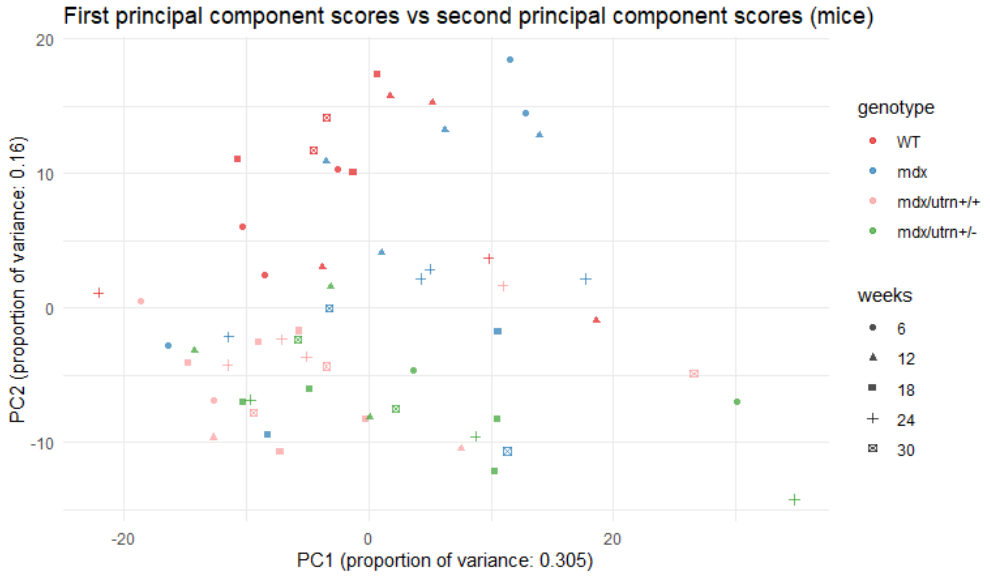



Figure 10: First two principal components of the mouse data capturing 46%. Samples of the data are coloured and shaped according their genotype and time point of measurements. Some clustering in the genotypes is visible; WT mice are clustered, while *mdx* mice are more or less scattered across the plot.

extremely high in any of the samples, as there are no metabolites having higher contribution than 7%. The second plot displays the relative abundance for each metabolite to the total abundance of the means of all metabolites (Figure 12). Each bar represents one metabolite and they are sorted in ascended order from top to bottom. From this plot it can again be seen that there are no metabolites that have extremely high measurements, because there are no metabolites with a significantly higher contribution in comparison with the rest.

### 4.1.2 Human data

The exploration as described in Section 3.1.2 is done on the human dataset. Also for this data, we made an initial boxplot for the human data (Figure 13). This one however, does show some difference between the control and Duchenne group. We can see that the average total values for the control group are somewhat higher than for the DMD group.

Then we performed a PCA, which resulted in a visualization of the first two principal components (Figure 14). These two components together explain about 41% of the human data in the plot. This plot does not show any patterns in the data, but it does reveal some outliers. We can see some points outside the region around zero along PC2, where the majority of the points are. To further study these outliers and what might have caused them, we look into the metabolite contribution and missing values of the dataset.

Next, we did the metabolite contribution analysis for the human data, making the same two plots as we created for the mouse data. The first plot, showing the metabolite contribution per sample, shows that the samples have an uneven partition of metabolites, because not all bars have similar colouring (Figure 15). This means that there is a difference is metabolite expression across samples, but there is no clear distinction between groups. The second plot, with the

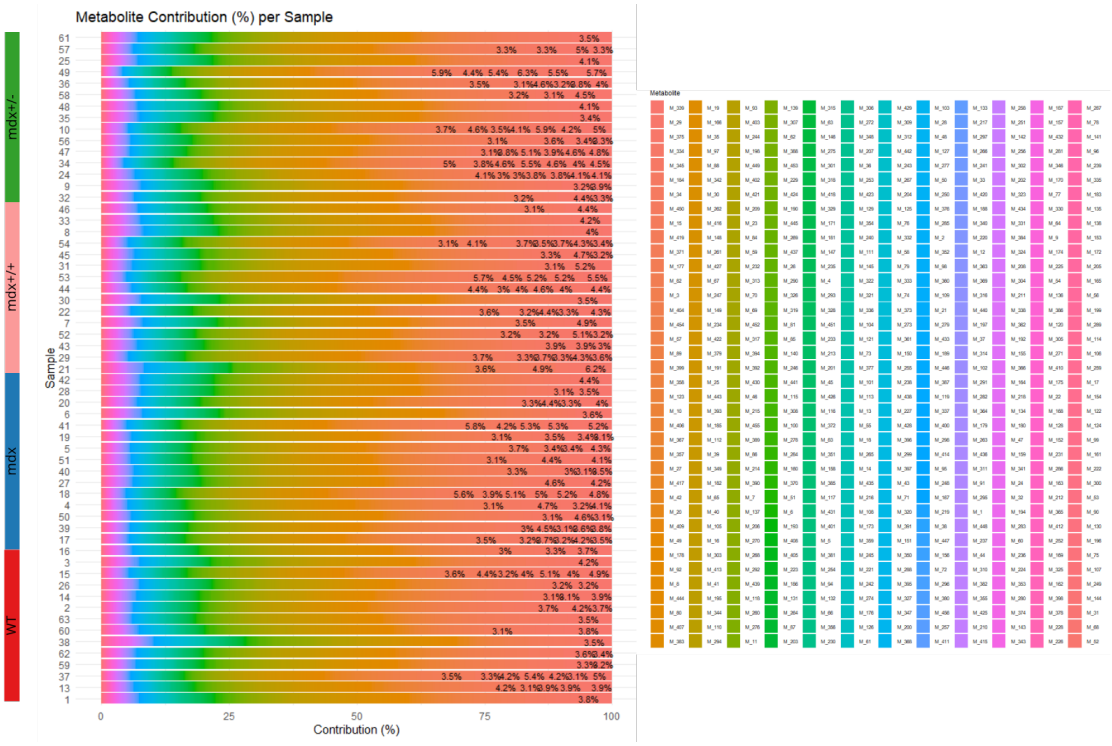



Figure 11: Metabolite contribution per sample for mouse data. Each bar represents a sample and is divided into segments that each represent one metabolite. The segments are sorted according to the mean abundance of the metabolites. There are no samples that have unexpected partition or extremely high metabolite contributions.

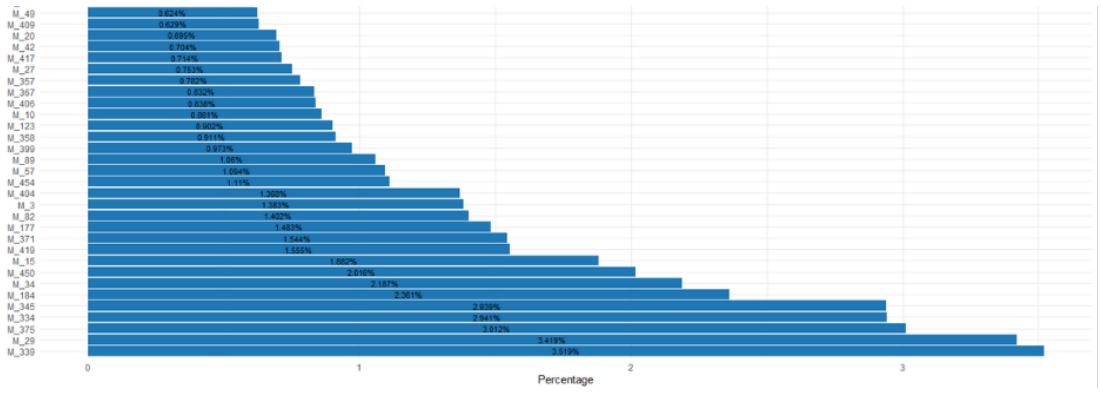



Figure 12: Bottom part of the plot of metabolite contribution per metabolite for mouse data. Each bar represents the relative contribution for one metabolite to the total of the means of all metabolites. The bars are sorted in ascending order based on this contribution. From this plot, it can be noted that there are no metabolites with a significantly high contribution.

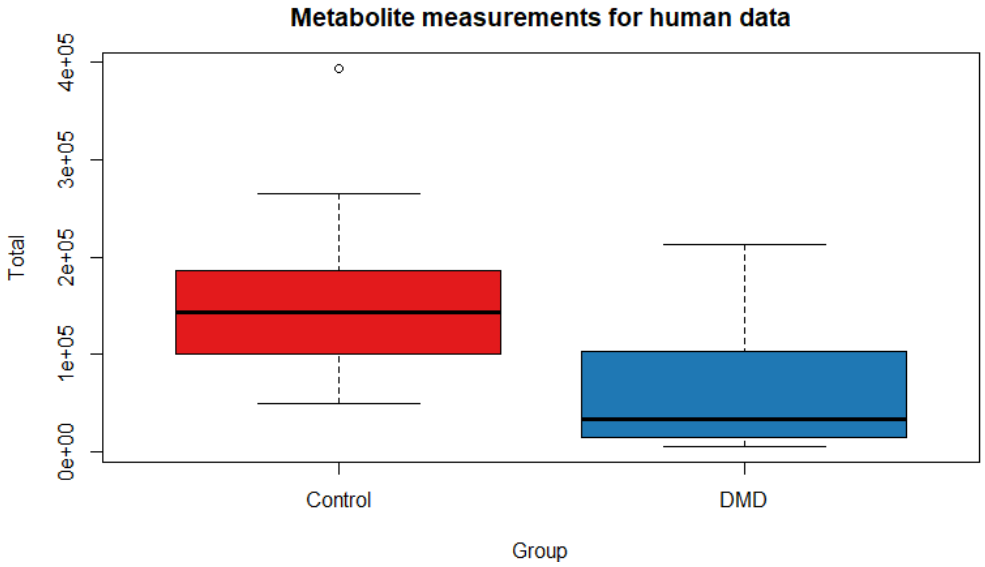



Figure 13: Boxplot of the human data. For each sample, the total value of all the metabolite measurements is calculated and this is what is displayed.

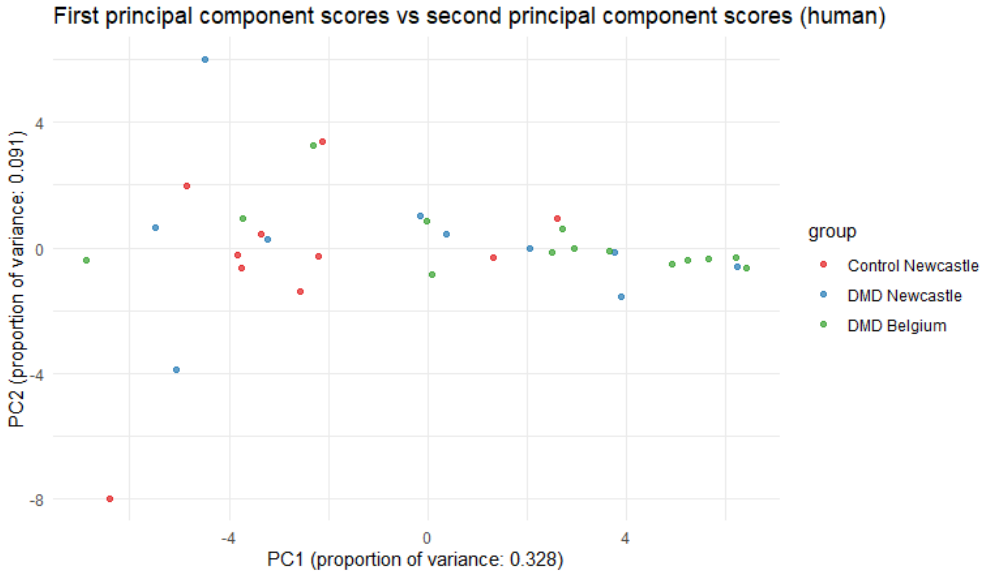



Figure 14: First two principal components of the human data capturing 41%. Samples of the data are coloured according their group. The plot reveals some outliers along the second principal component.


Figure 15: Metabolite contribution per sample for human data. Each bar represents a sample and is divided into segments that each represent one metabolite. The segments are sorted according to the mean abundance of the metabolites. Some uneven colouring is displayed across the samples, meaning that their metabolite expression differs.

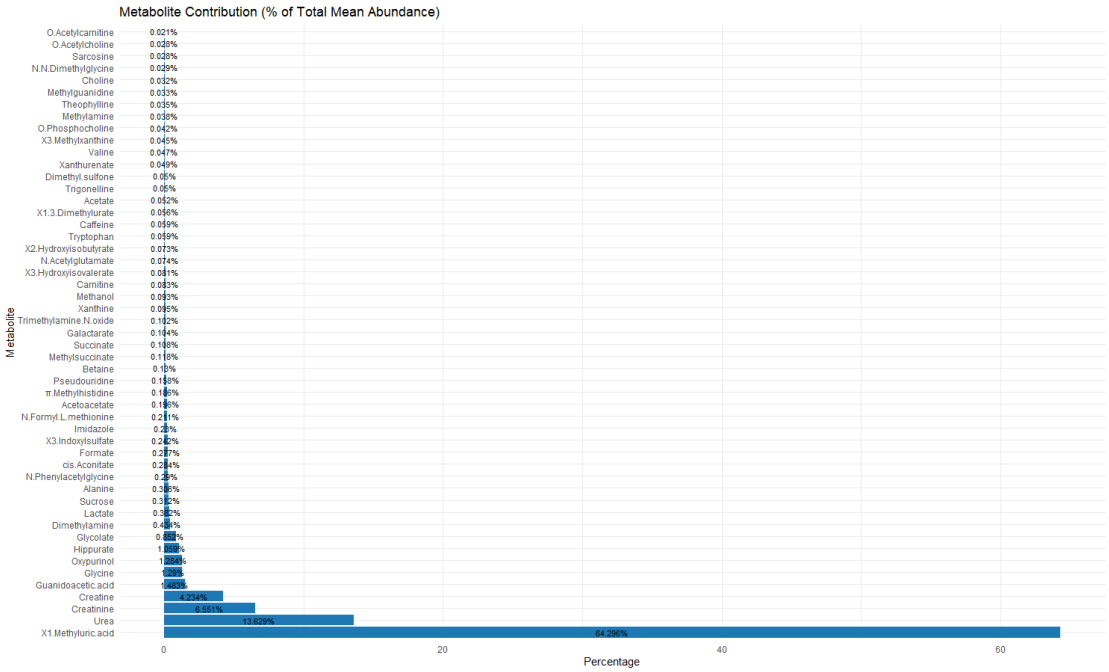
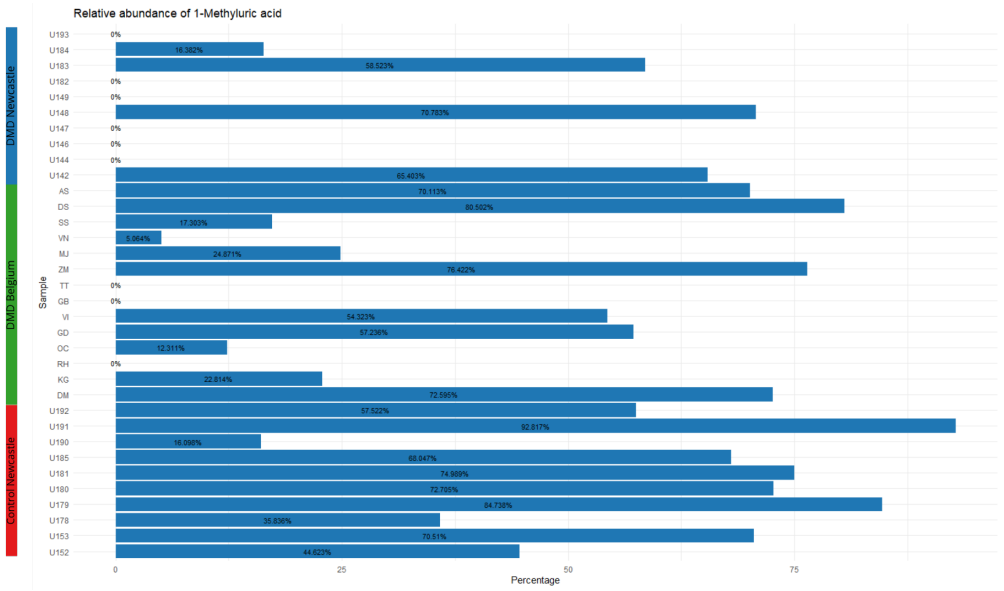


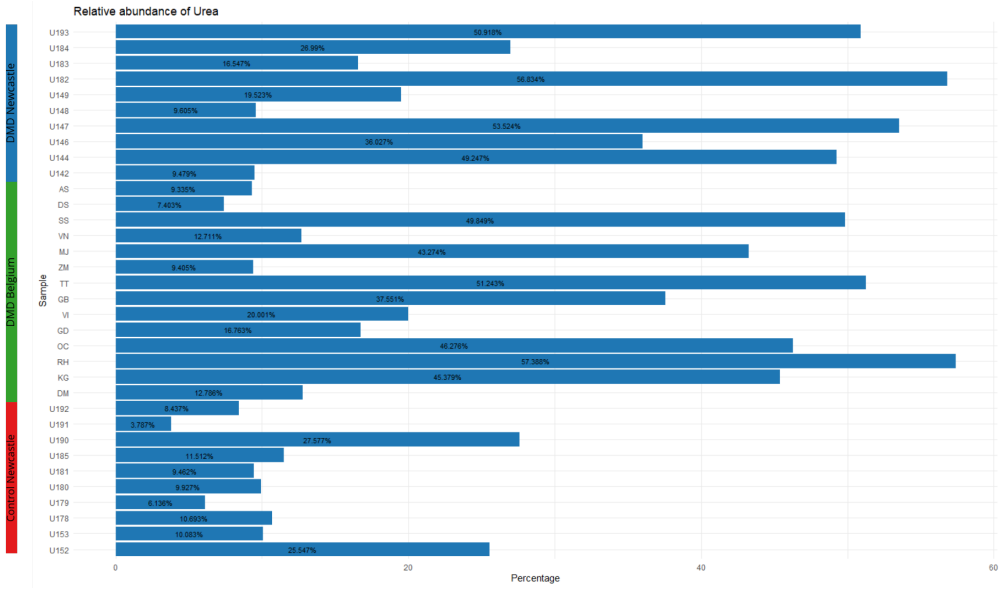

Figure 16: Metabolite contribution per metabolite for human data. Each bar represents the relative contribution for one metabolite to the total of the means of all metabolites. The bars are sorted in ascending order based on this contribution. This plot shows four metabolites having a significantly higher contribution than the rest.

relative contribution to the total contribution of all metabolites, shows four metabolites with a significantly larger contribution than the rest of the metabolites (Figure 16). 1-Methyluric acid, urea, creatinine and creatine all contribute more than 4%, while all the others contribute less than $1,5\%$. 1-Methyluric acid even has a contribution of 64%, meaning that it accounts for more than half of the average sample's urine. For these four metabolites, we looked more into the relative abundance for each sample. To do that, we made four plots for these metabolites with the contribution of that metabolite for each sample with respect to the total value of all metabolite measurements (Figure 17). 1-Methyluric acid, urea and creatinine show no clear distinction between groups; some groups have samples that match samples in other groups, which makes the group show no obvious difference. However, the plot for creatine does show a pattern. The control group shows significantly less abundance than the Duchenne group, meaning that creatine is mostly present in sick patients. Creatine is also a proven to significantly differ between healthy and DMD patients in blood, as well as its ratio with creatinine (Boca et al., 2016).

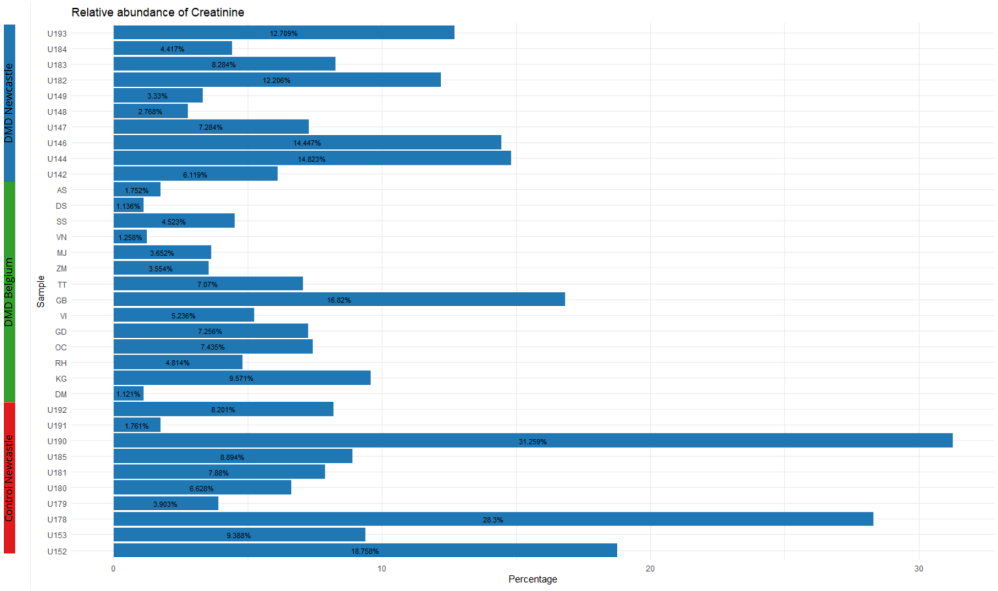
Furthermore, we did a missing value analysis on the human data, because we are interested in whether the missing values are not missing at random. We introduced two explanations for a possible pattern in Section 3.1.3. For the first explanation, we made a scatter plot. This plot shows the number of NA's against the median measurement value on a logarithmic scale per metabolite (Figure 18). It can be seen that there is a slight trend that the lower the mean abundance, the higher the number of missing values. This might be an indicator that there is a certain small threshold for the measuring equipment. But, there is also a point, 1-methyluric acid, that does not fit in this trend. This might just be random or, again, be an indicator for


Relative abundance of 1-Methyluric acid
DMD Newcastle
DMD Belgium
Control Newcastle
Sample
U193 0%
U184 16.382%
U183 58.523%
U182 0%
U149 0%
U148 79.783%
U147 0%
U146 0%
U144 0%
U142 65.403%
AS 70.113%
DS 80.502%
SS 17.302%
VN 5.064%
MJ 24.871%
ZM 76.422%
TT 0%
GB 0%
VI 54.323%
GD 57.236%
OC 12.311%
RH 0%
KG 22.814%
DM 72.595%
U192 57.522%
U191 92.617%
U190 16.098%
U185 68.047%
U181 74.989%
U180 72.705%
U179 84.738%
U178 35.836%
U153 70.51%
U152 44.623%
0
25
50
75
Percentage

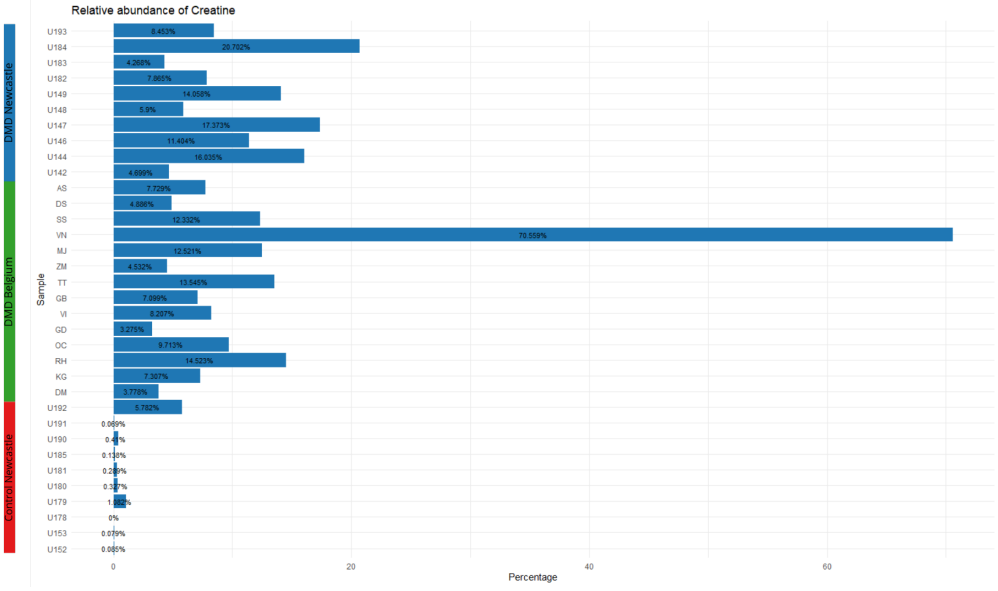
(a)


Relative abundance of Urea
DMD Newcastle
DMD Belgium
Control Newcastle
Sample
U193 50.918%
U184 26.99%
U183 16.547%
U182 56.834%
U149 19.523%
U148 9.605%
U147 53.524%
U146 36.027%
U144 49.247%
U142 9.479%
AS 9.335%
DS 7.403%
SS 49.849%
VN 12.711%
MJ 43.274%
ZM 9.405%
TT 51.243%
GB 37.551%
VI 20.001%
GD 16.763%
OC 46.276%
RH 57.388%
KG 45.379%
DM 12.786%
U192 8.437%
U191 3.787%
U190 27.577%
U185 11.512%
U181 9.462%
U180 9.927%
U179 6.136%
U178 10.693%
U153 10.083%
U152 25.547%
0
20
40
60
Percentage

(b)




(c)




(d)

Figure 17: Relative abundance per sample for the four metabolites that contribute most to the average urine sample. (a) 1-Methyluric acid, (b) urea and (c) creatinine do not show any clear distinction in abundance for the different groups. (d) Creatine, however, does show a clear difference. The healthy group shows significantly less abundance than the DMD group. This corresponds with the findings from research to biomarkers for DMD in blood.

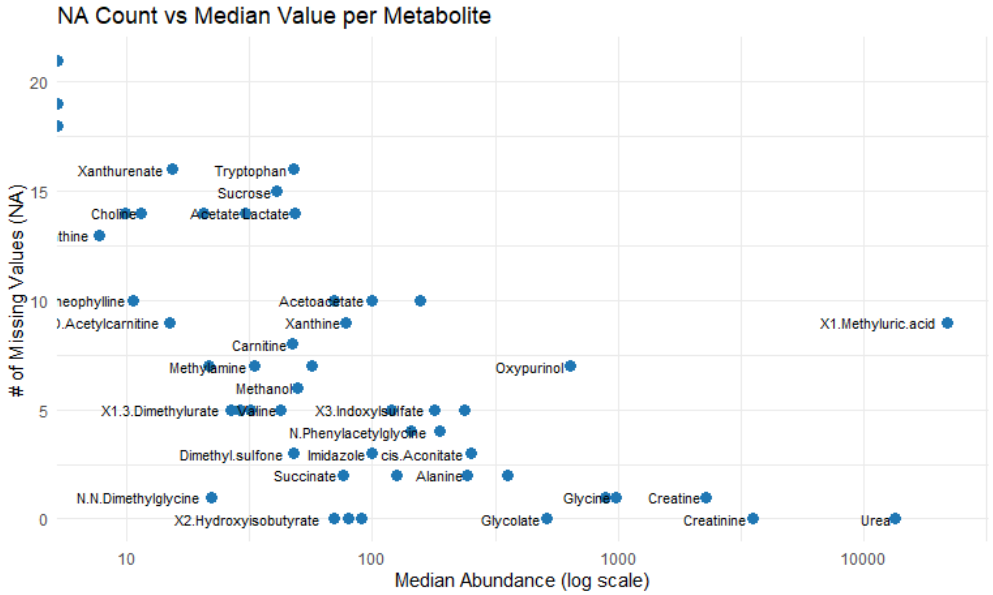



Figure 18: Scatter plot of number of missing values and median value per metabolite. We can see a general trend that the smaller the median value, the more missing values.

the other side of the equipment threshold.

For the second explanation, we performed two chi-square goodness-of-fit tests per metabolite; we test the goodness-of-fit of number of NA's for that metabolite based on sample size per group and the goodness-of-fit of number of NA's for that metabolite based on total number of NA's per group. For both test, only 1-methyluric acid gives a significant result, $p = 0.0298$ and $p = 0.0239$ respectively. This means that 1-methyluric acid is the only metabolite where NA's may be distributed according to the groups. However, after doing a multiple testing correction, we find no significant metabolites. This means that we cannot reject the null hypotheses and thus that the NA's are not distributed according to the sample sizes or total number of NA's per group.

In combination with the metabolite contribution analysis done earlier, we decided not to continue analysing the human data further. This is because for the metabolite that accounts for most of the average urine sample, there are relatively a lot of missing values; we can not find a solid reason for why 1-methyluric acid has so many NA's in our data. Furthermore, for urea, the second most abundant metabolite, there are no missing values. But, its relative abundance per sample plot does not have any clear distinction, as there are some samples with low abundance in every group (Figure 17b). The same goes for the third metabolite creatinine (Figure 17c). This means that for roughly 83% of the average urine sample, we would have to make a lot of assumptions before properly modelling the data. This would make the analysis very limited in terms of validity.

Despite not analysing the data further, the metabolite contribution analysis shows that the metabolite creatine is significantly more abundant in urine of Duchenne patients. Looking from a biological perspective, this makes a lot of sense. Creatine is a metabolite stored in muscle tissue. When the muscles break down, due to Duchenne, creatine enters the circulation of the

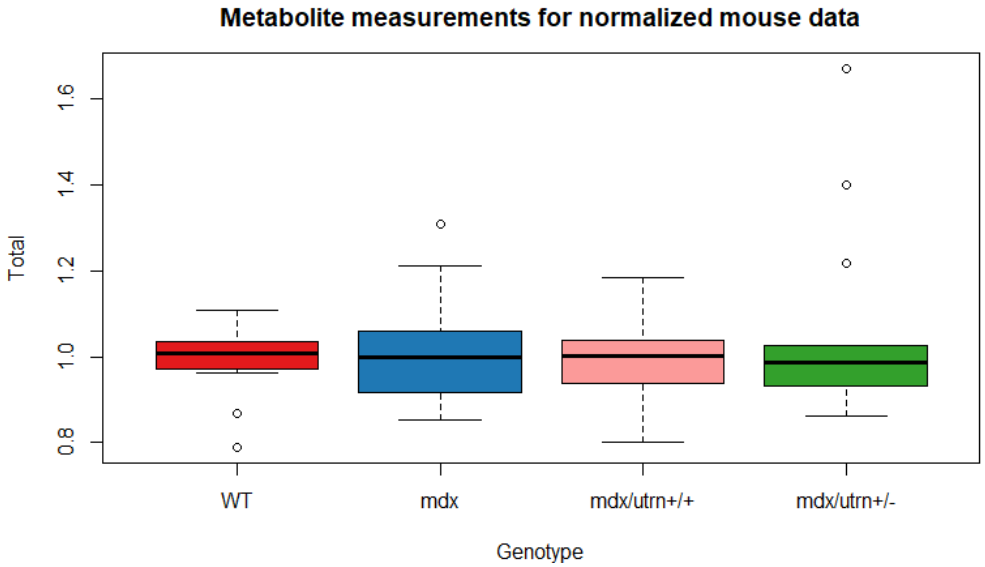



Figure 19: Boxplot for the normalized mouse data. For each sample, the total value of all the metabolite measurements is calculated and this is what is displayed.

body and ends up in urine, leading to elevated levels in DMD patients. Creatine is therefore a highly potential biomarker for Duchenne in urine.

## 4.2 Normalization

We normalized the mouse data using PQN and to have a look at this scaled data, we recreated some of the exploration plots. First, we made a boxplot (Figure 19). In comparison with the boxplot before normalization, the averages are even closer together. However, there are still outliers in the *mdx+/-* group (Figure 8).

We also recreated the principal component plot (Figure 20). This one looks very similar in comparison with the plot before normalization, so the pattern we observed earlier still is visible (Figure 10). Thus, there is some clustering of the genotypes; WT mice clustered and *mdx* mice are on the other hand distributed across the whole plot.

## 4.3 Hypothesis testing on mouse data

As described in Section 3.4, we tested some hypotheses hierarchically on the mouse data. These hypotheses were tested in order according to the tree (Figure 7). We first look into the results of the left main branch and then at right main branch.

The first hypothesis to test on the left branch of the tree is hypothesis 1, testing for a difference between WT and *mdx* mice. We find five significant metabolites for this hypothesis after correcting. These are metabolites 130, 264, 274, 296 and 381. For these five metabolites, we test the next nodes being its sub-hypotheses and hypotheses 2 and 3. Metabolite 130 is only significant for sub-hypothesis 1b, meaning that this metabolite has a significant difference in slope between WT and *mdx* mice (Figure 21a). Metabolites 264, 274, 296 and 381 are only significant for sub-hypothesis 1a, meaning that these metabolites have a significant difference in

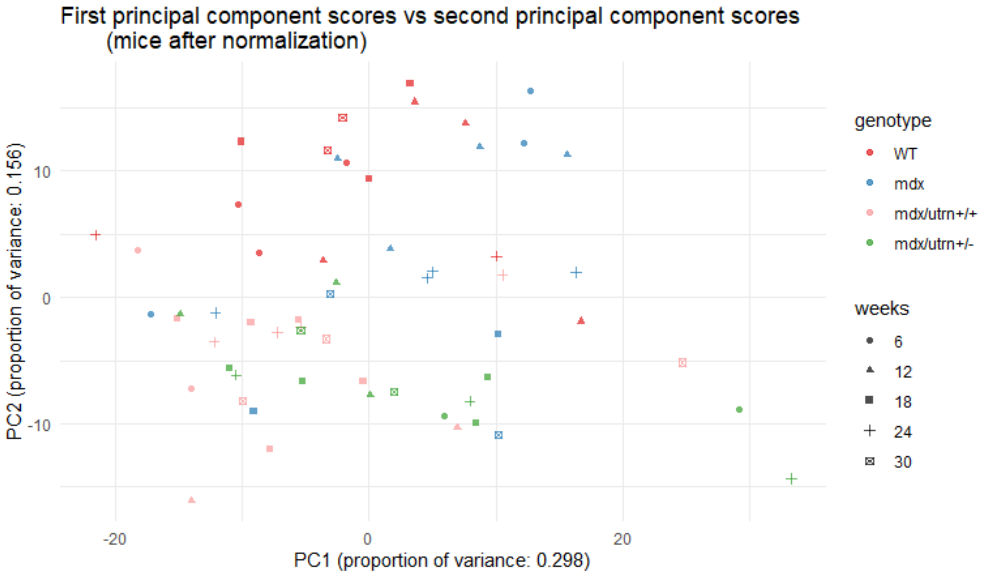



Figure 20: First two principal components of the mouse data capturing 34%. Samples of the data are coloured and shaped according their genotype and time point of measurements. There is some clustering in the genotypes; WT mice are clustered, while mdx mice are more or less scattered across the plot. This plot is very similar to the same plot with the data before normalization.

intercept between WT and *mdx* mice (Figures 21b, 21c, 21d and 21e). Since these metabolites do not show significant difference in slope, the difference shown at the intercept is also present over time during the whole experiment. The precise p-values can be found in Appendix A.2.

Because these five metabolites are significant for hypothesis 1, we also continue testing hypotheses 2 and 3. These hypotheses test the difference between WT and *mdx+/+* and *mdx+/-* respectively. After correcting for multiple testing, we find that metabolite 296 is not significant for either test and the other four metabolites, 130, 264, 274 and 381 are significant for both tests. This means that these four metabolites not only have a significant difference between just WT and *mdx* mice, but also between WT and *mdx+/+* and *mdx+/-* mice. So, for these four metabolites we continue to test for their sub-hypotheses. We find that metabolite 130 is only significant for sub-hypothesis 2b, meaning that this metabolite has a significantly different slope between WT and *mdx+/+* mice (Figure 22a). Metabolites 264 and 381 are significant for only sub-hypothesis 2a, meaning that they have a significant difference in distance between the trajectories of the two groups (Figures 22b and 22d). Finally, metabolite 274 is not significant for both sub-hypotheses (Figure 22c). So we did find a difference but we cannot exactly say what this difference is, probably due to the subtests not being powerful enough. The precise p-values can be found in Appendix A.2.

We also test for the sub-hypotheses of hypothesis 3. From this we find the same results for all interesting metabolites for sub-hypotheses 3a and 3b as for sub-hypotheses 2a and 2b (Figure 23). The complete results, with all p-values and their adjusted values for these five metabolites for the tests of the left branch, can be found in the Appendix A.2.

The right branch of hypotheses starts with testing hypothesis 6, which tests for the difference between *mdx+/+* and *mdx+/-* mice. After applying the multiple testing correction, none of the

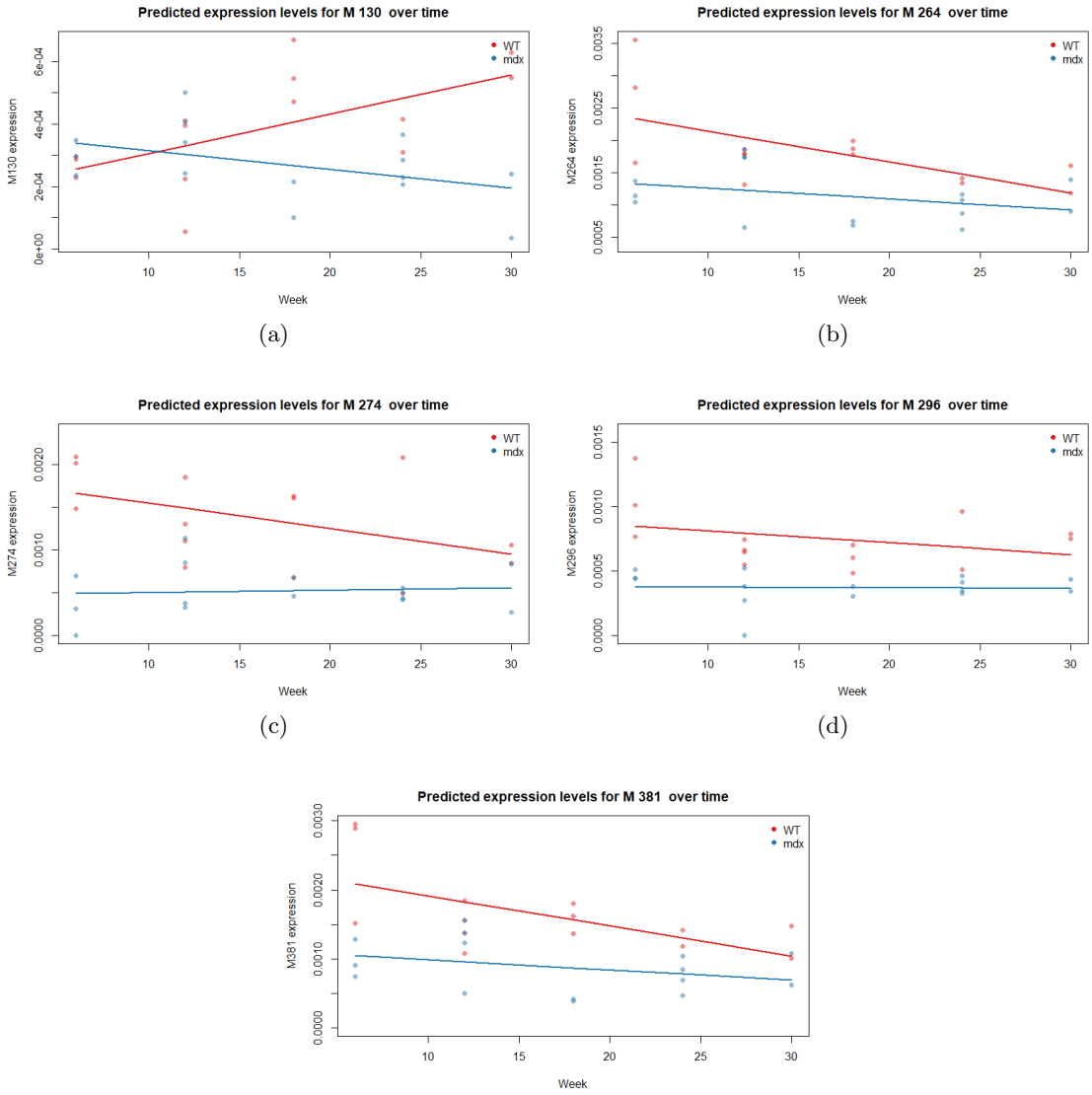
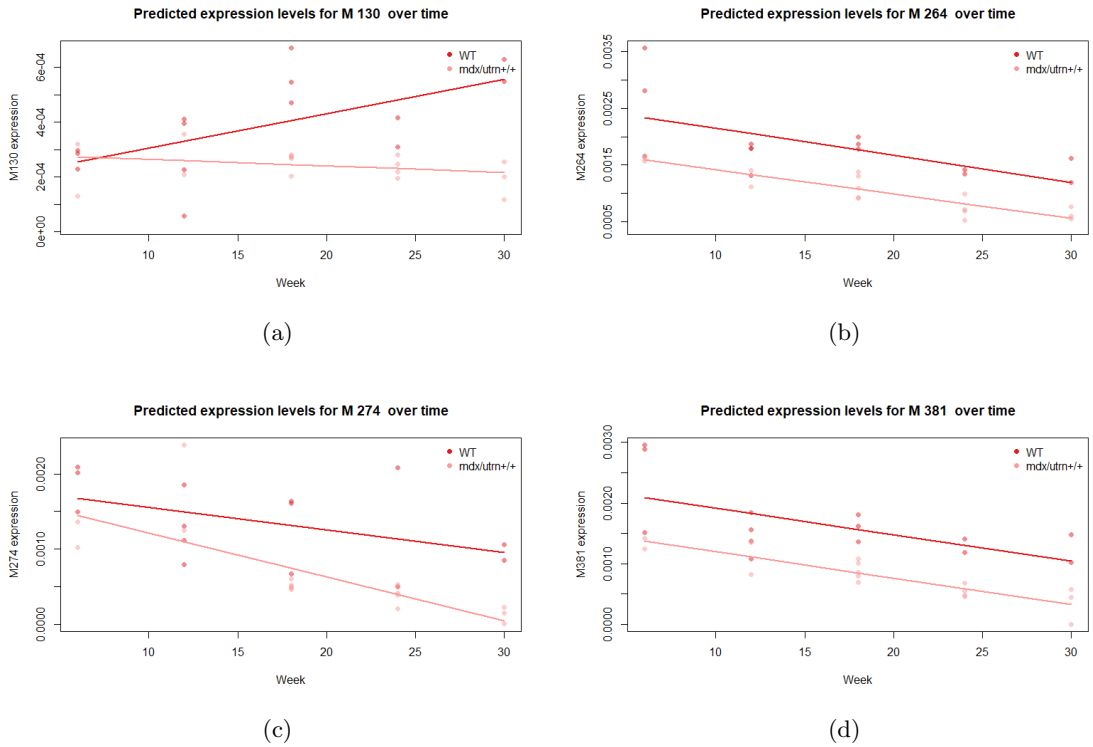



Figure 21: Predicted expressions for WT and *mdx* mice for metabolites that have significant difference between these two groups. (a) Metabolite 130 has significant difference in slope. (b-e) Metabolites 264 (b), 274 (c), 296 (d) and 381 (e) have significant difference in intercept.




(a) (b)

(c) (d)

Figure 22: Predicted expressions for WT and *mdx+/+* mice for metabolites that have significant difference between WT and *mdx* mice, as well as these two groups. (a) Metabolite 130 has a significantly different slope. (b,d) Metabolites 264 (b) and 381 (d) have a significantly different intercept. (c) Metabolite 274 does have a significant difference, but we can not find what it is exactly.

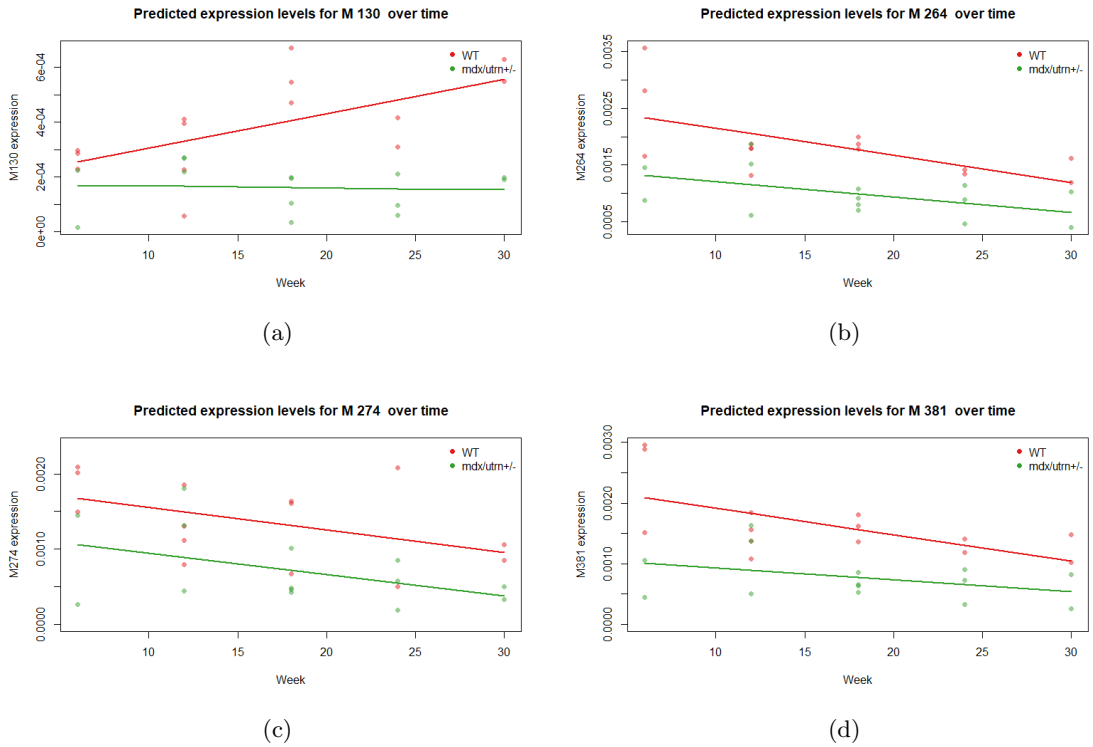



Figure 23: Predicted expressions for WT and *mdx+/-* mice for metabolites that have significant difference between WT and *mdx* mice, as well as these two groups. (a) Metabolite 130 has a significantly different slope. (b,d) Metabolites 264 (b) and 381 (d) have a significantly different intercept. (c) Metabolite 274 does have a significant difference, but we can not find what it is exactly.

metabolites give a significant result. This means that are no metabolites that have significantly different expression between these mouse groups. As this test is at the top of this branch, we do not test the other hypotheses below this node.

# 5 Conclusion

The goal of this thesis is to assess the feasibility of urine to identify potential biomarkers related to DMD and trajectories over time. To do this, we formulated research questions that we want to answer, as described in Section 2.3. The main question to be answered reads, "Are there metabolites whose expression level differ between healthy and dystrophic individuals?"

With our analysis on the mouse data using LMMs, we found some noteworthy results that could indicate urine's potential. We found five interesting metabolites that show significant difference in predicted expression between WT and *mdx* mice. From these five metabolites, four also have a significant difference between the WT group and the other two *mdx* groups. One metabolite, metabolite 130, shows for all three group comparisons that the difference is in the slope and two other metabolites, metabolites 264 and 381, show that the difference is in the intercept. These last two metabolites show no difference in slope according to the predictions, meaning that the distance between the two trajectories can be found over the whole timespan of the experiment and not only at the start. Furthermore, one metabolite, 296, only shows a difference between WT and *mdx* mice and one metabolite, 274, does show a difference between WT and all three *mdx* mice, but we can not find what it is exactly.

So, answering Research Question 2, we find three metabolites, 130, 264 and 381, that consistently show a clear difference in expression between healthy and sick mice. As metabolite 130 shows a significant difference in slope, this means that this metabolite is suitable for monitoring biomarkers for DMD in mice. On the other hand, metabolite 264 and 381 show significant difference in intercept, which means that these are better as diagnostic biomarkers.

We were not able to properly analyse the human data, but did find a relevant result for our research. We found that creatine is significantly more abundant in urine from patients with DMD, which makes this metabolite a potential biomarker. Considering that this data has its limitations, such as values missing at random and not having a control group for both countries the data is from, it might be an even stronger indicator that creatine would work well as a biomarker. A more comprehensive study might lead to even stronger indications and more promising results in terms of biomarkers. All in all, to answer Research Question 3; yes, there are metabolites that differ between healthy and DMD patients, namely creatine.

All in all, with the answers to Research Questions 2 and 3, we can confidently conclude that there is potential in using urine as a source for biomarkers for DMD, although there are issues in measuring metabolites in urine.

# 6 Discussion

This thesis researched the potential in using urine for obtaining biomarkers for Duchenne Muscular Dystrophy in both patients and mouse models. We concluded that there is in fact potential and found promising diagnostic and monitoring biomarkers.

However, this research also has its limitations. First of all, urine can have a lot variation due to the high dependence on an individuals water intake and the interval since last urination. For

example, morning urine is usually more concentrated compared to samples collected later on the day (Bottin et al., 2016).

For the human data, we did not have a Belgian control group available. This led to an asymmetry that might have introduced some bias in our analysis. Furthermore, the data has a number of missing values. This is likely a result of the high variability of urine, which obviously complicates a valid analysis.

Moreover, the available mouse data was collected during an experiment were also blood samples were obtaining alongside of urine samples. This means that the urine samples might be compromised as a result of repeatedly obtaining blood samples, possibly causing stress. While we normalized the data, this could still be an issue during analysis. Additionally, as the focus was not on the mice's urine, it could have happened that there was less pressure to actually obtain the urine. So, the data contains *only* 59 urine samples, which could be why we were not always able to fit a model. Furthermore, the mouse data did not have any labels for the metabolites, meaning that we were not able to couple the mouse and human data. If these datasets would have matching metabolites, this could be been a great step to translate our findings from the mouse data to the human data, which would ultimately help research for DMD patients.

For future research, we think it would be insightful to set up a more detailed and comprehensive experiment for obtaining the data. This would lead to more complete data that is easier and better to analyse. With labels for the metabolites, we could interpret more meaningful results in a biological way and conclude stronger results regarding the biomarkers.

# References


Asher, D., Thapa, K., Dharia, S., Khan, N., Potter, R., and Rodino-Klapac, L. (2020). Clinical development on the frontier: Gene therapy for duchenne muscular dystrophy. *Expert Opinion on Biological Therapy*, 20.

Boca, S. M., Nishida, M., Harris, M., Rao, S., Cheema, A. K., Gill, K., Seol, H., Morgenroth, L. P., Henricson, E., McDonald, C., et al. (2016). Discovery of metabolic biomarkers for duchenne muscular dystrophy within a natural history study. *PloS one*, 11(4):e0153461.

Bottin, J. H., Lemetais, G., Poupin, M., Jimenez, L., and Perrier, E. T. (2016). Equivalence of afternoon spot and 24-h urinary hydration biomarkers in free-living healthy adults. *European Journal of Clinical Nutrition*, 70(8):904–907.

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J., Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., and Wishart, D. S. (2013). The human urine metabolome. *PLoS One*, 8(9):e73076.

Bucciolini Di Sagni, M., Vannelli Gori, G., and Oriolo, R. (1982). [structural and ultrastructural changes in the skeletal muscles of patients in the early stages of duchenne muscular dystrophy and "possible" carriers]. *Bollettino della Societa italiana di biologia sperimentale*, 58(10):632—638.

Chandel, N. S. (2021). Basics of metabolic reactions. *Cold Spring Harbor Perspectives in Biology*, 13(8):a040527.

Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H. (2006). Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. application in 1h nmr metabonomics. *Analytical Chemistry*, 78(13):4281–4290. PMID: 16808434.

Duan, D., Goemans, N., Takeda, S., Mercuri, E., and Aartsma-Rus, A. (2021). Duchenne muscular dystrophy. *Nature Reviews Disease Primers*, 7(1):13.

Duchenne Parent Project (2024). Duchenne parent project. `https://www.duchenne.com`.

FDA-NIH Biomarker Working Group (2016). *BEST (Biomarkers, EndpointS, and other Tools) Resource*. National Institutes of Health (US), Bethesda (MD). https://www.ncbi.nlm.nih.gov/books/NBK326791/.

Gerothanassis, I. P., Troganis, A., Exarchou, V., and Barbarossou, K. (2002). Nuclear magnetic resonance (nmr) spectroscopy: basic principles and phenomena, and their applications to chemistry, biology and medicine. *Chemistry Education Research and Practice*, 3(2):229–252.

Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., and Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2(1):100.

Li, A. and Barber, R. F. (2018). Multiple testing with the structure-adaptive benjamini–hochberg algorithm. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 81(1):45–74.

National Cancer Institute (2024). National cancer institute. `https://www.cancer.gov`.

Perkins, K. J. and Davies, K. E. (2002). The role of utrophin in the potential therapy of duchenne muscular dystrophy. *Neuromuscular Disorders*, 12:S78–S89. ENMC Special Centennial Work-

shop Supplement: Therapeutic Possibilities in Duchenne Muscular Dystrophy, Naarden, The Netheralnds, 30 November-2 December 2001.

Pitt, J. J. (2009). Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clinical Biochemist Reviews*, 30(1):19–34.

Signorelli, M., Ebrahimpoor, M., Veth, O., Hettne, K., Verwey, N., García-Rodríguez, R., Tanganyika-deWinter, C. L., Lopez Hernandez, L. B., Escobar Cedillo, R., Gómez Díaz, B., Magnusson, O. T., Mei, H., Tsonaka, R., Aartsma-Rus, A., and Spitali, P. (2021). Peripheral blood transcriptome profiling enables monitoring disease progression in dystrophic mice and patients. *EMBO Molecular Medicine*, 13(4):e13328.

Spitali, P., Hettne, K., Tsonaka, R., Charrout, M., van den Bergen, J., Koeks, Z., Kan, H. E., Hooijmans, M. T., Roos, A., Straub, V., Muntoni, F., Al-Khalili-Szigyarto, C., Koel-Simmelink, M. J., Teunissen, C. E., Lochmüller, H., Niks, E. H., and Aartsma-Rus, A. (2018). Tracking disease progression non-invasively in duchenne and becker muscular dystrophies. *Journal of Cachexia, Sarcopenia and Muscle*, 9(4):715–726.

Szigyarto, C. A.-K. and Spitali, P. (2018). Biomarkers of duchenne muscular dystrophy: current findings. *Degenerative Neurological and Neuromuscular Disease*, 8:1–13.

Timonen, A., Lloyd-Puryear, M., Hougaard, D. M., Meriö, L., Mäkinen, P., Laitala, V., Pölönen, T., Skogstrand, K., Kennedy, A., Airenne, S., Polari, H., and Korpimäki, T. (2019). Duchenne muscular dystrophy newborn screening: Evaluation of a new GSP® neonatal creatine kinase-MM kit in a US and danish population. *International Journal of Neonatal Screening*, 5(3):27.

Turk, R., Sterrenburg, E., van der Wees, C. G. C., de Meijer, E. J., de Menezes, R. X., Groh, S., Campbell, K. P., Noguchi, S., van Ommen, G. J. B., den Dunnen, J. T., and Hoen, P. A. C. t. (2006). Common pathological mechanisms in mouse models for muscular dystrophies. *The FASEB Journal*, 20(1):127–129.

# A Appendices

## A.1 List of all null hypotheses

The main hypotheses tests are done to find whether there is any difference between the four different mouse groups. For all six unique pairs, the null hypothesis can be found here.

T1 WT and *mdx*

$$H_0 : \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

T2 WT and *mdx+/+*

$$H_0 : \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

T3 WT and *mdx+/-*

$$H_0 : \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

T4 *mdx* and *mdx+/+*

$$H_0 : \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

T5 *mdx* and *mdx+/-*

$$H_0 : \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

T6 *mdx+/+* and *mdx+/-*

$$H_0 : \begin{pmatrix} 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \beta_{r2} \\ \beta_{r3} \\ \beta_{r4} \\ \beta_{r5} \\ \beta_{r6} \\ \beta_{r7} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

## A.2 Complete list of results from hypothesis testing

For the interesting five metabolites, all the exact p values and their corrected values can be found here. Values printed bold are significant.

| Metabolite | Test | P-value | Adjusted p-value |
|---|---|---|---|
| | T1 | **0.0003394** | **0.0417752** |
| | T1a | 0.2046586 | 0.2046586 |
| | T1b | **0.0001963** | **0.0003926** |
| | T2 | **0.0000186** | **0.0000373** |
| 130 | T2a | 0.2314925 | 0.2314925 |
| | T2b | **0.0090792** | **0.0181583** |
| | T3 | **0.0003074** | **0.0003074** |
| | T3a | 0.7938534 | 0.7938534 |
| | T3b | **0.0027451** | **0.0054901** |
| | T1 | **0.0004476** | **0.0417752** |
| | T1a | **0.0002197** | **0.0004395** |
| | T1b | 0.0799461 | 0.0799461 |
| | T2 | **0.0001583** | **0.0003166** |
| 264 | T2a | **0.0003651** | **0.0007303** |
| | T2b | 0.2627978 | 0.2627978 |
| | T3 | **0.0009023** | **0.0009023** |
| | T3a | **0.0070919** | **0.0141839** |
| | T3b | 0.7811791 | 0.7811791 |
| | T1 | **0.0000499** | **0.0144687** |
| | T1a | **0.0000560** | **0.0001119** |
| | T1b | 0.0974594 | 0.0974594 |
| | T2 | **0.0042741** | **0.0042741** |
| 274 | T2a | **0.0350605** | 0.0701210 |
| | T2b | 0.9460573 | 0.9460573 |
| | T3 | **0.0028998** | **0.0042741** |
| | T3a | 0.4245344 | 0.4245344 |
| | T3b | 0.1556083 | 0.3112165 |
| | T1 | **0.0005009** | **0.0417752** |
| | T1a | **0.0013976** | **0.0027953** |
| 296 | T1b | 0.3974060 | 0.3974060 |
| | T2 | 0.0507540 | 0.0507540 |
| | T3 | **0.0288780** | 0.0507540 |
| | T1 | **0.0000694** | **0.0144687** |
| | T1a | **0.0000578** | **0.0001155** |
| | T1b | 0.0876068 | 0.0876068 |
| | T2 | **0.0000268** | **0.0000535** |
| 381 | T2a | **0.0000763** | **0.0001527** |
| | T2b | 0.1778301 | 0.1778301 |
| | T3 | **0.0001937** | **0.0001937** |
| | T3a | **0.0060909** | **0.0121818** |
| | T3b | 0.9954081 | 0.9954081 |

Table 3: P-values and adjusted p-values for the five metabolites that were significant for hypothesis test 1 and are therefore interesting in the left branch of tests. Values printed in bold are significant.