

# UNIVERSITY OF VERONA



**Department of**

Diagnostics and Public Health

**PhD School of**

Natural Sciences and Engineering

**PhD in**

Nanoscience and Advanced Technologies

XXXIII cycle (2017)

## **NEW METHODOLOGIES FOR THE DETERMINATION OF AMMONIUM AND SYNTHETIC CANNABINOIDS IN FORENSIC SAMPLES**

S.S.D. MED/43

Coordinator: Prof. Adolfo Speghini

Tutor: Prof. Franco Tagliaro

PhD student: Covadonga Palacio Gutiérrez

UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT OF  
DIAGNOSTICS AND PUBLIC HEALTH

GRADUATE SCHOOL OF  
NATURAL SCIENCES AND ENGINEERING

DOCTORAL PROGRAM IN  
NANOSCIENCE AND ADVANCED TECHNOLOGIES

WITH THE FINANCIAL CONTRIBUTION OF  
(NAME IF THE FUNDING INSTITUTION)

University of Verona

Cycle / year (1° year of attendance) 33/2017

TITLE OF THE DOCTORAL THESIS

NEW METHODOLOGIES FOR THE DETERMINATION OF AMMONIUM AND  
SYNTHETIC CANNABINOIDS IN FORENSIC SAMPLES

S.S.D. MED/43

(Please complete this space with the S.S.D. of your thesis – mandatory information)\*

Coordinator: Prof. Adolfo Speghini

Signature

Tutor: Prof. Franco Tagliaro

Signature

Doctoral Student: Dott.ssa Covadonga Palacio Gutierrez

Signature

For the list of S.S.D. please refer to the Ministerial Decree of 4<sup>th</sup> October 2000, Attachment A "Elenco dei Settori Scientifico – disciplinari" available at: [http://www.miur.it/atti/2000/alladm001004\\_01.htm](http://www.miur.it/atti/2000/alladm001004_01.htm)

*To my grandmothers, Carmen and Teresita.*

*“Happiness can be found even in the darkest of times, if one only remembers to turn on the light.”*

*J.K. Rowling*

# **TABLE OF CONTENTS**

Framework and aims .....	5
1. Capillary electrophoresis.....	7
1.1 Fundamentals .....	8
1.1.1 Electrophoresis and Electroosmotic Flow.....	8
1.1.2 Sample injection .....	9
1.1.3 Peak efficiency.....	9
1.2 Analytical conditions .....	9
1.3 Separation techniques .....	9
1.3.1 Capillary Zone Electrophoresis (CZE).....	10
1.3.2 Micellar Electrokinetic Chromatography (MEKC) .....	10
1.3.3 Capillary Gel Electrophoresis (CGE).....	10
1.3.4 Capillary Isotachopheresis (CITP).....	10
1.3.5 Capillary Electrochromatography (CEC) .....	11
1.4 Forensic applications of CE .....	11
2. Statistics and chemometrics in forensic science .....	13
2.1 Artificial Neural Networks (ANNs) .....	13
2.1.1 Network training.....	15
2.1.2 Software .....	17
2.1.3 Applications of ANNs in forensic science .....	17
3. Estimation of the post-mortem interval (PMI) through the detection of $\text{NH}_4^+$ and $\text{K}^+$ in vitreous humour with CE .....	19
3.1 Introduction .....	19
3.2 Materials and methods.....	21
3.2.1 Chemicals.....	21
3.2.2 CE system and software .....	21
3.2.3 Electrophoretic conditions.....	21
3.2.4 Preparation of standards and samples.....	22
3.2.5 Method validation.....	22

3.2.6	Statistical analysis .....	23
3.3	Results and discussion .....	24
3.3.1	Precision: intra- and inter-day studies .....	24
3.3.2	Bias .....	25
3.3.3	Linearity, LOD, and LOQ .....	25
3.3.4	Discussion of the preliminary study (linear approach) .....	26
3.3.5	Discussion of the logarithmic approach.....	28
3.3.6	Discussion of the neural networks approach.....	36
3.3.7	Discussion of the three approaches .....	44
3.3.8	Ammonium concentration differences in the two eyes.....	45
3.4	Conclusions .....	45
3.5	Future research .....	46
1.	Liquid chromatography coupled with mass spectrometry .....	47
1.1	Liquid Chromatography (LC).....	47
1.1.1	Reverse phase chromatography .....	47
1.1.2	Ultra-Performance Liquid Chromatography (UPLC) .....	48
1.2	Mass Spectrometry (MS) .....	48
1.2.1	LC–MS Ion sources .....	48
1.2.2	Mass Analyser .....	50
1.2.3	Detector .....	52
1.3	Forensic applications .....	52
2.	Elucidation of the 5F- APINAC metabolic pathway.....	54
2.1	Introduction.....	54
2.2	Materials and methods.....	56
2.2.1	Reagents .....	56
2.2.2	In vitro & In vivo studies .....	56
2.2.3	Preparation of rat urine samples .....	56
2.2.4	Tentative theoretical metabolic pathway .....	57
2.2.5	LC–IT–MS (Toxtyper) analyses .....	57
2.2.6	LC–QTOF analyses .....	57

2.3	Results.....	58
2.3.1	LC–IT–MS and LC–QTOF analyses .....	58
2.3.2	HLM incubation of 5F-APINAC .....	62
2.3.3	Metabolic pathway.....	63
2.4	Discussion .....	65
2.5	Conclusion.....	66
2.6	Limitations and Future research.....	67
3.	Development and validation of a new method for the determination of 13 second-wave synthetic cannabinoids in hair.....	68
3.1	Introduction.....	68
3.2	Materials and methods.....	69
3.2.1	Reagents .....	69
3.2.2	Instrumentation.....	72
3.2.3	Solutions and calibration curve preparation .....	73
3.2.4	Hair sample preparation.....	73
3.2.5	Method validation.....	73
3.2.6	Real cases.....	74
3.3	Results and discussion .....	74
3.3.1	Comparison of the two columns.....	74
3.3.2	Comparison of ESI, iB, and APCI sources .....	75
3.3.3	Calibration, LLOQ, LOD, and selectivity.....	77
3.3.4	Precision and bias .....	77
3.3.5	Matrix effect.....	79
3.3.6	Authentic hair samples.....	80
3.4	Conclusion.....	80
4.	Abbreviations.....	81
5.	References .....	84
6.	Appendices.....	95
6.1	Appendix I. NH <sub>4</sub> <sup>+</sup> in vitreous humour.....	95
6.1.1	List of vitreous humour samples used for the study .....	95

6.1.2	Train test split and data pre-processing .....	96
6.1.3	Code for Adam solver .....	96
6.1.4	Code for SGD solver.....	102
6.1.5	Code for LBFGS .....	106
6.1.6	Code for untrained real sample predictions.....	108
6.2	Appendix II. Elucidation of the 5f- Apinac metabolic pathway.....	109
6.2.1	Extracted ion chromatograms for the metabolites associated with 5f-apinac found <i>in vitro</i> and <i>in vivo</i> .....	109
7.	Scientific publications.....	111
8.	Oral and poster communications.....	112
9.	Acknowledgements.....	113



## FRAMEWORK AND AIMS

Forensic science is a discipline that combines the knowledge of chemistry, biology, pharmacology, statistics, mathematics, physics, and medicine, among others. Statistics is used for different types of analysis regarding the certainty of results, regressions, predictions, and probabilities. Mathematics and physics are used, for instance, in the determination of bullet trajectories or calculating the likelihood of one fingerprint belonging to one person or another. Medicine is applied in the autopsy room to determine the time and cause of death, and any other queries related to the body that doctors and scientists might need for parallel studies. Chemistry is extensively used in forensic science. From the use of infrared (IR) and nuclear magnetic resonance (NMR) for structural characterisation of drugs, to the use of chromatographic techniques coupled with mass detection for the elucidation of metabolic pathways, or identification and determination of drugs in several biological matrices. Also, to establish the elemental composition of gunshot residue (GSR) for identifying the firearm used to shoot a person and the distance of the shooter, techniques such as scanning electron microscope (SEM) with energy dispersive X-ray spectroscopy (SEM-EDS) are used. Spectrophotometry is another important technique that is used, for example, to determine the concentration of compounds, such as poisons, dissolved in other liquids. In the area of arson, gas chromatography coupled with mass detection (GC-MS/MS) is used, for instance, for the analysis and identification of volatile liquids that might have been used as accelerants in arson cases. Moreover, techniques such as surface-enhanced Raman spectroscopy (SERS) can be used for the identification of explosives from crime scenes debris. In the discipline of environmental forensics, techniques such as X-Ray diffraction and ion mobility spectrometry (IMS) can be used for the characterisation of soil samples in order to compare them with a soil sample from a cadaver. Moreover, IMS is currently used in airport security for the detection of explosives. Entomology and botany, two biological disciplines, are also heavily used in forensic science. For example, through the study of insects, bugs, plants, and trees, the time of death could be established, as well as the discovery of clandestine graveyards. And last but not least, identification of the victim through DNA or fingerprints. Forensic science truly is a highly multidisciplinary field.

This is only a glimpse of all that forensic science encompasses as a discipline. Currently, all areas of forensic science require the attention of scientists for improvement, and this doctoral thesis was focused on two of them. The first is the estimation of the time of death, which is extremely important when bodies are disposed and hidden, where the main methodologies for the estimation of the time of death are only useful up to 48 hours. The second one is related to the emerging threat of a new type of illicit drugs called novel psychoactive substances (NPS) and the aim is two-fold: metabolic pathway elucidation, which is relevant for toxicological risk assessment in order to develop drug screening procedures for the detection of drugs and their metabolites in blood or urine and to correctly administer an antidote; and the development and validation of a new method for the

determination of NPS in hair, which is relevant when some other biological matrices might not be available or might not give the necessary information.

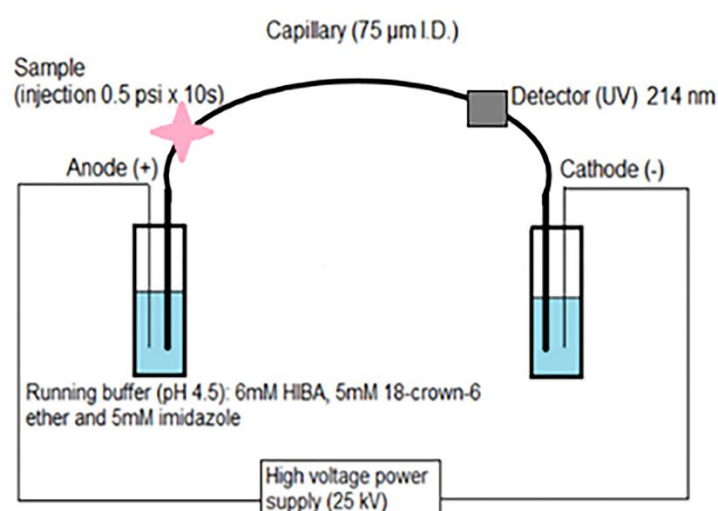
Given all the aforementioned considerations and the emphasis on the multidisciplinary nature of forensic science, the aims of the present doctoral thesis are 3-fold:

- ✚ Validation of a capillary electrophoresis (CE) method for the determination of ammonium in vitreous humour. Furthermore, to explore the use of potassium and ammonium in vitreous humour with different statistical analyses in order to improve the estimation of the post-mortem interval (PMI)
- ✚ Metabolic pathway elucidation of 5F-APINAC (adamantan-1-yl 1-(5-fluoropentyl)-1H-indazole-3-carboxylate), a synthetic cannabinoid which is one of the emerging NPS
- ✚ Development and validation of a fast and sensitive method for the screening of 13 new synthetic cannabinoids in human hair

# CAPILLARY ELECTROPHORESIS (CE) APPLIED TO FORENSIC SCIENCE

## 1. CAPILLARY ELECTROPHORESIS

Capillary electrophoresis (CE) is an analytical technique that separates charged molecules in small sample volumes quickly and efficiently. Separations take place inside a small capillary due to the ion's differences in electrophoretic mobilities in an electrophoretic environment. Historically, the first trial for electrophoretic separation took place in 1886 when Lodge tried to migrate a  $H^+$  in a tube of phenolphthalein "jelly". Due to continuous research in chromatography and electrophoresis, the technique has evolved to its present form [1].



**Figure 1.** Capillary electrophoresis instrumental set-up

The instrumental set-up is depicted in Figure 1. It generally consists of a capillary (where the separation takes place) that goes through the optical centre of a detector; two electrodes made of platinum, the anode and cathode; an autosampler; a high voltage power supply; and two buffer reservoirs to submerge the electrodes and capillary. For separation to take place, the capillary is filled with the buffer solution, then the sample is injected, and voltage is applied. This causes electroosmotic and electrophoretic movements which result in the sample moving towards the cathode passing through the detector [2].

The most used capillaries are fused silica with an external polyamide coating to protect them and with internal diameters between 20-100 µm. On the inner wall silanol groups are present. One of the main advantages of the CE is that only very little sample is injected (to the order of nL) due to small volume of the capillary column and this represents a great advantage particularly when dealing with low amount of samples, such as in vitreous humour analysis [3].

## 1.1 FUNDAMENTALS

### 1.1.1 ELECTROPHORESIS AND ELECTROOSMOTIC FLOW

The electrophoresis movement is the process by which ions move from one electrode to the other due to the applied voltage. The ion's mobility (migration rate) is ruled by the number of ionic charges and its size (Eq.(1)).

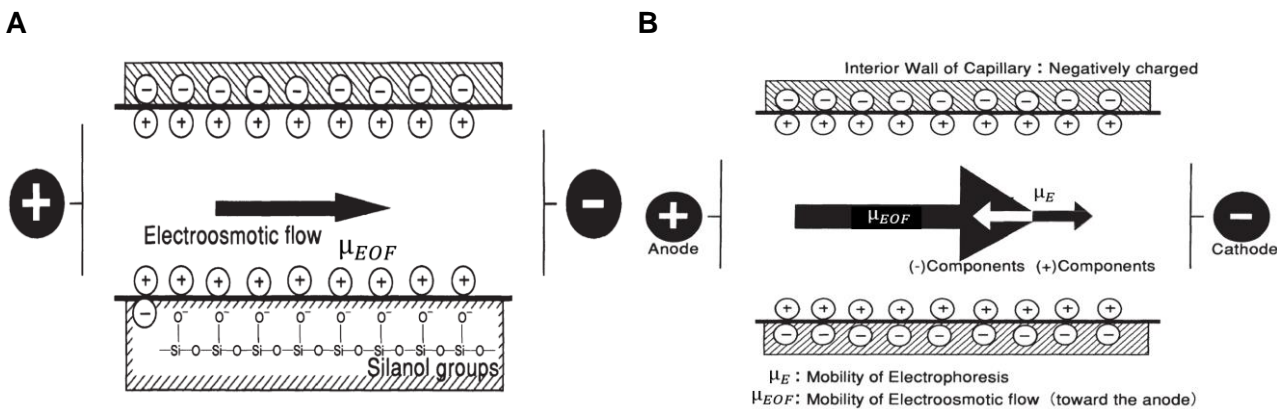
$$(1) \quad \mu_E = \left(\frac{q}{6\pi}\right)\eta r$$

where  $\mu_E$  is electrophoretic mobility,  $q$  is number of charges,  $\eta$  is solution viscosity, and  $r$  is the radius of the ion. Therefore, when separating a mixture of ions, the larger ones with less charge will be the slowest to reach the detector, whereas the smaller, highly charged ions will be the first ones to be detected. As an example, if two ions have the same number of charges, the bigger ion will move slower than the small one [2].

Another phenomenon caused by the application of high voltage into the capillary is electroosmotic flow (EOF). When the acidic silanol groups inside the capillary come in contact with the buffer solution, they are ionised and EOF is created, driving the solute towards the detector (Eq.(2)).

$$(2) \quad \mu_{EOF} = \frac{\varepsilon\zeta}{\eta}$$

where  $\mu_{EOF}$  is EOF mobility,  $\varepsilon$  is the buffer dielectric constant,  $\zeta$  is the charge on the surface of the capillary (zeta potential), and  $\eta$  is solution viscosity.



**Figure 2.** **2A:** Schematic diagram of capillary surface and of EOF generation. **2B:** Migration of analyte in the presence of EOF. The migration velocity of a given species can be expressed as a vector sum of the EOF ( $\mu_{EOF}$ ) and the electrophoretic mobility ( $\mu_E$ ) of the solute (adapted from [3]).

When the separation buffer's pH is higher than 4, the silanol groups start to be ionised, leading to the formation of a negatively charged layer. In order to maintain electroneutrality, a cation layer is formed, and when voltage is applied, they migrate towards the cathode generating an EOF (Figure 2).

The main parameter affecting the level of EOF is the pH of the electrolyte. At high pH (above 9), the EOF is strong because the silanol groups are fully ionised, whereas at low pH (below 4), the EOF is weak because they are barely ionised. The overall migration time of an ion (apparent mobility  $\mu_A$ ) is a sum of the EOF ( $\mu_{EOF}$ ) and the electrophoretic mobility ( $\mu_E$ ) of the solute (Eq.(3)) [2],[3].

$$(3) \quad \mu_A = \mu_E + \mu_{EOF}$$

### 1.1.2 SAMPLE INJECTION

There are three mechanisms for introducing the sample: electrokinetic, gravity, and hydrodynamic. All three involve the application of a force to inject the sample in the capillary while it is immersed in the sample solution.

In electrokinetic injection, both the capillary and electrode are placed in the sample vial. It is based on the application of a high voltage at the tip of the capillary causing solute ions to enter by EOF and electrophoretic migration. With this method a greater number of ions enter the capillary which is an advantage when trying to quantify trace levels of small ions, but it can also be disadvantageous since it can cause sample bias effects (more mobile species are injected in larger quantities).

Meanwhile, gravity injection is based on mechanically raising the capillary above the height of the detector electrolyte vial.

Finally, the hydrodynamic method is based on a pressure differential. While the capillary is immersed in the sample solution, a pressure difference is applied either in the form of vacuum or positive pressure. In order to ensure sample injection reproducibility, maintaining a constant temperature is needed [1,2].

### 1.1.3 PEAK EFFICIENCY

One mayor advantage of CE is the minimisation of band broadening effects, which represent a major issue in HPLC and conventional electrophoresis. Sample dispersion along the capillary is minimised due to the EOF flow dynamics in comparison with the laminar flow of pump systems; convection related band broadening (typical in conventional electrophoresis) is reduced because the heat generated due to the voltage application is dissipated through the capillary walls (Joule heating); and post-separation broadening is eliminated due to on-column detection [2].

## 1.2 ANALYTICAL CONDITIONS

### 1.3 SEPARATION TECHNIQUES

As mentioned before, the separation in CE takes place due to electric charge differences in the ions of interest. It may seem that CE is only suitable for ionic molecules, however, by including proper additives into the background electrolyte it is possible to convert non-ionic analytes into charged

species, thus allowing for their separation in CE. The five most common CE separation modes are explained below.

### **1.3.1 CAPILLARY ZONE ELECTROPHORESIS (CZE)**

CZE is the most common separation technique. In this technique, separation is achieved through the solutes' migration in the buffer inside the capillary, which depends on their own electric mobility. The main factors affecting separation are: pH, conductivity, temperature, ionic strength of the solution, presence of additives and length of the capillary. UV detection can be done indirectly or directly, with indirect detection offering higher sensitivity for some compounds. For indirect detection of cations, such as ammonium, a strong absorbing molecule like imidazole can be used as a UV absorbent. In addition, chelating agents such as crown ethers can be added to the background electrolyte to improve separation since they can interact with the ions [3].

### **1.3.2 MICELLAR ELECTROKINETIC CHROMATOGRAPHY (MEKC)**

Aiming to tackle a serious limitation of CE, which is the necessity of the substance to be charged, MEKC was created allowing neutral and charged analytes to be separated. This is achieved through the use of micellar solutions of ionic surfactants. An ionic surfactant solution with a concentration higher than its critical micelle concentration (CMC) is added to the buffer so that micelles are formed. The surfactant added can be either cationic or anionic. If a cationic surfactant is added, the EOF migrates towards the positive electrode (anode) and the micelle migrates to the negative electrode (cathode). If the surfactant added is anionic, the micelle migrates to the positive electrode by electrophoresis and the EOF flows toward the negative electrode due to the negative charge of the capillary surface [2].

### **1.3.3 CAPILLARY GEL ELECTROPHORESIS (CGE)**

In CGE, the analytes migrate through polymer networks (acting as sieves) at different speeds depending on their molecular sizes. CGE possesses a number of potential advantages such as the prevention of analyte adsorption in the capillary walls, minimisation of solute diffusion, and elimination of electroosmosis. Altogether, these increase column resolution in short columns and can be a highly sensitive method if used with laser-induced fluorescence detectors. The most widely used capillaries are polyacrylamide and agarose filled capillaries [1,3].

### **1.3.4 CAPILLARY ISOTACHOPHORESIS (CITP)**

In this technique, a leading electrolyte is added to the background buffer. First, a large volume of sample dissolved in the leading electrolyte is injected into the column. Then, the injection end of the capillary is submerged in a solution containing a terminating electrolyte and then voltage is applied during a set amount of time in order to focus the sample in bands (analyte zones) which are continuously formed. Finally, the capillary is submerged in the support buffer solution and voltage is

applied to separate the sample by standard CE separation. The separated zones can be detected by using a potential gradient or a conductivity detector [2,3].

### **1.3.5 CAPILLARY ELECTROCHROMATOGRAPHY (CEC)**

Capillary electrochromatography was developed as a hybrid between chromatography and electrophoresis. In CEC, a chromatographic stationary phase is introduced in an electrophoretic environment either as a film in the walls of the capillary or as a packed bed. In this technique, neutral species will be separated without the necessity for any additives, whereas ionic species will be superimposed on the electrophoretic separation [2].

## **1.4 FORENSIC APPLICATIONS OF CE**

CE has gained popularity in the forensic field due to its clear advantages, such as minimal sample injection, speed, high mass sensitivity, low consumption of solvents, low cost, and its suitability to be coupled with mass detection. Substances analysed with CE techniques vary from DNA fragments to inorganic ions [4]. Some examples of its broad applicability in this field are discussed.

In the area of illicit drugs, Porpiglia et al.[5], used CZE for the chiral determination of ketamine and norketamine (its major metabolite) in hair. They used cyclodextrins and electrokinetic injection. Airado-Rodríguez et al.[4], developed a method for the analysis of lysergic acid diethylamide and its C-8 isomer in hair samples with CZE. Evans et al.[6], used a portable CE instrument to determine the inorganic ionic profiles of three pharmaceutical samples and precursors of two illicit drugs: methyldone and para-methoxymethamphetamine. Gottardo et al.[7], developed a CE-ESI-TOF method for the determination of illicit and controlled drugs in blood using CZE and electrokinetic injection. Concerning fire arms debris analysis, Northrop et al.[8], used MECK for the separation and identification of organic gunshot residue and explosive constituents. Erol et al.[9], determined amounts of nitrite and nitrate in gunshot residue by CZE. Regarding forensic toxicology, Tagliaro et al.[10], developed and validated a new method for the determination of potassium in vitreous humour and its use as an estimator of the post-mortem interval (PMI) with CZE. Furthermore, the determination of Carbohydrate-Deficient Transferrin (CDT), which is a marker of alcohol abuse, is one of the major applications of CZE. A lot of research has been done by the research group of F. Tagliaro on the determination of CDT with CZE with improvements in sample treatment and cost-effectiveness in comparison with HPLC determination [11–14]. Regarding counterfeit investigations of documents, ink and dyes can be analysed by MECK. CE replaced HPLC due to its very good performance distinguishing between water-soluble inks that contain ionic dyestuffs. Inks coming from different manufacturers and different countries give different patterns in the electropherograms [15]. CE can also be applied to investigations of clandestine laboratories. Hauser et al.[16], developed and validated a CE method for the quantification of ammonium, chloride, sodium, sulphate and formate ions, which are present in clandestine production of amphetamine. Wright et al.[17], recently

developed a CE with laser-induced fluorescence detection methodology for the fast and reliable analysis of sexual assault kits (SAKs).

Further applications include analysis of forensic DNA with multicapillary electrophoresis with fluorescent detection and analysis of ions and small molecules for investigations in intentional poisoning (cyanide) or fire and arson investigations [18].



## 2. STATISTICS AND CHEMOMETRICS IN FORENSIC SCIENCE

In forensic science, where the results of the examination of complicated evidence must lead to unbiased verdicts in legal proceedings, it is of high importance to correctly identify and assess samples. To do so, analytical methods such as liquid and gas chromatography or capillary electrophoresis are used in combination with statistical analysis. For instance, in clinical chemistry, many analytes can be determined in a sample of vitreous humour, blood, or urine yielding multivariate data for one specimen. Hence, some samples might produce more complex data than others which requires more sophisticated statistical methods. To overcome such difficulty, advanced chemometric methods can deal with complex and large datasets. By definition, chemometrics *“extract the maximum valuable information from the dataset by using the best measurement techniques/ optimal procedures and acquires most of the chemical information from the sample data. It correlates quality parameters or physical properties of the data”*. Indeed, chemometrics, or multivariate statistics, extracts the information from spectra obtained by analytical methods and it analyses the data through chemical pattern recognition. This supervised pattern recognition method constructs a model based on information of known samples and uses the model to further predict unknown samples. For each sample, there are two groups of variables: predictors and response. Finding a relationship between these two variables is key to obtain information from unknown samples. There are different types of analysis to apply, such as k-nearest neighbour (kNN), partial-least-squares discriminant analysis (PLS-DA) and artificial neural networks (ANN) [19,20].

### 2.1 ARTIFICIAL NEURAL NETWORKS (ANNs)

ANNs is a genetic algorithm, an attempt to make a computer model of the brain. The algorithm works by processing a selection of initial data then generating an output that is tested for its accuracy. This process undergoes multiple iterations until an acceptable outcome is produced. Once the data is in the system, ANNs imitate the neurons in the brain (Figure 3) trying to figure out the relationships within the data. Through the dendrites, neurons receive input signals and send them out through the axon, and release neurotransmitters to pass them on to other neurons. If the strength of the inputs is not enough to overcome a threshold value, a signal is not produced in the axon. Similarly, ANNs have layers of interlinked artificial neurons (nodes), akin to processors that work in parallel. The measured data is given to the input layer, and through a series of mathematical operations in the hidden layers, the output is generated. In more detail, the network is interactively trained with a set of data.

In the current project, using the peak information of the analytes and the known post-mortem interval (PMI), the network will predict a PMI for each sample and compare it to the known PMI. Any differences between the known PMI and the network's output will be used to adjust the internal parameters (in the hidden layer). This process will continue until the required degree of accuracy is achieved, which is evaluated with a test set. The network can be over-fitted if one forgets the fact

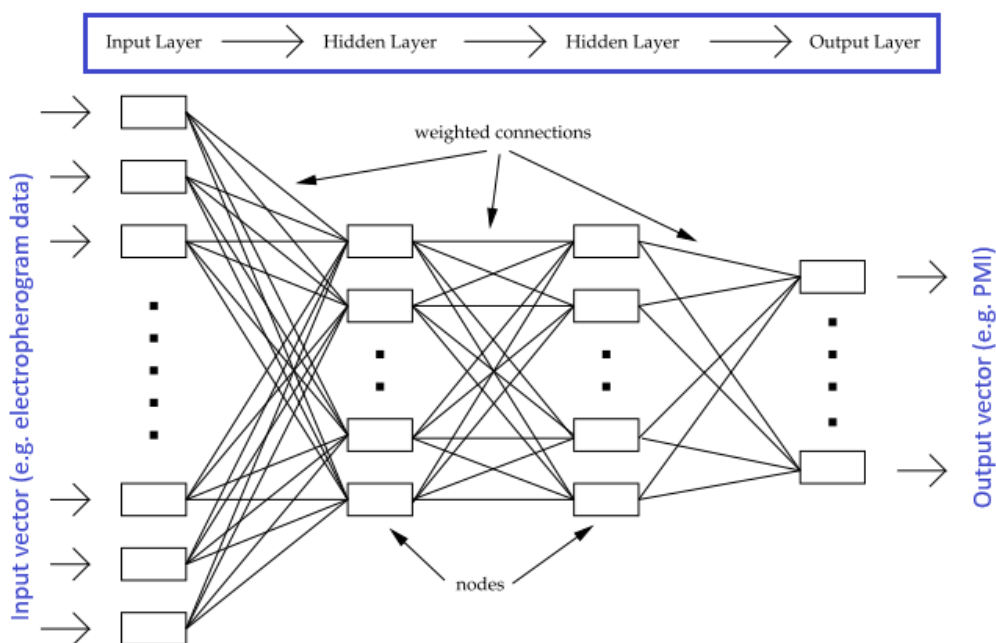
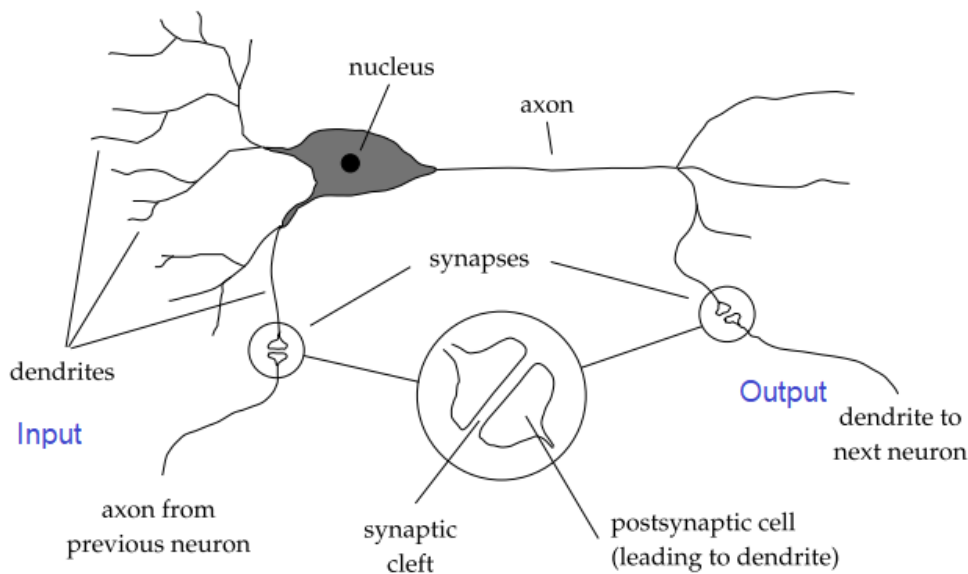
that the test and training sets will differ to some degree. If the network is over-fitted, it will perform extremely well in the training set, not so well in the test set, and badly in the unknown samples. Thus, the aim of most networks is to perform some type of data modelling and they can do it in two ways: under supervised and unsupervised conditions. Supervised networks attempt to learn a relationship between data and a parameter, whereas unsupervised networks try to find natural clusters irrespective of external restrictions [19–21].

The most important type of supervised learning networks is a feedforward multilayer perceptron or multilayer perceptron (MLP), shown in Figure 3 . The data flows forward and there is a defined input and output. As seen in Figure 3, the input data is fed to the nodes in the first hidden layer, which forms a weighted sum of the inputs and passes that sum through a non-linear transfer function, and the values are then sent to the nodes in the second hidden layer which then performs a similar operation, and the outputs are generated. In the end, the network output is the sum of its inputs [21]. The focus of the project will be in MLP, using it as a regression tool.

Neural networks can be used for pattern recognition, modelling, classification, and multivariate data analysis (i.e. regression) purposes [22]. As it can be seen, ANNs are great tools for multivariate data analysis since they possess extremely good information processing features, as presented in Table 1.

**Table 1.** Advantages and limitations of ANNs

<b>Advantages</b>	<b>Limitations</b>
<ul style="list-style-type: none"> <li>✚ Generalisation: the application of the created model to untrained (or unknown) data</li> <li>✚ Learning adaptability: the network changes its structure in response to a shifting environment</li> <li>✚ It is not necessary to know the relationship between input and output data, the network learns that relationship through training</li> <li>✚ Better fitting of the data: the network decides the function that relates the input and output data, either linear or non-linear. It decides based on the amount of training received</li> <li>✚ Noise insensitivity: It can make accurate predictions in the case of measurement errors and uncertain data</li> <li>✚ High parallelism (nodes) implies fast processing</li> </ul>	<ul style="list-style-type: none"> <li>✚ The quality of the network depends on the amount and quality of the input data. Large datasets are desirable since it provides more information for training and testing, leading to better final network quality</li> <li>✚ There is no way to estimate confidence intervals</li> <li>✚ The computing power depends on the complexity and abundance of the input data</li> </ul>



**Figure 3.** Comparison between a neuron and ANNs. The ANN has two layers, with 4 nodes each and several inputs and outputs. Adapted from Bailer-Jones et al.[21].

Even keeping into consideration their limitations, ANNs still represent a powerful tool in many fields [19,22]

### 2.1.1 NETWORK TRAINING

It is important to point out that the network does not create a new regression, rather, it interpolates the training data to generalise a relationship between the input data and the target (known outputs). In order for the network to create accurate outputs, the weights (free parameters of the network) need to be set to fitting values and this is done by training the network. MLP progresses by

minimising the error function with regard to all the network weights. For each node  $h$ , the sum value is obtained, according to Eq. 4:

$$(4) \quad sum_h = \sum_f x_f w_{fh} + \gamma_h$$

Where,  $w_{fh}$  is the product of each connection weighting from node  $f$  to node  $h$ ,  $x_f$  is the input, and  $\gamma_h$  is the bias value. This is later transformed with a transfer function (identity, sigmoid, hyperbolic tangent, and rectified linear unit) which is then used for the output. Since the aim of training is to minimise the error by changing the weights between layers, the back-propagation algorithm is used for this purpose, i.e. to adjust the weights according to error (E), Eq. 5:

$$(5) \quad E = \frac{1}{2} \sum_p \sum_f (y_{pf} - t_{pf})^2$$

Where  $E$  is the error of the network,  $y$  are the sum of the output values over  $p$  training pattern and  $f$  output nodes, and  $t_{pf}$  are the target values (desired output). Then, the error is minimised according to a function for weight optimisation (stochastic gradient descent, stochastic gradient descent-based optimiser and Broyden–Fletcher–Goldfarb–Shanno). This represents the learning process, which is iterative, and goes on until the network converges, i.e. the error reaches a minimum. As mentioned before, the aim of the network is to generalise an answer, not to memorise the training data (over-fit). In order to solve this, a regularisation technique is used [21,23]. The most common technique used is regularisation by weight decay (denominated as alpha,  $\alpha$ ), which penalises large weights to prevent them from counting more than smaller weights. A larger weight might be mistaken as more important, which can lead to over-fitting [24].

### 2.1.1.1 ACTIVATION FUNCTIONS FOR THE HIDDEN LAYERS

The activation function transforms  $sum_h$  to determine the outputs. The activation functions for the hidden layer are:

- ⇒ Identity:  $F(x) = x$ , it takes the arguments and returns them unchanged
- ⇒ Sigmoid:  $F(x) = \frac{1}{(1+e^{-x})}$ , it takes the arguments and transforms them in the range of (0,1)
- ⇒ Hyperbolic tangent:  $F(x) = \tanh(x)$ , it takes the arguments and transforms them in the range (-1,1)
- ⇒ Rectified linear unit:  $F(x) = \max(0, x)$ , used for categorical variables.

Conversely, the advantages of choosing one function over another is not yet understood. Hence, the way of knowing which function better suits the data is by trial and error [22,25].

### 2.1.1.2 SOLVERS FOR WEIGHT OPTIMISATION

The error in Eq.5 is minimised using algorithms. There are three possibilities:

- ⇒ Stochastic Gradient Descent (SGD). Using gradient descent for weight optimisation, the network stops once it reaches a minimum. However, this could be a local minimum and not a global one. An element called momentum ( $\mu$ ) is added to the weight updates, which is a parameter that allows to break from the local minima. Conversely, it might lead to slower convergence. Another element is the learning rate ( $\eta$ ) which controls the weight's update speed. If a small learning rate is selected, the search towards the global minimum will be slow but steady, and the network might take a very long time to converge. If the learning rate is high, the weight vector will vary greatly from one cycle to another and might cause the network to waver around the global minimum but never converge. Thus, during learning, it is key to observe the effects of learning rate and momentum. Gradient descent is the most frequently used solver due to ease of coding, but might not always be the most suitable option [22].
- ⇒ Adam: Stochastic Gradient Descent-based optimiser, based on adaptive moment estimation. It combines the advantages of two extensions of the SGD algorithm and the little tuning needed by the hyperparameters. Based on how quickly the weights are changing, the parameters' learning rates are adapted [26].
- ⇒ Broyden–Fletcher–Goldfarb–Shanno (L-BFGS). Used for solving nonlinear optimisation problems. Instead of using a gradient descent direction, the best descent direction is used [27].

## 2.1.2 SOFTWARE

Several possibilities exist for the creation of ANNs. The easiest one is through the use of statistical software which has pre-installed ANN analysis, such as SPSS. The main advantage is the user-friendly display when it comes to adjusting the settings of the networks and the analysis of the results. Conversely, not all parameters can be selected/unselected and it might not be updated with the latest developments in machine learning, which can be a major drawback. In addition, this statistical software uses fit-them-all basic settings which may not always suit the needs of the user [28]. Another option is MATLAB, a hybrid between a statistical program and pure coding, with cost being the main limitation [29]. Another possibility is Python [30]. Its main advantage is that it is open-source, and all additions, libraries, and extensions are free as well. In addition, it has great support and an online database for coders (Stackoverflow).

## 2.1.3 APPLICATIONS OF ANNS IN FORENSIC SCIENCE

In the forensic toxicology area, two studies combined the use of ANN with instrumental analytical methods. Bocaz-Beneventi et al.[23], optimised a CE method with ANN for the accurate determination of the post-mortem interval using the data from ammonium, potassium, and sodium as input variables, with PMI as the output. They used the TRAJAN programme, which is an ANN simulator, and the network was built for regression purposes. Butcher et al.[31], used GC-MS to

analyse larvae and a classification neural network to estimate the age of the larvae to help determine the PMI. Broséus et al.[32], used a neural network classifier to determine the ability of the ANN to classify different cannabis seeds. In the field of digital forensics, Rodríguez et al.[33], compared the performance of facial comparison software with experts to assess whether the use of software can aid experts in court. They used convolutional neural networks through open-source software. Del Espiritu et al.[34], used ANN for the identification of total and partial fingerprints. They built a classification ANN using SGD as a solver. Grantham et al.[35], developed a new algorithm for use with classifier neural networks for helping with geolocation in a crime scene. Forensic anthropology is also a field with some ANN applications. For example, Hemalatha and Rajkumar [36], developed an approach for dental age estimation, and Prescher et al.[37], developed one for the characterisation of anthropological features of nose variability, with both using classifier neural networks. In fire investigations, Zong et al.[38], were able to identify soot sources analysed by GC-MS using principal component analysis (PCA) and inputting the data to a classifier neural network. Casamento et al.[39], optimised a CE methodology with ANN for the separation of organic explosives. The pH and concentration of SDS were the input variables and the explosive's mobility were the output variables. As it can be seen, the classification ANNs are more used in comparison with regression.

### **3. ESTIMATION OF THE POST-MORTEM INTERVAL (PMI) THROUGH THE DETECTION OF $\text{NH}_4^+$ AND $\text{K}^+$ IN VITREOUS HUMOUR WITH CE**

#### **3.1 INTRODUCTION**

In modern forensic medicine, the estimation of time since death has been and still is one of the major open issues, making it one of the most attractive research subjects. In criminal investigation, the accurate estimation of the post-mortem interval (PMI), and consequently, of the time of death, often close to the time of the murder, helps to reduce the number of suspects, to support (or deny) the witnesses' statements, and to verify the reported actions of the suspects of a crime. Because of the highest importance of an accurate estimation of the time of death by objective methods, a large body of literature has accumulated in the past decades, as witnessed by numerous scientific papers and reviews [40–45] and specific books [40,46].

The determination of the PMI is traditionally based on the evaluation of post-mortem physical changes in the body, and more recently, on post-mortem chemistry. Post-mortem changes in the so-called 'early post-mortem period' (usually defined as the time between death and the appearance of generalised putrefaction) comprise rigor mortis, lividities, body cooling, stomach emptying, and in some cases, the so-called supravital phenomena. This approach, however, has important flaws, including high inter-individual variability, and a time window of application limited to 24–48 hours since death. In particular, lividities and rigor mortis also suffer from subjectivity of recording, which largely depends on the experience of the forensic pathologist. Concerning body cooling, the environmental temperature, clothing, posture, shielding against the supporting surface, body temperature before death, body dimensions, etc. are all factors that influence the phenomenon and limit substantially its reliability in real cases [45]. As a matter of fact, notwithstanding a plethora of papers on the optimisation of this approach, the calculation of PMI based on body temperature still shows important limits in terms of accuracy [40]. Moreover, the equilibration of the body temperature with the environment is almost completed within the first 20 hours after death, which limits the time window of applicability of the method. In relation to stomach emptying, the nature of the food, systemic or nervous shock, or stress and the continuation of digestion after death are all factors that make it a weak indicator. Moreover, the chronology of stomach emptying can be used to determine the time of death only if the time of the last ingestion of food is precisely known [40,45].

In contrast, post-mortem chemistry, known as thanatochemistry, is based on the analysis of the modifications of the concentrations of endogenous compounds, which are released, produced, or transformed in the body after death. Contrary to the approach based on body cooling, thanatochemistry allows for the estimation of the time of death for up to 100 hours and more [40]. Among the biological samples used so far, vitreous humour is, in principle, the most promising due to its isolation and protection from the environment and easy sampling when compared to blood or

cerebrospinal fluid [41]. Furthermore, it was found that vitreous humour, in most instances, is barely contaminated by external agents even in the late PMIs. Numerous post-mortem chemical compounds such as potassium, chloride, sodium, magnesium, calcium, phosphate, creatinine, lactate, and urea were investigated in the vitreous humour [42]. Among these different parameters, the most studied by far is potassium. Its concentration in the ocular fluid increases after death due to the passive release of intracellular potassium into the vitreous body (substantially an extracellular fluid), until the equilibrium between the extra and intracellular compartments is reached [40]. Since the early '60s until present, many studies have identified this increase as one of the main parameters for inferring the PMI [10,43,44,47,48]. However, discrepancies among researches have been reported when examining the linear relationship potassium-PMI with regard to the intercept and the slope of the regression lines, which have severely limited the practical application of the method [10]. These discrepancies can be attributed to several factors such as different causes of death, different storage conditions of the bodies, different modes of sample collection, but above all, to the different analytical techniques used by the authors, most of which were validated for clinical chemistry, but not for post-mortem chemistry [49].

The analytical techniques applied to vitreous humour can be divided into two categories: non-separation and separation methods. Non-separation methods include flame photometry and ion-selective electrodes, whereas separation methods include ion chromatography (IC) and capillary electrophoresis (CE) [10,47,50]. Non-separation analytical techniques are widely accepted in the clinical environment for the analysis of serum and urine (mostly based on ion selective electrodes), but are more susceptible to interferences due to the matrix composition, as compared to separation techniques [48]. This limitation applies particularly to cadaveric fluids, the composition of which is typically non-standardised and affected by numerous variables, the most important of which is putrefaction. On the other hand, among separation techniques, capillary electrophoresis (CE) is known for its selectivity, low sample consumption, high efficiency, resolution, and speed. Furthermore, CE, lacking a packed column (as IC) which can easily be clogged or contaminated, can tolerate "dirty" samples, such as cadaveric fluids. In fact, CE was used for the analysis of potassium in vitreous humour with very encouraging results, allowing for the direct injection of diluted vitreous humour micro samples [10,48]. On the other hand, Zhou et al.[47], using low pressure IC fitted with a self-made conductivity detector, also reported promising results, but the application to real cases was limited to PMIs up to 27 hours.

On the basis of the above considerations and personal experience with potassium analysis, CE was chosen in an attempt to study other ions in the vitreous humour, specifically ammonium. This ion, a typical product of post-mortem protein deamination, has long been almost neglected, probably due to the well-known problems for its analysis in clinical chemistry [51]. To the best of our knowledge, only a single paper appeared in 1978 reporting ammonium determination in the vitreous humour.



For this purpose, a colorimetric method was used based on Nessler's reagent. The results, although obtained with an outdated non-separation method, look quite promising [52].

Taking into account the above-mentioned considerations, a CE method for the determination of ammonium in the vitreous humour has been developed and validated with the aim of providing a new and reliable tool for thanatochemical studies. Further aims of the research were to carry out a preliminary investigation on the behaviour of ammonium concentrations in the vitreous humour of real cases, and given positive results, to further study its applicability and usefulness alone and in combination with potassium to achieve more reliable estimation, by thorough statistical analysis of the post-mortem interval, and consequently, of the time of death. In addition, the present work also includes a specific study aimed at verifying the consistency of the post-mortem increase of ammonium concentrations in the vitreous humour by comparing the results from the two eyes at the same time after death.

## **3.2 MATERIALS AND METHODS**

### **3.2.1 CHEMICALS**

All chemicals were analytical-reagent grade. Imidazole (99% pure) was obtained from Sigma (St. Louis, MO, USA), 18-crown-6 ether (99% pure) and  $\alpha$ -hydroxybutyric acid (HIBA) (99% pure) from Aldrich (Milan, Italy). Standard solutions of ammonium ( $\text{NH}_4^+$ ) and barium ( $\text{Ba}^{2+}$ ) were prepared from AnalaR salts ( $\text{NH}_4\text{Cl}$  and  $\text{BaCl}_2$ ) (Merck, Darmstadt, Germany). The electrophoretic buffer pH was adjusted to the desired pH with glacial Acetic Acid (Merck, Germany). Ultrapure water (milliQ) was obtained using a water purification system Purelab Chorus (Elga Veolia, High Wycombe, UK).

### **3.2.2 CE SYSTEM AND SOFTWARE**

All experiments were performed using a P/ACE MDQ Capillary Electrophoresis System (Beckman, Fullerton, CA, USA) equipped with a UV filter detector. The capillary was thermostated with a perfluorinated coolant flowing in the capillary cartridge. During all experiments, untreated fused-silica capillaries (75  $\mu\text{m}$  I.D., 50 cm effective length; Beckman) were used with a detection window of 200x100  $\mu\text{m}$ . Beckman P/ACE Station (version 8.0) was used for instrument control, data acquisition, and processing.

### **3.2.3 ELECTROPHORETIC CONDITIONS**

Electrophoretic separations were performed using a running buffer consisting of 6 mM HIBA, 5 mM 18-crown-6 ether, and 5 mM imidazole adjusted to pH 4.5. Constant voltage runs were performed in all experiments by applying a voltage of 500 V/cm at 25°C with a resulting current of about 18  $\mu\text{A}$ . The separation took place at 25 kV for 4 minutes. To overcome the lack of optical absorbance of ammonium, the ions were detected using indirect UV detection at 214 nm. The analytes were injected at the anodic end of the capillary for 10 s at 0.5 psi. The capillary was conditioned every day

with a solution containing 1 M NaOH (20 min) followed by water (15 min) and finally with the running buffer (10 min). Between consecutive runs the capillary was washed with water (3 min) and then with the running buffer (2 min). Water blanks were routinely checked for contamination with trace amounts of cations.

### **3.2.4 PREPARATION OF STANDARDS AND SAMPLES**

All solutions were prepared in polypropylene vials. A 100 mM stock solution of  $\text{NH}_4^+$  was prepared by dissolving adequate amount of  $\text{NH}_4\text{Cl}$  in milliQ water. Ammonium working standard solution was prepared daily to the desired concentration by diluting with water.  $\text{BaCl}_2$  internal standard (IS) solution was prepared by dissolving the appropriately weighed amount to achieve a final concentration of 40  $\mu\text{g/mL}$ .

Vitreous humour samples were collected from a total of 38 medico-legal autopsies or external examination of corpses of violent or sudden deaths, in which the time of death was exactly known (Appendix I. 6.1.1 List of vitreous humour samples used for the study). The causes of death included mostly road accidents and other traumas. For the linear and polynomial approaches, samples with low PMIs were not considered. The vitreous humour samples were prepared by diluting 1:20 with the IS solution. Samples were analysed in triplicate. Because of the lack of recent deaths at the time of the present work and considering only the preliminary application of the method to the study of real forensic cases, most of the samples analysed had been stored at  $-24^\circ\text{C}$  for several months (up to 5 months) before analysis.

### **3.2.5 METHOD VALIDATION**

The analytical method presented was developed based on a previous study on the determination of potassium in vitreous humour [10], where a peak attributed to ammonium could only be identified, but never formally studied [47]. Therefore, the present work was conducted for the optimisation and validation of ammonium determination by considering the following parameters: precision, bias, linearity, sensitivity and selectivity.

The optimization and validation of methods is a way to establish if they meet the analytical requirements for which they were designed. It evaluates the method's capability to be used in a specific application. Method development and validation are intertwined since the method performance characteristics are also considered when a method is validated. Method validation is a necessity since many methods are used to make conclusions, for example how long a person has been dead. Therefore, it is of high importance to produce accurate results. The performance parameters evaluated in method optimization and validation for quantitative methods are as follow: bias (trueness), precision, limit of detection (LOD), limit of quantification (LOQ), working range (linearity) and selectivity [53]. Precision and bias studies can be carried out simultaneously since both should be assessed at three different concentration levels (low, middle and high) and be

evaluated for at least 3 days. Precision results are expressed by the analytical reproducibility (RSD%) and are an indication of how reproducible the results are in the same day on different runs and over time (5 days or more). The RSD values should be within  $\pm 15\%$ . Bias is expressed with the analytical bias (%) and it is an indication of the error between the measured value and the true value. The values should be between  $\pm 15\%$  for methods using biological matrices. The linearity is the range of analyte concentrations in which the method will be used. It is highly important to correctly establish the working limits of this linearity range, otherwise accurate results might not be obtained. The most common approach is a linear regression. LOD and LOQ are also evaluated parallel to linearity. LOD is the lowest concentration of analyte that can be measured with statistical certainty and the LOQ is the lowest concentration of analyte that can be quantified with a suitable level of trueness and precision. Selectivity indicates the ability of the method to correctly identify the analyte, in the presence of other compounds. There are several ways to evaluate it, but the relevant one for this work is using complexing agents [54,55].

In CE, complexing agents are substances added in the electrolyte solution to shift migration times in the case of cations with similar electrophoretic mobilities. Lactic acid and crown ethers are two common examples. Specifically, crown ethers have been used successfully for the electrophoretic separation of  $\text{NH}_4^+$  and  $\text{K}^+$  [56–58]. The crown ether interacts with  $\text{K}^+$  and slows its migration time. The specific complexing mechanism is based on the relationship between the cavity size of the crown ether and the radius of the cations. Precisely, the 18-Crown-6 ether has a radius of 1.58 Å and the ionic radius of  $\text{K}^+$  is 1.58 Å, whereas the radius of  $\text{NH}_4^+$  is 1.51 Å. Therefore, potassium fits perfectly well inside the crown ether and its migration time is slowed allowing for a separation from the  $\text{NH}_4^+$  [56]. In this work, two complexing agents have been combined, namely HIBA and 18-Crown-6 ether. In this case HIBA does not complex  $\text{NH}_4^+$  and  $\text{K}^+$  but modifies the mobility of  $\text{Ba}^{2+}$  [59]. For the present study, the constituents of the background electrolyte have been adapted from a previously developed CE methodology for the determination of  $\text{K}^+$  in vitreous humour [10].

### **3.2.6 STATISTICAL ANALYSIS**

To better tune the usefulness and accuracy of the present validated method, three different statistical methods were performed. Firstly, preliminary results were fitted to a linear method. Secondly, to increase correlation and accuracy of predictions, the logarithmic correlation was explored. Finally, to further improve PMI estimation, ANNs were performed in the final data (obtaining a multivariate statistical analysis). Linearity and logarithmic analysis were done using Microsoft Excel 2013 software. The ANNs analysis was performed using Scikit-learn module (0.23.2) MLPRegressor [25] in Jupyter Notebook environment (version 6.0.3), from the Anaconda Navigator (anaconda3) [60]. Data was manipulated with Pandas and Numpy [25]. The code was written with Python 3.8.

### 3.3 RESULTS AND DISCUSSION

The previously described procedures have been applied to the data. The most relevant results are discussed.

#### 3.3.1 PRECISION: INTRA- AND INTER-DAY STUDIES

The intra-day precision was verified at three concentration levels (low, middle, and high). The standards were diluted 1:20 with IS. Seven injections were made for each concentration level. To calculate the inter-day precision, determinations were performed on three different non-consecutive days. In addition, the stability and reproducibility in authentic vitreous humour matrix was studied.

**Table 2.** Reproducibility (RSD%) of the CZE analysis of ammonium in standard solutions (n=3) tested in three different days.

Concentration (mM)	Relative area RSD (%)				Relative migration time RSD (%)			
	Intra-day			Inter-day	Intra-day			Inter-day
	Day 1	Day 2	Day 3		Day 1	Day 2	Day 3	
5	2.12	0.74	1.64	7.32	0.54	0.48	0.81	2.44
1.25	8.28	3.53	5.82	11.43	0.56	0.46	0.20	3.99
0.31	3.28	6.78	5.48	18.51	0.44	0.61	0.39	3.81

Three real samples were selected and analysed over a period of 5 days with 5 injections per sample. In all the studies, the concentration of the  $\text{NH}_4^+$ , as well as the migration time of this ion were recorded and statistically analysed in terms of Relative Standard Deviation (%RSD). The results of the Intra- and inter-day precision studies carried out on the ammonium standard solutions are summarised in Table 2, and those from real samples are presented in Table 3. From both the tables, the results show that in all cases, the analytical precision was acceptable, with %RSDs well below 20%, even in real biological samples.

**Table 3.** Analytical reproducibility (RSD%) of CZE analysis of ammonium in vitreous humour (n=3) tested on five different days.

Sample number	Relative area RSD (%)						
	Intra-day					Inter-day	
	Day 1	Day 2	Day 3	Day 4	Day 5		
1	7.03	5.89	6.12	4.14	5.02	12.90	
2	8.40	2.91	6.50	7.52	4.62	4.40	
3	9.93	4.51	5.38	2.68	1.34	4.88	
Sample number	Relative migration time RSD (%)						
	1	1.70	0.28	0.49	0.57	0.16	1.04
	2	0.48	0.20	0.18	3.47	0.20	1.98
	3	0.57	0.64	0.99	1.26	0.78	1.64

### 3.3.2 BIAS

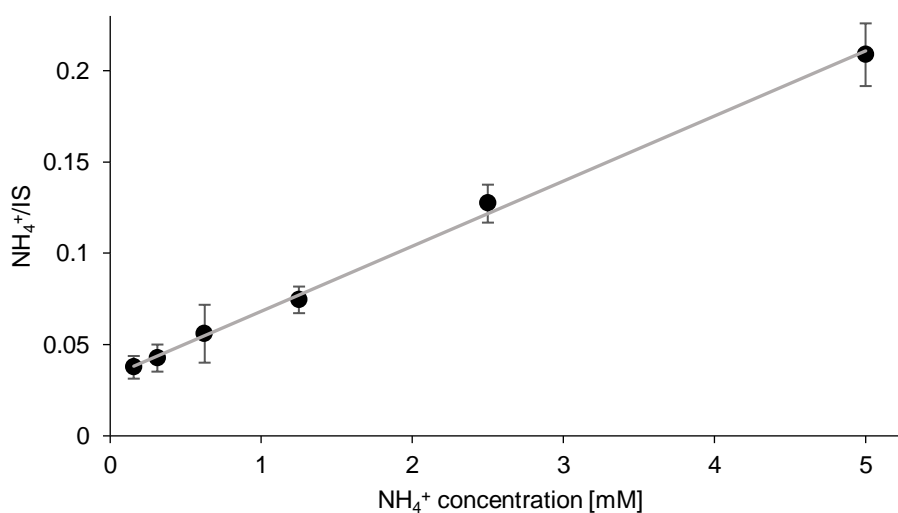
Bias was calculated at three different concentrations (0.312, 1.25, and 5.0 mM). However, in order to also study the analytical bias in real matrix, a standard addition experiment was performed, by spiking known amounts of  $\text{NH}_4^+$  standard (15 and 10 mM) into real vitreous humour samples. Bias was calculated according to the formulas reported by the Scientific Working Group for Forensic Toxicology (SWGTOX) [54]. The bias results from the spiking test were -1.24 % and 7.13 % for 15 and 10 mM, respectively. Overall, the bias was fairly acceptable as shown in Table 4.

**Table 4.** Analytical bias (%) of CZE analysis of ammonium in standard solutions (n=3)

Bias (%) of the relative area			
Concentration (mM)	Day 1	Day 2	Day 3
5	0.93	0.75	-0.49
1.25	-1.65	-0.45	-5.37
0.31	-5.64	-0.21	1.36

### 3.3.3 LINEARITY, LOD, AND LOQ

For the linearity studies, 5 vials with increasing  $\text{NH}_4^+$  concentrations were measured at the following concentrations: 0.16, 0.31, 0.63, 1.25, 2.5, and 5.0 mM. Each concentration level was measured in duplicates. A calibration curve was obtained by plotting the relative peak area ( $\text{NH}_4^+/\text{IS}$ ) against the concentration of ammonium (Figure 4).



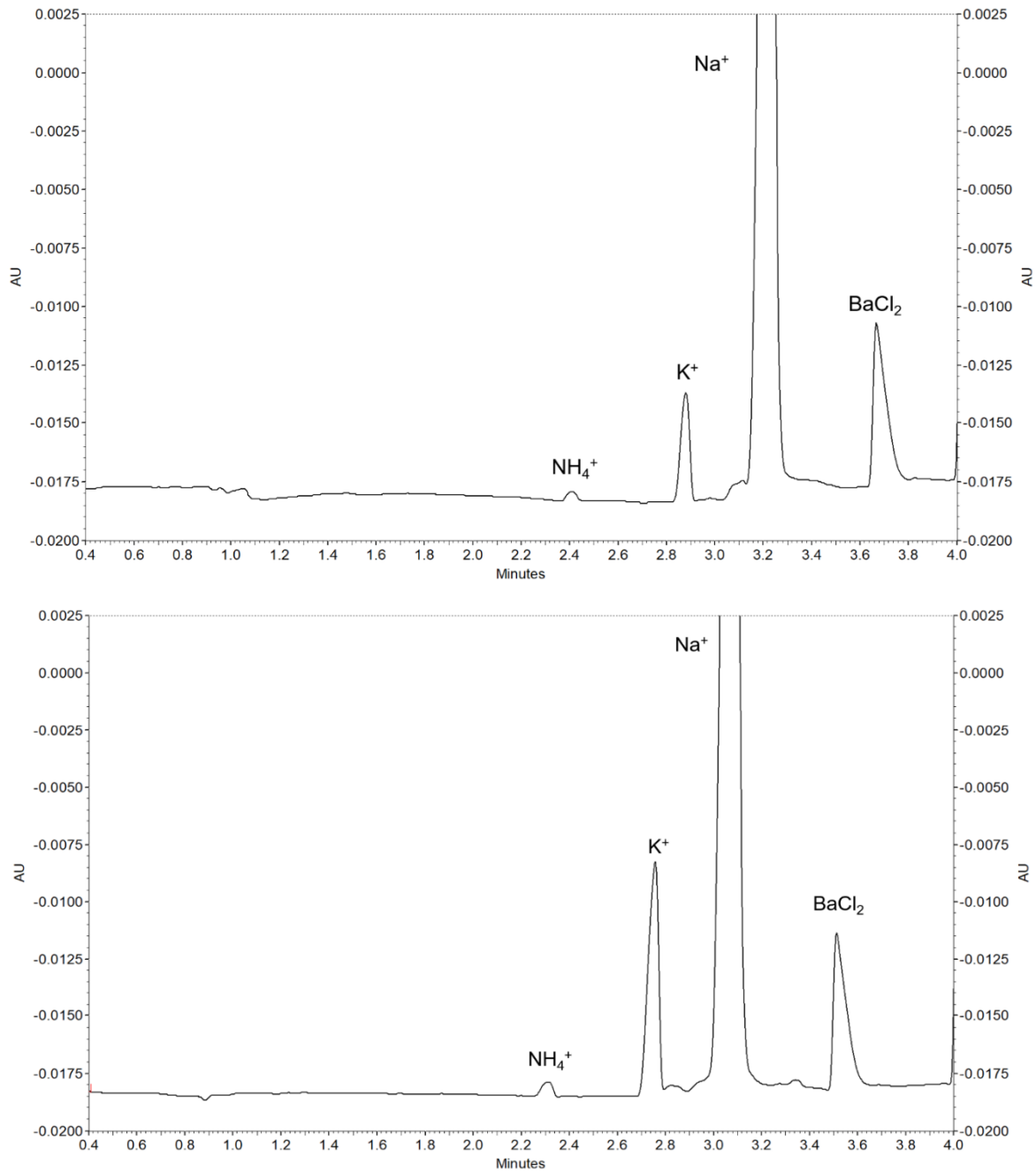
**Figure 4.** Linear calibration curve obtained by plotting the relative peak area ( $\text{NH}_4^+/\text{IS}$ ) against the concentration of  $\text{NH}_4^+$  (mM).

The results were linearly correlated according to the equation  $y = (0.0357 \pm 0.0008 \text{ SD}) x + (0.033 \pm 0.002 \text{ SD})$ ;  $r^2 = 0.998$  ( $y$  = relative peak area;  $x$  = concentration of  $\text{NH}_4^+$  in mM). The limit of detection (LOD) and the limit of quantification (LOQ) were determined to be 0.039 mM and 0.31 mM, respectively.

### **3.3.4 DISCUSSION OF THE PRELIMINARY STUDY (LINEAR APPROACH)**

Strangely enough, ammonium, a typical product of post-mortem protein deamination, has rarely been considered as a potentially useful indicator of the time elapsed since death. This can probably be justified by the known variability of putrefaction phenomena occurring in corpses, which are responsible for most of the release of ammonium. Also, ammonium analysis in clinical chemistry is generally seen as a very problematic issue, with a large uncertainty caused by pre-analytical and analytical problems [51,61]. These considerations are presumably due to a general lack of interest for the determination of this analyte in forensic pathology as well. However, vitreous humour is typically sterile and is well protected from the environment by the sclera and the cornea, and therefore, less susceptible to putrefaction. Therefore, in the field of forensic pathology, it appears to be a suitable analyte for the evaluation of a time-dependent post-mortem increase in ammonium concentration.

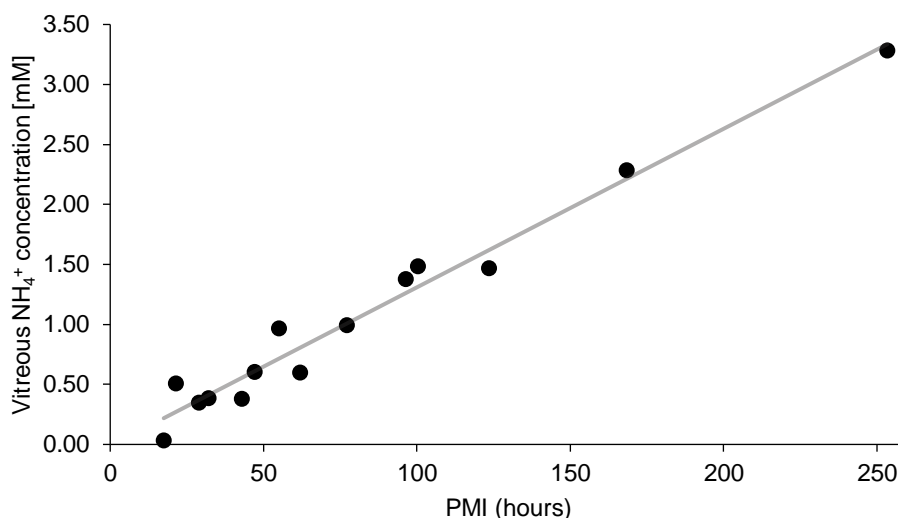
For this purpose, in order to avoid possible interferences originating from a non-standardised matrix such as a forensic sample, a separation method was preferred. On the grounds of previous research [10], CE was chosen as the most suitable analytical technique. Because of the lack of UV absorbance of the ammonium ion and in order to avoid the use of a conductivity detector, indirect detection using imidazole as the UV absorbing additive was adopted. Moreover, in order to improve the separation between ammonium and potassium peaks, 18-crown ether was added to the background electrolyte as a complexing agent, selectively retarding the mobility of the different cations. The validation parameters, according to the SWGTOX criteria, have been fully satisfied [54] even when performing minimal sample pre-treatment (limited to a 1:20 dilution). Because of the selectivity of indirect detection, extremely clean electropherograms were obtained even with injection of biological samples without any pre-treatment, but instead diluted with the I.S. solution (Figure 5).



**Figure 5.** Typical electropherograms of human vitreous humour (Top:  $\text{NH}_4^+ = 0.39$ ;  $\text{K}^+ = 8.62$ ; PMI= 32 h; bottom:  $\text{NH}_4^+ = 1.18$ ;  $\text{K}^+ = 20.81$ ; PMI= 99 h)

### 3.3.4.1 APPLICATION TO REAL CASES

A preliminary study to investigate an expected correlation between ammonium concentration in the vitreous humour and PMI was conducted on 14 subjects whose time since death was well known on the basis of circumstantial information. The results are shown in Figure 6. This shows that ammonium concentrations are well correlated with the PMI ( $r^2 = 0.970$ ).



**Figure 6.** Correlation between PMI (hours) and vitreous NH<sub>4</sub><sup>+</sup> concentration (mM) determined in 14 forensic deaths with known time since death. Equation  $y = (0.0132)x + (-0.012)$ ;  $r^2 = 0.970$ , in which  $y = \text{NH}_4^+$  concentration (mM);  $x = \text{PMI}$  (hours).

Finally, as already reported by Tagliaro et al. [10], under the reported analytical conditions, other ions are also detectable, including sodium and potassium [23]. This opportunity may offer the possibility of grounding PMI evaluations onto two substantially independent parameters, such as ammonium and potassium, with a high expected increase in accuracy and reliability. This preliminary approach shows a new and simple application of CE. The next step is to fully validate the method with more cases.

### 3.3.5 DISCUSSION OF THE LOGARITHMIC APPROACH

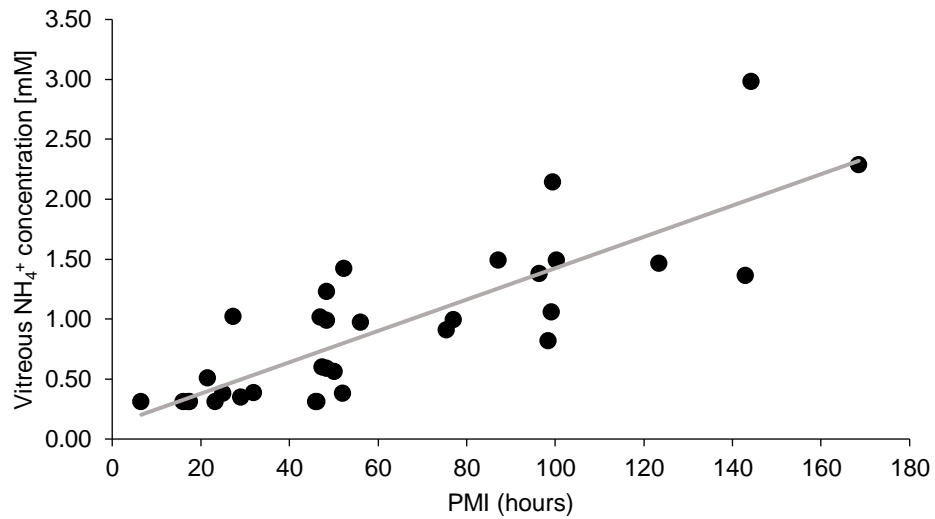
As the number of samples increased (from 14 to 33), the correlation of the NH<sub>4</sub><sup>+</sup> concentration with the PMI was not as accurate as expected (Figure 7). Therefore, with a total of 33 forensic cases (Appendix I. 6.1.1 List of vitreous humour samples used for the study) a different statistical approach was explored. Since this approach was aimed at improving the correlation and the validation of the simultaneous determination of the two ions in post mortem cases, in a first step, calibration curves for ammonium and potassium were produced based on 33 medico-legal autopsies; the results are displayed in Figure 8 and Figure 9, respectively, showing a statistically significant relationship between both ions and PMI. The correlation is better described with polynomial equations as follows:

$$\text{Ammonium (NH}_4^+): y = 2 \times 10^{-6}x^2 + 0.0127x + 0.1461$$

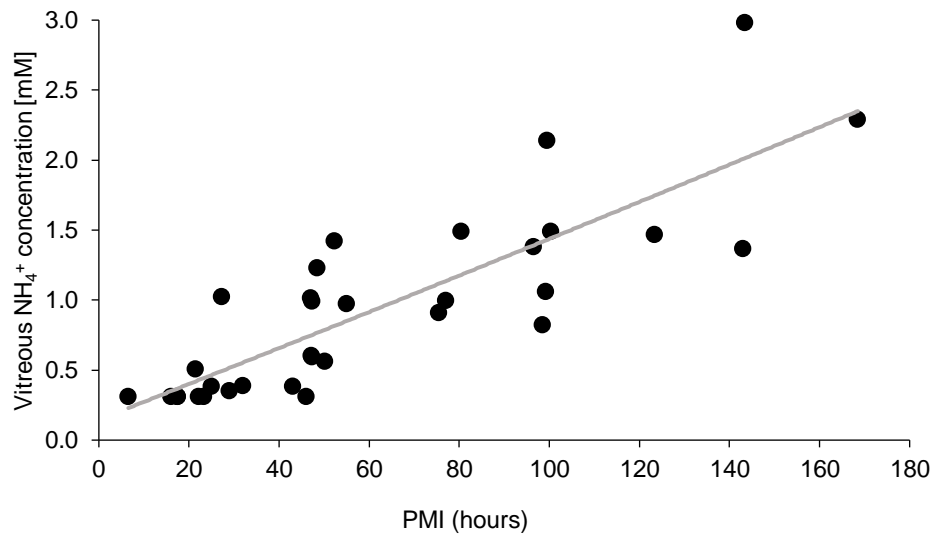
$$\text{Potassium (K}^+): y = -0.0005x^2 + 0.2018x + 6.173$$

The correlation coefficient for vitreous humour ammonium vs. PMI is  $r^2 = 0.70$  and that for potassium is  $r^2 = 0.75$ , which is an improvement over the linear correlation ( $r^2 = 0.67$ , Figure 7)

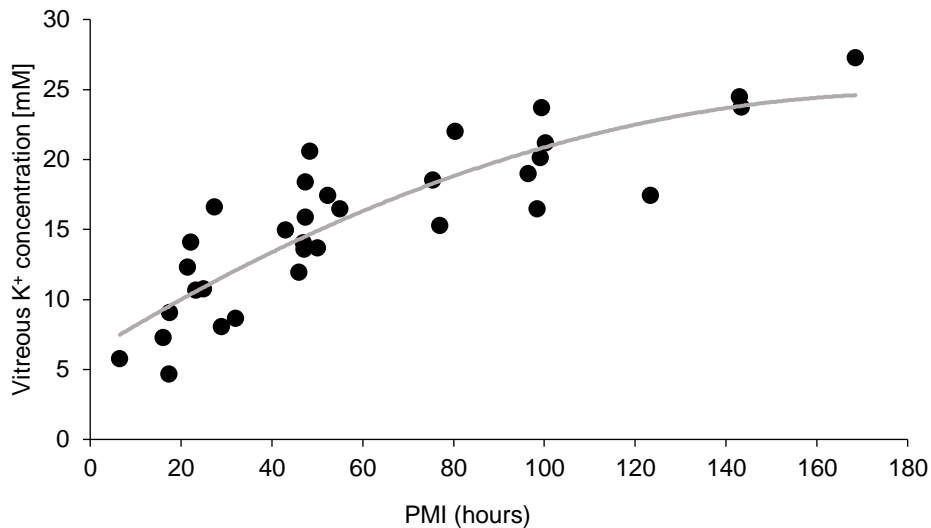




**Figure 7.** Linear correlation of the concentration of ammonium against the PMI of 33 cases. Equation:  $y = 0.0131x + 0.1172$ ,  $r^2=0.67$ .

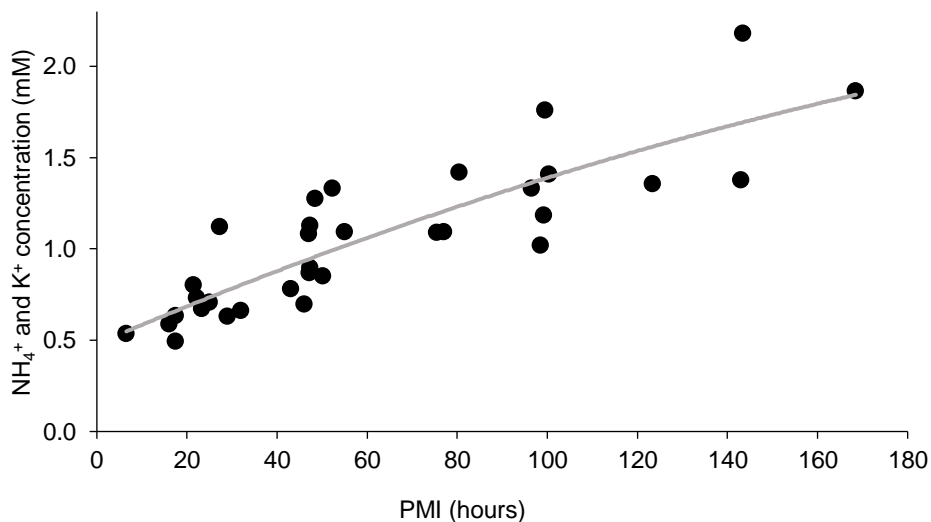


**Figure 8.** Second degree polynomial correlation between PMI (hours) and vitreous  $\text{NH}_4^+$  concentration (mM) determined in 33 forensic cases with known time of death. Equation  $y = 2 \times 10^{-6}x^2 + 0.0127x + 0.1461$ ;  $r^2 = 0.70$  in which  $x = \text{PMI (hours)}$ ; and  $y = \text{NH}_4^+$  concentration (mM).



**Figure 9.** Correlation between PMI (hours) and vitreous K<sup>+</sup> concentration (mM) determined in 33 forensic cases with known time of death. Equation  $y = -0.0005x^2 + 0.2018x + 6.173$ ;  $r^2 = 0.75$  in which  $x = \text{PMI (hours)}$ ; and  $y = \text{K}^+$  concentration (mM).

In a second step, an integrated use of these two parameters by using the average of potassium and ammonium concentrations was also tested, as shown in Figure 10. With the aim of balancing the weight of the two parameters in the calculation, the logarithm of the potassium concentration was used. This approach led to a small increase in the correlation coefficient ( $r^2=0.74$ ).



**Figure 10.** Correlation between PMI (hours) and combined vitreous NH<sub>4</sub><sup>+</sup> and K<sup>+</sup> concentrations (mM) determined in 33 forensic cases with known time of death. Equation  $y = -0.00001x^2 + 0.0106x + 0.4799$ ;  $r^2 = 0.74$  in which  $x = \text{PMI (hours)}$ ; and  $y = \text{NH}_4^+$  and K<sup>+</sup> concentrations combined.

In order to calculate the experimental error during practical use of the method, each case was treated, one-by-one, as “unknown” and the theoretical PMI was calculated after having excluded its ion concentrations from the calibration curve. The theoretical PMIs were compared to the observed

values, and the differences recorded (Table 5, Table 6, Table 7). The average absolute error of estimation for potassium and ammonium was 17.8 hours (SD = 15.3) and 20.2 hours (SD = 15.6), respectively.

**Table 5.** Estimation of error for potassium. The coefficients of the curve are denoted by a,b,c, ( $y = ax^2 + bx + c$ ). Estimated PMI is obtained as the solution to the quadratic equation.

VH sample	[K <sup>+</sup> ] (mM)	PMI (hours)	a	b	c	Estimated PMI (hours)	Absolute error (hours)	Relative error	Error as % PMI
1	15.28	77.1	-0.0006	0.2130	5.9396	51.3	-25.8	-0.34	-33.5
2	14.03	47	-0.0005	0.2023	6.1764	43.5	-3.5	-0.08	-7.5
3	12.30	21.5	-0.0005	0.2078	5.8904	33.6	12.1	0.56	56.1
4	24.47	143	-0.0005	0.2030	6.1570	135.2	-7.8	-0.05	-5.4
5	11.94	46	-0.0006	0.2044	6.2006	30.9	-15.1	-0.33	-32.9
6	22.01	80.45	-0.0005	0.1904	6.4179	119.2	38.7	0.48	48.1
7	4.67	17.45	-0.0005	0.1830	6.9891	-12.3	-29.7	-1.70	-170.3
8	17.43	123.45	-0.0006	0.2093	5.9239	68.4	-55.1	-0.45	-44.6
9	8.62	32	-0.0005	0.1986	6.4336	11.3	-20.7	-0.65	-64.6
10	16.47	55	-0.0005	0.2000	6.1979	60.5	5.5	0.10	10.0
11	14.95	43	-0.0005	0.2010	6.1465	50.0	7.0	0.16	16.4
12	27.23	168.5	-0.0009	0.2412	5.3116	162.5	-6.0	-0.04	-3.6
13	7.27	16.1	-0.0005	0.1930	6.5324	3.9	-12.2	-0.76	-76.0
14	13.56	47.2	-0.0006	0.2029	6.1795	41.5	-5.7	-0.12	-12.2
15	21.18	100.3	-0.0005	0.2008	6.1971	99.0	-1.3	-0.01	-1.3
16	8.04	29	-0.0005	0.1967	6.4968	8.0	-21.0	-0.72	-72.4
17	18.98	96.45	-0.0006	0.2072	6.0412	81.8	-14.6	-0.15	-15.1
18	17.43	52.3	-0.0005	0.1978	6.1987	68.7	16.4	0.31	31.4
19	23.66	99.45	-0.0005	0.1922	6.4117	142.8	43.3	0.44	43.6
20	16.43	98.45	-0.0006	0.2165	5.8112	58.5	-39.9	-0.41	-40.5
21	20.11	99.16	-0.0006	0.2041	6.1162	95.2	-3.9	-0.04	-4.0
22	23.73	143.4	-0.0005	0.2016	6.1756	127.2	-16.2	-0.11	-11.3
23	10.75	25	-0.0005	0.2015	6.1875	24.1	-0.9	-0.04	-3.7
24	14.08	22.24	-0.0006	0.2120	5.6808	45.5	23.2	1.04	104.5
25	10.66	23.35	-0.0005	0.2020	6.1643	23.6	0.3	0.01	1.1
26	9.03	17.55	-0.0005	0.1998	6.2599	14.4	-3.2	-0.18	-18.1
27	18.39	47.4	-0.0005	0.1968	6.1494	77.4	30.0	0.63	63.4
28	15.86	47.4	-0.0005	0.2001	6.1652	56.4	9.0	0.19	19.0
29	13.67	50.15	-0.0006	0.2037	6.1687	42.0	-8.1	-0.16	-16.2
30	20.58	48.45	-0.0005	0.1935	6.1602	100.7	52.3	1.08	107.9
31	5.74	6.55	-0.0005	0.1893	6.6474	-4.7	-11.3	-1.72	-172.3
32	18.51	75.5	-0.0005	0.2010	6.1899	75.5	0.0	0.00	-0.1
33	16.57	27.35	-0.0006	0.2110	5.6377	63.2	35.8	1.31	130.9

**Table 6.** Estimation of error for potassium. The coefficients of the curve are denoted by a,b,c, ( $y = ax^2 + bx + c$ ). Estimated PMI is obtained as the solution to the quadratic equation.

VH sample	[NH <sub>4</sub> <sup>+</sup> ] (mM)	PMI (hours)	a	b	c	Estimated PMI (hours)	Absolute error (hours)	Relative error	Error as % PMI
1	0.99	77.1	-5.00E-07	0.0132	0.1356	65.2	-11.88	-0.15	-15.41
2	1.01	47	5.00E-06	0.0123	0.1440	68.7	21.74	0.46	46.26
3	0.51	21.5	1.00E-06	0.0129	0.1340	28.8	7.35	0.34	34.16
4	1.36	143	2.00E-05	0.0115	0.1617	90.3	-52.71	-0.37	-36.86
5	0.31	46	1.00E-06	0.0131	0.1510	12.1	-33.87	-0.74	-73.64
6	1.49	80.45	9.00E-06	0.0115	0.1702	106.1	25.66	0.32	31.90
7	0.31	17.45	4.00E-06	0.0124	0.1560	12.4	-5.08	-0.29	-29.11
8	1.47	123.45	2.00E-06	0.0131	0.1330	100.2	-23.24	-0.19	-18.83
9	0.39	32	3.00E-06	0.0125	0.1590	18.0	-13.99	-0.44	-43.73
10	0.97	55	4.00E-06	0.0124	0.1487	64.7	9.74	0.18	17.72
11	0.38	43	3.00E-07	0.0129	0.1537	17.6	-25.41	-0.59	-59.10
12	2.29	168.5	1.00E-05	0.0117	0.1667	159.5	-9.01	-0.05	-5.35
13	0.31	16.1	3.00E-06	0.0125	0.1535	12.5	-3.62	-0.22	-22.47
14	0.60	47.2	1.00E-06	0.0129	0.1472	35.3	-11.87	-0.25	-25.15
15	1.49	100.3	3.00E-06	0.0125	0.1502	104.2	3.89	0.04	3.88
16	0.35	29	3.00E-06	0.0124	0.1614	15.2	-13.85	-0.48	-47.74
17	1.38	96.45	2.00E-06	0.0127	0.1453	95.8	-0.67	-0.01	-0.70
18	1.42	52.3	1.00E-05	0.0116	0.1532	100.5	48.20	0.92	92.16
19	2.14	99.45	1.00E-05	0.0103	0.2057	162.2	62.79	0.63	63.14
20	0.82	98.45	-8.00E-06	0.0147	0.0958	50.7	-47.79	-0.49	-48.54
21	1.06	99.16	-4.00E-06	0.0139	0.1154	69.3	-29.82	-0.30	-30.07
22	2.98	143.4	2.00E-05	0.0145	0.1207	161.3	17.90	0.12	12.49
23	0.38	25	3.00E-06	0.0125	0.1559	17.9	-7.15	-0.29	-28.59
24	0.31	22.24	4.00E-06	0.0123	0.1621	12.0	-10.26	-0.46	-46.14
25	0.31	23.35	4.00E-06	0.0123	0.1630	11.9	-11.44	-0.49	-49.01
26	0.31	17.55	4.00E-06	0.0124	0.1561	12.4	-5.19	-0.30	-29.56
27	0.99	47.4	5.00E-06	0.0124	0.1447	66.4	18.99	0.40	40.07
28	0.59	47.4	9.00E-07	0.0129	0.1472	34.2	-13.16	-0.28	-27.76
29	0.56	50.15	1.00E-07	0.0130	0.1453	31.9	-18.26	-0.36	-36.41
30	1.23	48.45	7.00E-06	0.0120	0.1451	86.1	37.64	0.78	77.68
31	0.31	6.55	-6.00E-07	0.0133	0.1241	14.0	7.44	1.14	113.53
32	0.91	75.5	-2.00E-06	0.0134	0.1315	58.6	-16.89	-0.22	-22.37
33	1.02	27.35	-1.00E-06	0.0136	0.0929	68.5	41.16	1.51	150.51

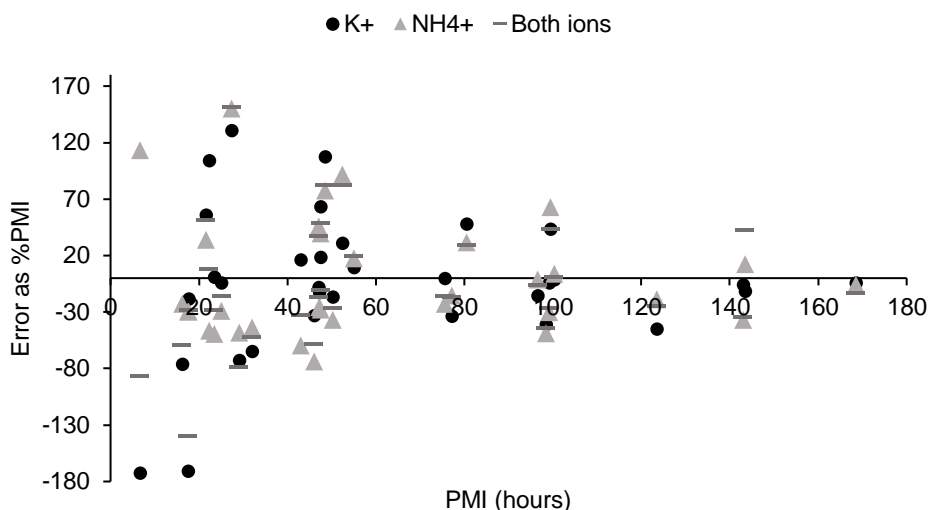
**Table 7.** Estimation of error for the combined ammonium and potassium. The coefficients of the curve are denoted by a,b,c, ( $y = ax^2 + bx + c$ ). Estimated PMI is obtained as the solution to the quadratic equation.

VH sample	$[\text{NH}_4^+] + \log[\text{K}^+]$ (mM)	PMI (hours)	a	b	c	Estimated PMI (hours)	Absolute error (hours)	Relative error	Error as % PMI
1	1.09	77.1	-2.00E-06	0.0092	0.5089	64.0	-13.13	-0.17	-17.03
2	1.08	47	8.00E-07	0.0087	0.5146	64.6	17.61	0.37	37.46
3	0.80	21.5	-2.00E-06	0.0091	0.5038	32.6	11.11	0.52	51.69
4	1.38	143	4.00E-06	0.0087	0.5152	94.8	-48.21	-0.34	-33.71
5	0.69	46	-2.00E-06	0.0091	0.5200	19.1	-26.85	-0.58	-58.38
6	1.42	80.45	3.00E-06	0.0082	0.5308	104.1	23.68	0.29	29.43
7	0.49	17.45	3.00E-06	0.0082	0.5460	-6.9	-24.34	-1.39	-139.47
8	1.35	123.45	-2.00E-06	0.0093	0.5017	93.5	-29.97	-0.24	-24.28
9	0.66	32	2.00E-08	0.0087	0.5271	15.3	-16.70	-0.52	-52.18
10	1.09	55	8.00E-07	0.0087	0.5180	65.6	10.64	0.19	19.35
11	0.78	43	-1.00E-06	0.0089	0.5197	29.1	-13.92	-0.32	-32.37
12	1.86	168.5	8.00E-06	0.0078	0.5390	147.3	-21.24	-0.13	-12.61
13	0.59	16.1	1.00E-06	0.0086	0.5291	6.6	-9.52	-0.59	-59.10
14	0.87	47.2	-1.00E-06	0.0089	0.5170	39.4	-7.80	-0.17	-16.52
15	1.41	100.3	-3.00E-07	0.0088	0.5171	101.6	1.29	0.01	1.28
16	0.63	68	-8.00E-05	0.0087	0.5146	15.1	-52.92	-0.78	-77.82
17	1.33	96.45	-1.00E-06	0.0090	0.5190	90.9	-5.51	-0.06	-5.72
18	1.33	52.3	3.00E-06	0.0083	0.5132	95.2	42.91	0.82	82.05
19	1.76	99.45	6.00E-06	0.0076	0.5477	143.0	43.53	0.44	43.77
20	1.02	98.45	-7.00E-06	0.0101	0.4849	54.8	-43.60	-0.44	-44.29
21	1.18	99.16	-4.00E-06	0.0096	0.4985	73.4	-25.75	-0.26	-25.97
22	2.18	143.4	-5.00E-06	0.0091	0.5170	205.7	62.35	0.43	43.48
23	0.71	25	-1.00E-07	0.0088	0.5199	21.1	-3.88	-0.16	-15.52
24	0.73	22.24	-6.00E-07	0.0089	0.5142	24.2	1.97	0.09	8.85
25	0.67	23.35	2.00E-07	0.0087	0.5231	16.7	-6.61	-0.28	-28.31
26	0.63	17.55	3.00E-07	0.0087	0.5227	12.7	-4.90	-0.28	-27.89
27	1.13	47.4	1.00E-06	0.0086	0.5144	70.7	23.29	0.49	49.13
28	0.90	47.4	-7.00E-07	0.0089	0.5167	42.7	-4.73	-0.10	-9.99
29	0.85	50.15	-1.00E-06	0.0090	0.5163	37.0	-13.16	-0.26	-26.23
30	1.27	48.45	2.00E-06	0.0084	0.5144	88.3	39.85	0.82	82.25
31	0.53	6.55	9.00E-07	0.0086	0.5268	0.9	-5.66	-0.86	-86.41
32	1.09	75.5	-2.00E-06	0.0092	0.5101	63.8	-11.72	-0.16	-15.53
33	1.12	27.35	-3.00E-06	0.0095	0.4797	68.9	41.51	1.52	151.78

When the concentrations of both ions were combined, the average absolute error was 21.5 hours (SD = 16.5). The average relative error, expressed as percentage of the PMI (%PMI), was 44%, 42%, and 43% for potassium, ammonium and the combined concentration of both, respectively. It is

clear that the relative errors decrease steadily in increasing PMIs, from about 200% in the first hours down to low percentages for PMIs above 100 hours (Figure 11).

This approach confirms the time-dependent increase of the concentration of potassium in vitreous humour, but more importantly, also shows that ammonium has a potential as a marker of the time elapsed since death, confirming the preliminary findings obtained with the linear correlation.



**Figure 11.** Estimation error (expressed as %PMI) with the PMI (hours) for potassium, ammonium, and the sum of the concentrations of both ions.

The comparative study of the correlation of the two parameters with the PMI and of the possibility of their use to infer the time of death in real forensic cases demonstrates the superiority of potassium in comparison with ammonium. However, ammonium can complement potassium for the estimation of PMI, particularly when potassium concentrations in the vitreous humour could have been affected by confounding factors, e.g. ocular traumas or drug overdoses [62] or the PMI is higher than 100 hours [40]. In fact, because of the different mechanisms involved in the post-mortem increase of both ions, carrying out of their determination in parallel can provide mutual corroboration in terms of the reliability of the estimated PMI's.

With regard to the error related to the determination, the data show that as the PMI increases, so does the relative accuracy of its estimation, as depicted in Figure 11. This confirms previous findings reporting the variability of PMI estimation for PMIs below 26 hours [40].

The main limitation of this approach is still the small number of the cases investigated, justified by obvious difficulties in the selection of traumatic or sudden deaths with known PMIs. A second limitation is related to the error of the determination. As it can be seen in Figure 11, at early PMIs (up to 60 hours) the error expressed as the % of the PMI is very high. However, as outlined in the introduction, for the determination of short PMIs, there are other methodologies based on the evaluation of physical changes in the body occurring during the 'early post-mortem period'.

### 3.3.5.1 APPLICATION TO REAL SAMPLES

The logarithmic approach was applied to 11 unknown samples in which a PMI interval is known, and the results are shown in Table 8. The main limitation of this approach is the concentration limits. In the lower limit, PMIs smaller than 48 hours will not give high ion concentrations, and thus, cannot be taken into consideration for the statistical model. In the upper limit, concentrations higher than 3.87 mM for  $\text{NH}_4^+$  and combined use (ammonium and potassium) do not give a solution to the second-degree equation, and so, the PMI cannot be estimated. This limitation is later addressed and solved with the ANNs. Overall, there is a potential usefulness of ammonium determination in the vitreous humour, in addition to potassium, as an objective biochemical parameter to infer the PMI.

**Table 8.** Estimation of the PMI of 5 unknown vitreous samples. The coefficients of the curve are denoted by a,b,c, ( $y = ax^2 + bx + c$ );  $c' = [\text{K}^+] - c$ . The estimated PMI is obtained as the solution to the quadratic equation using  $c'$ . NS stands for No Solution to the second-degree equation. **Green** indicates that they fall within the PMI interval, and **yellow**, that they fall very close.

VH code	$[\text{K}^+]$ (mM)	a	b	c	$c'$	Estimated PMI (prediction $\pm$ 15.3)	Real PMI (hours.min)
3	12.82	-0.0005	0.2018	6.17	-6.65	51.5	48-72
7	6.34	-0.0005	0.2018	6.17	-0.17	16.1	12.3
8	15.91	-0.0005	0.2018	6.17	-9.74	71.4	80
15	9.62	-0.0005	0.2018	6.17	-3.45	33.2	PMI $\geq$ 21.20
25	11.98	-0.0005	0.2018	6.17	-5.81	46.5	42-50
30	11.66	-0.0005	0.2018	6.17	-5.49	44.6	24-48
46	9.44	-0.0005	0.2018	6.17	-3.27	32.2	24-29
51	9.76	-0.0005	0.2018	6.17	-3.59	34.0	PMI < 24
65	8.96	-0.0005	0.2018	6.17	-2.79	29.6	PMI $\geq$ 30
67	19.65	-0.0005	0.2018	6.17	-13.48	99.8	71.20-73.20

VH code	$[\text{NH}_4^+]$ (mM)	a	b	c	$c'$	Estimated PMI (prediction $\pm$ 15.6)	Real PMI (hours.min)
3	0.92	2.00E-06	0.0127	0.1461	-0.77	44.8	48-72
7	0.00	2.00E-06	0.0127	0.1461	0.15	0	12.3
8	2.67	2.00E-06	0.0127	0.1461	-2.53	177.4	80
15	0.44	2.00E-06	0.0127	0.1461	-0.30	7.7	PMI $\geq$ 21.20
25	0.36	2.00E-06	0.0127	0.1461	-0.22	32.6	42-50
30	6.71	2.00E-06	0.0127	0.1461	-6.56	464.7	24-48
46	0.22	2.00E-06	0.0127	0.1461	-0.07	21.0	24-29
51	0.29	2.00E-06	0.0127	0.1461	-0.15	27.1	PMI < 24
65	1.04	2.00E-06	0.0127	0.1461	-0.90	54.4	PMI $\geq$ 30
67	1.27	2.00E-06	0.0127	0.1461	-1.13	71.9	71.20-73.20

VH code	[NH <sub>4</sub> <sup>+</sup> ] + log[K <sup>+</sup> ] (mM)	a	b	c	c'	Estimated PMI (prediction ± 16.5)	Real PMI (hours.min)
3	2.03	-1.00E-05	0.0106	0.4799	-1.55	158.4	48-72
7	0.80	-1.00E-05	0.0106	0.4799	-0.32	14.8	12.3
8	3.87	-1.00E-05	0.0106	0.4799	-3.39	NS	80
15	1.43	-1.00E-05	0.0106	0.4799	-0.95	81.9	PMI ≥ 21.20
25	1.44	-1.00E-05	0.0106	0.4799	-0.96	83.6	42-50
30	7.77	-1.00E-05	0.0106	0.4799	-7.29	NS	24-48
46	1.19	-1.00E-05	0.0106	0.4799	-0.71	55.3	24-29
51	1.28	-1.00E-05	0.0106	0.4799	-0.80	65.5	PMI < 24
65	2.00	-1.00E-05	0.0106	0.4799	-1.52	154.1	PMI ≥ 30
67	2.57	-1.00E-05	0.0106	0.4799	-2.09	244.7	71.20-73.20

### 3.3.6 DISCUSSION OF THE NEURAL NETWORKS APPROACH

As mentioned in the introduction (2.1 Artificial Neural Networks (ANNs)), ANNs are a great multivariate analysis tool since it generalises, and therefore, it is not necessary to know the connection between input and output data and it deals well with measurement errors. Given all these advantages and the fact that the correlation and prediction of the previous regressions could be improved, in addition to the fact that lower PMIs could be used, this approach was explored.

Input data considered ammonium and potassium concentrations and the sodium ratio (peak area of sodium/peak area IS), (total of three inputs); and PMI was the output. The accuracy of the training and test ( $r^2$  score) was the main metric used for the evaluation of the network. If accuracies were similar, the mean absolute error (MAE) was used as well. The  $r^2$  score was expected to be as close to 1 as possible, whereas the MAE must be as low as possible, expecting a higher error in the test set than in the training set (because the test set has less samples). The networks were built based on the different solvers for the weight optimisation: Adam, SGC, and LBFGS. For each solver, the different parameters were optimised and once the final network was tuned and validated, 11 untrained samples (not used in the learning process of the network) were used to test it.

Firstly, the train/test split must be decided, that is, to decide how many samples are used for training and how many for testing. Ideally, the samples in the training split should cover the entire domain of the problem, i.e. a wide selection of different PMIs. Hence the more samples in the training set, the better trained the network is. However, enough samples should be left to test the network [22]. A good compromise is a 75/25 split, 75% of samples are used for training and 25% for testing. With a total of 38 available samples (Appendix I. 6.1.1 List of vitreous humour samples used for the study), 28 were used for training, and 10 for testing. Secondly, the data should be normalised, which was done through the MinMaxScaler of scikit-learn (Appendix I. 6.1.2 Train test split and data pre-processing). Thirdly, the evaluation of the network with different solvers was carried out.



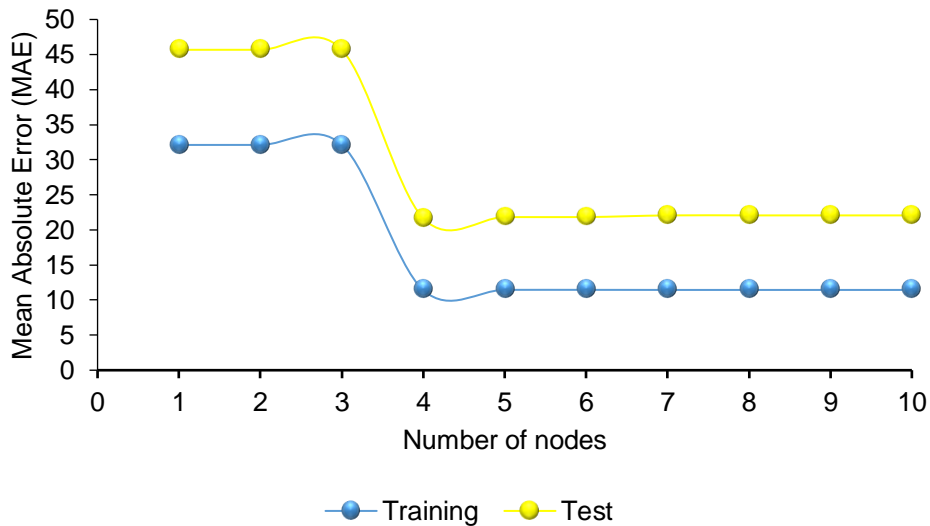
### 3.3.6.1 ADAM SOLVER

#### 3.3.6.1.1 Network training

The code for the Adam solver can be found in Appendix I (6.1.3 Code for Adam solver). Several parameters were optimised. Table 9 summarises the optimisation results of different network parameters. The number of hidden layers was set to 1 because according to literature [24] more hidden layers are for larger and more complicated datasets. The first parameter to optimise was the number of nodes. For this parameter, the accuracy and the MAE must be taken into account [23]. The ideal number of nodes in the hidden layer is when the MAE is kept to the minimum value without losing accuracy. As seen in Figure 12, four nodes yielded the best accuracy and lowest error.

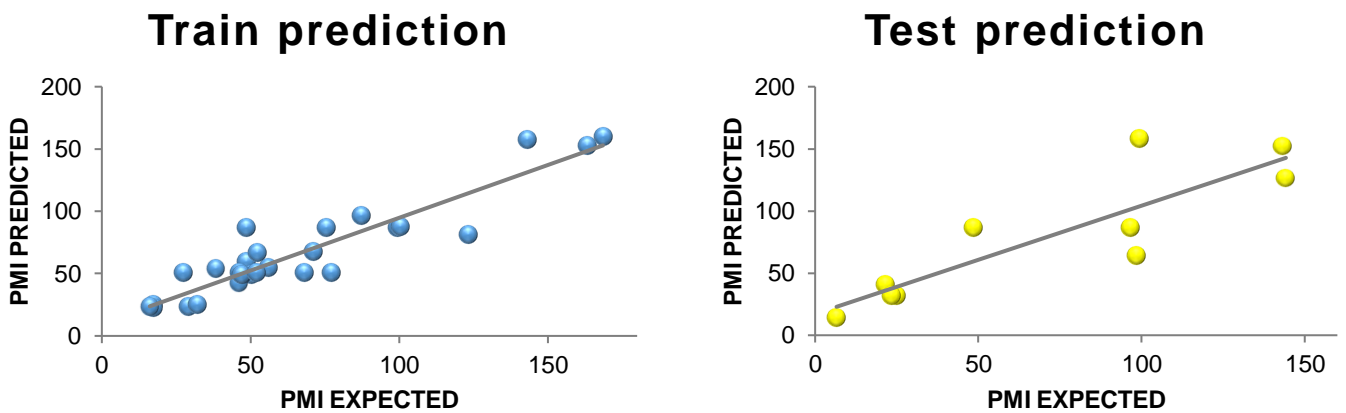
**Table 9.** Parameter optimisation summary with the Adam solver. The names in *cursive* indicate the part of the code belonging to that parameter. The option highlighted in grey indicates the best value for the corresponding parameter.

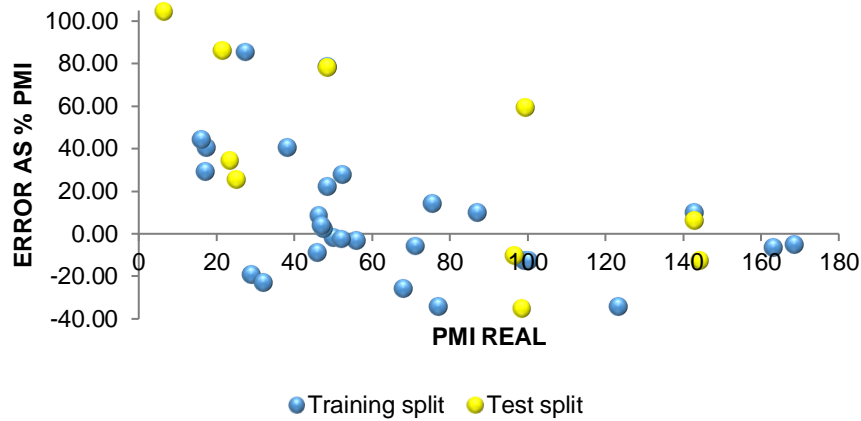
Number of nodes in hidden layer ( <i>hidden_layer_sizes</i> )										
Accuracy of prediction ( $r^2$ )	1	2	3	4	5	6	7	8	8	10
Training set	0	0	0	0.86	0.85	0.86	0.86	0.86	0.86	0.86
MAE Training	32.1	32.1	32.1	11.4	11.5	11.5	11.4	11.4	11.4	11.4
Test set	0	0	0	0.701	0.69	0.69	0.69	0.69	0.69	0.69
MAE Test	45.7	45.7	45.7	21.7	21.8	21.8	22.1	22.1	22.1	22.1
Activation function for the hidden layer ( <i>activation</i> )										
Accuracy of prediction ( $r^2$ )	identity	logistic	tanh	relu						
Training set	0.75	0.86	0.88	0.78						
Test set	0.82	0.70	0.67	0.80						
Regularisation factor ( $\alpha$ ) ( <i>alpha</i> )										
Accuracy of prediction ( $r^2$ )	0.10	0.01	0.001	0.2	0.3					
Training set	0.86	0.87	0.88	0.85	0.00					
Test set	0.70	0.68	0.67	0.71	0.01					
Learning rate ( <i>learning_rate_init</i> )										
Accuracy of prediction ( $r^2$ )	0.0001	0.001	0.01	0.10	0.20	0.30	0.40	0.50	0.60	0.70
Training set	0.00	0.86	0.86	0.82	0.00	0.00	0.00	0.00	0.00	0.00
Test set	0.00	0.70	0.70	0.71	0.00	0.00	0.00	0.00	0.00	0.00
Random seed in MLP regressor ( <i>random_state</i> )										
Accuracy of prediction ( $r^2$ )	0	17	25	28	88	99				
Training set	0.86	0.86	0.86	0.86	0.86	0.86				
Test set	0.70	0.70	0.70	0.70	0.70	0.70				



**Figure 12.** Number of nodes in the hidden layer against MAE for Adam solver

The next parameter was the activation function, where the best option was the logistic activation function. The network was also tested with the identity activation function, but it yielded the worst accuracies towards the end. The regularisation factor and learning rate were optimised and the ones with higher accuracy in both tests were selected. Finally, the random state in the MLPRegressor was explored. Random seed randomly generates the weights, initialises the bias and shuffles the data in the Train/Test split. Effectively, different models can be obtained when using the exact parameters but different random seeds. This factor is highly important with smaller datasets, since if a model is robust, changing the random state should not greatly affect accuracy [24]. Six random numbers were tested and as seen from Table 9, the accuracy is still high for all case scenarios. Hence, the network is robust. The accuracy graphs and the error as a function of the PMI are shown in Figure 13. Much like with the logarithmic approach, the error (expressed as %PMI) is high at low PMIs and decreases with increasing PMIs.





**Figure 13.** Training and test graphs for the final network ( $r^2_{\text{training}} = 0.86$  and  $r^2_{\text{test}} = 0.70$ ). Estimation error (expressed as %PMI) with the PMI (hours) for the training and test samples.

### 3.3.6.1.2 Validation

Finally, the model was validated with two different methods (Code can be found in Appendix I. 6.1.3.6 Validation). Validation is needed to see whether the network is over or under fitted, whether it generalises well, and to evaluate the model’s performance on untrained data. The first validation was performed through k-fold cross-validation. It is the most common method for validating neural networks. K stands for the number of folds or “divisions” of all data (training and testing). If k is 5, the data is divided in five folds, where four folds are used for training and one for testing; this allows for a different selection of random states in the Train/Test split. This process is repeated until all 5 folds have been used as testing. For each run, the accuracy of the network is measured with the  $r^2$  parameter. For models that are not computationally costly, three to five different cross validation runs are recommended [63]. However, NNs are computationally costly, so only one run is performed. The value of k is selected to be 5, as recommended for moderately sized datasets.

**Table 10.** Summary of the 5-fold validation process.

Cross-validation score (5-fold)							
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	STD
Accuracy of the network ( $r^2$ )	0.76	0.72	0.90	0.88	0.83	0.82	0.08

A second validation was performed using the leave-one-out (LOO) method, using the MAE as the accuracy metric. LOO is used to test the model’s performance on untrained samples. It is a form of cross-validation where k equals the number of samples, which is this case is 38. For each fold (or run) it takes one sample out (used for testing), whereas the remaining 37 are used for training. Usually, this method is extremely computationally expensive, and it is not used in NNs unless a supercomputer can be used [64]. However, since the present network is smaller than usual, the LOO analysis was performed. Drawing from Table 10, the network generalises well since the accuracy is

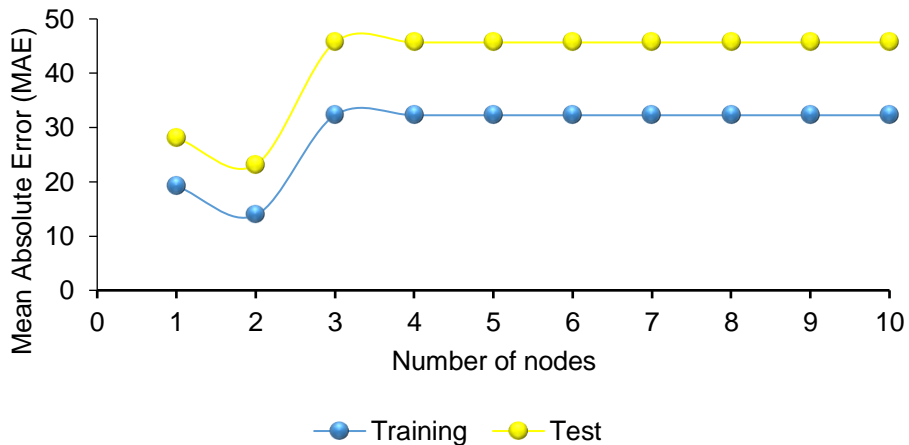
very good in all five folds. In addition, the MAE is 18.5, which agrees with the optimisation process results where MAE is between 11 and 21 (Table 9).

### 3.3.6.2 SGD SOLVER

#### 3.3.6.2.1 Network training

The code for the SGD solver can be found in Appendix I. (6.1.4 Code for SGD solver). The results of the optimisation of parameters are found in Table 11.

Based on Figure 14, two nodes were selected. The regularisation factor was chosen to be 0.01 even when the accuracy was very similar to 0.001 because a higher alpha means a simpler model. If the problem is not too complex, choosing a lower alpha might compromise generalisation [24]. Regarding momentum, values between 0.1-0.8 have similar accuracies. Since the momentum is the parameter that helps to break from the local minima, a high value increases the chances of missing the solution whereas a lower value might not be enough to break from the local minima [22]. As a compromise, 0.3 was chosen. Nesterov momentum is a different type of momentum, which guarantees a better convergence rate [65]. The two options were whether to activate (True) or deactivate (False); accuracy was better without the Nesterov momentum.



**Figure 14.** Number of nodes in the hidden layer against MAE for SGD solver

As seen in Table 11, the network looks promising, with accuracies similar to the one with the Adam solver. Therefore, it was validated similarly. The mean 5-fold cross-validation accuracy was -0.29, and thus, no further validation was performed, and this solver was discarded for the prediction of untrained samples.

**Table 11.** Parameter optimisation summary with the SGD solver. The names in *cursive* indicate the part of the code belonging to that parameter. The option highlighted in **grey** indicates the best value for the corresponding parameter.

<b>Number of nodes in hidden layer (<i>hidden_layer_sizes</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>8</b>	<b>10</b>
<b>Training set</b>	0.63	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>MAE Training</b>	19.2	13.8	32.3	32.3	32.3	32.3	32.3	32.3	32.3	32.3
<b>Test set</b>	0.65	0.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>MAE Test</b>	28.1	23.1	45.7	45.7	45.7	45.7	45.7	45.7	45.7	45.7
<b>Activation function for the hidden layer (<i>activation</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>identity</b>	<b>logistic</b>	<b>tanh</b>	<b>relu</b>						
<b>Training set</b>	0.75	0.80	0.59	0.73						
<b>Test set</b>	0.82	0.71	0.68	0.76						
<b>Regularisation factor (<math>\alpha</math>) (<i>alpha</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>0.10</b>	<b>0.01</b>	<b>0.001</b>	<b>0.2</b>	<b>0.3</b>					
<b>Training set</b>	0.75	0.75	0.75	0.75	0.75					
<b>Test set</b>	0.82	0.82	0.82	0.82	0.82					
<b>Learning rate (<i>learning_rate_init</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>0.0001</b>	<b>0.001</b>	<b>0.01</b>	<b>0.1</b>	<b>0.2</b>	<b>0.3</b>	<b>0.4</b>	<b>0.5</b>	<b>0.6</b>	<b>0.7</b>
<b>Training set</b>	0.71	0.73	0.50	0.09	0	0	0	0	0	0
<b>Test set</b>	0.76	0.81	0.62	0.18	0	0	0	0	0	0
<b>Momentum (<i>momentum</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>0.01</b>	<b>0.1</b>	<b>0.2</b>	<b>0.3</b>	<b>0.4</b>	<b>0.5</b>	<b>0.6</b>	<b>0.7</b>	<b>0.8</b>	<b>0.9</b>
<b>Training set</b>	0.81	0.81	0.81	0.82	0.82	0.82	0.82	0.82	0.82	0.82
<b>Test set</b>	0.72	0.72	0.72	0.71	0.71	0.71	0.70	0.70	0.70	0.69
<b>Nesterovs momemtum</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>False</b>	<b>TRUE</b>								
<b>Training set</b>	0.82	0.79								
<b>Test set</b>	0.71	0.74								
<b>Random seed in MLP regressor (<i>random_state</i>)</b>										
<b>Accuracy of prediction (<math>r^2</math>)</b>	<b>0</b>	<b>17</b>	<b>25</b>	<b>28</b>	<b>88</b>	<b>99</b>				
<b>Training set</b>	0.82	0.82	0.82	0.82	0.82	0.82				
<b>Test set</b>	0.71	0.71	0.71	0.71	0.71	0.71				

### 3.3.6.3 LBFGS SOLVER

#### 3.3.6.3.1 Network training

The code for the LBFGS solver can be found in Appendix I. (6.1.5 Code for LBFGS). The results of the optimisation of parameters are found in Table 12. As seen in Figure 15, the optimal number of nodes was three. The activation function and the regularisation factor were identity and 0.001, respectively.

**Table 12.** Parameter optimisation summary with the LBFGS solver. The names in *cursive* indicate the part of the code belonging to that parameter. The option highlighted in grey indicates the best value for the corresponding parameter.

Number of nodes in hidden layer ( <i>hidden_layer_sizes</i> )										
Accuracy of prediction ( $r^2$ )	1	2	3	4	5	6	7	8	9	10
Training set	0.79	0.82	0.81	0.00	0.00	0.86	0.86	0.86	0.90	0.86
MAE Training	13.6	12.1	12.7	32.2	32.2	11.4	11.4	11.4	10.6	11.4
Test set	0.71	0.70	0.75	0.00	0.00	0.69	0.69	0.69	0.72	0.69
MAE Test	23.2	22.2	20.4	45.7	45.7	22.0	22.0	22.0	21.2	22.0

Activation function for the hidden layer ( <i>activation</i> )				
Accuracy of prediction ( $r^2$ )	Identity	Logistic	tanh	relu
Training set	0.75	0.81	0.49	0.78
Test set	0.82	0.75	0.40	0.80

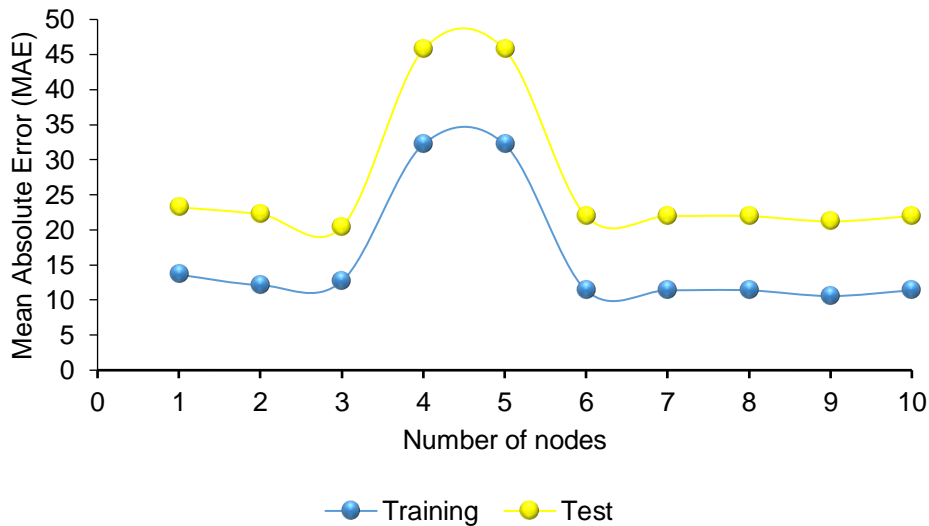
  

Regularisation factor ( $\alpha$ ) ( <i>alpha</i> )					
Accuracy of prediction ( $r^2$ )	0.1	0.01	0.001	0.2	0.3
Training set	0.81	0.83	0.88	0.82	0.82
Test set	0.75	0.36	0.66	0.71	0.72

Random seed in MLP regressor ( <i>random_state</i> )						
Accuracy of prediction ( $r^2$ )	0	17	25	28	88	99
Training set	0.81	0.79	0.79	0.82	0.79	0.82
Test set	0.75	0.71	0.71	0.40	0.71	0.40

As shown in Table 12, the change in random seed did not achieve very good results. Random seed 28 and 99 had very low accuracies for the test set. In other words, the network does not generalise well, i.e., it is not robust enough. For that reason, this network was not further validated nor used for testing real unknown samples.



**Figure 15.** Number of nodes in the hidden layer against MAE for LBFGS solver

### 3.3.6.4 APPLICATION TO REAL UNKNOWN SAMPLES

Due to the low number of samples, only 11 untrained samples were left to validate the model (Code in Appendix I. 6.1.6) where some information about the PMI or its range was available. The predictions are summarised in Table 13. For the predictions, the Adam-solver-tuned network with seven different Train/Test splits was used. As seen from Table 13, the predictions are generally good except for samples 51 and 67. The incorrect predictions of these two samples might be due to the data set not covering a wider and more comprehensive PMI range.

**Table 13.** Prediction of PMI for real unknown samples. The prediction results are given taking into account the standard deviation (the values rom column labelled as SD ( $n=7$ )). **Green** indicates that they fall within the PMI range, and **yellow** that they fall close.

Real unknown samples (untrained)			
VH sample	PMI predicted with NN	SD ( $n=7$ )	Real PMI (hrs.min)
3	51	7	48-72
7	13	4	12.30
8	71	9	80
15	22.3	5	PMI $\geq$ 21.20
25	56	19	42-50
30	31	19	24-48
33	40	11	32
46	29	4	24-29
51	27	15	PMI < 24
65	31	5	PMI $\geq$ 30
67	120	16	71.20-73.20

Hence, the main limitation of this study is the size of the dataset. In ANNs, the size of the data set is directly related to the capacity of the network to learn and generalise. In addition, the dataset should cover the widest possible range of PMIs [22]. In the present study, a total of 38 samples with known cause of death and PMI (7–168h) were used. In comparison, a study carried out by Bocaz-Beneventi et al.[23], with the same purpose but a total number of 72 samples and a slightly different network, achieved more accurate estimations. This extra 34 samples increased the accuracy of the network, and thus, the accuracy of predictions.

### 3.3.7 DISCUSSION OF THE THREE APPROACHES

With the linear and logarithmic approach, samples with small PMIs (lower than 48h) can skew the correlation. Hence, more accurate methods for estimating the time of death at lower intervals are used in these cases. Thus, these cases are usually not taken into account. However, with ANNs, all PMIs are valid. Since networks learn by iterative analysis, they are unbiased by small PMIs. In comparison with the logarithmic and linear approaches, ANNs are superior.

**Table 14.** Summary of predictions of untrained samples with all three statistical approaches. Values highlighted in blue indicate the most accurate predictions for each sample. NS stands for No Solution to the second-degree equation.

VH	Real PMI (hrs)	[NH <sub>4</sub> <sup>+</sup> ] (mM)	NH <sub>4</sub> <sup>+</sup> PMI predictions (hrs.min)		[K <sup>+</sup> ] (mM)	K <sup>+</sup> PMI predictions (hrs.min)		Combined use	PMI predictions with NN
			linear	logarithmic		linear	logarithmic	logarithmic	
3	48-72	0.92	70.6	44.8	12.82	37.5	51.5	158.4	51
7	12.3	<i>bellow LOD</i>	0.0	-27.1	6.34	7.4	16.1	14.8	13
8	80	2.67	203.3	177.4	15.91	51.9	71.4	NS	71.4
15	PMI ≥ 21.20	0.44	34.4	7.7	9.62	22.6	33.2	81.9	22.3
25	42-50	0.36	28.4	32.6	11.98	33.6	46.5	83.6	56.0
30	24-48	6.71	509.0	464.7	11.66	32.1	44.6	NS	31.0
46	24-29	0.22	17.2	21.0	9.44	21.8	32.2	55.3	29.0
51	PMI < 24	0.29	23.1	27.1	9.76	23.3	34.0	65.5	27.0
65	PMI ≥ 30	1.04	80.1	54.4	8.96	19.5	29.6	154.1	31.0
67	71.20-73.20	1.27	97.4	71.9	19.65	69.3	99.8	244.7	120.0

Table 14 gathers the predictions of the three different approaches: linear, logarithmic and NN. For example, in VH 7, which has a small PMI, the best prediction was with the ANN. This proves that the network has learned small PMIs, and thus, there is no need to eliminate them. The incorrectly predicted samples VH 51 and 67 is due to the lack of additional samples with similar input data values, hindering the network from learning the different variations in combinations of potassium and ammonium concentrations and sodium ratio. The main limitation of the linear and polynomial models



is the removal of small PMIs. In addition, in the polynomial model, concentrations higher than 3.87 for  $\text{NH}_4^+$  and combined use do not give a solution to the second-degree equation, so there is no estimated PMI. Regarding neural networks, their main limitation is the size of the data set. The more data available, the more information the network can learn allowing it to make more accurate predictions.

### **3.3.8 AMMONIUM CONCENTRATION DIFFERENCES IN THE TWO EYES**

Another aim of the project was aimed at verifying the consistency of ion concentrations between the two eyes. A similar study was carried out by Tagliaro et al.[48] on potassium concentrations and reported no significant difference.

The statistical difference between the ammonium concentrations in the two eyes of the same subject was evaluated in 20 forensic autopsies. The mean concentrations of  $\text{NH}_4^+$  ranged from 0.07 to 6.45 mM with PMI's varying from 31 up to 89 h. The differences in concentration between the two eyes were evaluated using a paired student's t-test ( $\alpha = 0.01$ ). The findings suggest that no statistically significant differences exist (p value for two tails = 0.87). In addition, a good correlation was found between the two eyes, expressed by  $r^2 = 0.94$ .

The present data show that there is also no substantial difference in the concentrations of ammonium between the two eyes. This evidence supports the hypothesis that there is an even progression of ammonium ion production in this body compartment.

## **3.4 CONCLUSIONS**

In summary, the present method represents a new, simple application of CE in the forensic environment, where separation methods are precious for dealing with the complexity and variability of post-mortem biological matrices. An increase in ammonium concentration was observed up to 168 hours after death, thus offering a new analytical tool for PMI evaluation that is useful for a much wider time span than that offered by potassium analysis (limited to about 90–100 hrs). After acquiring more real vitreous samples and further validating the method, ammonium analysis can become, "per se" or in combination with potassium, a powerful tool for unravelling one of the most challenging issues of forensic investigation, such as the accurate determination of the time of death.

In conclusion, the findings support the potential usefulness of ammonium determination in vitreous humour, in addition to potassium, as an objective biochemical parameter for inferring the PMI, when combined with ANN multivariate statistical analysis. In addition, the present work shows the suitability of capillary electrophoresis for performing rapid simultaneous determinations.

### **3.5 FUTURE RESEARCH**

Future research should therefore concentrate on the acquisition of more vitreous humour samples with a wide range of different PMIs and further analysis to build a more reliable neural network that can be used in the analysis of real unknown samples.

# **LIQUID CHROMATOGRAPHY (LC) APPLIED TO FORENSIC SCIENCE**

## **1. LIQUID CHROMATOGRAPHY COUPLED WITH MASS SPECTROMETRY**

### **1.1 LIQUID CHROMATOGRAPHY (LC)**

Liquid chromatography (LC) is a widely used technique due to its accuracy, sensitivity, adaptability, possibility of automation, suitability for the separation of thermally fragile or non-volatile species, and its extensive applicability in industry and science. It allows for separating the components of a mixture on the bases of their distribution between a mobile phase, which is usually a solvent or a buffer, and a stationary phase containing small, packed particles. The differences type of interaction of the analytes with the two phases determines their separation [66].

A liquid chromatograph is constituted of four main parts: an inlet system, a high-pressure pump, a separation compartment and a detector. The inlet system allows the efficient introduction of the sample with minimal sample loss. Once the sample is introduced, it goes to the separation compartment where the column is located. Then, the samples are separated and go to the detector. Depending on which property of the analyte to focus on and the required sensitivity, there is a great variety of commercial detectors such as Ultraviolet-visible absorption (UV-Vis), also in the diode array (DAD) configuration, electrochemical, fluorescence, conductivity, Fourier transform infrared (FRTIR), and mass spectrometry (MS) [67,68]. According to the principle of separation, LC has different modes, such as normal phase, reverse phase, ion exchange, size exclusion, bio-affinity, and chirality [66]. The focus here will be on reverse phase separations.

#### **1.1.1 REVERSE PHASE CHROMATOGRAPHY**

Chromatographic separations take place in a chromatographic column. Columns are made of stainless steel with alumina or silica gel as a substrate and the stationary phase. In reverse phase, columns are packed with silica gel particles with their surfaces covered by alkyl groups bound to silanol (Si-OH) groups. Long alkyl chains (C18) bind stronger to low polarity organic molecules, whereas highly polar analytes pass through. Shorter alkyl columns (C8) have lower retention for low polarity analytes [69]. Nowadays, the ability to produce highly spherical, pure, and homogeneous silica particles has increased the reproducibility and resolution of analysis; and the high coverage of the stationary phase has led to improved retention times. The speed and efficiency of separation has improved in the past years, as well as their suitability for a greater range of analytical samples, due to the use of new stationary phases [68]. Usually, a guard (or pre-column) is placed before the analytical column in order to extend the life of the analytical column by protecting it from contaminants from the solvents and/or sample components that bind irreversibly to the stationary phase. Regarding the liquid phase, they have to be aqueous mixtures with methanol, acetonitrile, and additives as buffers [66].

## 1.1.2 ULTRA-PERFORMANCE LIQUID CHROMATOGRAPHY (UPLC)

A column's separation power increases with decreasing particle size which is unfortunately bound to a high pressure increase to allow enough flow rate. So far, a compromise was reached with a particle size of 5  $\mu\text{m}$ . With the intention of shortening analysis time and increasing resolution, the column particle size was decreased below 2  $\mu\text{m}$ , thus developing a new technique called Ultra Performance Liquid Chromatography (UPLC) or fast chromatography [70]. For flows of 0.5-1.0 mL/min, the pressure can reach up to 1200 bar. Due to the fast analysis in UPLC, the detectors need to have a sufficiently high sampling rate to obtain enough data points for reproducible results. There are currently only three detectors in the market that meet the condition: DAD, UV/VIS, and MS [68].

## 1.2 MASS SPECTROMETRY (MS)

The basic principle of mass spectrometry is the generation of ions, separation of these ions based on their mass-to-charge ratio ( $m/z$ ), and detection based on their respective  $m/z$  abundance. A mass spectrometer consists of three parts: the ionization source, which generates the ions by electric field and thermally by impacting energetic electrons, ions, and neutral atoms of photons; the mass analyser which separates the ions by magnetic, electric, or field-free regions; and the detector [69]. The analyser must be kept under vacuum to allow a free path for the ions. A low vacuum has a positive effect in mass resolution (ability to separate close ions) and sensitivity [71].

### 1.2.1 LC-MS ION SOURCES

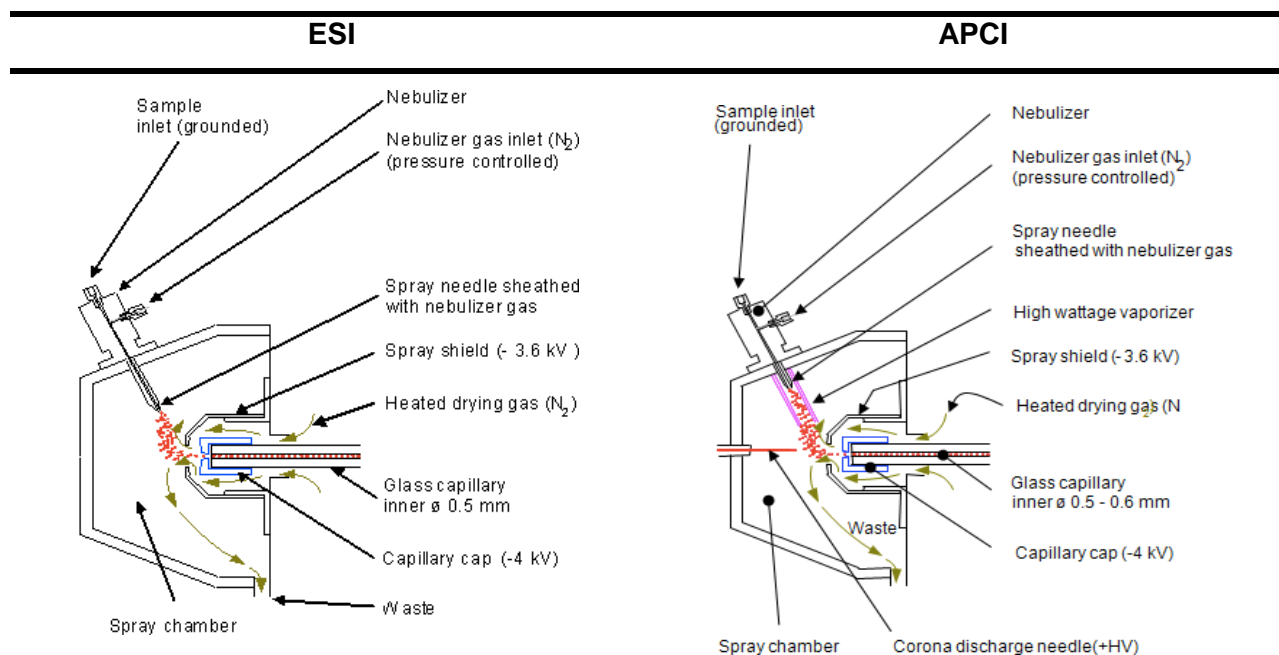
Since LC is unable to provide unequivocal identification of different analytes in a mixture based only on retention times, coupling it with MS can really help the identification of species as they elute from the column. The mass spectrometer combines high selectivity with a universal detector. In addition, the obtained mass spectra are unequivocal proof of the identification of the unknown substance. Also, it increases the sensitivity of the analysis and reduces analysis time [66,72].

There is a wide variety of sources depending on the substances of interest. Examples of said sources are electrospray ionisation (ESI), atmospheric pressure photoionisation (APPI), and atmospheric pressure chemical photoionisation (APCI) are the most popular [69]. The focus of this thesis is directed towards ESI, IonBooster™ (iB), and APCI.

#### 1.2.1.1 ESI, iB AND APCI

ESI is a soft ionisation technique that takes place at atmospheric pressure. It transforms the ions from solution into the gas phase through the nebulisation of the liquid into highly charged aerosol droplets, dissolution of the charged droplets, and ionisation of the analyte. This process takes place by applying an electrical potential to a liquid in an electrospray needle (Figure 16). For ESI to work, the solvent containing the analyte needs to be volatile. As mentioned above, it starts the ionisation at atmospheric pressure, slowly reaching the high vacuum necessary for the mass analyser. With

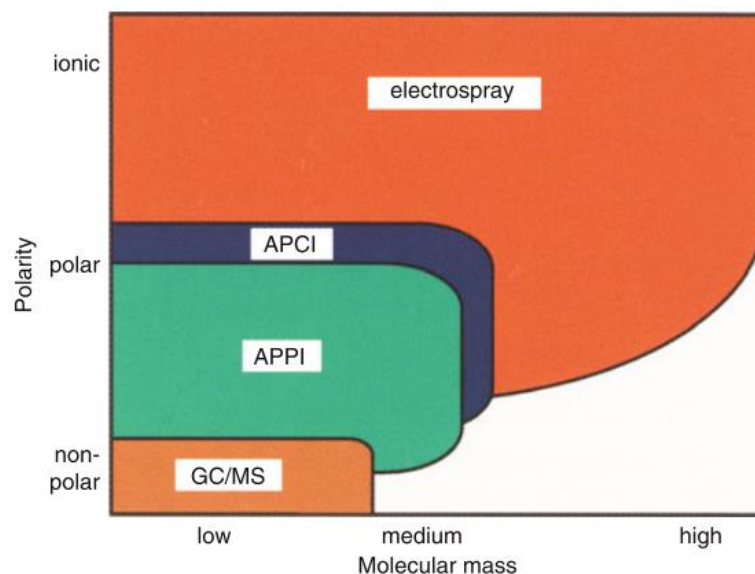
this source, the analytes are ionised in solution. Generally, the best sensitivity is achieved using low flow rates. ESI is superior to APCI due to its broad range of analyte tolerance in terms of mass and polarity (Figure 17) [67,69].



**Figure 16.** Fundamental principles of ESI and APCI sources. The green arrows indicate how the heated drying gas moves inside the source. The small red spheres represent the sample. The black arrows indicate the part of the source named. The sample enters through the sample inlet, the analytes are ionised and go to the mass analyser through the glass capillary. Adapted from [73,74].

The iB was developed by Bruker Daltonics and can be defined as a high-temperature ESI source. It is heated in the spray zone and results in a highly efficient desolvation process when using flow rates higher than 200  $\mu\text{L}/\text{min}$ . For this source, a sheath gas (nitrogen) must be used to prevent the thermal degradation of the analytes prior to the electrospray process. By increasing the ionisation efficiency, it lowers detection limits and enhances sensitivity. In comparison with ESI, it offers greater sensitivity under high flow rates [73,75].

APCI is also a soft ionisation technique where the analyte solution is nebulised by means of a pneumatic nebuliser. The aerosol is then vaporised in a heater chamber and analytes are ionised by solvent ions created in a corona discharge. Lastly, they go to the high vacuum of the mass analyser through a system of nozzles and skimmers and a drying gas (Figure 16). With this source, the analytes are ionised in the gas phase. In terms of technical conditions, APCI can be used with small amounts of non-volatile and volatile buffers and flow rates no higher than 2  $\text{mL}/\text{min}$ . As a prerequisite, analytes should be thermally stable. APCI does not have the same potential as ESI but it allows the sensitive determination of analytes of moderate mass and polarity (Figure 17) [67,69,72].



**Figure 17.** Comparison between ESI and APCI [69].

### 1.2.2 MASS ANALYSER

A mass analyser separates the ions according to their  $m/z$  ratio. It can operate in two modes: Scan mode (SCAN), which monitors all the  $m/z$ s in the sample and Selected Ion Monitoring (SIM) which monitors a targeted ion. It operates under vacuum (as the detector) in order to allow the analyte ions to travel from one side to the other without being hindered by air molecules. Different types of mass analysers can be used depending on the accuracy and/or sensitivity desired. They are summarised in Table 15. High Resolution MS is able to separate ions that differ in mass by 0.0001 atomic mass units (amu), whereas low resolution ones are able to separate them when they differ by at least 1 amu. In Multiple stage MS, both molecular and fragment ions are produced and detected; whereas in single step MS, only molecular ions are produced and detected [69,72].

**Table 15.** Different types of mass analysers. Adapted from [76].

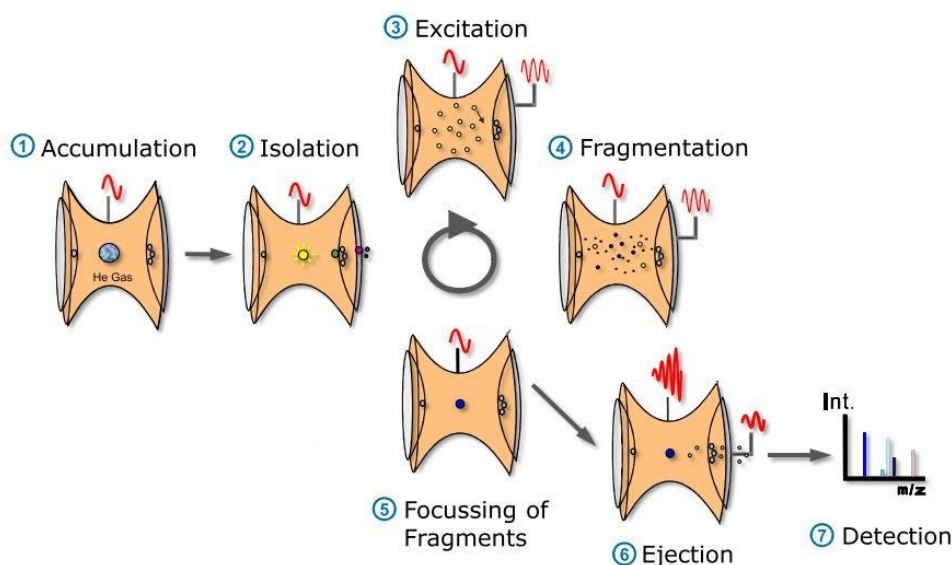
Criterion		Mass analyser
Low resolution MS	Single step MS	Quadrupole (Q) Ion trap (IT)
Low resolution MS	Multiple step MS	Triple quadrupole (QQQ) (MS/MS) Ion trap (MS <sup>n</sup> ) Quadrupole Ion Trap (QTRAP) (MS <sup>n</sup> )
High resolution MS	Single step MS	Time of flight (TOF)
High resolution MS	Multiple step MS	Quadrupole TOF (QTOF)

A quadrupole (Q) is made of four parallel rod electrodes, with each pair of opposite rods held at the same potential composed of AC (alternating current) and DC (direct current). A resonant frequency is created for each  $m/z$  and only ions of a specific  $m/z$  pass through the quadrupole, whereas others

are discarded. If the resonant frequency is changed, a full mass range can be scanned. The quadrupole can be used in SIM mode: selection of a number of masses (high selectivity); or SCAN mode: an ion is only detected a fraction of the total scan, and the masses reach the detector sequentially (low sensitivity) [69].

An Ion Trap (IT) is made by two hyperbolic electrodes and a ring electrode. Radio frequency (RF) and DC potentials are applied between them. The ions are introduced and accumulated in the centre of the IT and then sequentially ejected to the detector due to the application of RF (mass SCAN). Only ions of a certain  $m/z$  will be ejected to the detector (those with stable trajectories), whereas unwanted ions will collide with the walls and will be discarded due to their unstable trajectories [69].

An MS/MS combination can also be achieved with an IonTrap ( $MS^n$ ) when a single ion trap serves as both the mass analyser and collision cell. Ions are accumulated, the precursor ion is isolated and specifically activated by using a specific fragmentation amplitude, and product ions are created. Then the fragments (product ions) are focused, ejected, and detected (Figure 18). A very interesting feature of this instrumentation is that only precursor ions are selected and excited in comparison with QTOF, where both precursor and fragment ions are selected and excited [74].



**Figure 18.** Fundamentals of IonTrap in  $MS^n$  mode [74].

It is possible to use two mass analysers in tandem ( $MS/MS$ ) to create new operational possibilities where both of them can be used in SCAN or SIM mode. This is the case of the triple quadrupole configuration (QQQ) where the first (Q1) and last (Q3) quadrupole can scan masses sequentially or do SIM; while the second quadrupole (Q2), the one in the middle, acts as a “collision cell” where mass fragmentation occurs. It does not filter the ions, but rather accepts the ones coming from Q1 and passes all ions formed by collision to Q3 [69,72].

The QTRAP instrument is in the QQQ geometry, but one quadrupole can also trap and store ions. Here, Q3 is the hybrid quadrupole denoted as Linear Ion Trap (LIT), which means it can function as

one or the other. It offers the properties of QQQ, and in addition, the ones from the LIT, such as combinations of ion accumulation and scanning (enhanced sensitivity) and higher order tandem MS [69].

The principles of the time of flight (TOF) analyser relies on the acceleration of the ions in an electric field until they reach the detector through a free vacuum path of known length. Speed is inversely proportional to the mass of the ion, so lighter ions will reach the detector earlier. It is key that the ions of all  $m/z$  coming from the source are transferred, simultaneously and instantaneously, into the mass analyser by means of a pulsed ion beam or by pulsing ion packages out of a continuous source. Since they offer high ion transmission, TOF analysers are highly sensitive [69].

The QTOF instrument is also in the QQQ geometry, but one quadrupole is replaced by an orthogonal TOF. Q1 is a mass analyser where the precursor ion of interest is selected, Q2 is the collision cell where the precursor ion undergoes fragmentation, and Q3 is a TOF analyser. As mentioned before, it is key that the ions are transferred, simultaneously and instantaneously, into the mass analyser. Therefore, after exiting the collision cell, the ions are focused into a parallel beam that enters the TOF. Hence, QTOF is very sensitive to the quality of the incoming ion beam. QTOF offers high mass resolution measurements, identification power, ability to record all ions in parallel without scanning, and increased selectivity which are important in forensic toxicology [72,77]

### **1.2.3 DETECTOR**

The mass detector is crucial in the instrument. After the analytes are separated in the mass analyser based on the  $m/z$ , the detectors receive the current signal generated from the incident ions which is a reflection of the absolute or relative concentration of the analyte. It amplifies and registers the signal. A good detector is defined on the basis of low noise and cost, high collection efficiency and amplification, fast response time, reproducibility, long life, and long-term stability. Examples of some commercial detectors are high mass detection detectors, channel electron multipliers, Faraday cups, electron multipliers, scintillation counters, photographic plates, resistive anode encoder image detectors, conversion dynodes, helium leak detectors, and cryogenic detectors [78].

## **1.3 FORENSIC APPLICATIONS**

In forensic toxicology, LC-MS/MS has gained momentum in the field of drug screening and quantification. The main advantage is that the use of reference standards (sometimes expensive or hard to acquire) is not needed since accurate molecular weight information can be obtained with techniques such as high-resolution mass spectrometry (HRMS). This allows the unequivocal and robust determination of unknown substances present in different biological matrices (blood, urine, serum, hair, and vitreous humour). In order to reliably use tandem MS in a forensic field, a thorough assessment of ion suppression and interferences must be performed. The correct selection of the ions to be monitored must be appropriate and the mass spectrometer must be optimised for each



analyte. For the unequivocal identification of compounds, two or more ion transitions must be selected. Since it is vital to obtain structural information, the TOF is an optimal choice for such task, as well as for the identification of unknown substances in any of the biological matrices mentioned above [79]. In this field, LC-MS/MS has been used for the determination of drugs of abuse (DoA) and novel psychoactive substances (NPS) such as lysergic acid diethylamide, (LSD), cocaine and its metabolites, opiates, amphetamines, cannabinoids, synthetic cannabinoids, phenylcyclohexyl piperidine (PCP), and g-hydroxybutyrate (GHB) in biological samples [72]. LC-MS has also been applied to the determination of illicit drugs and its metabolites on serum samples of subjects driving under the influence (DUI) [80].

LC-MS has also been used as a predictive tool for raw materials information in home-made synthesis of nitrate ester explosive erythritol tetranitrate (ETN) [81] and as a tool for the detection of explosives and quantification among different types of explosives [82]. Also, LC-MS has been used to test the prevalence of GSR in police cars [83]. Wastewater epidemiology is a discipline of forensic science which estimates drug consumption in an area or community through the analysis of drugs and its metabolites in wastewater, where LC-MS/MS is the technique of choice. For instance, in the study by Centazzo et al., [84], they analysed wastewater in different areas of New York with LC-QQQ, finding amphetamines, cocaine, opioids, and cannabis. In the field of trace chemicals analysis, such as chemical warfare agents, LC-MS is a very useful technique for the determination of the metabolites of sarin, sulphur mustard, and ethyl methylphosphonothioate in urine [85].

Finally, another important application of LC-MS(MS) is the elucidation of the metabolic pathways of NPS. NPS are chemical analogues or derivatives of classic illicit substances designed to produce effects similar to that of the illegal drugs they imitate [86]. Several metabolic pathways have recently been elucidated, such as 5F-AKB-48 [87], THJ-018 and THJ-2201 [87], FUBIMINA [87], and APINAC [88] done through LC-QTOF or orbitrap.

## 2. ELUCIDATION OF THE 5F- APINAC METABOLIC PATHWAY

### 2.1 INTRODUCTION

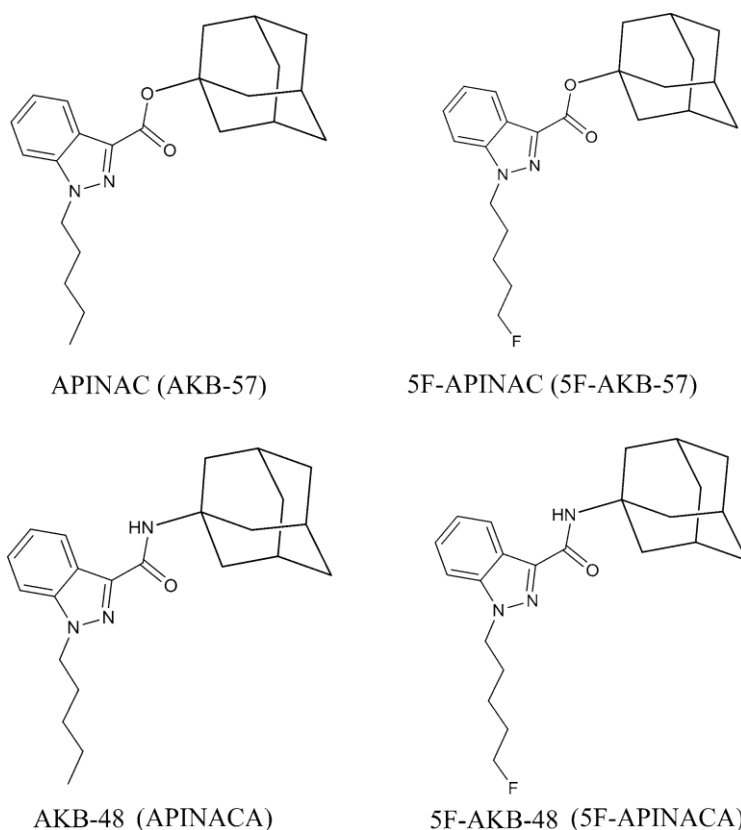
New psychoactive substances (NPS) or “designer drugs” are the trending drugs of today’s drug market. They are chemical derivatives or analogues of classic illicit substances designed to produce effects similar to that of illegal drugs [86]. The relentless modification of chemical structures by clandestine laboratories presents a challenge to legal processes, as they always manage to be one step ahead. In addition, it presents a threat to public health. Even small modifications in the chemical structures of these drugs can result in differences in biological activity and pharmacokinetic properties, and represent a challenge for identification [89]. Moreover, the psychoactive effects of terminally fluorinated analogues of synthetic cannabinoids (SCs) generally showed CB<sub>1</sub> receptor potency higher than that of non-fluorinated analogues (~ 2–5 times), and thus, may have a higher impact on health risks [90]. There is also a need to develop highly sensitive and accurate analytical technologies for the identification and quantitation of metabolites associated with NPS [91]. Advances in laboratory methodologies and scientific international collaboration are needed to tackle this problem [92,93]. Elucidating the metabolic pathways of these new drugs is highly important for toxicological risk assessment in order to develop drug screening procedures for the detection of drugs and their metabolites in blood or urine [94]. Since drug concentrations in the body depend on metabolism, more information about it would help in the medical treatment of poisoning or intoxication cases. From an analytical perspective, knowing the fragmentation and specific product ions helps towards the faster elucidation of similar, but unknown related compounds. For instance, in distinguishing between different compounds of the same family of NPS such as synthetic cannabinoids (SC). Metabolic path elucidation is very relevant since NPS are fast appearing in the market and scientists are trying to identify and characterise them at the same rate in which they appear. This speedy appearance is due to the relentless modification of chemical structures by clandestine laboratories [95,96].

SCs were originally developed for research purposes, but rapidly turned out to be used as drugs of abuse, mostly as an alternative to cannabis. To achieve psychoactive effects, chemical additives are sprayed onto dried plant materials to be sold as a legal alternative to cannabis [97]. To avoid regulations, the products are labeled as “not for human consumption“. These substances have structural features similar to tetrahydrocannabinol (THC), which enables them to bind to CB<sub>1</sub> and CB<sub>2</sub> receptors, triggering cannabis-like effects (i.e., paranoia, sedation, anxiety, euphoria, and impaired sense of time as demonstrated in *in vivo* studies [98]). However, they are more powerful receptor agonists than THC, thereby increasing the intensity of psychosis, agitation, and sympathomimetic effects [99,100]. The clinical toxicity of SCs is partly similar to the one provoked by amphetamines [101,102]. Typical symptoms of SC consumption are: vomiting, hallucinations, tachycardia, excitement, hypertension, and drowsiness [99,102]. Poisonings with SCs usually

manifest brain edema due to hyponatremia, hyperthermia, seizures, and serotonin syndrome [103]. Usually, SC intake occurs by smoking using a cigarette or a cannabis pipe. They are extensively metabolised, and their main metabolites are hydroxyl or carboxyl derivatives of the aromatic ring or the N-alkyl chain. JWH-018 was the most popular first-generation SC [90], closely followed by second-generation SCs, such as AB-PINACA, AKB-48, and ADB-PINACA [93,104]

Finally, the collaboration between countries in tracking trends in the illicit market is of paramount importance for tackling the NPS problem. The EU Early Warning System (EUEWS), set by the European Monitoring Center for Drugs and Drug Addiction (EMCDDA), has been internationally recognized for its capability to identify and respond to the emergence of NPS [105].

In the present work, a new SC, 5F-APINAC (adamantan-1-yl 1-(5-fluoropentyl)-1H-indazole-3-carboxylate) which is the fluorinated analogue of APINAC (adamantan-1-yl 1-pentyl-1H-indazole-3-carboxylate) was studied [93]. The fluorine substitution at the 5-pentyl position of pentylindazole/pentylindole is a popular modification that enhances the drug's stability, potency, and half-life [93,104]. APINAC and its metabolic reactions have already been characterised [88,106] Because one of the most common metabolic routes is defluorination, several metabolites of 5F-



**Figure 19.** Structures of APINAC, 5F-APINAC, AKB-48, and 5F-AKB-48

APINAC are expected to be similar to those of APINAC. APINAC and 5F-APINAC are structurally similar to AKB-48 (APINACA) and 5F-AKB-48 (5F-APINACA) [87], respectively (Figure 19). The unequivocal identification of 5F-APINAC-associated metabolites as potential markers of its intake is

critical in the detection of suspected SC consumption [93]. To our knowledge, no data related to 5F-APINAC metabolism has been reported. The aim of the present study was to investigate the metabolism of 5F-APINAC in human liver microsomes (HLMs) (*in vitro* study) and in a rat model (*in vivo* study) using liquid chromatography–ion trap mass spectrometry (LC–IT-MS) and liquid chromatography–quadrupole time- of-flight tandem mass spectrometry (LC–QTOF-MS/MS).

## **2.2 MATERIALS AND METHODS**

### **2.2.1 REAGENTS**

Ultra-pure water was purchased from Biosolve (Valkenswaard, the Netherlands); diethyl ether and acetone from Medkhimprom (Moscow, Russia); acetonitrile from J.T. Baker (Deventer, the Netherlands); methanol and ethyl acetate from Merck (Darmstadt, Germany);  $\beta$ -glucuronidase from *E. coli* K12 (solution in 50% glycerol) from Roche Diagnostics (Mannheim, Germany); and 5F-APINAC (purity  $\geq$  98%) from Cayman Chemical (Ann Arbor, MI, USA). All other chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO, USA). Solid-phase extraction (SPE) was performed using Oasis HLB columns (60 mg  $\times$  3 mL; Waters, Milford, MA, USA).

### **2.2.2 IN VITRO & IN VIVO STUDIES**

For the *in vitro* study, the incubation procedure was performed in accordance with the manufacturer's protocol for Human Liver Microsomes (HLMs) [105]. For the *in vivo* study, urine from male Sprague–Dawley rats was collected 3, 6, and 24 h after 5F-APINAC administration. Samples were stored at  $-80^{\circ}\text{C}$  until the analysis was performed. Both samples were prepared by an experienced pharmacist with experience in working with animals and *in vitro* studies. *All animal procedures were approved by the Institutional Animal Care and Ethical Committee of the Research Center for Molecular Diagnostics and Therapy at Sechenov University, Moscow, in accordance with principles of good laboratory practice [OECD Principles on GLP. C (97)186 Final], following the European Union directive principles of laboratory animal care guidelines (2010/63/EU).*

### **2.2.3 PREPARATION OF RAT URINE SAMPLES**

The urine samples were divided into two aliquots. A 1 mL volume of 0.8 M phosphate buffer (pH 6.5) containing 30  $\mu\text{L}$   $\beta$ -glucuronidase was added to 0.5 mL of urine (first aliquot). After incubation at  $37^{\circ}\text{C}$  for 60 min and pH adjustment (pH 2.0–2.5) using hydrochloric acid, the samples were extracted with 1 mL of ethyl acetate. For the second aliquot, the phosphate buffer did not contain  $\beta$ -glucuronidase and the addition of the enzyme and incubation steps were omitted. Following the evaporation of the organic layer at  $70^{\circ}\text{C}$  in a solid-state heater, the residue was reconstituted in 100  $\mu\text{L}$  of MeOH/  $\text{H}_2\text{O}$  (1:1, v/v) for instrumental analyses.

## 2.2.4 TENTATIVE THEORETICAL METABOLIC PATHWAY

Theoretical predictions of the possible metabolites were made based on the similarity of structure between APINAC and 5F-APINAC, AKB-48 (APINACA) and 5F-AKB-48 (5F-APINACA) [87], and the already elucidated metabolic pathway of APINAC [88]. With help of the ChemDraw Ultra software (Version 12.0, Cambridge Software) theoretical fragmentations of the metabolites were predicted. The theoretically predicted metabolites were also drawn, and using the fragmentation tool, different fragmentation possibilities were explored. Together with the formula and exact mass of the said fragments, all the theoretical proposals were compared with the experimental data.

## 2.2.5 LC–IT–MS (TOXTYPER) ANALYSES

The LC–IT–MS analyses were performed using ultra high-performance liquid chromatography (UHPLC) Dionex UltiMate 3000 (Thermo Scientific, Waltham, MA, USA) coupled with amaZon speed mass spectrometer (Bruker, Billerica, MA, USA). Chromatographic separation of the analytes was achieved on a Thermo Scientific Acclaim RSLC 120 C18 of 120Å column (100 × 2.1 mm, i.d., particle size 2.2 µm) (Thermo Scientific) guarded by a Vanguard BEH C18 column (20 × 2.1 mm. i.d.; Waters). The column temperature was set at 40°C and the autosampler at 8°C. Mobile phase A consisted of ultra-pure H<sub>2</sub>O with 2 mM ammonium formate, 1% acetonitrile, and 0.1% formic acid. Mobile phase B consisted of acetonitrile, 2 mM ammonium formate, 0.1% formic acid, and 1% H<sub>2</sub>O. The flow rate was set to 0.5 mL/min. Gradient elution was performed as follows: 0.0–1.0 min 1% B; 1.0–8.0 min 1–95% B, linear; 8.0–9.0 min 95% B; 9.0–9.06 min 95–1% B, linear; 9.06–11 min 1% B. The total run time was 11 min. Electrospray ionisation (ESI) was performed using positive mode. Scan range was set at *m/z* 70–800. Data were collected in data-dependent MS<sup>n</sup> acquisition mode.

## 2.2.6 LC–QTOF ANALYSES

LC–QTOF experiments were performed for confirmation purposes on a Bruker Elute series of UHPLC coupled with quadrupole time-of-flight mass spectrometer UHR- QTOF (Ultra-high resolution QQ-time-of-flight) maXis impact (Bruker). Chromatographic separation of the analytes was achieved using an Intensity Solo 2 C18 1.8 µm, 100 × 2.1 mm column (BRU-18C182-100 µm; Bruker) and a 1.7 µm ACQUITY UPLC BEH C18 pre-column (Waters). The column oven temperature was set at 40°C and the autosampler's at 4°C. Mobile phase A consisted of 1% methanol in H<sub>2</sub>O, 5 mM ammonium formate, and 0.01% formic acid. Mobile phase B consisted of methanol, 5 mM ammonium formate, and 0.01% formic acid. The flow rate was 0.2 mL/min. Gradient elution was performed as follows: 0.0–1.0 min 4% B; 1.0–6.0 min 50% B; 6.0–10.0 min 50–99.9% B; 10.0–10.05 min 99.9–4% B; 10.05–14.0 min 4% B. Ionisation was performed using ESI-source operation in positive mode and mass scans ranged from *m/z* 30–1000. The ion source parameters were set as follows: capillary voltage 4500 V; N<sub>2</sub> temperature (drying gas) 220°C; flow rate 8 L/min; collision energy 7.0 eV. The total run time of the analysis was 20 min. Metabolite search was based on the

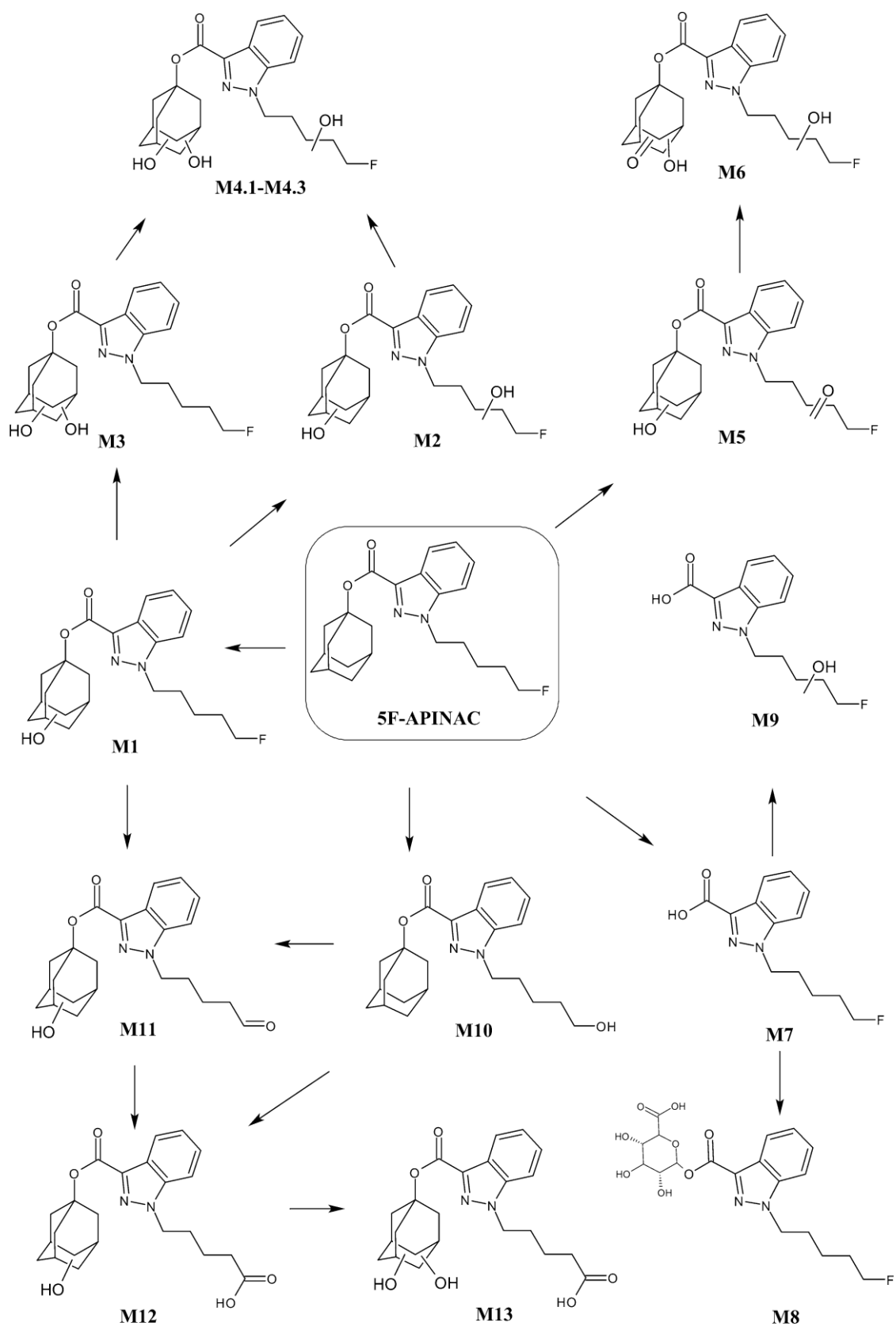
exact masses of the precursor ions of putative compounds and on typical product ions in auto MS/MS and target MS/MS modes.

## 2.3 RESULTS

The appreciable metabolites were found up to 6 h after administration. No metabolites were found 24 h after, which is in accordance with the extensive metabolism of SCs [104]. An overview of the proposed metabolic reactions associated with 5F-APINAC is presented in Figure 20. Extracted ion chromatograms for the metabolites associated with 5F-APINAC found *in vitro* and *in vivo* are presented in Appendix II Supplementary Figs. 1 and 2, respectively.

### 2.3.1 LC-IT-MS AND LC-QTOF ANALYSES

Fifteen 5F-APINAC-associated metabolites, including a phase II metabolite M8, were tentatively identified by LC-IT-MS and LC-QTOF-MS/MS following 5F-APINAC administration to rats (Figure 21, Figure 22). The structures of the metabolites were predicted with high accuracy and later corroborated by studying the fragmentation pattern of the product ions. Table 16 shows the proposed biotransformations, retention time,  $m/z$  of precursor and product ions, molecular formulae, and the mass errors of the proposed metabolites. 5F-APINAC eluted at 13.3 min and all the metabolites eluted between 5.1 and 12.5 min. 5F-APINAC had a protonated molecular ion at  $m/z$  385 with its consequent product ion at  $m/z$  135, corresponding to an adamantyl ring residue generated by the cleavage of the ester group (Figure 21 a). Identified metabolites were results of carboxylation, hydroxylation, ester hydrolysis, oxidation, or glucuronidation. Hydroxylation was the most common biotransformation. Identified metabolites were: 1-adamantanol metabolites (M1–M6), oxidative defluorination metabolites (M10–M13), and N-pentylindazole-3-carboxylic acid metabolites (M7–M9).

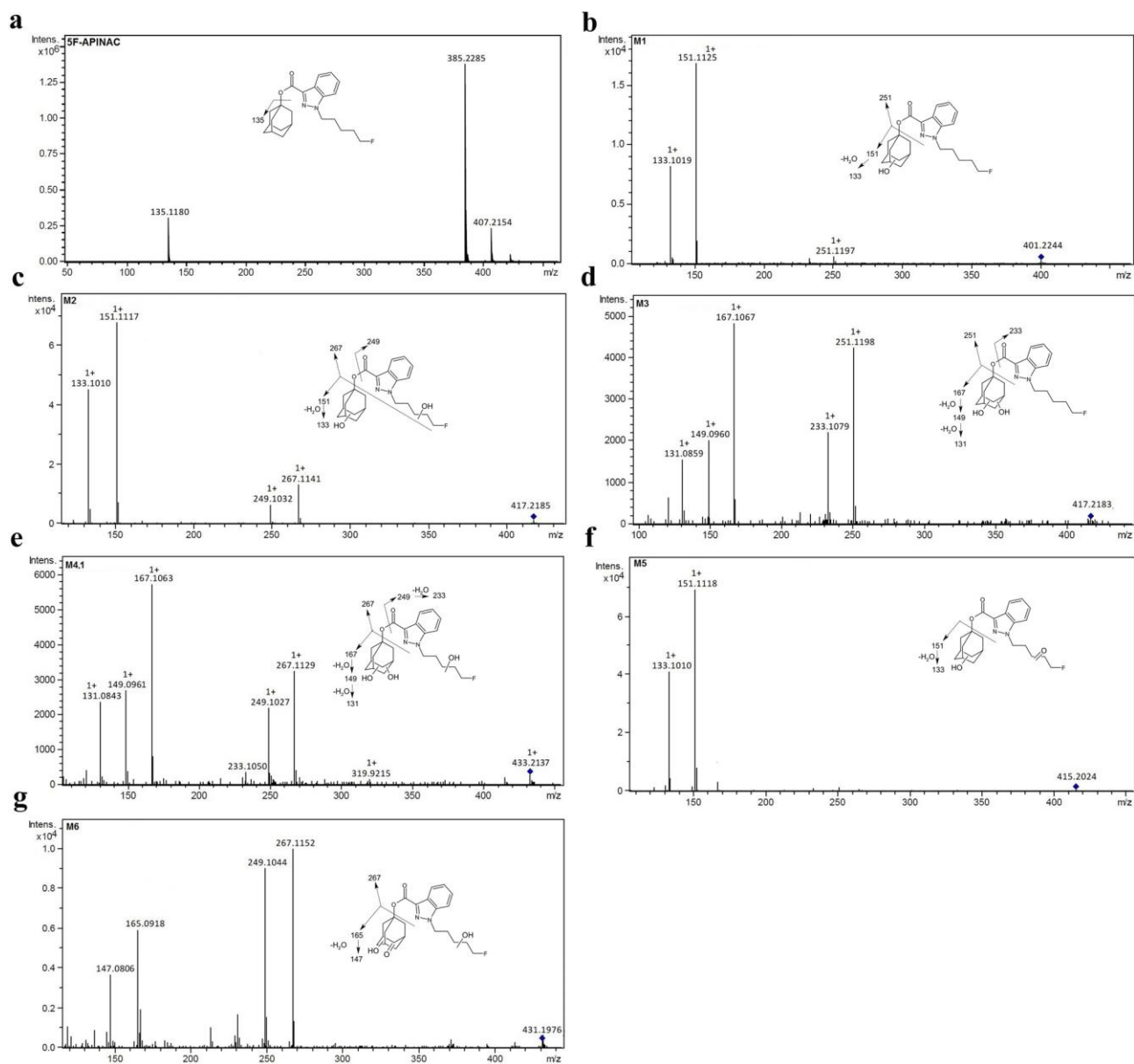


**Figure 20.** Proposed metabolic pathways of 5F-APINAC.

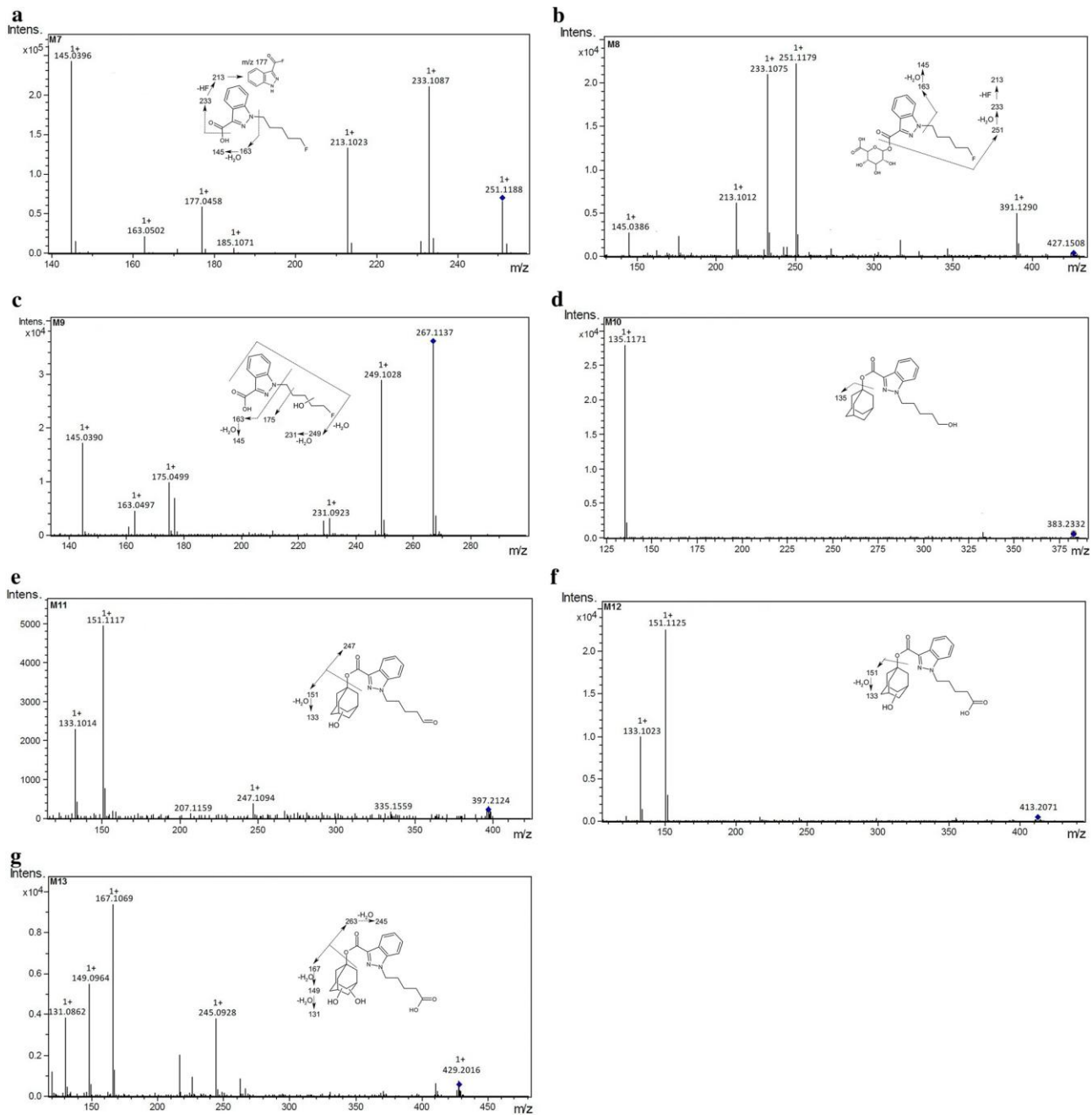
**Table 16.** Metabolites associated with 5F-APINAC identified in *in vivo* and *in vitro*.

Experiment	Metabolite	Biotransformation	RT (min)	Precursor ion (m/z)	Calc. (m/z)	Molecular formula	Error (ppm)	Product (m/z)	Product ion formula	Error (ppm)
In vitro In vivo	<b>5F-APINAC</b>	Parent compound	13.3	385.2285	385.2291	C <sub>23</sub> H <sub>30</sub> FN <sub>2</sub> O <sub>2</sub> <sup>+</sup>	4.4	135.1180	C <sub>10</sub> H <sub>15</sub>	4.4
In vitro In vivo	<b>M1</b>	Monohydroxylation	10.9	401.2244	401.2240	C <sub>23</sub> H <sub>30</sub> FN <sub>2</sub> O <sub>3</sub> <sup>+</sup>	1.0	133.1019 151.1125 251.1197	C <sub>10</sub> H <sub>13</sub> C <sub>10</sub> H <sub>15</sub> O C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>2</sub>	1.5 1.32 0.4
In vitro	<b>M2</b>	Dihydroxylation	9.1	417.2185	417.2190	C <sub>23</sub> H <sub>30</sub> FN <sub>2</sub> O <sub>4</sub> <sup>+</sup>	-1.2	133.1010 151.1117 249.1032 267.1141	C <sub>10</sub> H <sub>13</sub> C <sub>10</sub> H <sub>15</sub> O C <sub>10</sub> H <sub>13</sub> O C <sub>10</sub> H <sub>15</sub> O <sub>2</sub>	-5.2 -4.0 -2.8 -1.5
In vitro In vivo	<b>M3</b>	Dihydroxylation	9.7	417.2183	417.2190	C <sub>23</sub> H <sub>30</sub> FN <sub>2</sub> O <sub>4</sub> <sup>+</sup>	-1.6	131.0859 149.0960 167.1067 233.1079 251.1198	C <sub>10</sub> H <sub>11</sub> C <sub>10</sub> H <sub>13</sub> O C <sub>10</sub> H <sub>15</sub> O <sub>2</sub> C <sub>13</sub> H <sub>14</sub> FN <sub>2</sub> O C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>2</sub>	-1.5 -4.0 0.0 -4.7 0.8
In vitro In vivo	<b>M4.1</b> <b>M4.2</b> <b>M4.3</b>	Trihydroxylation	6.8 7.7 8.2	433.2137 433.2130 433.2134	433.2139	C <sub>23</sub> H <sub>30</sub> FN <sub>2</sub> O <sub>5</sub> <sup>+</sup>	-0.5 -2.1 -1.1	149.0961 167.1063 249.1027 267.1129	C <sub>10</sub> H <sub>13</sub> O C <sub>10</sub> H <sub>15</sub> O <sub>2</sub> C <sub>13</sub> H <sub>14</sub> FN <sub>2</sub> O <sub>2</sub> C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>3</sub>	-3.3 -2.4 -4.8 -3.7
In vitro	<b>M5</b>	Monohydroxylation + oxidation	8.9	415.2024	415.2033	C <sub>23</sub> H <sub>28</sub> FN <sub>2</sub> O <sub>4</sub> <sup>+</sup>	-2.2	133.1010 151.1118	C <sub>10</sub> H <sub>13</sub> C <sub>10</sub> H <sub>15</sub> O	-5.3 -3.3
In vitro	<b>M6</b>	Dihydroxylation + oxidation	7.6	431.1976	431.1976	C <sub>23</sub> H <sub>28</sub> FN <sub>2</sub> O <sub>5</sub> <sup>+</sup>	0.0	147.0806 165.0918 267.1152	C <sub>10</sub> H <sub>11</sub> O C <sub>10</sub> H <sub>13</sub> O <sub>2</sub> C <sub>13</sub> H <sub>14</sub> FN <sub>2</sub> O <sub>3</sub>	-2.7 4.8 2.6
In vitro In vivo	<b>M7</b>	Ester hydrolysis	7.2	251.1188	251.1190	C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>2</sub> <sup>+</sup>	-0.8	145.0396 163.0502 233.1087	C <sub>8</sub> H <sub>5</sub> N <sub>2</sub> O C <sub>8</sub> H <sub>7</sub> N <sub>2</sub> O <sub>2</sub> C <sub>13</sub> H <sub>14</sub> FN <sub>2</sub> O	0.0 0.0 -1.3
In vivo	<b>M8</b>	Ester hydrolysis + glucuronidation	5.6	427.1508	427.1511	C <sub>19</sub> H <sub>24</sub> FN <sub>2</sub> O <sub>8</sub> <sup>+</sup>	-0.7	145.0386 233.1075 251.1179 391.1290	C <sub>8</sub> H <sub>5</sub> N <sub>2</sub> O C <sub>13</sub> H <sub>14</sub> FN <sub>2</sub> O C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>2</sub> C <sub>19</sub> H <sub>20</sub> FN <sub>2</sub> O <sub>6</sub>	-6.9 -6.4 -6.8 -3.8
In vivo	<b>M9</b>	Ester hydrolysis + monohydroxylation	6.7	267.1137	267.1139	C <sub>13</sub> H <sub>16</sub> FN <sub>2</sub> O <sub>3</sub> <sup>+</sup>	-0.7	145.0390 175.0499 231.0923 249.1028	C <sub>8</sub> H <sub>5</sub> N <sub>2</sub> O C <sub>9</sub> H <sub>7</sub> N <sub>2</sub> O <sub>2</sub> C <sub>13</sub> H <sub>12</sub> FN <sub>2</sub> O C <sub>13</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub> F	-8.3 -4.6 -4.8 -4.4
In vitro	<b>M10</b>	Defluorination + monohydroxylation	12.5	383.2332	383.2329	C <sub>23</sub> H <sub>31</sub> N <sub>2</sub> O <sub>3</sub> <sup>+</sup>	0.8	135.1171	C <sub>10</sub> H <sub>15</sub>	-5.2
In vitro	<b>M11</b>	Oxidative defluorination (aldehyde formation) + monohydroxylation	9.3	397.2124	397.2127	C <sub>23</sub> H <sub>29</sub> N <sub>2</sub> O <sub>4</sub> <sup>+</sup>	-0.75	133.1014 151.1117 247.1094	C <sub>10</sub> H <sub>13</sub> C <sub>10</sub> H <sub>15</sub> O C <sub>13</sub> H <sub>15</sub> N <sub>2</sub> O <sub>3</sub>	-2.2 -4.0 -4.4
In vitro In vivo	<b>M12</b>	Oxidative defluorination (carboxyl formation) + monohydroxylation	9.0	413.2071	413.2071	C <sub>23</sub> H <sub>29</sub> N <sub>2</sub> O <sub>5</sub> <sup>+</sup>	0.0	133.1023 151.1125	C <sub>10</sub> H <sub>13</sub> C <sub>10</sub> H <sub>15</sub> O	4.5 1.3
In vitro In vivo	<b>M13</b>	Oxidative defluorination (carboxyl formation) dihydroxylation	7.9	429.2016	429.2020	C <sub>23</sub> H <sub>29</sub> N <sub>2</sub> O <sub>6</sub> <sup>+</sup>	-0.9	131.0862 149.0964 167.1069 245.0928	C <sub>10</sub> H <sub>11</sub> C <sub>10</sub> H <sub>13</sub> O C <sub>10</sub> H <sub>15</sub> O <sub>2</sub> C <sub>13</sub> H <sub>13</sub> N <sub>2</sub> O <sub>3</sub>	0.8 -1.3 1.2 0.8





**Figure 21.** Product ion spectra of metabolites associated with 5F-APINAC recorded by LC-QTOF (part 1). The diamond shows the location of each protonated molecular ion of the metabolites. **a** 5F-APINAC, **b** M1, **c** M2, **d** M3, **e** M4.1, **f** M5, **g** M6.



**Figure 22.** Product ion spectra of metabolites associated with 5F-APINAC recorded by LC-QTOF (part 2). **a** M7, **b** M8, **c** M9, **d** M10, **e** M11, **f** M12, **g** M13.

### 2.3.2 HLM INCUBATION OF 5F-APINAC

5F-APINAC was rapidly ester-hydrolysed when incubated with HLMs. It was fully converted after 3h of incubation. Utilisation of the NADPH-dependent system presumably resulted in the formation of 11 active metabolites related to 5F-APINAC, formed predominately by hydroxylation and oxidation transformations.

## 2.3.3 METABOLIC PATHWAY

### Monohydroxylation

M1 was found in the *in vitro* and in *in vivo* experiments up to 3h. It was proposed as a monohydroxylated metabolite. The protonated molecular ion was found at  $m/z$  401 and eluted at 10.9 min. The product ions at  $m/z$  151 indicate that the monohydroxylation was on its adamantyl ring. The ion at  $m/z$  133 indicates the loss of water from the adamantyl ring (Figure 21.b).

### Dihydroxylation

M2 and M3 were proposed as dihydroxylated metabolites. The protonated molecular peak was found at  $m/z$  417. M2 was only found *in vitro*, whereas M3 was found both *in vitro* and *in vivo* up to 6h after administration. Regarding M2, two product ions were found at  $m/z$  151 and 133. This supports monohydroxylation occurring in the adamantyl ring and dehydration from the product ion at  $m/z$  151. The ion at  $m/z$  267 indicates the presence of the hydroxyl group on the N-fluoropentyl side chain, while  $m/z$  249 corresponds to its dehydrated derivative (Figure 21.c).

M3 product ion spectrum showed fragments at  $m/z$  167, 149, and 131. The product ions support that two hydroxylations occurred in the adamantyl ring. The fragment at  $m/z$  167 correspond to adamantanediol, and the product ions at  $m/z$  149 and 131 indicate two consecutive dehydrations from the product ion at  $m/z$  167. At the same time, the product ion at  $m/z$  251 and the ion at  $m/z$  233 prove the absence of hydroxylation at N-fluoropentyl side chain (Figure 21.d).

### Trihydroxylation

M4.1, M4.2, and M4.3 were proposed as the trihydroxylation metabolites. They were found both *in vitro* and *in vivo* up to 6h after administration. Their protonated molecular ions appeared at  $m/z$  433, and their retention times were 6.8, 7.7 and 8.2 min, respectively (Appendix II Fig. S1). All metabolites had the same spectrum, which suggests that the hydroxylation on the N-fluoropentylindazole moiety chain occurs in three different places, and that the adamantyl moiety undergoes two hydroxylations. The product ions at  $m/z$  267 and 167 may characterise the breakdown of the ester bond. The one at  $m/z$  167 is thought to belong to adamantanediol, and the product ions at  $m/z$  149 and 131 indicate two consecutive dehydrations from the product ion at  $m/z$  167. The ion at  $m/z$  267 is thought to belong to the 3-carboxy-1-(5-fluorohydroxypentyl)-1H-indazole. The ions at  $m/z$  249 and 233 indicate two consecutive dehydrations from  $m/z$  267 (Figure 21.e).

### Monohydroxylation and oxidation

M5 was only found *in vitro* up to 3h after administration. It eluted at 8.9 min and had its protonated molecular ion at  $m/z$  415. Its spectrum showed two product ions at  $m/z$  151 and 133 belonging to the monohydroxylated adamantyl ring and its dehydration, respectively (Figure 21.f).

### **Dihydroxylation and oxidation**

M6 was found *in vitro* with a protonated molecular ion at  $m/z$  431. This metabolite indicates the presence of carbonyl and hydroxyl groups at the adamantyl ring ( $m/z$  165) and hydroxylation at the 5-fluoropentyl side chain ( $m/z$  267). Moreover, the ion at  $m/z$  147 is related to dehydration from the ion at  $m/z$  165 (Figure 21.g).

### **Ester hydrolysis**

Because it has been reported that SCs containing an ester group (i.e., BB-22, PB-22, 5F-PB-22, APINAC, and NM-2201) are extensively hydrolysed [107,108], 5F-APINAC was expected to undergo a similar process. M7 was found in the *in vitro* and *in vivo* experiments up to 6h after administration. Moreover, it was one of the principal metabolites found in rat urine. It eluted at 7.2 min, and its precursor ion was identified at  $m/z$  251. The ion at  $m/z$  233 corresponds to the dehydration of the protonated molecular M7, while  $m/z$  163 and 145 are associated with the consistent loss of the pentyl side chain and dehydration, respectively. Natural loss of hydrogen fluoride at  $m/z$  233 resulted in the formation of a product ion at  $m/z$  213. At the same time, an ion at  $m/z$  177 was formed by fluorine rearrangement [109] (Figure 22.a).

### **Glucuronidation**

M8 was identified as a phase II metabolite in the *in vivo* study, which is a glucuronide conjugate. Its protonated molecular ion was found at  $m/z$  427. M8 underwent several fragmentations in the ion source, which resulted in the appearance of product ion  $m/z$  251 corresponding to the protonated molecule of M7. As reported by Savchuk et al. [88], this is a property common to phase II metabolites (ester of glucuronic acid). Further product ions were found at  $m/z$  233, 213, 177 and 145, as described for M8. Another fragment with  $m/z$  391 could be associated with two dehydrations of the M8 molecular ion (Figure 22.b).

### **Ester hydrolysis followed by monohydroxylation**

M9 was identified only in the *in vivo* study up to 6h after administration. It eluted at 6.7 min and was identified from its protonated molecular ion at  $m/z$  267. The product ion spectra contained ions at  $m/z$  249, 231, 175, and 145. The product ion with the highest intensity at  $m/z$  249 corresponds to the metabolite after dehydration. The fragment at  $m/z$  231 was generated by dehydration from its precursor ion at  $m/z$  249. The ion at  $m/z$  175 strongly suggests a cleavage between the N- $\alpha$  and  $\beta$  carbons of the hydroxylated N-fluoropentyl group in this metabolite (Figure 22.c)

### **Oxidative defluorination (hydroxyl formation)**

M10 was only found *in vitro* and had its protonated molecular ion at  $m/z$  383. M10 was proposed as the metabolite after a loss of fluorine and the consecutive addition of a hydroxyl group. The

adamantyl cation at  $m/z$  135 was the only fragment in the product ion spectra which suggests an unmodified adamantyl ring. Such a process has also been reported in 5F-AKB-48, XLR-11, and AM-2201 [98] (Figure 22.d).

#### **Oxidative defluorination (aldehyde or carboxyl formation) and monohydroxylation**

M11 (only found *in vitro*) showed a protonated molecular ion at  $m/z$  397 that was associated with consequent oxidation of a hydroxyl group to aldehyde and the addition of a hydroxyl group on the adamantyl ring ( $m/z$  151). The fragment at  $m/z$  133 corresponds to dehydration from  $m/z$  151, and the ion at  $m/z$  247 to the breakdown of the ester bond (Figure 22.e).

M12 found *in vitro* and *in vivo* showed a protonated molecular ion at  $m/z$  413, probably followed by oxidation to carboxyl group via the aldehyde (M11) and the addition of a hydroxyl group on the adamantyl ring ( $m/z$  151). The fragment at  $m/z$  133 corresponds to dehydration from  $m/z$  151 (Figure 22.f).

#### **Oxidative defluorination (carboxyl formation) and dihydroxylation**

M13 was found *in vitro* and *in vivo* up to 6h after administration and had its protonated molecular ion at  $m/z$  429 which corresponds to the loss of fluorine followed by introduction of a carboxyl group and addition of two hydroxyl groups on the adamantyl ring. The fragment at  $m/z$  167 is thought to be due to adamantanediol, and the ions at  $m/z$  149 and 131 indicate two consecutive dehydrations from the product ion at  $m/z$  167. The less intense fragment at  $m/z$  245 could belong to the fragment appearing after cleavage of the ester bond (Figure 22.g).

#### **Metabolites found in negative ionisation mode**

M10 had two product ions corresponding to the indazole moiety ( $m/z$  117) and the 1-fluoro-5-(1H-indazol-1-yl)pentan-3-ol fragment ( $m/z$  221). M7 had its protonated molecular ion at  $m/z$  249. The fragment corresponding to indazole appear at  $m/z$  117 and the fragment at  $m/z$  205 is thought to belong to the C-C cleavage between the indazole moiety and the carboxyl group. M13 ( $m/z$  427) had only one fragment in its product ion spectrum. It is suggested that  $m/z$  327 is a consequence of C-N cleavage, detaching the carbon chain from the rest of the molecule.

## **2.4 DISCUSSION**

Although the *in vivo* metabolism of 5F-APINAC in humans was not characterised, utilisation of *in vitro* models of HLMs together with *in vivo* experiments in rats is considered the first step for characterising the metabolism of new drugs. The non-fluoro counterpart APINAC (AKB57) is well known to have circulated and was first detected in smoking mixtures on the black market in South Korea [106]. In September 2016, a powdery substance (1.97 g, purity 72% investigated by high-performance liquid chromatography) containing APINAC was confiscated in Saratov Region

(Russia) in the course of a police investigation [88]. However, to our knowledge, at the international level, APINAC is subjected to control only in the Republic of Belarus and Japan.

As for the fluorinated analogue of APINAC, it is easily expected that the mode of drug action is enhanced, because the fluorine substitution at the 5-pentyl position of pentyl-indazoles/pentylindoles was reported to enhance the drug's stability, potency, and half-life [93,104]. Because of the worldwide lack of regulation and the high potency of 5F-APINAC, it is probable that this drug will circulate widely. Thus, it is important to characterise the metabolism of 5F-APINAC.

Due to the unavailability of authentic human urine specimens with 5F-APINAC and its metabolites, experiments with HLMs *in vitro*, and with rat urine *in vivo* collected 3, 6, and 24h after administration of 5F-APINAC were carried out instead. There are two papers describing the metabolism of APINAC, the non-fluoro counterpart of 5F-APINAC [88,107]. If the fluorine at the terminus of the pentyl side chain is intensively attacked for defluorination at the initial/early stage of metabolism, the metabolites of 5F-APINAC common to those of APINAC will appear. However, in contrast to our expectation, we found that M12 was the only metabolite common to both SCs. The metabolic profiles for 5F-AKB-48 and AKB-48, the compounds analogous to 5F-APINAC and APINAC, respectively (Figure 19), were also studied by Vikingsson et al. [110]. They concluded that only a few identified metabolites were shared between AKB-48 and 5F-AKB-48. Therefore, it is easy to discriminate between 5F-APINAC and APINAC, and also between 5F-AKB-48 and AKB-48 by checking multiple metabolites for each SC. In addition, it should be pointed out that N-5-fluoropentylindazole-3-carboxylic acid (M7), which appeared as an abundant peak in both *in vitro* and *in vivo* experiments, is not a metabolite specific to 5F-APINAC, but can also be most probably produced from 5F-SDB-005, an indazole carboxylate SC with a 5-fluoropentyl side chain.

Based on the experiments, the most prevalent metabolite, which will be easily detected in human urine, is M7. With this metabolite alone, it can be concluded that the SC has a 5-fluoropentyl side chain and is suggestive of the consumption of 5F-APINAC or 5F-SDB-005, although the use of a new/unknown 5-fluoropentylindazole carboxylate SC cannot be ruled out. In addition to M7, M4.1, or M13 having an adamantyl ring, it is essential to definitively prove 5F-APINAC consumption.

## 2.5 CONCLUSION

In the present study, 15 metabolites associated with the novel SC 5F-APINAC were described. Tentative identification of the metabolites was performed through *in vitro* incubation with HLMs and *in vivo* with rat urine after intravenous administration of the drug. The predominant metabolic reactions were ester hydrolysis remaining in both *in vitro* and *in vivo* experiments and resulting in the formation of M7 (5-fluoropentylindazole-3-carboxylic acid). Other metabolic transformations included mono-, di- and trihydroxylation of the adamantyl ring and N-fluoropentylindazole moiety, oxidation of the N-fluoropentyl side chain, oxidative loss of fluorine and glucuronidation, as well as

combinations thereof. The discovered metabolites may serve for future studies and are likely to be incorporated into routine analytical screening methods as urine markers of 5F-APINAC consumption.

## **2.6 LIMITATIONS AND FUTURE RESEARCH**

The metabolism of 5-F APINAC was not characterised in humans. Utilisation of *in vitro* models of HLM together with *in vivo* experiments in rats is considered an optimal choice for characterising the metabolism of new drugs. However, *in vitro* incubation simulates the metabolic transformation only in the liver, whereas metabolism in rats may differ from humans. Rat models cannot fully replace controlled human studies. Although the experiments were not performed in humans, the use of the murine model presented the advantage of being able to study the metabolism of the drug in time series, which is difficult to conduct in humans due to ethical considerations. It is important to point out that the research was limited in terms of the lack of identification of the exact positions of the modifications. It was only possible to deduce whether a modification was located on the adamantyl group or on the pentylindazole moiety. Future studies should ideally include experiments based on analyses performed in specimens collected from humans.

### **3. DEVELOPMENT AND VALIDATION OF A NEW METHOD FOR THE DETERMINATION OF 13 SECOND-WAVE SYNTHETIC CANNABINOIDS IN HAIR**

#### **3.1 INTRODUCTION**

Novel psychoactive substances (NPS) have been designed to mimic the psychoactive effects of common illicit drugs. However, their different chemical structures hinder their detection during standard toxicological screenings. The consumption of these psychoactive substances is known as a “legal high” since they escape current law restrictions. In recent years, the EMCDDA has reported an increased use of these substances throughout the European Union. The main concern is the increased psychoactive effects compared to classical drugs, which have life-threatening side effects. This increased potency is a result of the higher binding capacity of these NPS to CB<sub>1</sub> and CB<sub>2</sub> receptors [111]. During the pandemic, the illicit drug market was disrupted but quickly recuperated, with an increase of sales in the dark markets once the restriction measures were relaxed. Moreover, some health conditions linked to drug use together with drug sharing, such as a sharing a cannabis joint, is now a greater health risk than before [112].

Synthetic cannabinoids (SCs) are one type of NPS that mimic the effects of  $\Delta^9$ -Tetrahydrocannabinol (THC). SCs are in the top two substances being seized and the largest group of NPS being monitored, as per the latest report of the EMCDDA [112]. The first SC detected was JWH-018, the most popular first-generation SC, followed by the second-generation of SCs, such as ADB-PINACA, UR-144, 5F-PB-22, AB-FUBINACA, AKB-48 [93]. They do not contain cannabis but produce similar effects due to the binding to the CB<sub>1</sub> and CB<sub>2</sub> receptors. They are produced by spraying chemical additives on dried plant products and are sold as a legal alternative to cannabis [113]. They are available as incenses, tea, food additives, herbal mixtures, energy drinks, and scents [114]. Specific effects of SCs are hard to predict due to the lack of pharmacokinetic information. However, it is very well-known that all SCs are extensively metabolised [115]. Typical intoxication symptoms include vomiting, abdominal or flank pain, nausea, and hypertension. There is little clinical information, but acute kidney injury has been associated with SC use [93].

On account of the fast emergence of new SCs where only a side chain is modified (large structural diversity) together with an increase in consumption of these herbal mixtures, there is a need to develop fast and high throughput analytical screening methods for their reliable detection. To date, SCs have been determined in several biological specimens such as serum, hair, urine, and blood [114–120]. The method of choice is liquid chromatography coupled with different types of mass detection, such as, QTOF, Qtrap, and QQQ [114–120]. Recently, Bruker developed a high throughput LC–MS system, called Toxtyper®, which operates as screening device using libraries



that contain more than 4000 compounds, including their main metabolites. These libraries are constantly updated with the emergent NPS substances [121].

As mentioned above, several biological specimens have potential to be used for the analysis of SCs. It is a well-known fact that keratin in the hair's root contains several xenobiotics that are also present in blood and are protected from environmental factors as well as metabolism. Therefore, it seems highly suitable to use hair as the chosen specimen for the identification of SCs [114]. To the best of our knowledge, the only method which uses this technology for the toxicological screening of hair does not include synthetic cannabinoids [122].

The aim of the present work is the development and validation of a screening and semi-quantitative method for 13 second-wave SCs in hair based on liquid chromatography coupled with IonTrap (LC–QIT–MS): (*R*)-5F-ADB, 5F-PB 22, 5CI-AB-PINACA, 5F-AKB-48, 5F-APP-PICA (PX-1), 5F-APP-PINACA (PX-2), 5F-CUMYL-PINACA, AB-CHMINACA, AB-FUBINACA, ADB-FUBINACA, MDMB-CHMICA, MMB-2201, and UR-144. The method was applied to hair samples from subjects who have tested positive for one or several other illegal drugs.

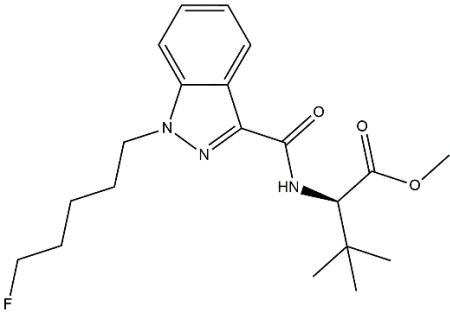
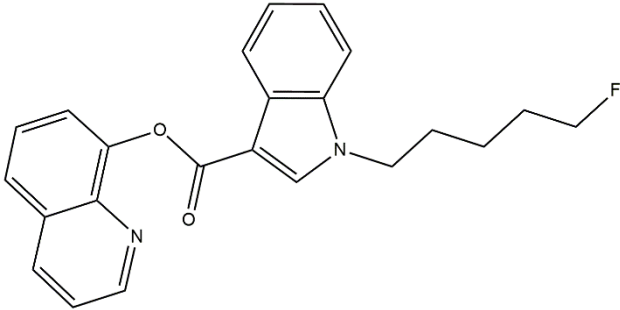
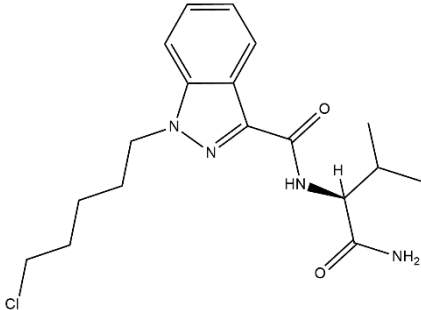
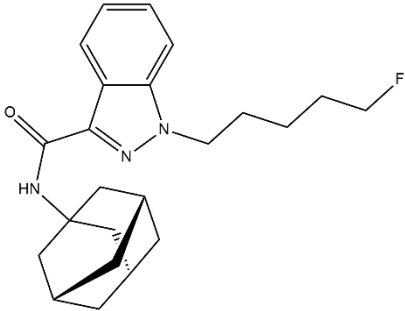
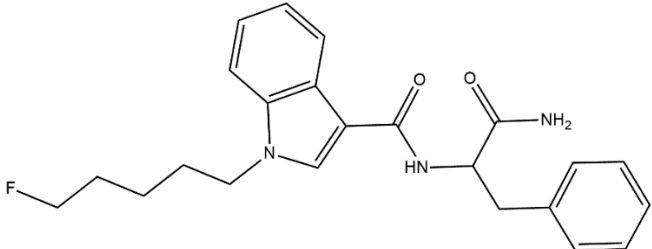
## **3.2 MATERIALS AND METHODS**

### **3.2.1 REAGENTS**

Standard solutions of (*R*)-5F-ADB, 5F-PB 22, 5CI-AB-PINACA, 5F-AKB48, 5F-APP-PICA (PX-1), 5F-APP-PINACA (PX-2), 5F-CUMYL-PINACA, AB-CHMINACA, AB-FUBINACA, ADB-FUBINACA, MDMB-CHMICA, MMB-2201, and UR-144 were purchased from Comedical (Trento, Italy). Their structures, MRM transitions, and retention times are displayed in Table 17. D<sub>5</sub>-diazepam was purchased from Cayman chemicals (Michigan, US) and used as IS.

Ammonium was provided by AnalaR in salt form (NH<sub>4</sub>Cl) (Merck, Darmstadt, Germany). Acetonitrile (ACN), methanol (MeOH), and 2-propanol for UHPLC were purchased from VWR Chemicals (Fontenay-Sous-Bois, France). 98% formic acid for LC–MS and Dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>) for HPLC were obtained from Merck KGaA (Darmstadt, Germany). Absolute ethanol was provided by Carlo Erba Reagenti (Milan, Italy). Ultrapure water was obtained from an ELGA VEOLIA (Lane End, High Wycombe, UK) water purification system.

**Table 17.** Structure, MRM transitions and retention time (RT) of all SCs researched.

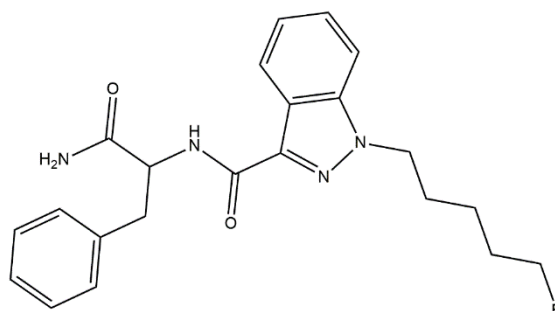
Synthetic cannabinoid	Precursor ( $m/z$ )	Fragment ( $m/z$ )	RT (min)	Structure
(R)-5F-ADB	378.2	318.2	4.13	
5F- PB 22	377.2	232.0	3.93	
5Cl- AB- PINACA	365.2	348.0	3.21	
5F-AKB-48	384.2	135.0	5.81	
5F-APP-PICA (PX-1)	396.2	379.2	3.12	

5F-APP-  
PINACA (PX-  
2)

397.2

380.1

3.19

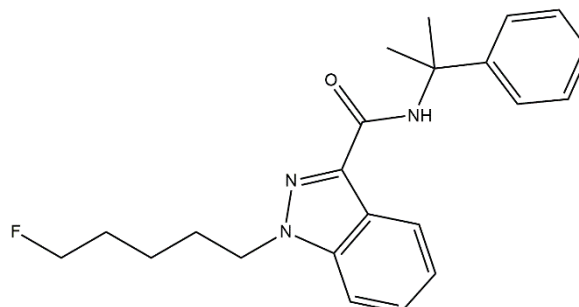


5F-CUMyL-  
PINACA

368.2

250.2

4.30

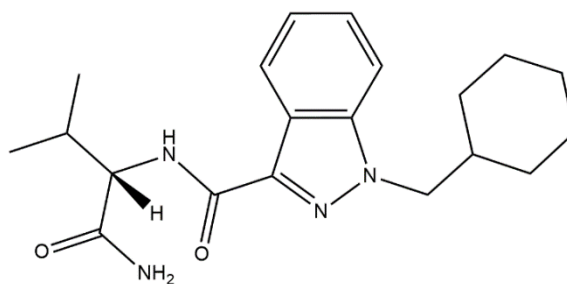


AB-  
CHMINACA

357.2

340.2

3.65

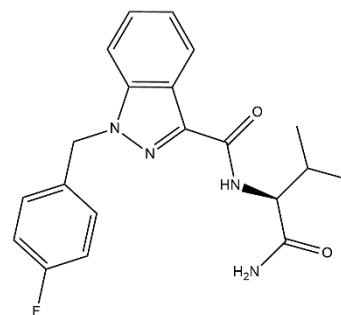


AB-FUBINACA

369.2

352.1

3.12

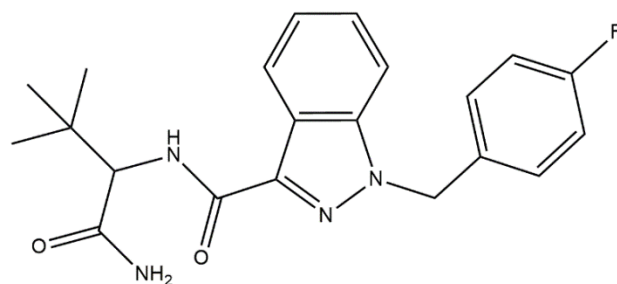


ADB-  
FUBINACA

383.2

366.1

3.16

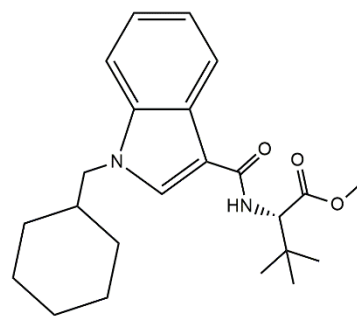


MDMB-  
CHMICA

386.2

326.2

4.90

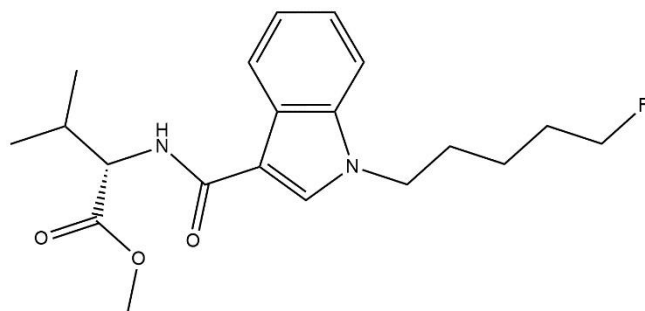


MMB-2201

363.2

232.1

3.51

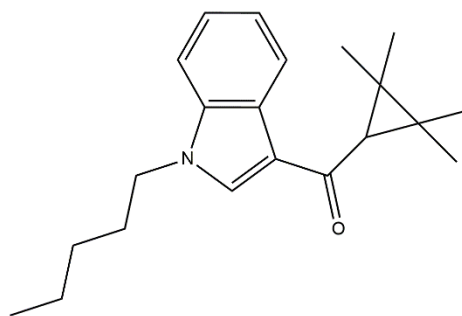


UR-144

312.2

214.0

6.55



### 3.2.2 INSTRUMENTATION

Analysis were performed on a Toxtyper<sup>®</sup> LC/IT-MS system. The instrument consisted of an Elute UHPLC system coupled with amaZon speed<sup>®</sup> ion trap with three different sources: electrospray ionisation sources (ESI), atmospheric pressure chemical ionisation (APCI), and IonBooster<sup>™</sup> (iB) (Bruker Daltonics, Bremen, Germany). Nitrogen was used as a nebulising and drying gas. The source settings were optimised in order to achieve the highest and most reproducible signal of the analytes.

Two analytical columns were used: Acclaim<sup>™</sup> RSLC 120 C18 (2.1x100 mm, 2.2  $\mu$ m, 120 Å) from Thermo Fisher (Idstein, Germany) and Kinetex C18 (2.1X100 mm, 2.6  $\mu$ m, 100Å) from Phenomenex (California, US). The column oven was kept at 40°C. Mobile phase A (aqueous) contained 1% ACN, 0.1% formic acid and 2 mM ammonium formate. Mobile phase B (organic) contained 1% H<sub>2</sub>O, 0.1% formic acid and 2 mM ammonium formate. Gradient elution was performed as follows: 0-1 min: 20% B, 1-2.50 min: 20-60% B linear, 2.50-4 min: 60-65 %B linear, 4-5.50 min: 65% B, 5.50-8 min: 65-99%, 8-10 min: 99%, 10-10.20 min: 99-20% B, 10.20-12 min: 20% B. The flow rate was 0.5 mL/min, and the injection volume was 5  $\mu$ L. The MS was operated in AutoMS<sup>n</sup> from 70–800 amu.

The method already provided by the company was optimised with the different sources. SC identification was done through the Toxtyper<sup>®</sup> software based on retention time (RT), full scan MS, MS<sup>2</sup>, and MS<sup>3</sup> using the synthetic cannabinoids spectral library. The acquired MS<sup>n</sup> spectra are compared against the library and a purity score is generated, which symbolises the degree of consistency between the library and the experimental spectrum. The purity has to be at least 700 for a positive identification and 1000 means 100% identity. In the present work, purity must be above 850 and only a  $\Delta_{RT} = \pm 0.3$  min was allowed for considering the SC identified by the Toxtyper<sup>®</sup> system.

### **3.2.3 SOLUTIONS AND CALIBRATION CURVE PREPARATION**

For the stock solution, a mix with all SCs was prepared by diluting adequate volumes of the SC with MeOH to a concentration of 1  $\mu\text{g/mL}$ . For the IS solution, d<sub>5</sub>-diazepam was diluted to 1  $\mu\text{g/mL}$  with MeOH. Both solutions were stored at  $-20$  °C. The hair calibration standards (3, 2, 1, 0.5, 0.3 ng/mg) were prepared by adding the adequate amount of SC and 50  $\mu\text{L}$  of IS solution (constant concentration of IS 1 ng/mg) to 50 mg of blank hair before extraction.

### **3.2.4 HAIR SAMPLE PREPARATION**

The method was developed using blank hair samples. The samples were spiked with the mix of SCs and the IS at the desired concentration levels. In addition, 11 samples from real cases were used for the validation of the methodology.

Sample preparation was adapted from Hutter et al.[116]. Briefly, 70 mg of blank hair were washed with dichloromethane for 5 minutes. They were left to dry for 10 minutes and afterwards they were cut and pulverised using the homogeniser instrument (Precellys<sup>®</sup> Evolution, Bertin instruments). Then, 50 mg of the blank hair was weighed in a glass vial and then extracted with 1,5 mL of EtOH for 3 hours in an ultrasonic bath. After extraction, the samples were centrifuged for 10 minutes. Successively, 1 mL of the solution was transferred to a glass vial and evaporated with N<sub>2</sub> at room temperature. The dry residue was reconstituted in 100  $\mu\text{L}$  of phases A/B 50/50 (v/v).

### **3.2.5 METHOD VALIDATION**

Firstly, the suitability of the two available columns (1 and 2) for the successful and unequivocal separation of all 13 SCs was evaluated. Secondly, the three different available sources were evaluated based on the intensity of the signal for all 13 SCs. For each source, their MS parameters were optimised and then the intensities of the 13 SCs were compared for each optimised source. Thirdly, after selecting the column and source, the method was validated according to method validation guidelines suggested by Poletini [71] for the quantification of SCs in human hair samples. Validation samples were prepared by spiking blank hair with the IS and SC mixture prior to extraction. Selectivity was assessed by measuring three blank hair samples in triplicate. A calibration curve was evaluated by six replicate measurements for each point. Accuracy and precision studies were carried out by analysing three samples at high (2 ng/mg), middle (1 ng/mg), and low (0.5 ng/mg)

concentrations for five days with six replicates for each concentration point. LOD was determined as the lowest concentration at which the analyte could be identified with  $S/N \geq 3$ .

Matrix effects, process efficiency, and extraction recovery were assessed following the guidelines of Matuszewski [123]. The matrix effect (ME) is calculated as the ratio between the peak areas determined by the analysis of hair samples fortified with the SCs after extraction and the standards in mobile phase (50/50, v/v). The process efficiency (PE) is calculated as the ratio between the peak areas determined by the analysis of hair samples fortified with the SCs before extraction and the standards in mobile phase (50/50, v/v). The extraction recovery (R) was calculated as the ratio between the peak areas determined by the analysis of hair samples fortified with the SCs before extraction and the peak areas determined by the analysis of hair samples fortified with the SCs after extraction.

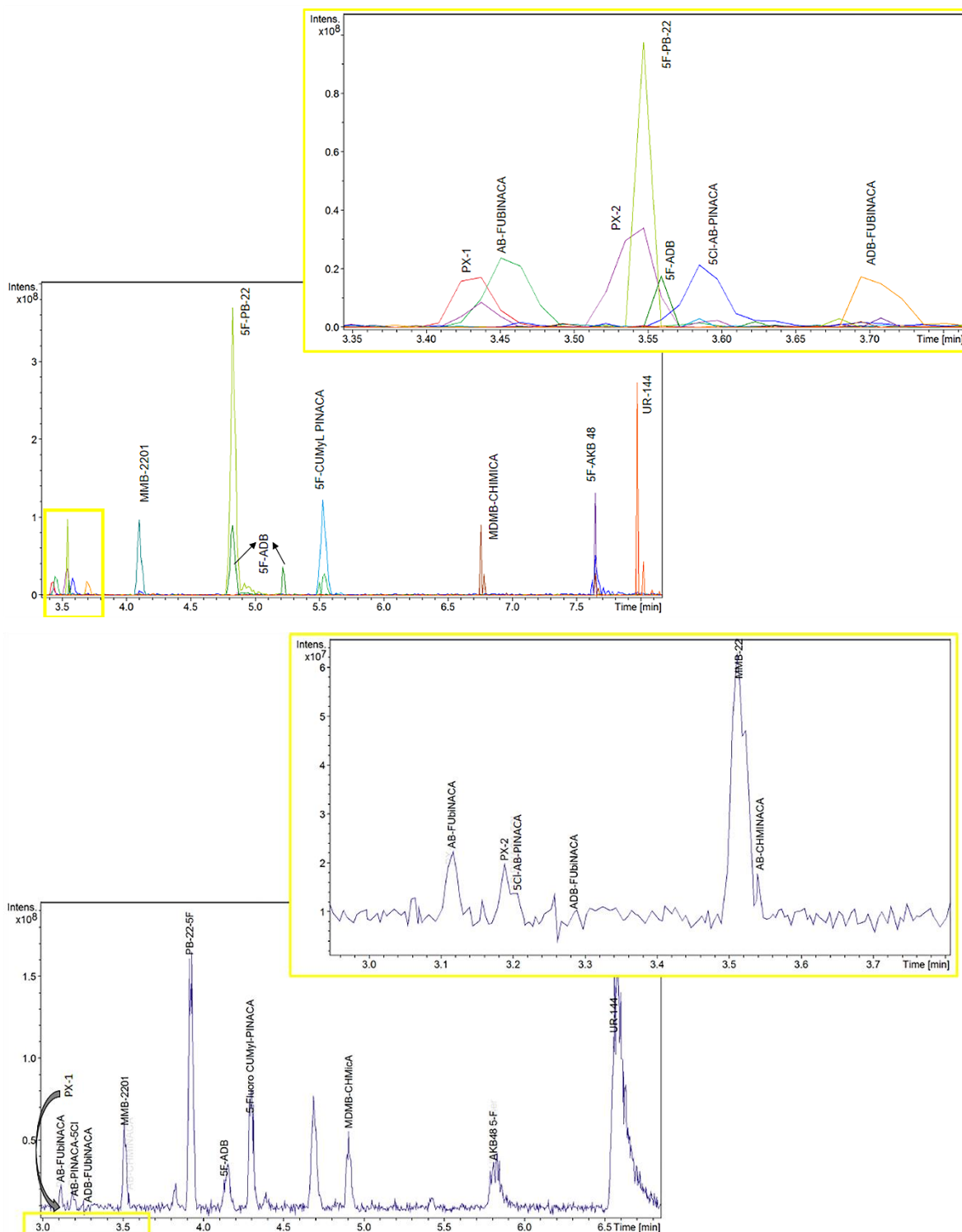
### **3.2.6 REAL CASES**

Authentic hair samples were collected from subjects tested positive for one or more of the following illegal drugs: cocaine and its metabolites cocaethylene, benzoylecgonine and ecgonine methyl ester, codeine, amphetamines, MDMA, MDA, ketamine, nor-ketamine, morphine, methadone and its metabolite EDDP, heroine metabolite 6-MAM, dihydrocodeine (DHC). No information about the use of SC was available. The samples were prepared following the developed method.

## **3.3 RESULTS AND DISCUSSION**

### **3.3.1 COMPARISON OF THE TWO COLUMNS**

Two columns were tested to evaluate their suitability for the best separation of all the SCs. In particular, the columns recommended by the equipment producer, Acclaim™ and Kinetex, were both evaluated (Column 1 and Column 2, respectively). They differ for particle sizes and pore sizes: 2.2  $\mu\text{m}$  (column 1) and 2.6  $\mu\text{m}$  (column 2), and 120  $\text{\AA}$  (column 1) and 100  $\text{\AA}$  (column 2), respectively. As shown in Figure 23, with column 1, separation occurs in 8 minutes but between 3.35–3.70 min many analytes elute in overlapping peaks. In addition, some compounds such as 5F-ADB and 5F-PB-22 show two peaks for the same compound. On the contrary, column 2 was the best choice since the separation takes place in a shorter time (7 minutes) and in column's 2 cluster area, only six compounds elute and are clearly very well separated. Therefore, column 2 was selected as the best choice for method development and validation.



**Figure 23.** Comparison of column 1 (top) and column 2 (bottom) for the separation of 13 SCs.

### 3.3.2 COMPARISON OF ESI, IB, AND APCI SOURCES

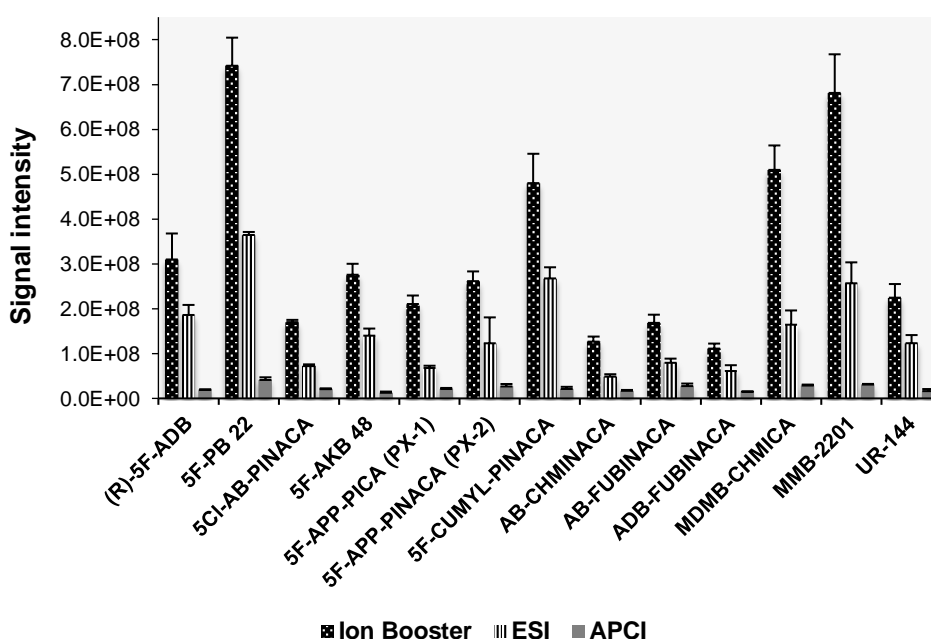
In this study, three different sources were tested in order to evaluate their suitability for the analysis of SCs. In addition to the common ESI and APCI atmospheric pressure sources, ionBooster (iB) was tested. The IB is a heated ESI source in the spray zone which increases the desolvation process when using flows rates higher than 200  $\mu\text{L}/\text{min}$ , resulting in improved detection limits.

The SCs were at a concentration of 200 ng/mL for all source tests. Information about the optimised parameters and the range of optimisation can be found in Table 18.

**Table 18.** Source settings for the optimisation of all three sources with the SynnCann method. The values in **bold** are the optimised values, (N/A not applicable).

	ESI	iB	APCI
Capillary voltage (V)	1000, <b>3000</b> , 4500	<b>1000</b> , 3000, 4000	2000, <b>4000</b> , 5500
End of plate offset (V)	100, <b>500</b> , 1000	<b>300</b> , 600, 900	<b>1000</b> , 2000, 3000
Corona (nA)	N/A.	N/A	5000, <b>15000</b> , 30000
Charging volt	N/A.	300, <b>600</b> , 900	N/A
Nebuliser (bar)	1, 3, <b>4</b>	1, <b>3</b> , 4	1, <b>3</b> , 4
Dry gas flow (L/min)	3, 7, <b>11</b>	<b>3</b> , 7, 11	<b>3</b> , 7, 11
Dry gas temp (°C)	60, 100, <b>160</b>	60, 100, <b>160</b>	60, <b>100</b> , 160
Vaporiser temp (°C)	N/A	<b>250</b> , 350, 450	250, <b>350</b> , 450

The criteria followed for choosing the best source was highest intensity and if intensities were similar, the source with lowest CV (CV<20%) was chosen. As shown in Figure 24 the Ion Booster (iB) source is the most adequate for the analysis, with the signal almost double of that obtained with ESI. This is in accordance with literature where the iB source has been previously proposed for the successful detection of the first wave of SCs [120]. As shown in Figure 24, the ionisation source strongly influences the signal intensity.



**Figure 24.** Influence of the ionization source on the signal intensities of the studied SCs.



### 3.3.3 CALIBRATION, LLOQ, LOD, AND SELECTIVITY

Selectivity was tested through the analysis of three different drug-free hair samples without IS. Blank samples did not reveal any interference with the MRM transition of the analytes.

For all compounds, linearity was assessed in the range of 0.3-3 ng/mg. The limit of detection (LOD), expressed as S/N>3, was 0.0625 ng/mg for all the compounds except for 5CI- AB-PINACA, 5F-APP-PICA (PX-1), 5F-APP-PINACA (PX-2) and AB-CHMINACA for which LODs were 0.125 ng/mg. The lower limit of quantification (LLOQ), expressed as S/N>10, was 0.3 ng/mg for all compounds, except for 5CI- AB-PINACA, for which the LLOQ is 0.5 ng/mg. The CVs in both LOD and LLOQ are all below 20%. As depicted in Table 19, all compounds show very good correlation coefficients, being always higher than 0.97.

**Table 19.** Validation figures obtained with Kinetex column and iB source: linearity curve, LOD and correlation coefficient.

Synthetic cannabinoid	LOD ng/mg	Linearity curve (0.3-3 ng/mg)	Correlation coefficient (r <sup>2</sup> )
(R)-5F-ADB	0.0625	$y = 1.35^8(\pm 4.5^6)x - 1.09^7(\pm 7.6^6)$	R <sup>2</sup> = 0.997
5F- PB 22	0.0625	$y = 4.51^8(\pm 6.7^7)x - 6.83^6(\pm 6.7^5)$	R <sup>2</sup> = 0.977
5CI- AB-PINACA	0.125	$y = 4.13^7(\pm 1.3^6)x + 7.28^5(\pm 2.3^4)$	R <sup>2</sup> = 0.997
5F-AKB-48	0.0625	$y = 1.51^8(\pm 9.7^6)x - 1.65^7(\pm 1.6^6)$	R <sup>2</sup> = 0.988
5F-APP-PICA (PX-1)	0.125	$y = 1.01^8(\pm 5.3^6)x - 1.6^7(\pm 8.9^6)$	R <sup>2</sup> = 0.992
5F-APP-PINACA (PX-2)	0.125	$y = 1.00^8(\pm 2.8^6)x - 9.08^6(\pm 4.7^6)$	R <sup>2</sup> = 0.998
5F-CUMyL-PINACA	0.0625	$y = 2.60^8(\pm 2.9^7)x - 1.50^7(\pm 4.9^6)$	R <sup>2</sup> = 0.970
AB-CHMINACA	0.125	$y = 6.76^7(\pm 3.6^6) - 4.92^6(\pm 6.1^5)$	R <sup>2</sup> = 0.992
AB-FUBINACA	0.0625	$y = 6.75^7((\pm 6.5^5)x + 1.67^6(\pm 1.1^6)$	R <sup>2</sup> = 0.999
ADB-FUBINACA	0.0625	$y = 5.90^7(\pm 6.2^5)x - 4.31^6(\pm 1.0^5)$	R <sup>2</sup> = 0.999
MDMB-CHMICA	0.0625	$y = 2.73^8(\pm 2.3^7)x - 4.0^7(\pm 3.9^6)$	R <sup>2</sup> = 0.979
MMB-2201	0.0625	$y = 3.21^8(\pm 2.5^7)x - 2.8^7(\pm 4.3^6)$	R <sup>2</sup> = 0.982
UR-144	0.0625	$y = 2.43^8(\pm 1.9^7)x - 3.0^7(\pm 3.2^6)$	R <sup>2</sup> = 0.982

### 3.3.4 PRECISION AND BIAS

Precision and bias results are summarised in Table 20 and Table 21. Intra- and inter-day precision values for all synthetic cannabinoids are below 25%. Bias values for all analytes are below 20%.

**Table 20.** Summary of intra- and inter-day studies.

Synthetic cannabinoid	Intra-day precision (%RSD)														
	2 ng/mg (n=6)					1 ng/mg (n=6)					0.5 ng/mg (n=6)				
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 1	Day 2	Day 3	Day 4	Day 5	Day 1	Day 2	Day 3	Day 4	Day 5
R-5F-ADB	20.1	19.3	23.6	14.6	20.4	16.4	15.9	21.9	17.3	15.5	22.0	21.6	20.3	13.9	13.2
5F- PB-22	7.1	15.8	12.7	15.6	17.5	13.9	9.7	15.2	6.7	11.3	18.3	15.1	11.6	16.6	20.5
5CI- AB-PINACA	12.6	16.0	13.3	12.6	15.8	5.1	12.6	18.5	4.5	9.9	13.2	19.4	12.3	23.3	19.4
5F-AKB-48	15.5	21.3	21.7	15.4	3.3	13.7	11.4	15.7	16.3	10.1	22.2	24.4	12.3	16.5	24.3
5F-APP-PICA (PX-1)	3.0	14.3	13.6	14.8	19.8	13.8	15.2	11.3	11.7	20.4	10.8	6.5	11.1	5.3	3.0
5F-APP-PINACA (PX-2)	14.6	4.2	6.6	25.0	19.9	7.8	15.6	14.1	19.3	17.9	8.4	21.4	19.2	23.0	21.3
5F-CUMYL-PINACA	22.5	16.2	20.0	21.6	11.9	12.6	13.4	17.1	20.6	12.3	9.9	19.1	20.2	15.2	22.2
AB-CHMINACA	17.8	19.9	23.5	21.5	14.0	12.1	18.3	11.9	12.2	18.1	12.5	20.5	20.9	21.8	19.3
AB-FUBINACA	5.3	12.2	8.7	16.6	11.4	20.7	5.4	12.3	17.1	18.5	24.6	16.4	18.1	22.4	18.5
ADB-FUBINACA	13.0	18.9	20.4	12.3	19.4	4.1	21.3	15.0	24.0	10.8	21.5	12.7	22.6	23.7	17.4
MDMB-CHMICA	18.2	15.4	20.3	18.8	21.1	14.3	23.3	22.9	13.6	9.6	9.7	20.3	20.3	9.1	32.2
MMB2201	11.0	10.4	8.1	18.4	16.7	21.7	17.0	17.7	9.9	9.9	19.6	11.5	10.4	5.3	44.1
UR-144	14.0	21.2	13.0	22.2	10.9	16.6	19.3	15.5	16.6	12.2	19.1	17.0	23.5	21.0	15.1
d <sub>5</sub> -Diazepam	14.0	11.6	9.9	9.5	8.6	14.0	15.4	8.0	47.5	9.2	17.8	11.1	22.2	5.6	14.0

Synthetic cannabinoid	Inter-day (5 days)		
	2 ng/mg	1 ng/mg	0.5 ng/mg
R-5F-ADB	16.0	18.5	11.7
5F- PB-22	3.0	11.9	12.7
5CI- AB-PINACA	23.4	24.4	13.8
5F-AKB-48	11.2	3.6	13.5
5F-APP-PICA (PX-1)	7.1	7.8	4.2
5F-APP-PINACA (PX-2)	5.1	19.6	32.6
5F-CUMYL-PINACA	20.7	15.0	2.6
AB-CHMINACA	11.1	23.2	19.1
AB-FUBINACA	1.1	13.7	9.6
ADB-FUBINACA	13.4	13.2	12.7
MDMB-CHMICA	17.1	21.1	11.7
MMB2201	6.9	12.5	3.3
UR-144	13.3	16.9	7.5
d <sub>5</sub> -Diazepam	10.7	9.0	8.2

**Table 21.** Bias results for the three validation concentrations (n=6).

Synthetic cannabinoid	Bias (%)		
	2 ng/mg	1 ng/mg	0.5 ng/mg
R-5F-ADB	9.4	-7.9	16.3
5F- PB-22	-20.8	-18.9	-19.9
5CI- AB-PINACA	-16.5	16.0	19.6
5F-AKB-48	-8.9	-9.3	-17.7
5F-APP-PICA (PX-1)	8.0	13.3	19.4
5F-APP-PINACA (PX-2)	5.0	19.8	9.0
5F-CUMYL-PINACA	-19.9	-0.5	18.6
AB-CHMINACA	13.1	7.5	18.2
AB-FUBINACA	19.3	17.9	5.2
ADB-FUBINACA	6.5	20.0	18.0
MDMB-CHMICA	-19.7	-5.4	18.2
MMB2201	19.8	13.6	17.3
UR-144	-2.6	15.6	19.4

### 3.3.5 MATRIX EFFECT

Matrix effect (ME), process efficiency (PE) and recovery (R) results are summarised in Table 22. Matrix effects at 2 ng/mg ranged from 25 to 56%; at 1 ng/mg from 16 to 50%; and at the lowest point (0.5 ng/mg) from 14 to 45%. Process efficiency at the highest point (2 ng/mg) ranged from 19 to 53%, at 1 ng/mg from 12 to 54%, and at the lowest point from 12 to 44%. Finally, recoveries at all concentrations were very good except for 5CI-AB-PINACA at 0.5 ng/mg, which might be due to sampling error at said concentration.

**Table 22.** Matrix effect, process efficiency, and recovery at three different concentrations: 2 ng/mg (high), 1 ng/mg (medium), and 0.5 ng/mg (low).

Synthetic cannabinoid	2 ng/mg			1 ng/mg			0.5 ng/mg		
	Matrix effect (%)	Process efficiency (%)	Recovery (%)	Matrix effect (%)	Process efficiency (%)	Recovery (%)	Matrix effect (%)	Process efficiency (%)	Recovery (%)
(R)-5F-ADB	53.7	53.1	99.0	16.0	12.2	75.9	37.0	33.9	91.6
5F-PB 22	25.2	19.4	77.1	17.7	14.5	81.9	14.3	13.7	96.3
5CI-AB-PINACA	44.3	27.6	62.2	40.1	37.3	93.2	44.5	11.8	26.5
5F-AKB48	40.7	36.4	89.6	14.2	16.9	118.8	14.6	12.6	86.6
5F-APP-PICA (PX-1)	48.4	33.4	69.1	42.3	42.4	100.3	32.3	34.1	105.6
5F-APP-PINACA (PX-2)	46.0	33.7	73.4	38.2	44.3	116.1	41.6	38.2	91.9
5F-CUMYL-PINACA	40.2	26.5	66.0	31.2	28.2	90.5	21.4	21.2	99.2
AB-CHMINACA	46.5	45.6	98.2	30.1	33.8	112.4	30.6	29.5	96.4
AB-FUBINACA	56.1	51.1	91.1	49.6	50.2	101.2	39.0	33.3	85.5
ADB-FUBINACA	39.1	31.7	81.1	26.5	28.8	108.8	27.0	40.1	148.8
MDMB-CHMICA	36.3	33.2	91.5	29.0	27.3	94.2	20.8	12.7	60.9
MMB-2201	34.0	28.7	84.4	32.1	31.0	96.6	45.4	44.7	98.5
UR-144	32.1	19.1	59.4	34.1	54.9	160.8	35.9	20.3	56.6

### 3.3.6 AUTHENTIC HAIR SAMPLES

Real hair samples from patients that have tested positive for other illegal drugs (Table 23) were analysed. No information was available regarding the intake of SC from the patients. However, all samples tested negative for SCs.

**Table 23.** Real hair samples obtained from 11 patients that have been tested for illegal drugs.

Hair sample	Positive for
1	Cocaine, Benzodiazepine
2	Cocaine, Benzodiazepine, morphine
3	Cocaine, Benzodiazepine
4	Cocaine, Benzodiazepine, Codeine
5	Cocaine, Benzodiazepine
6	Cocaine, Benzodiazepine, methadone
7	Cocaine, Benzodiazepine
8	Cocaine, Benzodiazepine
9	Cocaine, Benzodiazepine
10	Cocaine, Benzodiazepine, morphine, methadone
11	Cocaine, Benzodiazepine, morphine, methadone

### 3.4 CONCLUSION

The developed LC–IonTrap methodology proved suitable for the screening of 13 second-wave synthetic cannabinoids: (R)-5F-ADB, 5F-PB 22, 5Cl-AB-PINACA, 5F-AKB48, 5F-APP-PICA (PX-1), 5F-APP-PINACA (PX-2), 5F-CUMYL-PINACA, AB-CHMINACA, AB-FUBINACA, ADB-FUBINACA, MDMB-CHMICA, MMB-2201, and UR-144 in human hair samples. The method has a simple and effective sample treatment and presents good selectivity, linearity precision, and bias. It is high throughput and can be easily updated with the appearance of new synthetic cannabinoids in the market.

## 4. ABBREVIATIONS

Abbreviations are not given in alphabetical order but by order of appearance.

<b>IR</b>	Infrared
<b>NMR</b>	Nuclear Magnetic Resonance
<b>GSR</b>	Gun Shot Residue
<b>SEM-EDS</b>	Scanning Electron Microscope with energy dispersive X-ray spectroscopy
<b>GC-MS/MS</b>	Gas Chromatography coupled with mass detection
<b>SERS</b>	Surface-Enhanced Raman Spectroscopy
<b>IMS</b>	Ion Mobility Spectrometry
<b>DNA</b>	Deoxyribonucleic acid
<b>NPS</b>	New Psychoactive Substances
<b>CE</b>	Capillary Electrophoresis
<b>PMI/PMIs</b>	Post-mortem interval/ Post-mortem intervals
<b>µm</b>	Micrometers
<b>ηL</b>	Nanoliters
<b>EOF</b>	Electroosmotic Flow
<b>CZE</b>	Capillary zone electrophoresis
<b>MEKC</b>	Micellar electrokinetic chromatography
<b>CGE</b>	Capillary gel electrophoresis
<b>CITP</b>	Capillary isotachopheresis
<b>CEC</b>	Capillary electrochromatography
<b>CDT</b>	Carbohydrate-Deficient Transferrin
<b>kNN</b>	k-nearest neighbour
<b>PLS-DA</b>	Partial-least-squares discriminant analysis
<b>ANN</b>	Artificial Neural Networks
<b>MLP</b>	Multilayer perceptron
<b>SGD</b>	Stochastic Gradient Descent
<b>Adam</b>	Adam: Stochastic Gradient Descent-based optimiser,
<b>L-BFGS</b>	Broyden–Fletcher–Goldfarb–Shanno
<b>IC</b>	Ion Chromatography
<b>HIBA</b>	α-hydroxybutyric acid
<b>UV</b>	Ultra-violet
<b>mM</b>	Millimolar
<b>V/cm</b>	Volts per centimetre
<b>°C</b>	Degrees centigrade
<b>µA</b>	Microamperes

<b>kV</b>	Kilovolts
<b>nm</b>	Nanometres
<b>s</b>	Seconds
<b>psi</b>	Pounds per square inch
<b>M</b>	Molar
<b>NaOH</b>	Sodium hydroxide
<b>min</b>	Minutes
<b>IS</b>	Internal standard
<b>µg/mL</b>	Micrograms per millilitre
<b>RSD</b>	Relative Standard Deviation
<b>SWGTOX</b>	Scientific Working Group for Forensic Toxicology
<b>LOD</b>	Limit of detection
<b>LOQ</b>	Limit of quantification
<b>SD</b>	Standard Deviation
<b>MAE</b>	Mean Absolute Error
<b>LC</b>	Liquid chromatography
<b>UV-Vis</b>	Ultraviolet-visible absorption
<b>DAD</b>	Diode Array
<b>FRTIR</b>	Fourier Transform Infrared
<b>MS</b>	Mass Spectrometry
<b>UPLC</b>	Ultra-Performance Liquid Chromatography
<b><i>m/z</i></b>	Mass-to-charge ratio
<b>ESI</b>	Electrospray Ionisation
<b>APPI</b>	Atmospheric Pressure Photoionisation
<b>APCI</b>	Atmospheric Pressure Chemical Photoionisation
<b>iB</b>	IonBooster™
<b>EI</b>	Electrical Ionisation
<b>CI</b>	Chemical Ionisation
<b>SCAN</b>	Scan mode
<b>SIM</b>	Selected Ion Monitoring
<b>amu</b>	Atomic mass units
<b>Q</b>	Quadrupole
<b>IT</b>	Ion Trap
<b>QQQ</b>	Triple quadrupole
<b>QTRAP</b>	Quadrupole Ion Trap
<b>TOF</b>	Time of flight
<b>QTOF</b>	Quadrupole TOF

<b>AC</b>	Alternating current
<b>DC</b>	Direct current
<b>RF</b>	Radio Frequency
<b>LC-MS/MS</b>	Liquid chromatography coupled with tandem mass spectrometric detection
<b>SCs or SC</b>	Synthetic cannabinoids
<b>CB<sub>1</sub> and CB<sub>2</sub></b>	Cannabinoid receptor type 1 and type 2
<b>THC</b>	Tetrahydrocannabinol
<b>EUEWS</b>	European Union Early Warning System
<b>EMCDDA</b>	European Monitoring Center for Drugs and Drug Addiction
<b>HLMs</b>	Human liver microsomes
<b>LC-QTOF-MS/MS</b>	Liquid chromatography–quadrupole time-of-flight tandem mass spectrometry
<b>μL</b>	Microlitres
<b>mL</b>	Millilitres
<b>v/v</b>	Volume per volume
<b>V</b>	Volts
<b>mL/min</b>	Millilitres per minute
<b>eV</b>	Electronvolts
<b>h</b>	Hours
<b>UHPLC</b>	Ultra-high-performance liquid chromatography
<b>mg</b>	Milligrams
<b>ng/mg</b>	Nanograms per milligram

## 5. REFERENCES

1. Li, S. F. Y. *Capillary Electrophoresis: Principles, Practice and Applications*. (Elsevier, 1992).
2. Altria, K. D. *Capillary Electrophoresis Guidebook. Principles, Operation and Applications*. 52, (Humana Press, 1996).
3. Shintani, H. and Polonský, J. *Handbook of Capillary Electrophoresis Applications. Handbook of Capillary Electrophoresis Applications* (Chapman & Hall, 1997). doi:10.1007/978-94-009-1561-9
4. Airado-Rodríguez, D., Cruces-Blanco, C. and García-Campaña, A. M. Ultrasensitive analysis of lysergic acid diethylamide and its C-8 isomer in hair by capillary zone electrophoresis in combination with a stacking technique and laser induced fluorescence detection. *Anal. Chim. Acta* 866, 90–98 (2015).
5. Porpiglia, N., Musile, G., Bortolotti, F., De Palo, E. F. and Tagliaro, F. Chiral separation and determination of ketamine and norketamine in hair by capillary electrophoresis. *Forensic Sci. Int.* 266, 304–310 (2016).
6. Evans, E., Costrino, C., do Lago, C. L., Garcia, C. D., Roux, C. and Blanes, L. Determination of Inorganic Ion Profiles of Illicit Drugs by Capillary Electrophoresis. *J. Forensic Sci.* 61, 1610–1614 (2016).
7. Gottardo, R., Poletini, A., Sorio, D., Pascali, J. P., Bortolotti, F., Liottra, E. and Tagliaro, F. Capillary zone electrophoresis (CZE) coupled to time-of-flight mass spectrometry (TOF-MS) applied to the analysis of illicit and controlled drugs in blood. *Electrophoresis* 29, 4078–4087 (2008).
8. Northrop, D. M., Martire, D. E. and MacCrehan, W. A. Separation and Identification of Organic Gunshot and Explosive Constituents by Micellar Electrokinetic Capillary Electrophoresis. *Anal. Chem.* 63, 1038–1042 (1991).
9. Erol, Ö., Erdoğan, B. Y. and Onar, A. N. Nitrate and Nitrite Determination in Gunshot Residue Samples by Capillary Electrophoresis in Acidic Run Buffer. *J. Forensic Sci.* 62, 423–427 (2017).
10. Tagliaro, F., Manetto, G., Cittadini, F., Marchetti, D., Bortolotti, F. and Marigo, M. Capillary zone electrophoresis of potassium in human vitreous humour: Validation of a new method. *J. Chromatogr. B Biomed. Sci. Appl.* 733, 273–279 (1999).
11. Porpiglia, N. M., De Palo, E. F., Savchuk, S. A., Appolonova, S. A., Bortolotti, F. and Tagliaro, F. A new sample treatment for asialo-Tf determination with capillary electrophoresis: an added value to the analysis of CDT. *Clin. Chim. Acta* 483, 256–262 (2018).



12. Musile, G., De Palo, E. F., Savchuk, S. A., Shestakova, K., Bortolotti, F. and Tagliaro, F. A novel low-cost approach for the semi-quantitative analysis of carbohydrate-deficient transferrin (CDT) based on fluorescence resonance energy transfer (FRET). *Clin. Chim. Acta* 495, 556–561 (2019).
13. Porpiglia, N. M., Savchuk, S. A., Appolonova, S. A., Bortolotti, F. and Tagliaro, F. Capillary Electrophoresis (CE) vs. HPLC in the determination of asialo-Tf, a crucial marker for the reliable interpretation of questioned CDT increases. *Clin. Chim. Acta* 486, 49–53 (2018).
14. Bertaso, A., Sorio, D., Vadoros, A., De Palo, E. F., Bortolotti, F. and Tagliaro, F. Use of finger-prick dried blood spots (fpDBS) and capillary electrophoresis for carbohydrate deficient transferrin (CDT) screening in forensic toxicology. *Electrophoresis* 37, 2867–2874 (2016).
15. Cruces-Blanco, C., Gámiz-Gracia, L. and García-Campaña, A. M. Applications of capillary electrophoresis in forensic analytical chemistry. *TrAC - Trends Anal. Chem.* 26, 215–226 (2007).
16. Hauser, F. M., Hulshof, J. W., Rößler, T., Zimmermann, R. and Pütz, M. Characterisation of aqueous waste produced during the clandestine production of amphetamine following the Leuckart route utilising solid-phase extraction gas chromatography–mass spectrometry and capillary electrophoresis with contactless conductivity dete. *Drug Test. Anal.* 10, 1368–1382 (2018).
17. Zlotnick, J. A. and Smith, F. P. Chromatographic and electrophoretic approaches in ink analysis. *J. Chromatogr. B Biomed. Sci. Appl.* 733, 265–272 (1999).
18. Pascali, J. P., Bortolotti, F. and Tagliaro, F. Recent advances in the application of CE to forensic sciences, an update over years 2009-2011. *Electrophoresis* 33, 117–126 (2012).
19. Kumar, R. and Sharma, V. Chemometrics in forensic science. *TrAC - Trends Anal. Chem.* 105, 191–201 (2018).
20. Miller, J. M. and Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry. Technometrics* (2010). doi:10.1198/tech.2004.s248
21. Bailer-Jones, C. A. L., Gupta, R. and Singh, H. P. An introduction to artificial neural networks. 1–18 (2001).
22. Basheer, I. . and Hajmeer, M. Artificial neural networks: fundamentals, computing, design, and application. *J. Microbiol. Methods* 43, 3–31 (2000).
23. Bocaz-Beneventi, G., Tagliaro, F., Bortolotti, F., Manetto, G. and Havel, J. Capillary zone electrophoresis and artificial neural networks for estimation of the post-mortem interval (PMI) using electrolytes measurements in human vitreous humour. *Int. J. Legal Med.* 116, 5–11

(2002).

24. Hunt, V. J., Helmicki, A. J. and Aktan, A. E. *Intro to Machine Learning with Python. Proceedings of the Speciality Conference on Infrastructure Condition Assessment: Art, Science, Practice* (1997).
25. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. of Machine Learn. Res.* 12, 2825–2830 (2011).
26. Kingma, D. P. and Ba, J. L. Adam: A method for stochastic optimization. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* 1–15 (2015).
27. Liu, C. S. Optimal algorithms and the bfgs updating techniques for solving unconstrained nonlinear minimization problems. *J. Appl. Math.* 2014, 14 (2014).
28. IBM. SPSS Neural Networks. Available at: <https://www.ibm.com/analytics/spss-statistics-software>
29. Ruder, S. An overview of gradient descent optimization algorithms. arXiv (Cornell Univ. (2017)).
30. Reference Python Language. Python Software Foundation. Version 3.8. (2020). Available at: <http://www.python.org>.
31. Butcher, J. B., Moore, H. E., Day, C. R., Adam, C. D. and Drijfhout, F. P. Artificial neural network analysis of hydrocarbon profiles for the ageing of *Lucilia sericata* for post mortem interval estimation. *Forensic Sci. Int.* 232, 25–31 (2013).
32. Broséus, J., Vallat, M. and Esseiva, P. Multi-class differentiation of cannabis seedlings in a forensic context. *Chemom. Intell. Lab. Syst.* 107, 343–350 (2011).
33. Macarulla Rodriguez, A., Geradts, Z. and Worring, M. Likelihood Ratios for Deep Neural Networks in Face Comparison. *J. Forensic Sci.* 65, 1169–1183 (2020).
34. Del Espiritu, J., Rolluqui, G. and Gustilo, R. C. Neural network based partial fingerprint recognition as support for forensics. *8th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag. HNICEM 2015* 4–8 (2016). doi:10.1109/HNICEM.2015.7393227
35. Grantham, N. S., Reich, B. J., Laber, E. B., Pacifici, K., Dunn, R. R., Fierer, N., Gebert, M., Allwood, J. S. and Faith, S. A. Global forensic geolocation with deep neural networks. *J. R. Stat. Soc. Ser. C Appl. Stat.* 909–929 (2020). doi:10.1111/rssc.12427

36. Hemalatha, B. and Rajkumar, N. A versatile approach for dental age estimation using fuzzy neural network with teaching learning - based optimization classification. *Multimed. Tools Appl.* 79, 3645–3665 (2020).
37. Prescher, A., Meyers, A. and von Keyserlingk, D. G. Neural net applied to anthropological material: A methodical study on the human nasal skeleton. *Ann. Anat.* 187, 261–269 (2005).
38. Zong, R., Zhi, Y., Yao, B., Gao, J. and Stec, A. A. Classification and identification of soot source with principal component analysis and back-propagation neural network. *Aust. J. Forensic Sci.* 46, 224–233 (2014).
39. Casamento, S., Kwok, B., Roux, C., Dawson, M. and Doble, P. Optimization of the Separation of Organic Explosives by Capillary Electrophoresis with Artificial Neural Networks. *J. Forensic Sci.* 48, (2003).
40. Pekka Sauco and Bernard Knight. *Knight's Forensic Pathology*. (2004).
41. Salam, H. A., Shaat, E. A., Aziz, M. H. A., MoneimSheta, A. A. and Hussein, H. A. S. M. Estimation of postmortem interval using thanatochemistry and postmortem changes. *Alexandria J. Med.* 48, 335–344 (2012).
42. Thierauf, A., Musshoff, F. and Madea, B. Post-mortem biochemical investigations of vitreous humor. *Forensic Sci. Int.* 192, 78–82 (2009).
43. Balasooriya, B. A. W., Hill, C. A. S. and Williams, A. R. The biochemistry of vitreous humour. A comparative study of the potassium, sodium and urate concentrations in the eyes at identical time intervals after death. *Forensic Sci. Int.* 26, 85–91 (1984).
44. Coe, J. I. Vitreous potassium as a measure of the postmortem interval: An historical review and critical evaluation. *Forensic Sci. Int.* 42, 201–213 (1989).
45. Hubig, M., Muggenthaler, H., Sinicina, I. and Mall, G. Temperature based forensic death time estimation: The standard model in experimental test. *Leg. Med.* 17, 381–387 (2015).
46. Madea, B. *Estimation of the Time Since Death*. (CRC Press, 2015).
47. Zhou, B., Zhang, L., Zhang, G., Zhang, X. and Jiang, X. The determination of potassium concentration in vitreous humor by low pressure ion chromatography and its application in the estimation of postmortem interval. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 852, 278–281 (2007).
48. Tagliaro, F., Bortolotti, F., Manetto, G., Cittadini, F., Pascali, V. L. and Marigo, M. Potassium concentration differences in the vitreous humour from the two eyes revisited by microanalysis with capillary electrophoresis. *J. Chromatogr. A* 924, 493–498 (2001).

49. Madea, B. Is there recent progress in the estimation of the postmortem interval by means of thanatochemistry? *Forensic Sci. Int.* 151, 139–149 (2005).
50. Hayman, J., Oxenham, M., Hayman, J. and Oxenham, M. *Algor Mortis and Temperature-Based Methods of Estimating the Time Since Death. Human Body Decomposition* (2016). doi:10.1016/B978-0-12-803691-4.00002-9
51. Nikolac, N., Omazic, J. and Simundic, A. M. The evidence based practice for optimal sample quality for ammonia measurement. *Clin. Biochem.* 47, 991–995 (2014).
52. van den Oever, R. Post-mortem vitreous ammonium concentrations in estimating the time of death. *Zeitschrift für Rechtsmedizin* 80, 259–263 (1978).
53. Magnusson, B. and Örnemark, U. The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics. *Eurachem Guid.* 1–70 (2014). doi:978-91-87461-59-0
54. Scientific Working Group for Forensic Toxicology. Standard practices for method validation in Forensic Toxicology. *SWGTOX Guidel.* 1–52 (2013). doi:10.1093/jat/bkt054
55. Taverniers, I., De Loose, M. and Van Bockstaele, E. Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance. *TrAC - Trends Anal. Chem.* 23, 535–552 (2004).
56. Oehrle, S. A. Controlled changes in selectivity of cation separations by capillary electrophoresis using various crown-ether additives. *J. Chromatogr. A* 745, 87–92 (1996).
57. Shi, Y. and Fritz, J. S. New electrolyte systems for the determination of metal cations by capillary zone electrophoresis. *J. Chromatogr. A* 671, 429–435 (1994).
58. Francois, C., Morin, P. and Dreux, M. Effect of the concentration of 18-crown-6 added to the electrolyte upon the separation of ammonium, alkali and alkaline-earth cations by capillary electrophoresis. *J. Chromatogr. A* 706, 535–553 (1995).
59. Yang, Q., Zhuang, Y., Smeyers-Verbeke, J. and Massart, D. L. Interpretation of migration behaviour of inorganic cations in capillary ion electrophoresis based on an equilibrium model. *J. Chromatogr. A* 706, 503–515 (1995).
60. Distribution Anaconda Software. Anaconda Navigator. (2016). Available at: <https://www.anaconda.com/>.
61. Schuff-Werner, P. and Steiner, M. Preanalytical ammonia generation: A race with time but not with temperature. *J. Clin. Pathol.* 68, 757–758 (2015).
62. F. Bortolotti, R. Gottardo, M. Trettene, F. Cittadini, F. Tagliaro, M. M. Increased postmortem

concentrations of K<sup>+</sup> in the vitreous humour in heroin overdose deaths. in *Annales de Toxicologie Analytique*, Vol. XIV, n°3 306 (2002).

63. Machine Learning Mastery. Repeated k-Fold Cross-Validation for Model Evaluation in Python. Available at: <https://machinelearningmastery.com/repeated-k-fold-cross-validation-with-python/>. (Accessed: 19th October 2020)
64. Machine Learning Mastery. LOOCV for Evaluating Machine Learning Algorithms. Available at: [https://machinelearningmastery.com/loocv-for-evaluating-machine-learning-algorithms/?fbclid=IwAR2i33cTXNSVlo1un02hVVqoaPFM\\_z0a-lfk4ycMFjCAKLGJ-jMot\\_x29HA](https://machinelearningmastery.com/loocv-for-evaluating-machine-learning-algorithms/?fbclid=IwAR2i33cTXNSVlo1un02hVVqoaPFM_z0a-lfk4ycMFjCAKLGJ-jMot_x29HA). (Accessed: 19th October 2020)
65. Sutskever, I., Martens, J., Dahl, G. and Hinton, G. On the importance of initialization and momentum in deep learning. *30th Int. Conf. Mach. Learn. ICML 2013* 2176–2184 (2013).
66. Skoog, D., Holler, J. and Crouch, S. R. *Principles of Instrumental Analysis*. (2007). doi:10.1016/S0003-2670(00)84936-3
67. Niessn, W. M. . and Tinkeb, A. . Liquid chromatography-mass spectrometry general principles and instrumentation. *J. Chromatogr. A* 703, 37–57 (1995).
68. Pragst, F. *Chapter 13 High performance liquid chromatography in forensic toxicological analysis. Handbook of Analytical Separations 6*, (Elsevier B.V., 2008).
69. Gross, J. H. *Mass Spectrometry*. (Springer International, 2017). doi:10.1007/978-3-319-54398-7
70. Taleuzzaman, M. and Ali. Ultra performance liquid chromatography (UPLC): A review. *Austin J. Anal. Pharm. Chem.* 2, 1–5 (2015).
71. Poletini, A. *Applications of LC-MS in Toxicology*. (Pharmaceutical Press, 2006).
72. Van Bocxlaer, J. F., Clauwaert, K. M., Lambert, W. E., Deforce, D. L., Van Den Eeckhout, E. G. and De Leenheer, A. P. Liquid chromatography-mass spectrometry in forensic toxicology. *Mass Spectrom. Rev.* 19, 165–214 (2000).
73. Bruker Daltonics. Atmospheric Pressure Ionization: Fundamentals of IonBooster, APCI, DIP and APPI. 1–25 (2016).
74. Bruker Daltonics. Introduction to Ion Trap Mass Spectrometry Toxyper Key Components. 1–30 (2017).
75. Bruker Daltonics. LC-MS, Ion-sources, IonBooster. Available at: <https://www.bruker.com/products/mass-spectrometry-and-separations/lc-ms/ion-sources/ionbooster/overview.html>.

76. Bravo, S. R. Advanced Instrumental Analysis: Chromatography-Mass Spectrometry (Powerpoint). (2012).
77. Chernushevich, I. V., Loboda, A. V. and Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom.* 36, 849–865 (2001).
78. Medhe, S. Mass Spectrometry: Detectors Review. *Chem. Biomol. Eng.* 3, 51–58 (2018).
79. Madea, B. *Handbook of Forensic Medicine. Handbook of Forensic Medicine* (2014). doi:10.1002/9781118570654
80. Gottardo, R., Murari, M., Bertaso, A., Bortolotti, F. and Tagliaro, F. Drug screening by using the Toxtyper™ LC-ion trap MSn: optimization of its application on serum samples in a DUID context. *Clin. Chim. Acta* 510, 537–543 (2020).
81. Bezemer, K., McLennan, L., van Duin, L., Kuijpers, C. J., Koeberg, M., van den Elshout, J., van der Heijden, A., Busby, T., Yevdokimov, A., Schoenmakers, P., Smith, J., Oxley, J. and van Asten, A. Chemical attribution of the home-made explosive ETN – Part I: Liquid chromatography-mass spectrometry analysis of partially nitrated erythritol impurities. *Forensic Sci. Int.* 307, 110102 (2020).
82. Ostrinskaya, A., Kunz, R. R., Clark, M., Kingsborough, R. P., Ong, T. H. and Deneault, S. Rapid Quantitative Analysis of Multiple Explosive Compound Classes on a Single Instrument via Flow-Injection Analysis Tandem Mass Spectrometry. *J. Forensic Sci.* 64, 223–230 (2019).
83. Gassner, A. L. and Weyermann, C. Prevalence of organic gunshot residues in police vehicles. *Sci. Justice* 60, 136–144 (2020).
84. Centazzo, N., Frederick, B. M., Jacox, A., Cheng, S. Y. and Concheiro-Guisan, M. Wastewater analysis for nicotine, cocaine, amphetamines, opioids and cannabis in New York City. *Forensic Sci. Res.* 4, 152–167 (2019).
85. Wood, M., Laloup, M., Samyn, N., Ramirez, M., Bruijn, E. A. De, Maes, R. A. A. and Boeck, G. De. Recent applications of liquid chromatography – mass spectrometry in forensic science. *J. Chromatogr. A* 1130, 3–15 (2006).
86. Liechti, M. E. Novel psychoactive substances (designer drugs): overview and pharmacology of modulators of monoamine signalling. *Swiss Med Wkly* 145, 1–12 (2015).
87. Holm, N. B., Pedersen, A. J., Dalsgaard, P. W. and Linnet, K. Metabolites of 5F-AKB-48, a synthetic cannabinoid receptor agonist, identified in human urine and liver microsomal preparations using liquid chromatography high-resolution mass spectrometry. *Drug Test. Anal.* 7, 199–206 (2015).

88. Savchuk, S., Appolonova, S., Pechnikov, A., Rizvanova, L., Shestakova, K. and Tagliaro, F. In vivo metabolism of the new synthetic cannabinoid APINAC in rats by GC–MS and LC–QTOF-MS. *Forensic Toxicol.* **35**, 359–368 (2017).
89. Dresen, S., Ferreirós, N., Pütz, M., Westphal, F., Zimmermann, R. and Auwärter, V. Monitoring of herbal mixtures potentially containing synthetic cannabinoids as psychoactive compounds. *J. Mass Spectrom.* **45**, 1186–1194 (2010).
90. Banister, S. D., Stuart, J., Kevin, R. C., Edington, A., Longworth, M., Wilkinson, S. M., Beinat, C., Buchanan, A. S., Hibbs, D. E., Glass, M., Connor, M., McGregor, I. S. and Kassiou, M. Effects of Bioisosteric Fluorine in Synthetic Cannabinoid Designer Drugs JWH-018, AM-2201, UR-144, XLR-11, PB-22, 5F-PB-22, APICA, and STS-135. *ACS Chem. Neurosci.* **6**, 1445–1458 (2015).
91. Znaleziona, J., Ginterová, P., Petr, J., Ondra, P., Válka, I., Ševčík, J., Chrastina, J. and Maier, V. Determination and identification of synthetic cannabinoids and their metabolites in different matrices by modern analytical techniques - a review. *Anal. Chim. Acta* **874**, 11–25 (2015).
92. European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). *European Drug Report. Trends and Developments.* (2015).
93. Zawilska, J. B. and Andrzejczak, D. Next generation of novel psychoactive substances on the horizon—a complex problem to face. *Drug Alcohol Depend.* **157**, 1–17 (2015).
94. Rodrigues, C., Alves, C., Santos-neto, A. J., Fernandes, C. and Lan, F. M. Analysis of tricyclic antidepressant drugs in plasma by means of solid-phase microextraction-liquid chromatography-mass spectrometry. *J. mass Spectrom.* 1342–1347 (2007). doi:10.1002/jms
95. Bijlsma, L., Sancho, J. V., Hernández, F. and Niessen, W. M. A. Fragmentation pathways of drugs of abuse and their metabolites based on QTOF MS/MS and MS E accurate-mass spectra. *J. Mass Spectrom.* **46**, 865–875 (2011).
96. Yamada, H., Ishii, Y. and Oguri, K. Metabolism of drugs of abuse: Its contribution to the toxicity and the inter-individual differences in drug sensitivity. *J. Heal. Sci.* **51**, 1–7 (2005).
97. Musah, R. A., Domin, M. A., Walling, M. A. and Shepard, J. R. E. Rapid identification of synthetic cannabinoids in herbal samples via direct analysis in real time mass spectrometry. *Rapid Commun. Mass Spectrom.* **26**, 1109–1114 (2012).
98. Diao, X. and Huestis, M. A. Approaches, Challenges, and Advances in Metabolism of New Synthetic Cannabinoids and Identification of Optimal Urinary Marker Metabolites. *Clin. Pharmacol. Ther.* **101**, 239–253 (2017).
99. Cohen, J., Morrison, S., Greenberg, J. and Saidinejad, M. Clinical presentation of intoxication

due to synthetic cannabinoids. *Pediatrics* 129, (2012).

100. Derungs, A., Schwaninger, A. E., Mansella, G., Bingisser, R., Kraemer, T. and Liechti, M. E. Symptoms, toxicities, and analytical results for a patient after smoking herbs containing the novel synthetic cannabinoid MAM-2201. *Forensic Toxicol.* 31, 164–171 (2013).
101. Carvalho, M., Carmo, H., Costa, V. M., Capela, J. P., Pontes, H., Remião, F., Carvalho, F. and De Lourdes Bastos, M. Toxicity of amphetamines: An update. *Arch. Toxicol.* 86, 1167–1231 (2012).
102. Hermanns-Clausen, M., Kneisel, S., Szabo, B. and Auwärter, V. Acute toxicity due to the confirmed consumption of synthetic cannabinoids: Clinical and laboratory findings. *Addiction* 108, 534–544 (2013).
103. Takematsu, M., Hoffman, R. S., Nelson, L. S., Schechter, J. M., Moran, J. H. and Wiener, S. W. A case of acute cerebral ischemia following inhalation of a synthetic cannabinoid. *Clin. Toxicol.* 52, 973–975 (2014).
104. Aldigan, A. A. and Torrance, H. J. Bioanalytical methods for the determination of synthetic cannabinoids and metabolites in biological specimens. *TrAC - Trends Anal. Chem.* 80, 444–457 (2016).
105. European Monitoring Centre for Drugs and Drug Adiction. *European Drug Report 2018: Trends and Developments.* (2018). doi:10.2810/88175
106. Lee, J. H., Park, H. N., Leem, T. S., Jeon, J. hyoung, Cho, S., Lee, J. and Baek, S. Y. Identification of new synthetic cannabinoid analogue APINAC (adamantan-1-yl 1-pentyl-1H-indazole-3-carboxylate) with other synthetic cannabinoid MDMB(N)-Bz-F in illegal products. *Forensic Toxicol.* 35, 45–55 (2017).
107. Hwang, J., Hwang, J., Ganganna, B., Song, I., Heo, M. Y., Ahn, S. H. and Lee, J. Metabolic and pharmacokinetic characterization of a new synthetic cannabinoid APINAC in rats. *Forensic Toxicol.* 36, 88–101 (2018).
108. Minakata, K., Hasegawa, K., Nozawa, H., Yamagishi, I., Saitoh, T., Yoshino, A., Suzuki, M., Kitamoto, T., Suzuki, O. and Watanabe, K. Sensitive quantification of BB-22 and its metabolite BB-22 3-carboxyindole, and characterization of new metabolites in authentic urine and/or serum specimens obtained from three individuals by LC–QTRAP-MS/MS and high-resolution LC–Orbitrap-MS/MS. *Forensic Toxicol.* 37, 164–173 (2019).
109. Diao, X., Wohlfarth, A., Pang, S., Scheidweiler, K. B. and Huestis, M. A. High-resolution mass spectrometry for characterizing the metabolism of synthetic cannabinoid THJ-018 and its 5-fluoro analog THJ-2201 after incubation in human hepatocytes. *Clin. Chem.* 62, 157–169



(2016).

110. Vikingsson, S., Josefsson, M. and Gréen, H. Identification of AKB-48 and 5F-AKB-48 metabolites in authentic human urine samples using human liver microsomes and time of flight mass spectrometry. *J. Anal. Toxicol.* 39, 426–435 (2015).
111. Appolonova, S. A., Palacio, C., Shestakova, K. M., Mesonzhnik, N. V., Brito, A., Kuznetsov, R. M., Markin, P. A., Bochkareva, N. L., Burmykin, D., Ovcharov, M., Musile, G., Tagliaro, F. and Savchuk, S. A. In vivo and in vitro metabolism of the novel synthetic cannabinoid 5F-APINAC. *Forensic Toxicol.* 38, 160–171 (2020).
112. European Monitoring Centre for Drugs and Drug Adiction. *Trends and developments.* (2020). doi:10.1093/tandt/ttm111
113. European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). Synthetic cannabinoids in Europe - update 2017. *Perspect. Drugs* 9 (2017).
114. Gottardo, R., Sorio, D., Musile, G., Trapani, E., Seri, C., Serpelloni, G. and Tagliaro, F. Screening for synthetic cannabinoids in hair by using LC-QTOF MS: A new and powerful approach to study the penetration of these new psychoactive substances in the population. *Med. Sci. Law* 54, 22–27 (2014).
115. Goebel, A., Boehm, M., Kirchherr, H. and Nikolaus Kuhn-Velten, W. Simultaneous identification and quantification of synthetic cannabinoids (cannabimimetics) in serum, hair, and urine by rapid and sensitive HPLC tandem mass spectrometry screenings: Overview and experience from routine testing. *LaboratoriumsMedizin* 37, 167–180 (2013).
116. Hutter, M., Kneisel, S., Auwärter, V. and Neukamm, M. A. Determination of 22 synthetic cannabinoids in human hair by liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 903, 95–101 (2012).
117. Saito, T., Sasaki, C., Namera, A., Kurihara, K. and Inokuchi, S. Experimental study on external contamination of hair by synthetic cannabinoids and effect of hair treatment. *Forensic Toxicol.* 33, 155–158 (2015).
118. Franz, F., Jechle, H., Angerer, V., Pegoro, M., Auwärter, V. and Neukamm, M. A. Synthetic cannabinoids in hair – Pragmatic approach for method updates, compound prevalences and concentration ranges in authentic hair samples. *Anal. Chim. Acta* 1006, 61–73 (2018).
119. Ong, R. S., Kappatos, D. C., Russell, S. G. G., Poulsen, H. A., Banister, S. D., Gerona, R. R., Glass, M., Johnson, C. S. and McCarthy, M. J. Simultaneous analysis of 29 synthetic cannabinoids and metabolites, amphetamines, and cannabinoids in human whole blood by liquid chromatography–tandem mass spectrometry – A New Zealand perspective of use in

2018. *Drug Test. Anal.* 12, 195–214 (2020).

120. Huppertz, L. M., Kneisel, S., Auwärter, V. and Kempf, J. A comprehensive library-based, automated screening procedure for 46 synthetic cannabinoids in serum employing liquid chromatography-quadrupole ion trap mass spectrometry with high-temperature electrospray ionization. *J. Mass Spectrom.* 49, 117–127 (2014).
121. Meyer, M., Drury, T. and Kiehne, A. *Comprehensive detection and identification of synthetic cannabinoids using the Toxtyper platform.*
122. Musile, G., Mazzola, M., Shestakova, K., Savchuk, S., Appolonova, S. and Tagliaro, F. A simple and robust method for broad range screening of hair samples for drugs of abuse using a high-throughput UHPLC-Ion Trap MS instrument. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 1152, 122263 (2020).
123. Matuszewski, B. K., Constanzer, M. L. and Chavez-Eng, C. M. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Anal. Chem.* 75, 3019–3030 (2003).

## 6. APPENDICES

### 6.1 APPENDIX I. NH<sub>4</sub><sup>+</sup> IN VITREOUS HUMOUR

#### 6.1.1 LIST OF VITREOUS HUMOUR SAMPLES USED FOR THE STUDY

**Table 24.** List of vitreous humour samples, their electrophoresis data and ion concentrations used in the PMI interval evaluation. PA stands for Peak Area, PH stands for Peak Height.

Code	PA <sub>NH<sub>4</sub><sup>+</sup></sub>	PH <sub>NH<sub>4</sub><sup>+</sup></sub>	[NH <sub>4</sub> <sup>+</sup> ] mM	PA <sub>K<sup>+</sup></sub>	PH <sub>K<sup>+</sup></sub>	[K <sup>+</sup> ] mM	PA <sub>NH<sub>4</sub><sup>+</sup> + PA<sub>K<sup>+</sup></sub></sub>	PH <sub>NH<sub>4</sub><sup>+</sup> + PH<sub>K<sup>+</sup></sub></sub>	PA <sub>IS</sub>	PH <sub>IS</sub>	PMI (hours.mins)
VH 1	1316	533	0.99	13700	6048	15.28	15016	6581	19326	6441	77.1
VH 2	1410	539	1.01	13348	5693	14.03	14758	6232	20492	6762	47
VH 3	1032	364	0.51	11677	4803	12.30	12709	5166	20399	6666	21.5
VH 4	2495	602	1.36	34647	9937	24.47	37142	10539	30697	7615	143
VH 5	1017	245	0.00	18775	5049	11.94	19792	5293	33774	8549	46
VH 6	1915	719	1.49	22530	8433	22.01	24445	9152	22195	6397	87.15
VH 7	929	400	0.03	6169	2800	4.67	7098	3201	27560	7293	17.15
VH 8	2509	595	1.47	24187	9178	17.43	26696	9773	29554	7280	123.44
VH 9	2757	583	1.20	42316	10580	25.11	45073	11164	36540	8362	143
VH 10	727	391	0.39	6520	3631	8.62	7247	4023	16140	5615	32
VH 11	1155	496	0.97	13131	5928	16.47	14286	6424	17202	5775	56.05
VH 12	773	344	0.38	11576	5284	14.95	12348	5628	16686	5443	52
VH 13	2172	788	2.29	23883	9210	27.23	26056	9999	19030	6188	168.5
VH 14	787	291	0.00	8610	3494	7.27	9398	3785	25119	6854	16.1
VH 15	1490	484	0.60	17338	6309	13.56	18828	6794	27512	7262	47.4
VH 16	1780	578	1.49	20349	7574	21.18	22129	8152	20793	6636	100.3
VH 17	1027	308	0.08	19201	6411	14.33	20228	6719	28896	6713	68
VH 18	713	332	0.35	6028	2786	8.04	6741	3118	15855	5122	29
VH 19	3768	1067	3.10	31437	10213	25.95	35205	11280	26275	7481	163.2
VH 20	2299	703	1.38	24644	8483	18.98	26943	9186	28090	7709	96.45
VH 21	2421	684	1.42	18290	7434	17.43	20711	8118	12409	7151	52.3
VH 22	3825	901	2.14	43168	11870	23.66	46993	12771	35130	8535	99.45
VH 23	1643	562	0.82	20230	7825	16.43	21873	8387	26578	7346	98.45
VH 24	1699	594	1.06	22439	8786	20.11	24138	9380	24115	7128	99.16
VH 25	4121	1198	2.98	32412	10812	23.73	36533	12011	29643	7717	144.23
VH 26	1284	428	0.38	14012	5096	10.75	15296	5524	27918	7564	25
VH 27	1087	345	0.14	19104	6643	14.08	20191	6989	29176	7117	46.24
VH 28	908	378	0.04	13229	5610	10.66	14137	5988	26594	6908	23.35
VH 29	654	272	0.00	10899	4687	9.03	11552	4958	25797	6920	17.55
VH 30	1802	693	0.99	22586	9142	18.39	24388	9835	26533	7126	48.4
VH 31	1482	578	0.59	20377	8492	15.86	21859	9070	27709	7276	48.4
VH 32	1377	540	0.56	16614	7181	13.67	17990	7721	26173	6943	50.15
VH 33	1198	409	0.24	22060	8668	16.33	23258	9077	29208	7605	71.15
VH 34	1777	547	1.23	22563	7869	20.58	24340	8416	23416	6243	48.45
VH 35	2015	804	2.27	14241	6392	17.37	16256	7195	17763	5273	38.3
VH 36	416	157	0.00	7517	3023	5.74	7933	3180	27657	7184	6.55

VH 37	1371	564	0.91	18031	8036	18.51	19402	8600	21135	6326	75.5
VH 38	978	446	1.02	14441	7205	17.58	15419	7651	19697	5753	27.35

**Pink** samples were used for the preliminary linear approach together with two more (a total of 14) but those two were later removed for further analysis since they were overdoses. For the logarithmic approach all samples except the **blue** ones (a total of 33) were used. For the neural networks, the same samples as the logarithmic approach and the **blue** ones were used (total of 38).

## 6.1.2 TRAIN TEST SPLIT AND DATA PRE-PROCESSING

```
import matplotlib.pyplot as plt
import pandas as pd
from sklearn import neighbors, datasets, preprocessing
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPRegressor
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error

vhdata = pd.read_csv('vhrawdata.csv')
X = vhdata[['CNH4', 'CK', 'rNA']]
y = vhdata['PMI']
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_train_norm = scaler.fit_transform(X_train)
X_test_norm = scaler.transform(X_test)
```

## 6.1.3 CODE FOR ADAM SOLVER

### 6.1.3.1 NUMBER OF NODES IN THE HIDDEN LAYER

```
for hn in [1,2,3,4,5,6,7,8,9,10]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,hn), activation = 'logistic',
                        solver = 'adam', alpha = 0.1, max_iter=400000,random_state=0)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} nodes: {:.3f}'.format(hn,
    r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with {} nodes:{:.3f}'.format(hn,
    r2_score(y_test,y_test_prediction)))

    print('Mean Absolute error Train (MSE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_train, y_train_prediction)))
    print('Mean Absolute error Test (MAE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_test, y_test_prediction)))
```

```
Training set prediction (R2 score) with 1 nodes: -0.000
Test set prediction (R2 score) with 1 nodes:-0.010
Mean Absolute error Train (MSE) with 1 nodes: 32.114
Mean Absolute error Test (MSE) with 1 nodes: 45.678
Training set prediction (R2 score) with 2 nodes: -0.000
Test set prediction (R2 score) with 2 nodes:-0.010
Mean Absolute error Train (MSE) with 2 nodes: 32.118
Mean Absolute error Test (MSE) with 2 nodes: 45.678
Training set prediction (R2 score) with 3 nodes: -0.000
Test set prediction (R2 score) with 3 nodes:-0.010
```

```

Mean Absolute error Train (MSE) with 3 nodes: 32.123
Mean Absolute error Test (MSE) with 3 nodes: 45.678
Training set prediction (R2 score) with 4 nodes: 0.860
Test set prediction (R2 score) with 4 nodes:0.701
Mean Absolute error Train (MSE) with 4 nodes: 11.424
Mean Absolute error Test (MSE) with 4 nodes: 21.667
Training set prediction (R2 score) with 5 nodes: 0.859
Test set prediction (R2 score) with 5 nodes:0.696
Mean Absolute error Train (MSE) with 5 nodes: 11.458
Mean Absolute error Test (MSE) with 5 nodes: 21.832
Training set prediction (R2 score) with 6 nodes: 0.858
Test set prediction (R2 score) with 6 nodes:0.694
Mean Absolute error Train (MSE) with 6 nodes: 11.483
Mean Absolute error Test (MSE) with 6 nodes: 21.911
Training set prediction (R2 score) with 7 nodes: 0.857
Test set prediction (R2 score) with 7 nodes:0.693
Mean Absolute error Train (MSE) with 7 nodes: 11.446
Mean Absolute error Test (MSE) with 7 nodes: 22.070
Training set prediction (R2 score) with 8 nodes: 0.858
Test set prediction (R2 score) with 8 nodes:0.693
Mean Absolute error Train (MSE) with 8 nodes: 11.436
Mean Absolute error Test (MSE) with 8 nodes: 22.072
Training set prediction (R2 score) with 9 nodes: 0.856
Test set prediction (R2 score) with 9 nodes:0.690
Mean Absolute error Train (MSE) with 9 nodes: 11.449
Mean Absolute error Test (MSE) with 9 nodes: 22.188
Training set prediction (R2 score) with 10 nodes: 0.856
Test set prediction (R2 score) with 10 nodes:0.690
Mean Absolute error Train (MSE) with 10 nodes: 11.449
Mean Absolute error Test (MSE) with 10 nodes: 22.189

```

### 6.1.3.2 ACTIVATION FUNCTION

```

for actfun in ['identity','logistic','tanh','relu']:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = actfun, solver
    = 'adam', alpha = 0.1, max_iter=400000, random_state=0)

    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} activation function:
    {:.3f}'.format(actfun, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with {} activation function:
    {:.3f}'.format(actfun, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with identity activation function: 0.753
Test set prediction (R2 score) with identity activation function:0.820
Training set prediction (R2 score) with logistic activation function: 0.860
Test set prediction (R2 score) with logistic activation function:0.701
Training set prediction (R2 score) with tanh activation function: 0.879
Test set prediction (R2 score) with tanh activation function:0.673
Training set prediction (R2 score) with relu activation function: 0.776
Test set prediction (R2 score) with relu activation function:0.803

```

### 6.1.3.3 ALPHA VALUE

```

for alphavalue in [0.1, 0.01, 0.001, 0.2, 0.3]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = 'logistic',
    solver = 'adam', alpha = alphavalue,
    learning_rate_init=0.001, max_iter=400000, random_state=0)

```

```

nnvh.fit(X_train_norm, y_train)
y_train_prediction= nnvh.predict(X_train_norm)
y_test_prediction= nnvh.predict(X_test_norm)

print('Training set prediction (R2 score) with alpha= {:.3f}'.format
(alphavalue, r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with alpha= {:.3f}'.format
(alphavalue, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with alpha= 0.1 : 0.860
Test set prediction (R2 score) with alpha= 0.1 :0.701
Training set prediction (R2 score) with alpha= 0.01 : 0.874
Test set prediction (R2 score) with alpha= 0.01 :0.678
Training set prediction (R2 score) with alpha= 0.001 : 0.877
Test set prediction (R2 score) with alpha= 0.001 :0.667
Training set prediction (R2 score) with alpha= 0.2 : 0.850
Test set prediction (R2 score) with alpha= 0.2 :0.711
Training set prediction (R2 score) with alpha= 0.3 : -0.000
Test set prediction (R2 score) with alpha= 0.3 :-0.010

```

### 6.1.3.4 LEARNING RATE (DEFAULT=0.001)

```

for lr in [0.0001,0.001,0.01,0.1,0.2,0.3,0.4,0.5,0.6,0.7]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = 'logistic',
                        solver = 'adam', alpha = 0.1,
                        learning_rate_init=lr, max_iter=400000, random_state=0)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with learning rate={:.3f}'.
format (lr, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with learning rate={:.3f}'.format
(lr, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with learning rate= 0.0001 : -0.000
Test set prediction (R2 score) with learning rate= 0.0001 :-0.011
Training set prediction (R2 score) with learning rate= 0.001 : 0.860
Test set prediction (R2 score) with learning rate= 0.001 :0.701
Training set prediction (R2 score) with learning rate= 0.01 : 0.861
Test set prediction (R2 score) with learning rate= 0.01 :0.700
Training set prediction (R2 score) with learning rate= 0.1 : 0.818
Test set prediction (R2 score) with learning rate= 0.1 :0.712
Training set prediction (R2 score) with learning rate= 0.2 : -0.000
Test set prediction (R2 score) with learning rate= 0.2 :-0.008
Training set prediction (R2 score) with learning rate= 0.3 : -0.001
Test set prediction (R2 score) with learning rate= 0.3 :-0.004
Training set prediction (R2 score) with learning rate= 0.4 : -0.009
Test set prediction (R2 score) with learning rate= 0.4 :-0.000
Training set prediction (R2 score) with learning rate= 0.5 : -0.023
Test set prediction (R2 score) with learning rate= 0.5 :-0.001
Training set prediction (R2 score) with learning rate= 0.6 : -0.042
Test set prediction (R2 score) with learning rate= 0.6 :-0.006
Training set prediction (R2 score) with learning rate= 0.7 : -0.060
Test set prediction (R2 score) with learning rate= 0.7 :-0.012

```

### 6.1.3.5 RANDOM SEED IN THE MLPREGRESSOR

```

for numrs in [0,17,25,28,88,99]:

```

```

nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = 'logistic',
                    solver = 'adam', alpha = 0.1, max_iter=400000,
                    random_state= numrs)
nnvh.fit(X_train_norm, y_train)
y_train_prediction= nnvh.predict(X_train_norm)
y_test_prediction= nnvh.predict(X_test_norm)

print('Training set prediction (R2 score) with random state {} : {:.3f}'.
      format(numrs, r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with random state {} : {:.3f}'.
      format(numrs, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with random state 0 : 0.860
Test set prediction (R2 score) with random state 0 :0.701
Training set prediction (R2 score) with random state 17 : 0.860
Test set prediction (R2 score) with random state 17 :0.701
Training set prediction (R2 score) with random state 25 : 0.860
Test set prediction (R2 score) with random state 25 :0.701
Training set prediction (R2 score) with random state 28 : 0.860
Test set prediction (R2 score) with random state 28 :0.700
Training set prediction (R2 score) with random state 88 : 0.860
Test set prediction (R2 score) with random state 88 :0.701
Training set prediction (R2 score) with random state 99 : 0.860
Test set prediction (R2 score) with random state 99 :0.701

```

### 6.1.3.6 VALIDATION

```

from sklearn.model_selection import cross_val_score
from sklearn.model_selection import LeaveOneOut

from numpy import mean
from numpy import std
from numpy import absolute

X = vldata[['CNH4', 'CK', 'rNA']]
y = vldata['PMI']
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=77*)

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_train_norm = scaler.fit_transform(X_train)
X_test_norm = scaler.transform(X_test)

nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = 'logistic', solver
= 'adam', alpha = 0.1, max_iter=400000, random_state= 0)

nnvh.fit(X_train_norm, y_train)

cv = KFold(n_splits=5, shuffle=True)
scores = cross_val_score(nnvh, X, y, cv=cv, n_jobs=-1)

print('Cross-validation score (5-fold): ', scores)
print('Accuracy: %.2f (%.2f)' % (mean(scores), std(scores)))

cv = LeaveOneOut()
# evaluate model
scores = cross_val_score(nnvh, X, y, scoring='neg_mean_absolute_error', cv=cv,
n_jobs=-1)
# force positive
scores = absolute(scores)

```

```
# report performance
print('MAE: {:.3f} % (mean(scores))')
```

```
Cross-validation score (5-fold): [0.75739 0.72023 0.9021909 0.88255 0.8
2953]
Accuracy: 0.82 (0.08)
MAE: 18.485
```

### 6.1.3.7 FINAL NETWORK

```
X = vldata[['CNH4', 'CK', 'rNA']]
y = vldata['PMI']
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=1)

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_train_norm = scaler.fit_transform(X_train)
X_test_norm = scaler.transform(X_test)

nnvh = MLPRegressor(hidden_layer_sizes = (1,4), activation = 'logistic', solver
= 'adam', alpha = 0.1, max_iter=400000, random_state= 0)

nnvh.fit(X_train_norm, y_train)

y_train_prediction= nnvh.predict(X_train_norm)
y_test_prediction= nnvh.predict(X_test_norm)

plt.figure()
plt.scatter(y_train,y_train_prediction, marker = 'o', color='grey')
plt.xlabel('PMI expected (hrs)')
plt.ylabel('PMI predicted (hrs)')
plt.title('Correlation of PMI predicted by MLPRegressor with the real PMI
(Training set)')
plt.show()
print('Training set prediction (R2 score):
{:.3f}'.format(r2_score(y_train,y_train_prediction)))
print('Mean Squared error Train (MSE):
{:.3f}'.format(mean_squared_error(y_train, y_train_prediction)))
print('Mean Absolute error Train (MSE):
{:.3f}'.format(mean_absolute_error(y_train, y_train_prediction)))

plt.figure()
plt.scatter(y_test,y_test_prediction, marker = 'o', color='blue')
plt.xlabel('PMI expected (hrs)')
plt.ylabel('PMI predicted (hrs)')
plt.title('Correlation of PMI predicted by MLPRegressor with the real PMI (Test
set)')
plt.show()
print('Test set prediction (R2 score):
{:.3f}'.format(r2_score(y_test,y_test_prediction)))
print('Mean Squared error Test (MSE): {:.3f}'.format(mean_squared_error(y_test,
y_test_prediction)))
print('Mean Absolute error Test (MSE):
{:.3f}'.format(mean_absolute_error(y_test, y_test_prediction)))

plt.figure()
plt.plot(nnvh.loss_curve_)
plt.xlabel('Iterations')
plt.ylabel('Loss')
plt.show()
print('Current loss : {:.2f}'.format(nnvh.loss_))
```



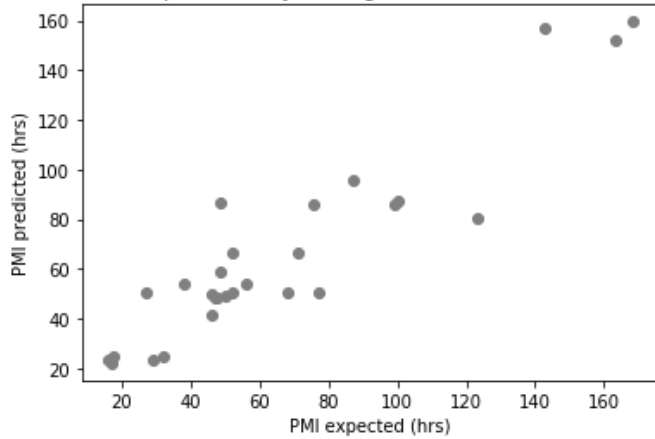
```

##For getting the predictions of the Train/Test split##
print('Train/Test predictions')

for j in range(y_test_prediction.shape[0]):print('{} {}{}'.format
(y_test.index[j],y_test[y_test.index[j]],y_test_prediction[j]))
for j in range(y_train_prediction.shape[0]):print('{} {} {}'.format
(y_train.index[j],y_train[y_train.index[j]],y_train_prediction[j]))

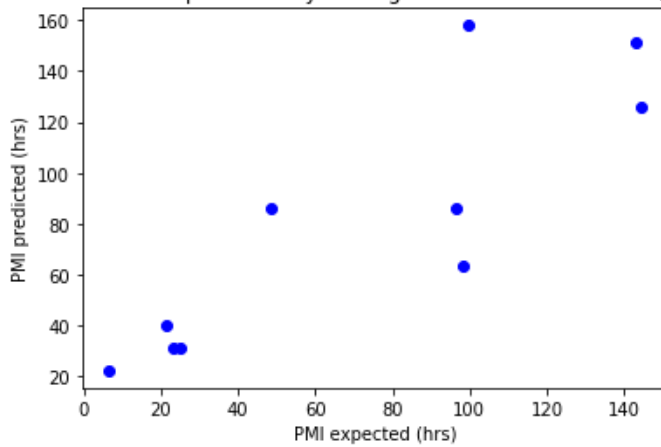
```

Correlation of PMI predicted by MLPregressor with the real PMI (Training set)

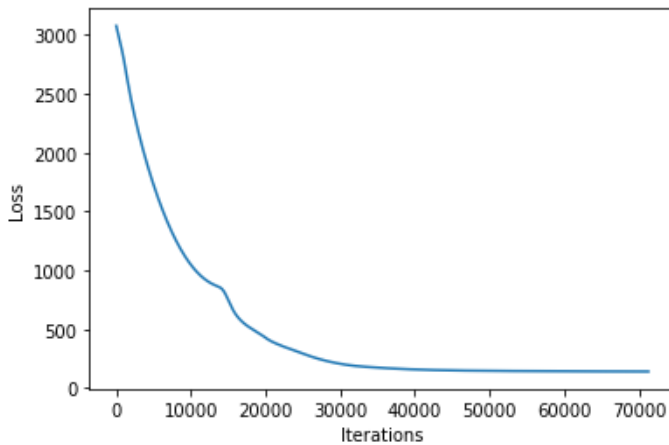


Training set prediction (R2 score): 0.860  
Mean Absolute error Train (MAE): 11.424

Correlation of PMI predicted by MLPregressor with the real PMI (Test set)



Test set prediction (R2 score): 0.701  
Mean Absolute error Test (MAE): 21.667



Current loss : 141.15

## 6.1.4 CODE FOR SGD SOLVER

### 6.1.4.1 NUMBER OF NODES IN THE HIDDEN LAYER

```
for hn in [1,2,3,4,5,6,7,8,9,10]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,hn), activation = 'logistic',
                        solver = 'sgd', alpha = 0.1, learning_rate= 'constant',
                        learning_rate_init= 0.001, max_iter=400000,
                        random_state=0, momentum= 0.3, nesterovs_momentum=
                        False)

    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} nodes: {:.3f}'.format(hn,
    r2_score(y_train,y_train_prediction))
    print('Test set prediction (R2 score) with {} nodes:{:.3f}'.format(hn,
    r2_score(y_test,y_test_prediction))

    print('Mean Absolute error Train (MSE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_train, y_train_prediction))
    print('Mean Absolute error Test (MSE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_test, y_test_prediction))
```

```
Training set prediction (R2 score) with 1 nodes: 0.629
Test set prediction (R2 score) with 1 nodes:0.648
Mean Absolute error Train (MSE) with 1 nodes: 19.231
Mean Absolute error Test (MSE) with 1 nodes: 28.079
Training set prediction (R2 score) with 2 nodes: 0.796
Test set prediction (R2 score) with 2 nodes:0.706
Mean Absolute error Train (MSE) with 2 nodes: 13.842
Mean Absolute error Test (MSE) with 2 nodes: 23.130
Training set prediction (R2 score) with 3 nodes: 0.000
Test set prediction (R2 score) with 3 nodes:-0.007
Mean Absolute error Train (MSE) with 3 nodes: 32.258
Mean Absolute error Test (MSE) with 3 nodes: 45.670
Training set prediction (R2 score) with 4 nodes: -0.001
Test set prediction (R2 score) with 4 nodes:-0.004
Mean Absolute error Train (MSE) with 4 nodes: 32.484
Mean Absolute error Test (MSE) with 4 nodes: 45.670
Training set prediction (R2 score) with 5 nodes: -0.003
Test set prediction (R2 score) with 5 nodes:-0.001
Mean Absolute error Train (MSE) with 5 nodes: 32.745
Mean Absolute error Test (MSE) with 5 nodes: 45.667
```

Training set prediction (R2 score) with 6 nodes: -0.008  
 Test set prediction (R2 score) with 6 nodes:0.000  
 Mean Absolute error Train (MSE) with 6 nodes: 33.107  
 Mean Absolute error Test (MSE) with 6 nodes: 45.666  
 Training set prediction (R2 score) with 7 nodes: -0.014  
 Test set prediction (R2 score) with 7 nodes:0.001  
 Mean Absolute error Train (MSE) with 7 nodes: 33.415  
 Mean Absolute error Test (MSE) with 7 nodes: 45.664  
 Training set prediction (R2 score) with 8 nodes: -0.020  
 Test set prediction (R2 score) with 8 nodes:0.000  
 Mean Absolute error Train (MSE) with 8 nodes: 33.752  
 Mean Absolute error Test (MSE) with 8 nodes: 45.655  
 Training set prediction (R2 score) with 9 nodes: -0.027  
 Test set prediction (R2 score) with 9 nodes:-0.001  
 Mean Absolute error Train (MSE) with 9 nodes: 34.112  
 Mean Absolute error Test (MSE) with 9 nodes: 45.646  
 Training set prediction (R2 score) with 10 nodes: -0.034  
 Test set prediction (R2 score) with 10 nodes:-0.002  
 Mean Absolute error Train (MSE) with 10 nodes: 34.392  
 Mean Absolute error Test (MSE) with 10 nodes: 45.636

### 6.1.4.2 ACTIVATION FUNCTION

```

for actfun in ['identity','logistic','tanh','relu']:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = actfun, solver
                        = 'sgd', alpha = 0.1, learning_rate='constant',
                        learning_rate_init= 0.001, max_iter=400000,
                        random_state=0, momentum= 0.3, nesterovs_momentum=
                        False)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} activation function:
    {:.3f}'.format(actfun, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with {} activation
    function:{:.3f}'.format(actfun, r2_score(y_test,y_test_prediction)))
  
```

Training set prediction (R2 score) with identity activation function: 0.753  
 Test set prediction (R2 score) with identity activation function:0.820  
 Training set prediction (R2 score) with logistic activation function: 0.796  
 Test set prediction (R2 score) with logistic activation function:0.706  
 Training set prediction (R2 score) with tanh activation function: 0.587  
 Test set prediction (R2 score) with tanh activation function:0.680  
 Training set prediction (R2 score) with relu activation function: 0.730  
 Test set prediction (R2 score) with relu activation function:0.755

### 6.1.4.3 ALPHA VALUE

```

for alphavalue in [0.1, 0.01, 0.001, 0.2, 0.3]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = 'logistic',
                        solver = 'sgd', alpha = alphavalue, learning_rate=
                        'constant', learning_rate_init= 0.001, max_iter=400000,
                        random_state=0, momentum= 0.3, nesterovs_momentum=
                        False)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)
  
```

```

print('Training set prediction (R2 score) with alpha= {} :
{: .3f}'.format(alphavalue, r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with alpha= {}
{: .3f}'.format(alphavalue, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with alpha= 0.1 : 0.796
Test set prediction (R2 score) with alpha= 0.1 :0.706
Training set prediction (R2 score) with alpha= 0.01 : 0.794
Test set prediction (R2 score) with alpha= 0.01 :0.740
Training set prediction (R2 score) with alpha= 0.001 : 0.794
Test set prediction (R2 score) with alpha= 0.001 :0.694
Training set prediction (R2 score) with alpha= 0.2 : 0.797
Test set prediction (R2 score) with alpha= 0.2 :0.742
Training set prediction (R2 score) with alpha= 0.3 : 0.796
Test set prediction (R2 score) with alpha= 0.3 :0.730

```

#### 6.1.4.4 LEARNING RATE

```

for lr in [0.0001,0.001,0.01,0.1,0.2,0.3,0.4,0.5,0.6,0.7]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = 'logistic',
                        solver = 'sgd', alpha = 0.01, learning_rate= 'constant',
                        learning_rate_init= lr, max_iter=400000,
                        random_state=0, momentum= 0.3, nesterovs_momentum=
                        False)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with learning rate= {:.3f}'.
          format(lr, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with learning rate= {:.3f}'.
          format(lr, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with learning rate= 0.0001 : 0.811
Test set prediction (R2 score) with learning rate= 0.0001 :0.722
Training set prediction (R2 score) with learning rate= 0.001 : 0.816
Test set prediction (R2 score) with learning rate= 0.001 :0.712
Training set prediction (R2 score) with learning rate= 0.01 : 0.498
Test set prediction (R2 score) with learning rate= 0.01 :0.628
Training set prediction (R2 score) with learning rate= 0.1 : 0.000
Test set prediction (R2 score) with learning rate= 0.1 :-0.009
Training set prediction (R2 score) with learning rate= 0.2 : 0.000
Test set prediction (R2 score) with learning rate= 0.2 :-0.009
Training set prediction (R2 score) with learning rate= 0.3 : -0.000
Test set prediction (R2 score) with learning rate= 0.3 :-0.009
Training set prediction (R2 score) with learning rate= 0.4 : -0.000
Test set prediction (R2 score) with learning rate= 0.4 :-0.009
Training set prediction (R2 score) with learning rate= 0.5 : -0.000
Test set prediction (R2 score) with learning rate= 0.5 :-0.009
Training set prediction (R2 score) with learning rate= 0.6 : -0.000
Test set prediction (R2 score) with learning rate= 0.6 :-0.009
Training set prediction (R2 score) with learning rate= 0.7 : -0.000
Test set prediction (R2 score) with learning rate= 0.7 :-0.009

```

#### 6.1.4.5 MOMENTUM

```

for m in [0.01,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = 'logistic',
                        solver = 'sgd', alpha = 0.01, learning_rate= 'constant',

```

```

        learning_rate_init= 0.001, max_iter=400000,
        random_state=0, momentum= m, nesterovs_momentum= False)
nnvh.fit(X_train_norm, y_train)
y_train_prediction= nnvh.predict(X_train_norm)
y_test_prediction= nnvh.predict(X_test_norm)

print('Training set prediction (R2 score) with momentum= {:.3f}'.format
(m, r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with momentum={:.3f}'.format
(m,r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with momentum= 0.01 : 0.811
Test set prediction (R2 score) with momentum= 0.01 :0.723
Training set prediction (R2 score) with momentum= 0.1 : 0.813
Test set prediction (R2 score) with momentum= 0.1 :0.718
Training set prediction (R2 score) with momentum= 0.2 : 0.814
Test set prediction (R2 score) with momentum= 0.2 :0.715
Training set prediction (R2 score) with momentum= 0.3 : 0.816
Test set prediction (R2 score) with momentum= 0.3 :0.712
Training set prediction (R2 score) with momentum= 0.4 : 0.817
Test set prediction (R2 score) with momentum= 0.4 :0.711
Training set prediction (R2 score) with momentum= 0.5 : 0.818
Test set prediction (R2 score) with momentum= 0.5 :0.708
Training set prediction (R2 score) with momentum= 0.6 : 0.819
Test set prediction (R2 score) with momentum= 0.6 :0.701
Training set prediction (R2 score) with momentum= 0.7 : 0.820
Test set prediction (R2 score) with momentum= 0.7 :0.698
Training set prediction (R2 score) with momentum= 0.8 : 0.821
Test set prediction (R2 score) with momentum= 0.8 :0.699
Training set prediction (R2 score) with momentum= 0.9 : 0.821
Test set prediction (R2 score) with momentum= 0.9 :0.687

```

#### 6.1.4.6 NESTEROVS MOMENTUM

```

nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = 'logistic', solver
= 'sgd', alpha = 0.01, learning_rate= 'constant',
        learning_rate_init= 0.001, max_iter=400000, random_state=0,
nesterovs_momentum= True)
nnvh.fit(X_train_norm, y_train)
y_train_prediction= nnvh.predict(X_train_norm)
y_test_prediction= nnvh.predict(X_test_norm)

print('Training set prediction (R2 score) with nesterovs momentum :
{:.3f}'.format(r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with nesterovs momentum
{:.3f}'.format(r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with nesterovs momentum : 0.794
Test set prediction (R2 score) with nesterovs momentum :0.740

```

#### 6.1.4.7 RANDOM STATE MLPREGRESSOR

```

for numrs in [0,17,25,28,88,99]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,2), activation = 'logistic',
        solver = 'sgd', alpha = 0.01, learning_rate= 'constant',
        learning_rate_init= 0.001, max_iter=400000,
        random_state=numrs, momentum= 0.4, nesterovs_momentum=
False)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

```

```

print('Training set prediction (R2 score) with random state {}: {:.3f}'.
      format(numrs, r2_score(y_train,y_train_prediction)))
print('Test set prediction (R2 score) with random state {}: {:.3f}'.
      format(numrs, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with random state 0 : 0.816
Test set prediction (R2 score) with random state 0 :0.712
Training set prediction (R2 score) with random state 17 : 0.816
Test set prediction (R2 score) with random state 17 :0.712
Training set prediction (R2 score) with random state 25 : 0.816
Test set prediction (R2 score) with random state 25 :0.712
Training set prediction (R2 score) with random state 28 : 0.816
Test set prediction (R2 score) with random state 28 :0.712
Training set prediction (R2 score) with random state 88 : 0.816
Test set prediction (R2 score) with random state 88 :0.712
Training set prediction (R2 score) with random state 99 : 0.816
Test set prediction (R2 score) with random state 99 :0.711

```

## 6.1.5 CODE FOR LBFGS

### 6.1.5.1 NUMBER OF NODES IN THE HIDDEN LAYER

```

for hn in [1,2,3,4,5,6,7,8,9,10]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,hn), activation = 'logistic',
                        solver = 'lbfgs', alpha = 0.1, max_iter=40000,
                        random_state=0)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} nodes: {:.3f}'.format(hn,
    r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with {} nodes: {:.3f}'.format(hn,
    r2_score(y_test,y_test_prediction)))
    print('Mean Absolute error Train (MAE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_train, y_train_prediction)))
    print('Mean Absolute error Test (MAE) with {} nodes: {:.3f}'.format(hn,
    mean_absolute_error(y_test, y_test_prediction)))

```

```

Training set prediction (R2 score) with 1 nodes: 0.786
Test set prediction (R2 score) with 1 nodes:0.710
Mean Absolute error Train (MAE) with 1 nodes: 13.612
Mean Absolute error Test (MAE) with 1 nodes: 23.232
Training set prediction (R2 score) with 2 nodes: 0.822
Test set prediction (R2 score) with 2 nodes:0.700
Mean Absolute error Train (MAE) with 2 nodes: 12.138
Mean Absolute error Test (MAE) with 2 nodes: 22.235
Training set prediction (R2 score) with 3 nodes: 0.813
Test set prediction (R2 score) with 3 nodes:0.747
Mean Absolute error Train (MAE) with 3 nodes: 12.681
Mean Absolute error Test (MAE) with 3 nodes: 20.407
Training set prediction (R2 score) with 4 nodes: -0.000
Test set prediction (R2 score) with 4 nodes:-0.009
Mean Absolute error Train (MAE) with 4 nodes: 32.180
Mean Absolute error Test (MAE) with 4 nodes: 45.678
Training set prediction (R2 score) with 5 nodes: -0.000
Test set prediction (R2 score) with 5 nodes:-0.009
Mean Absolute error Train (MAE) with 5 nodes: 32.180
Mean Absolute error Test (MAE) with 5 nodes: 45.678

```

Training set prediction (R2 score) with 6 nodes: 0.861  
 Test set prediction (R2 score) with 6 nodes:0.692  
 Mean Absolute error Train (MAE) with 6 nodes: 11.388  
 Mean Absolute error Test (MAE) with 6 nodes: 21.968  
 Training set prediction (R2 score) with 7 nodes: 0.861  
 Test set prediction (R2 score) with 7 nodes:0.691  
 Mean Absolute error Train (MAE) with 7 nodes: 11.390  
 Mean Absolute error Test (MAE) with 7 nodes: 21.985  
 Training set prediction (R2 score) with 8 nodes: 0.861  
 Test set prediction (R2 score) with 8 nodes:0.692  
 Mean Absolute error Train (MAE) with 8 nodes: 11.390  
 Mean Absolute error Test (MAE) with 8 nodes: 21.978  
 Training set prediction (R2 score) with 9 nodes: 0.896  
 Test set prediction (R2 score) with 9 nodes:0.716  
 Mean Absolute error Train (MAE) with 9 nodes: 10.576  
 Mean Absolute error Test (MAE) with 9 nodes: 21.210  
 Training set prediction (R2 score) with 10 nodes: 0.860  
 Test set prediction (R2 score) with 10 nodes:0.690  
 Mean Absolute error Train (MAE) with 10 nodes: 11.408  
 Mean Absolute error Test (MAE) with 10 nodes: 22.013

### 6.1.5.2 ACTIVATION FUNCTION

```

for actfun in ['identity','logistic','tanh','relu']:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,3), activation = actfun, solver
                        = 'lbfgs', alpha = 0.1, max_iter=400000,
                        random_state=0)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with {} activation function:
    {:.3f}'.format(actfun, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with {} activation function:{:.3f}'.
    format(actfun, r2_score(y_test,y_test_prediction)))
  
```

Training set prediction (R2 score) with identity activation function: 0.753  
 Test set prediction (R2 score) with identity activation function:0.820  
 Training set prediction (R2 score) with logistic activation function: 0.813  
 Test set prediction (R2 score) with logistic activation function:0.747  
 Training set prediction (R2 score) with tanh activation function: 0.485  
 Test set prediction (R2 score) with tanh activation function:0.398  
 Training set prediction (R2 score) with relu activation function: 0.777  
 Test set prediction (R2 score) with relu activation function:0.803

### 6.1.5.3 ALPHA VALUE

```

for alphavalue in [0.1, 0.01, 0.001, 0.2, 0.3]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,3), activation = 'logistic',
                        solver = 'lbfgs', alpha = alphavalue, max_iter=400000,
                        random_state=0)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with alpha= {} : {:.3f}'.
    format(alphavalue, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with alpha= {} : {:.3f}'.
    format(alphavalue, r2_score(y_test,y_test_prediction)))
  
```



```

Training set prediction (R2 score) with alpha= 0.1 : 0.813
Test set prediction (R2 score) with alpha= 0.1 :0.747
Training set prediction (R2 score) with alpha= 0.01 : 0.828
Test set prediction (R2 score) with alpha= 0.01 :0.358
Training set prediction (R2 score) with alpha= 0.001 : 0.881
Test set prediction (R2 score) with alpha= 0.001 :0.656
Training set prediction (R2 score) with alpha= 0.2 : 0.819
Test set prediction (R2 score) with alpha= 0.2 :0.710
Training set prediction (R2 score) with alpha= 0.3 : 0.816
Test set prediction (R2 score) with alpha= 0.3 :0.724

```

#### 6.1.5.4 RANDOM STATE MLPREGRESSOR

```

for numrs in [0,17,25,28,88,99]:
    nnvh = MLPRegressor(hidden_layer_sizes = (1,3), activation = 'logistic',
                        solver = 'lbfgs', alpha = 0.1, max_iter=400000,
                        random_state= numrs)
    nnvh.fit(X_train_norm, y_train)
    y_train_prediction= nnvh.predict(X_train_norm)
    y_test_prediction= nnvh.predict(X_test_norm)

    print('Training set prediction (R2 score) with random state {} : {:.3f}'.
          format(numrs, r2_score(y_train,y_train_prediction)))
    print('Test set prediction (R2 score) with random state {} : {:.3f}'.
          format(numrs, r2_score(y_test,y_test_prediction)))

```

```

Training set prediction (R2 score) with random state 0 : 0.813
Test set prediction (R2 score) with random state 0 :0.747
Training set prediction (R2 score) with random state 17 : 0.785
Test set prediction (R2 score) with random state 17 :0.710
Training set prediction (R2 score) with random state 25 : 0.785
Test set prediction (R2 score) with random state 25 :0.710
Training set prediction (R2 score) with random state 28 : 0.824
Test set prediction (R2 score) with random state 28 :0.403
Training set prediction (R2 score) with random state 88 : 0.785
Test set prediction (R2 score) with random state 88 :0.710
Training set prediction (R2 score) with random state 99 : 0.824
Test set prediction (R2 score) with random state 99 :0.402

```

#### 6.1.6 CODE FOR UNTRAINED REAL SAMPLE PREDICTIONS

```

vhdatapred = pd.read_csv('vhrawdatapredictions5.csv')*
X = vhdatapred[['PA NH4', 'PH NH4', 'PA K', 'PH K', 'PA NH4 + PA K', 'PH NH4 +
PH K', 'PA IS', 'PH IS']]

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_norm = scaler.fit_transform(X)

y_prediction_pmi= nnvh.predict(X_norm)
y_prediction_pmi= nnvh.predict(X_norm)
for j in range(y_prediction_pmi.shape[0]):
    print('{}'.format(y_prediction_pmi[j]))
*The file with the inputs (X)

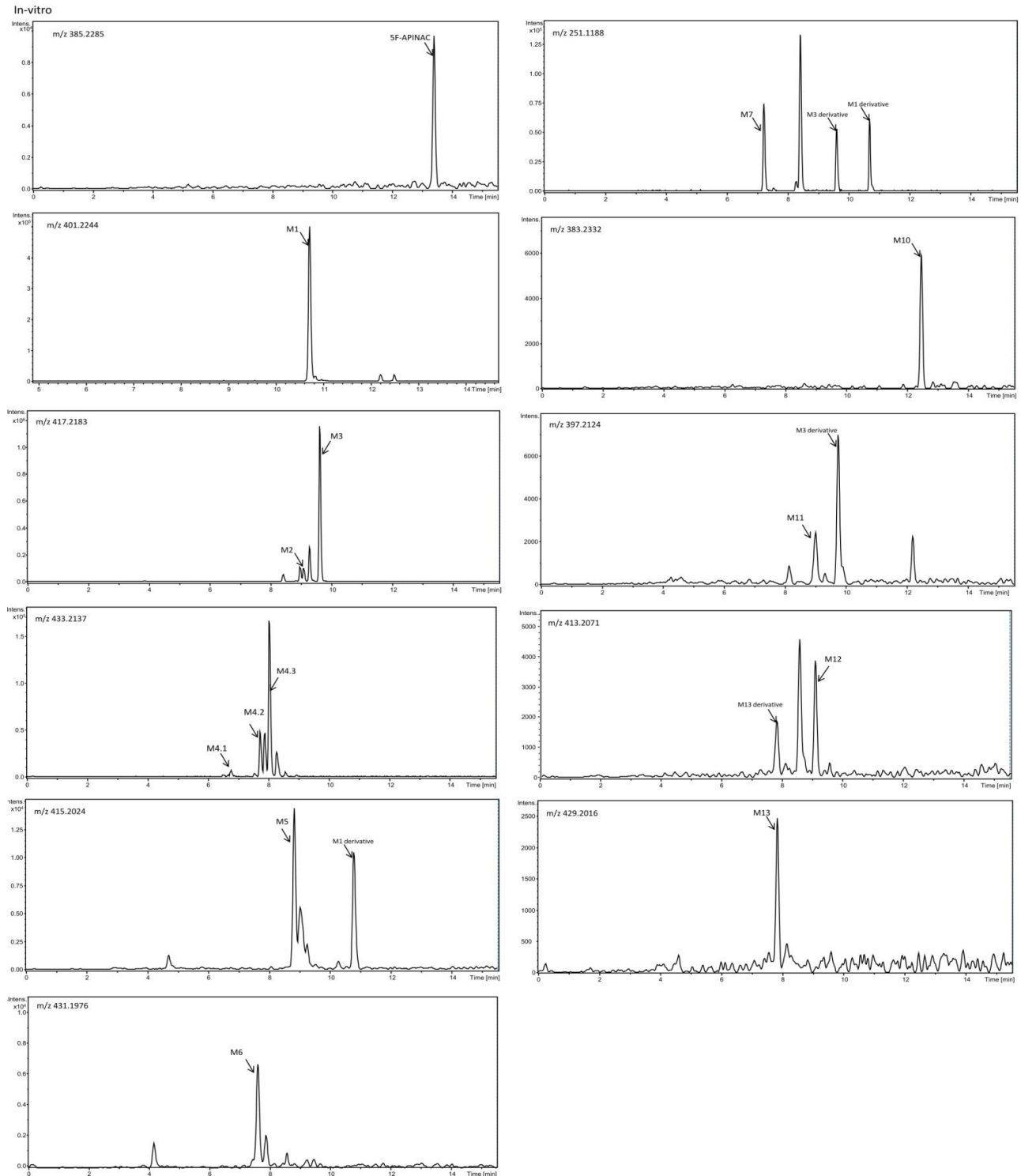
```

For each random seed (1, 77, 34, 99, 70, 88, 100), this code is written afterwards to predict the untrained samples in each network.

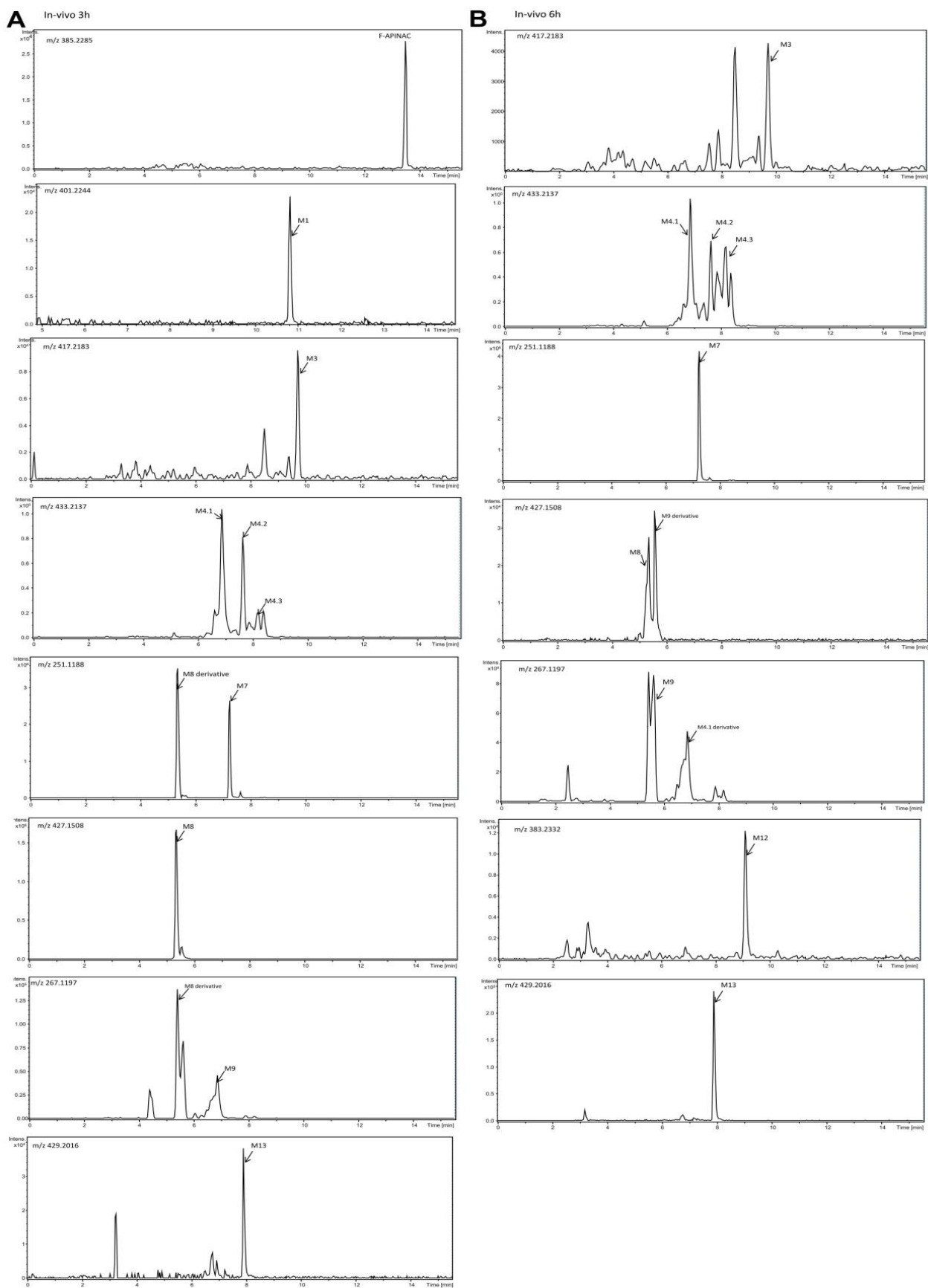


## 6.2 APPENDIX II. ELUCIDATION OF THE 5F- APINAC METABOLIC PATHWAY

### 6.2.1 EXTRACTED ION CHROMATOGRAMS FOR THE METABOLITES ASSOCIATED WITH 5F-APINAC FOUND *IN VITRO* AND *IN VIVO*



Supplementary Figure 1: Chromatograms for the metabolites associated to 5F-APINAC found *in vitro*



**Supplementary Figure 2:** Chromatograms for the metabolites associated to 5F-APINAC found *in vivo*. Panel A: 3 hours after drug administration. Panel B: 6 hours after drug administration

## 7. SCIENTIFIC PUBLICATIONS

Sergey A. Savchuk, Covadonga Palacio, Artyom Gil, Franco Tagliaro, Roman M. Kuznetsov, Alex Brito, Svetlana A. Appolonova - **Determination of the chemical composition of alcoholic beverages by gas chromatography-mass spectrometry**. *Journal of Food Processing and Preservation* (2020). DOI: [doi/10.1111/jfpp.14676](https://doi.org/10.1111/jfpp.14676) (On-line only)

Covadonga Palacio, Rossella Gottardo, Vito Cirielli, Giacomo Musile, Yvane Agard, Federica Bortolotti, Franco Tagliaro - **Simultaneous analysis of potassium and ammonium ions in the vitreous humour by capillary electrophoresis and their integrated use to infer the post mortem interval (PMI)**. *Science, Medicine and the Law* (2020). DOI: [10.1177/0025802420934239](https://doi.org/10.1177/0025802420934239) (On-line only)

Svetlana A. Appolonova, Covadonga Palacio, Ksenia M. Shestakova, Natalia V. Mesonzhnik, Alex Brito, Roman M. Kuznetsov, Pavel A. Markin, Natalia L. Bochkareva, Dmitry Burmykin, Maxim Ovcharov, Giacomo Musile, Franco Tagliaro, Sergey A. Savchuk - **In vivo and in vitro metabolism of the novel synthetic cannabinoid 5F-APINAC**. *Forensic Toxicology* (2019), volume 38, pages 160-161. DOI: [doi.org/10.1007/s11419-019-00503-z](https://doi.org/10.1007/s11419-019-00503-z)

Rossella Gottardo, Covadonga Palacio, Kseniia M. Shestakova, Natalia E. Moskaleva, Federica Bortolotti and Franco Tagliaro - **A new method for the determination of ammonium in the vitreous humour based on capillary electrophoresis and its preliminary application in thanatochemistry**. *Clinical Chemistry and Laboratory Medicine* (2018), volume 57: issue 4, pages 504-509. DOI: [doi.org/10.1515/cclm-2018-0384](https://doi.org/10.1515/cclm-2018-0384)

Giacomo Musile, Rossella Gottardo, Covadonga Palacio, Ksenia Shestakova, Dario Raniero, Elio F. De Palo, Franco Tagliaro - **Development of a low-cost gas diffusion device for ammonia detection in the vitreous humour and its preliminary application for estimation of the time since death**. *Forensic Science International* (2018), volume 295, pages 150-156. DOI: [10.1016/j.forsciint.2018.12.012](https://doi.org/10.1016/j.forsciint.2018.12.012)

## 8. ORAL AND POSTER COMMUNICATIONS

Rossella Gottardo and Covadonga Palacio – **A new method for the determination of ammonium in the vitreous humour based on capillary electrophoresis and its preliminary application in thanatochemistry.** New methods of sample preparation and identification of psychoactive substances. *Moscow, 29<sup>th</sup> October 2018.* Oral presentation

Sergey Savchuk and Covadonga Palacio – **Fast LLE method for the preparation of cadaver tissue for GC-MS and HPLC-MS/MS analysis. Case study. Detection of carfentanyl in putrid tissue** – NPS Maastricht 2019: Carfentanyl poster presentation VI International Conference on Novel Psychoactive Substances. *Maastricht, 8-9 April 2019.* Poster

Covadonga Palacio, Rossella Gottardo and Franco Tagliaro – **Determination of ammonium in the vitreous humour based on capillary electrophoresis: preliminary application in thanatochemistry** – 25<sup>th</sup> Latin-American Symposium on Biotechnology, Biomedical, Biopharmaceutical, and Industrial Applications of Capillary Electrophoresis and Microchip Technology (LACE). *Madrid, September 29<sup>th</sup> - October 2<sup>nd</sup>, 2019.* Oral presentation and poster.

## 9. ACKNOWLEDGEMENTS

Throughout the writing of this dissertation, I have received a great deal of support and assistance.

I would first like to thank my two supervisors, Professor Franco Tagliaro and Prof.ssa Rossella Gottardo, whose expertise was invaluable in formulating the research questions and methodology. Also, for their valuable guidance throughout these three years. You provided me with the tools that I needed to choose the right direction and successfully complete my dissertation.

I would also like to extend my deepest gratitude to all my colleagues from Medicina Legale. To Prof.ssa Federica Bortolotti for her help, assistance and advice through different times of my PhD. To Prof. Elio de Palo, Dr. Nadia Porpiglia, Dr. Anna Bertaso, Dr. Daniela Sorio and Dr. Giacomo Musile, Dr. Francesco Taus, Matilde, Marco and Sara for all their help, support, supervision, teaching, fun days at the laboratory, friendship and their valuable input on the do's and don'ts of the Italian culture. I gratefully acknowledge the assistance and guidance of Dr. Musile in the last project of my thesis. I would like to further thank Matilde for the extra help with the Toxyper and the Kinder Buenos when I most needed them.

Special thanks to my PhD companions and friends Ksenia Shestakova and Yvane Agard. For all the conversations, advices, help, support and fun that we had together. I will treasure your friendship forever.

Thanks, should also go to Svetlana A. Appolonova and Sergey A. Savchuk for the opportunity given to work in Sechenov University for three months. I also wish to thank everyone in the laboratory for making me feel welcomed and helping with the research.

I would like to extend my sincere thanks to Carl Jackson for his mentoring in Neural Networks and initiation to programming in Python. You provided me with the tools that I needed to conclude the first project for my thesis. Also, many thanks to Jeremy Vicencio for taking his time to proofread my dissertation.

In addition, I would like to thank my parents for their emotional and financial support. Finally, I could not have completed this dissertation without the support of my close friends, Inés, Valle and Olaya, who supported and cheered me up, planned trips, excursions and birthdays and many other happy distractions to rest my mind outside of my research in Italy and when I travelled back home.

Thank you all for these last three years.