Breath Pharmacometabolomics for Therapeutic Drug Monitoring of Epilepsy and type-1 Diabetes

Inaugural dissertation

to

be awarded the degree of Dr. sc. med.

presented at

the Faculty of Medicine

of the University of Basel

by

Mohamad Awchi

From Kirkuk, Kirkuk

Basel 2023

Approved by the Faculty of Medicine
On application of
Prof. Pablo Sinues, *First Supervisor*
Prof. Urs Frey, *Second Supervisor*
Prof. Renato Zenobi, *Further Advisor*
Prof. Patrik Španěl, *External Expert*

Basel, 19.06.2023

Dean
Prof. Primo Schär

Acknowledgements

I owe the completion of this doctoral thesis to the invaluable support and guidance of a few key people who have been instrumental in my academic journey.

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Pablo Sinues, for granting me the opportunity to pursue this PhD. I am truly grateful for all the scientific insights, help, and support that you have provided me with over the years, without which this work would not have been possible.

I am also grateful to my second supervisor, Professor Urs Frey, for facilitating our research at the UKBB. His support and contribution to various research projects have been invaluable.

I extend my heartfelt appreciation to my other supervisors, Professor Renato Zenobi, whose inspirational lecture back in 2018 in Stockholm sparked my interest in breath research, and Professor Patrik Španěl, for being an external expert for my thesis, I am grateful to have you on board. I am also thankful to Professor Marc Donath for chairing my Ph.D. defense.

I am deeply indebted to my collaborators, Prof. Gabor Szinnai, PD Alexandre Datta, Prof. Diego García-Gómez, PD Patricia Dill, Dr. med. Sara Bachmann, Dr. med. Marie-Anne Burckhardt, and Dr. med. Melanie Hess, for their significant contributions to my research. Furthermore, I would like to thank all the study participants for their invaluable contributions to my research, and for the pleasant interactions that we had. I would also like to acknowledge Fiona Beck and Madelon Bakx for editing and proofreading my thesis to ensure its quality.

My gratitude goes to the team at the SinuesLab, Jiafa Zeng, Kim Arnold, Isabel Gonzales, Noemi Künstle, Kapil dev Singh, Melina Richard, Jakob Usemann, and Zhihong Yin, thanks for the interesting conversations and discussions throughout my PhD, you were a big part of my journey here. I would like to extend my gratitude to Axel and Chris for their intriguing conversations and discussions about my research, their research, and everything else. I am particularly grateful to Axel for being my sounding board. My friends and family back in the Netherlands have been a constant source of support throughout my academic journey, and I am deeply grateful to them.

Lastly, I am forever indebted to my partner Yoëlle, who has been a pillar of strength throughout my Ph.D. Your unwavering support, encouragement, and assistance have been priceless.

Abbreviations

Acetoacetate – AcAc

Anti-seizure medication – ASM

Base excess – BE

Collision induced dissociation – CID

Concordance correlation coefficient - CCC

Diabetic ketoacidosis – DKA

Electrospray ionization – ESI

Entropy similarity – ES

Exhaled breath condensate – EBC

Gas chromatography - mass spectrometry – GC-MS

Gaussian process regression – GPR

Higher-energy C-trap dissociation – HCD

Hydrophilic lipophilic balance – HLB

Hyperglycemia – HG

Liquid chromatography tandem mass spectrometry – LC-MS/MS

Intensive care unit – ICU

Mass-to-charge – m/z

Principal component analysis – PCA

Proton transfer reaction mass spectrometry – PTR-MS

Ribonucleic acid – RNA

Secondary electrospray ionization - high resolution mass spectrometry – SESI-HRMS

Selected ion flow tube mass spectrometry – SIFT-MS

Solid phase extraction – SPE

Support vector machine – SVM

Therapeutic drug monitoring – TDM

Tricarboxylic acid – TCA

Ultra-high-performance liquid chromatography – UHPLC

Valproic acid - VPA

Volatile organic compounds - VOC

Summary

Background: Pharmacometabolomics is the study of pharmaceuticals and their effect on metabolites in the human body. By studying pharmacometabolomics, one can gain insights into the mode of action and effectiveness of pharmaceuticals, which is patient-dependent, and thereby infer information for personalized medicine purposes. This study can be done on various bodily fluids. One of the fastest ways to obtain this information is by analyzing breath. It provides a window into human metabolism in a non-invasive, fast, and cost-effective way. One novel approach is the use of secondary electrospray ionization coupled with high-resolution mass spectrometry (SESI-HRMS). Within this thesis, we will look at two applications of personalized medicine utilizing breath by SESI-HRMS. Namely, for epilepsy and type-1 diabetes.

Aim: Epilepsy and type-1 diabetes are two of the most common chronic diseases worldwide and are associated with tremendous medical costs and their monitoring entails invasive methods. Furthermore, the exact pathophysiology of these diseases is still debated. Previous work in our lab reported the therapeutic drug monitoring (TDM) for anti-seizure medication (ASM) valproic acid (VPA) and in addition identified metabolites to be involved in response to and side effects from ASMs by breath analysis. The identification however, was only on a low level of confidence. Therefore, in this thesis we aim to unambiguously confirm these reported metabolites. In addition, the method was done in a real-time fashion, meaning that patients have to directly exhale into an instrument, which for (cognitively) impaired patients is cumbersome. Secondly, we aim to convert this real-time method to a patient friendly off-line method. Lastly, we aim to study the anabolic effect of insulin on patients undergoing diabetic ketoacidosis (DKA) for pathophysiological understanding and monitoring.

Methods: For breath measurements, SESI-HRMS was utilized. For patients incapable of performing real-time measurements or for evaluation purposes, Nalophan bags were used to capture breath, after which their contents were infused into the SESI-HRMS. For the confirmation of metabolites in response to and side effects from ASMs, exhaled breath condensate (EBC) was used, which was subsequently analyzed by ultra-high-performance liquid chromatography coupled with HRMS. Hereafter, the data was analyzed using statistical and machine learning methods.

Results: We successfully unambiguously confirmed six amino acids in breath by utilizing EBC with a “level 1” confidence, which were previously reported to be involved in response to and side effects from ASMs. Second, the feasibility of transitioning an existing TDM approach by breath to an off-line approach was successfully assessed. Measurements taken over the course of 3.5 years showed good agreement in terms of accuracy but showed an overall decrease in signal intensity when utilizing a bag system. Furthermore, free VPA and total VPA correlated with a Lin’s concordance correlation coefficient (CCC) of 0.60 and 0.47 respectively. Metabolites reported to be involved in response to and side effects from ASM showed poor Lin’s CCC values, indicating that this information is largely lost by the bag system. Lastly, we were able to show clear metabolic trajectories in patients undergoing DKA after insulin onset. Breath analysis revealed that the generally accepted altered pathways were measurable by breath and were, to a large extent, in agreement with previously reported observations. Acetone and acetoacetate were identified as metabolites with the ability to monitor DKA progression, with breath acetoacetate being a novel finding. Finally, DKA patients demonstrated trajectories towards those of patients stratified as “in control” or “non-diabetic” after insulin onset.

Conclusion: This thesis provided novel insights into the capabilities of breath analysis to perform pharmacometabolomics in a real clinical setting. First, we identified amino acids reported to be involved in response to and side effects from ASM. This allows to confidently infer information from these metabolites in the future. Second, we showed the ability to measure free VPA and total VPA by offline SESI-HRMS, allowing bedside and (cognitively) impaired individuals to benefit from the technique. Lastly, we provided a framework for disease monitoring and pathophysiological understanding of DKA. This holds the potential to be used bedside in the Intensive care unit (ICU) for DKA disease progression as well as a further pathophysiological understanding of DKA. Overall, we showed the power of breath analysis by SESI-HRMS as a tool to perform pharmacometabolomics studies.

Zusammenfassung

Hintergrund: Pharmakometabolomik ist die Erforschung von Arzneimitteln und ihrer Wirkung auf Metaboliten im menschlichen Körper. Durch die Beobachtung der Pharmakometabolomik kann man Kenntnisse über die patientenabhängige Wirkungsweise und Wirksamkeit von Arzneimitteln gewinnen und so Rückschlüsse auf die personalisierte Medizin ziehen. Diese Analyse kann in verschiedenen Körperflüssigkeiten durchgeführt werden. Einer der schnellsten Wege, diese Informationen zu erhalten, ist jedoch die Analyse der Atemluft. Sie bietet einen Einblick in den menschlichen Stoffwechsel auf nicht-invasive, schnelle und kostengünstige Weise. Ein neuartiger Ansatz ist die Verwendung der sekundären Elektrospray-Ionisierung in Verbindung mit der hochauflösenden Massenspektrometrie (SESI-HRMS). Im Rahmen dieser Arbeit werden wir zwei Anwendungen der personalisierten Medizin untersuchen, bei denen der Atem durch SESI-HRMS genutzt wird. Dies sind Epilepsie und Diabetes Typ 1.

Ziel: Epilepsie und Diabetes Typ 1 gehören zu den häufigsten chronischen Krankheiten weltweit, sind mit hohen medizinischen Kosten verbunden und ihre Überwachung erfordert invasive Methoden. Darüber hinaus ist die genaue Pathophysiologie dieser Krankheiten immer noch umstritten. Frühere Arbeiten unserer Forschungsgruppe berichteten über das therapeutische Arzneimittelmonitoring (TDM) für das Anti-Epileptikum (ASM) Valproinsäure (VPA) und identifizierten durch Atemanalyse Metaboliten, die an der Wirkung und an den Nebenwirkungen von ASMs beteiligt sind. Die Identifizierung erfolgte jedoch nur mit einem geringen Maß an Vertrauen. In dieser Arbeit wird daher versucht, diese Metaboliten eindeutig zu bestimmen. Außerdem wurde die Methode in Echtzeit durchgeführt, was bedeutet, dass die Patienten direkt in ein Gerät ausatmen müssen, was für (kognitiv) eingeschränkte Patienten schwierig ist. Zudem wollen wir diese Echtzeit-Methode in eine patientenfreundliche Offline-Methode umwandeln. Des Weiteren wollen wir die anabole Wirkung von Insulin bei Patienten mit diabetischer Ketoazidose (DKA) untersuchen für ein besseres Verständnis der Pathophysiologie und um DKA überwachen zu können.

Methoden: Für die Atemmessungen wurde SESI-HRMS verwendet. Bei Patienten, die nicht in der Lage waren Echtzeitmessungen durchzuführen oder zu Auswertungszwecken wurden Nalophan-Beutel zum Auffangen der Atemluft verwendet, deren Inhalt anschließend in das SESI-

HRMS eingespeist wurden. Zur Bestätigung der Messbarkeit von Metaboliten, die der Wirkung und Nebenwirkung von ASMs beteiligt sind, wurde Ausatemkondensat (EBC) verwendet, das anschließend durch Ultra-Hochleistungs-Flüssigkeitschromatographie, gekoppelt mit HRMS, analysiert wurde. Anschließend wurden die Daten mit statistischen und Machine Learning Methoden analysiert.

Ergebnisse: Durch die Verwendung von EBC konnten wir sechs Aminosäuren in der Atemluft eindeutig mit einer "Level 1"-Konfidenz bestätigen, von denen zuvor berichtet wurde, dass sie an der Wirkung und an den Nebenwirkungen von ASMs beteiligt sind. Zudem wurde die Durchführbarkeit der Umstellung eines bestehenden TDM-Ansatzes durch Atemluft auf eine Offline-Methode erfolgreich geprüft. Die über einen Zeitraum von 3,5 Jahren durchgeführten Messungen zeigten eine gute Übereinstimmung in Bezug auf die Genauigkeit, wiesen jedoch eine allgemeine Abnahme der Signalintensität bei Verwendung der Nalophan-Beutel auf. Darüber hinaus korrelierten freies VPA und Gesamt-VPA mit einem Lin's Konkordanzkorrelationskoeffizienten (CCC) von 0,60 bzw. 0,47. Metaboliten, von denen berichtet wurde, dass sie an der Wirkung und an den Nebenwirkungen von ASM beteiligt sind, zeigten schlechte Lin's CCC-Werte, was darauf hindeutet, dass diese Metaboliten durch die Nalophan-Beutel weitgehend verloren gehen. Des Weiteren waren wir in der Lage, klare Stoffwechselwege bei DKA-Patienten nach Insulineinleitung aufzuzeigen. Die Analyse der Atemluft zeigte, dass die bekannten veränderten Stoffwechselwege in der Atemluft messbar waren und weitgehend mit zuvor beschriebenen Beobachtungen übereinstimmten. Aceton und Acetoacetat wurden als Metaboliten identifiziert, mit denen das Fortschreiten der DKA überwacht werden kann, wobei Acetoacetat in der Atemluft noch nicht beschrieben wurde. Ausserdem wiesen die DKA-Patienten nach Beginn der Insulinbehandlung einen Stoffwechselweg auf, der dem von Patienten entsprach, die als "kontrollierter Diabetes" oder "nicht-diabetisch" eingestuft wurden.

Schlussfolgerung: Diese Arbeit lieferte neue Einblicke in die Möglichkeiten der Atemanalyse zur Durchführung von Pharmakometabolomik in einer realen klinischen Umgebung. Wir haben Aminosäuren identifiziert, die Berichten zufolge an der Wirkung und an den Nebenwirkungen von ASMs beteiligt sind. Dies ermöglicht es, in Zukunft zuverlässig Informationen aus diesen Metaboliten abzuleiten und mehr Konfidenz in eine zuvor entwickelte Risikobewertung zu haben. Zudem haben wir gezeigt, dass freie VPA und Gesamt-VPA mit dem Offline-SESI-HRMS gemessen werden können, so dass auch bettlägerige und (kognitiv) eingeschränkte Personen von

dieser Technik profitieren können. Ausserdem haben wir einen Rahmen für die Krankheitsüberwachung und das pathophysiologische Verständnis der DKA geschaffen. Dieses Verfahren könnte in Zukunft auf der Intensivstation am Krankenbett eingesetzt werden, um den Krankheitsverlauf der DKA zu überwachen und das pathophysiologische Verständnis der DKA zu vertiefen. Zusammenfassend konnten wir die Leistungsfähigkeit der Atemanalyse mit SESI-HRMS als Instrument zur Durchführung von Pharmakometabolomik-Studien zeigen.

Contents

Chapter 1: Introduction

1.1 Evolution of Medicine

Throughout history, medicine has taken many forms. The first evidence of medicine dates back to the Neanderthals, who showed evidence of recovering from injuries that would not have been possible without care[1]. Other examples include primitive dental care from 3000 BCE[2] and a tablet containing medical prescriptions from 1800 BCE[3]. A big turning point, however, came at roughly around 400 BCE, because of a shift in conceptual thinking started by Hippocrates, credited as the father of modern medicine. Hippocrates stated that diseases could originate from natural causes rather than supernatural causes[4]. More promising historical results originated in the 15th century, when Leonardo Da Vinci studied embryos and contributed greatly to the further understanding of human anatomy[5]. Throughout history, albeit some forms of medicine were different to modern methods, and not always optimal, bloodletting[6] being a great example here, the goal has always remained the same: to relieve pain and suffering, promote health, prevent—and cure—diseases whenever possible, and care for those who cannot be cured. This fundamental goal, together with technological advancements, has shifted modern medicine towards evidence-based medicine, that: the process through which caring for patients creates the need for clinically important information about the diagnosis, prognosis, therapy and healthcare issues of patients[7].

One branch of evidence-based medicine is commonly performed by the so-called -omic fields. The -omic fields want to understand a person's phenotype by studying their genetic information, ribonucleic acids (RNA), proteins, and metabolites amongst others, to make a data-based decision about their health status and possible treatment plan. While genomics focuses on a person's genetics, which provides information that encodes RNA, transcriptomics deals with the information obtained from RNA, which encodes the synthesis of proteins. Proteomics, on the other hand, is concerned with the information available in the proteins present in the human body. Finally, metabolomics studies small molecules called metabolites (small molecules with a molecular weight <2000 Da) that are present in the human body, with the ultimate goal of identifying these metabolites, understanding their specific functions, and determining their relationship to a particular phenotype[8]. The function of metabolites is to make energy and make up the materials needed for the maintenance, reproduction, and growth of cells[9]. Individuals with a specific

phenotype or disease may have different concentrations of certain metabolites active in their body and these can therefore be used as a source of information for diagnosis or the improvement of treatment plans.

When the study of metabolomics is done in the context of medicinal use, the subfield is called pharmacometabolomics[10]. This field looks at the interactions of pharmaceuticals with their surrounding metabolites and aims to infer information about the effect of those pharmaceuticals on the metabolites present, which will ultimately aid evidence-based medicine.

1.2 Pharmacometabolomics

Although pharmacometabolomics is a field of study of modern times, the effects of pharmaceuticals on the human body have been studied since the Middle Ages. During those times, smelling, tasting, and viewing bodily fluids were used to infer information about the effect of treatment[11]. Nowadays, highly sophisticated instrumentation, such as mass spectrometry (which we will touch upon in chapter 2), is used instead. Within this thesis, we will focus on pharmacometabolomics utilizing a special body fluid: breath. Breath is a very attractive bodily fluid to study since it is available in almost unlimited quantities, its analysis is non-invasive, it does not require sample pretreatment—depending on the analysis—, and its analysis is relatively cheap[12]. It provides a window into the human metabolism with the aforementioned advantages. Many successes have been reported on its applications[13 14 15 16 17 18].

Figure 1 gives an overview of the scientific location we are situated in. As mentioned before, the study of the human body for evidence-based medicine falls under the realm of the -omic fields and can be attributed to the study of the phenotype. Metabolomics is a subfield thereof. We perform breath analysis, a subfield of metabolomics, with the goal of pharmacometabolomics in mind.

The goal of pharmacometabolomics is to aid individualized—or personalized—medicine. The goal of personalized medicine has a few main components, namely to avoid adverse drug reactions, maximize drug efficacy and select responsive patients[19]. This means that some drugs are very selective, other drugs may be less selective and can be harmful. With personalized medicine we can minimize these unwanted interactions by dosing optimally, and this can be achieved by performing pharmacometabolomics. We ultimately want to infer information from breath to advance our understanding of the effect of drugs on the metabolism. However, before we can do that, we have to first infer information out of breath itself.

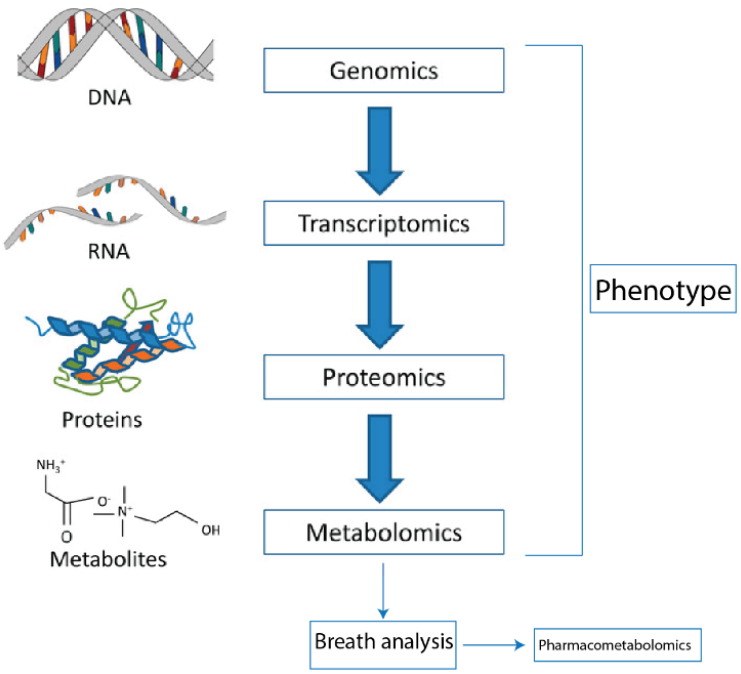


Figure 1 Cascade of different -omic levels. Breath analysis is a subfield of metabolomics. This thesis focuses on pharmacometabolomics within the realm of breath analysis. Adapted from Haukaas et al. [20]

1.3 Breath as a source of information

Obtaining clinically relevant information from breath presupposes the use of a detector, such as the human nose. The human nose can be considered a "detector" because it is able to analyze and identify substances. The earliest probable records of the human nose being used to perform identification is circa 400 BCE, when Hippocrates uttered that the entire youth of ancient Greece needed to have a pleasant breath, as this was believed to be related to a kind and gentle heart and pure soul[21]. The first time that actual instrumentation was used to extract information about the composition of breath was by the French chemist Antoine Lavoisier. He was able to extract CO_2 out of breath and by doing so prove that the body produces CO_2 and consumes O_2[22]. Roughly 200 years later, it was double Nobel laureate Linus Pauling who showed that breath contains a lot more information than solely CO_2[23]. Albeit no further conclusion could be drawn from these observations; breath analysis was added to the list of bodily fluids that can be studied and used for clinical purposes.

We now know that breath contains more information than previously assumed. Modern breath analysis focuses on what kind of information breath contains and for what purpose this information can be used. Therefore, we need to know what kind of information breath contains in general. Most

of its composition reflects the composition of air. It roughly consists of 75% N_2, 13% O_2, 6% H_2O, 5% CO_2, and 1% volatile organic compounds (VOCs)[24]. These VOCs are a product of the metabolic processes in the body, which is where the focus of modern breath analysis lies. By analyzing these VOCs, we can identify metabolic markers that can be associated with certain diseases, conditions, or the effect of a certain treatment. For example, we can identify markers associated with colorectal cancer[18], diabetes[25], and chronic obstructive pulmonary disease[26].

To analyze these VOCs in breath, high-end instrumentation is used, such as selected ion flow tube mass spectrometry (SIFT-MS)[27], proton transfer reaction mass spectrometry (PTR-MS)[28], secondary electrospray high-resolution mass spectrometry (SESI-HRMS)[29], gas chromatography-mass spectrometry (GC-MS)[30], and electric noses[31]. We will briefly touch upon these instrumentations and their successes as it is relevant to put the thesis into the context of the field of breath analysis. However, it must be noted that in this thesis only SESI-HRMS has been used which will be discussed in detail in chapter 2.

1.4 Overview of commonly used techniques

VOCs in our breath are metabolites and contain information about the biological processes that underly our bodily function. Several techniques have been developed to tap into the information of these metabolites. In this section we will discuss those most relevant to breath analysis.

1.4.1 Real-time and off-line analysis

Within the breath analysis realm, two approaches are being used for detecting breath: real-time (i.e., online) analysis[32 33 34 35 36] and off-line analysis[30 37]. Real-time analysis requires patients to exhale into a breath analyzer, which includes methods such as SESI-HRMS[38 39], SIFT-MS[27 40 41], PTR-MS[28 42], and electric noses[31]. Off-line methods, on the other hand, capture breath in e.g., bags (made of various materials), and, after possible pre-treatment, infused into a mass spectrometer with a specific ionization source. Off-line methods range from applications such as GC-MS[43 44], to methods that utilize filter substrates which are then detected by low-temperature plasma ionization coupled to mass spectrometric technique (LTP-MS)[45].

The application in real-time is not always feasible in clinical settings with patient's incapable of performing the required exhalation maneuvers. These include immobile (e.g., geriatric or too weak) patients and infants. The benefits include fast turnaround times, and there is no need for

sample preparation. Limitations for off-line analysis often include the necessity for sample preparation which can compromise sample integrity, and long analysis runs.

Within this thesis, we will consider both approaches. In study II we look at the feasibility of transitioning an existing real-time method to an offline method, because of the reasons mentioned before. What's more, in study III, we were not able to perform real-time measurements on patients residing in the ICU.

Lastly, it must be noted, that the golden standard is real-time analysis because, the overall ability to capture metabolic information decreases significantly when utilizing a bag system to capture the information[46 47].

1.4.2 SIFT-MS

One well-established and most commonly used technique to perform real-time breath analysis is SIFT-MS. It works by generating reagent ions by a microwave-induced plasma, in turn; these reagent ions ionize breath VOCs. Commonly used reagent ions include H_3O^+, NO^+, and O^+. The benefit of these reagent ions is that they do not react rapidly with the major breath constituents (N_2, O_2, CO_2). A downstream mass spectrometer is then used to detect the ionized VOCs and infer information about their abundance[40]. The gas-phase ion chemistry is well-understood with SIFT-MS, which allows for absolute quantification, as opposed to other real-time techniques with poorly understood quantification, such as SESI. The downside of SIFT-MS is, however, that for exhaled breath, humidity has to be taken into consideration which makes absolute quantification more cumbersome[40]. Lastly, SIFT-MS is not an ambient ionization technique, meaning it is not adaptable with current high-resolution instruments such as an Orbitrap. The low-resolution makes it nearly impossible for untargeted metabolomics due to the inability to separate isobaric compounds. Reported successes with SIFT-MS range from the study of aspartame metabolism[48] to the catabolism of lipids[49].

1.4.3 PTR-MS

PTR was initially mainly used for environmental research with the use of bags. Due to its advantages, such as high sensitivity, reproducibility, and ability to (semi)quantify, it has transitioned into the field of breath analysis. As the name implies, the ionization mechanism relies on protons transferring to an analyte. The most common reagent ion used is H_3O^+, which, in turn,

ionizes analytes. These ions are generated by H_2O in a hollow cathode discharge. Then, this reagent ion plume is introduced into a drift tube in which the ionization of gaseous analytes occurs. Unfortunately, PTR-MS is considered a hard ionization technique leading to in-source fragmentation caused by energetic collision in the drift tube[50]. The compounds that are ionized are limited by PTR's own ionization mechanism; only analytes that have a higher proton affinity than that of H_2O will be ionized. To overcome this, reagent ions such as NO^+ and O_2^+ can be used[51]. Despite these limitations, PTR-MS is extremely sensitive and in this regard remains superior to SIFT-MS. Successes of PTR-MS range from the quantification of propofol[52] to the detection of lung cancer[53].

1.4.4 SESI-HRMS

In this section, we will only briefly touch upon SESI-HRMS because it will be more extensively covered in the method section. SESI is a derivative method from electrospray ionization (ESI). It was first introduced by H. Hill and coworkers[29 54] when, back in 1989, Nobel laureate John Fenn realized that gas phase molecules can be ionized with ESI[55]. Subsequent work from R. Zenobi and P. Sinues further explored the feasibility to apply SESI-HRMS to breath analysis[56 57]. In SESI-HRMS, ions are created by a nano-electrospray made of acidified water. The electrospray fume is mixed with gaseous samples such as breath, after which ionization occurs by ligand-switching reactions[58]. SESI-HRMS offers great sensitivity and modularity with downstream high-resolution mass spectrometry, —plus, as it is a soft ionization tool—it allows for predominately protonated species, simplifying compound identification. In addition, the adoptability with HRMS with fragmentation capabilities allows for the use of exhaled breath condensate (EBC) as a complementary technique to unequivocally identify compounds—an aspect we will see in study I. SESI-HRMS has successfully been used in pharmacokinetic studies[59] and for therapeutic drug management of anti-seizure medication[60].

1.4.5 GC-MS

GC-MS is a commonly used off-line breath collection method and is considered the gold standard for off-line breath analysis. The first time that solely GC was used for identification of breath was when the before-mentioned Linus Pauling identified that breath contains much more information than previously thought. Due to its upstream chromatographic separation, it offers the possibility

for unambiguous compound identification as well as great selectivity. However, as mentioned, off-line techniques come with their drawbacks, such as long analysis run-time and a necessity for sample preparation, and storage, which can compromise sample integrity[61]. The applications of GC-MS' range from drug monitoring of propofol[62] to the detection of lung cancer[63].

1.4.6 Electric noses

Electric noses are an upcoming technology. The detection method relies on sensors and for each application different sensors can be used. These sensors range from mini mass spectrometers to conducting polymers. They are trained to detect diseases from a distinguished breath-print. Although they have the advantage of being small and portable, they lack resolving power for compound identification and need considerable method development time[64 65]. Successes have been reported from monitoring ketosis to gastritis[66].

1.5 Motivation and aim

Breath contains valuable information, specifically within the field of metabolomics. The study of specific drug interactions and their effects on metabolites falls within the realm of pharmacometabolomics, with the ultimate goal being to enhance personalized medicine. This thesis focuses on two different applications of personalized medicine, namely epilepsy and type-1 diabetes. The pursuit of personalized medicine is of significant public interest because epilepsy and diabetes are among the most common diseases in the world and are accompanied by tremendous medical costs[67]. Yet the current treatment plans for diabetes and epilepsy management are invasive, expensive, and time-intensive[68]. In addition, the exact pathophysiology of these diseases is still poorly understood[69 70]. Breath analysis has the potential to resolve these limitations since a wide array of metabolites have been identified in breath[71 72 73], and, in diabetes, these metabolites play a key role in enabling pathophysiological understanding. Furthermore, as previously mentioned, analyzing breath is non-invasive, quick, and cheap in comparison with other techniques, and breath itself being available in almost unlimited quantities, makes it a very attractive bodily fluid to study.

That being said, unambiguous compound identification is a bottleneck within metabolomics studies, including breath, and many identified compounds remain at a low level of confidence. For

this reason, in study I, we aim to univocally confirm metabolites previously reported[60] to be involved in response to and side effects from ASM.

In addition, this study[60] reports the ability to perform TDM on ASM. It is, however, done in a real-time fashion, making it cumbersome for patient's incapable of performing exhalation maneuvers to benefit from the technique. Therefore, in study II we will assess the feasibility of transferring this technique to a patient friendly off-line approach. Within type-1 diabetes, study III, we look at how exhaled metabolites respond to insulin after patients have fallen into a diabetic ketoacidosis state in order to advance our understanding of the capability to monitor and to further understand its pathophysiology.

1.6 Thesis Summary

In chapter two, we will outline the specific objectives and the overall goals of this thesis with regards to the existing gaps in the literature. Chapter three will focus on the methods and information that are needed to understand the chapters following. Chapter four, study I, focuses on the confirmation of metabolites previously identified to be involved in response to and side effects from ASM. This identification is necessary in the context of trying to translate breath analysis into clinical practice. In chapter 5, study II, we will continue within the realm of epilepsy and look at one specific ASM, valproic acid (VPA), and the TDM aspect of it. In chapter 6, study III, we will transition to the usefulness of breath in the context of diabetes type 1. We will explore how breath analysis, using SESI-HRMS, can be beneficial for monitoring and understanding the pathophysiology of type-1 diabetes. In the last chapter, chapter 7, we will conclude and discuss the thesis.

Chapter 2: Objectives

The overall aim of this thesis was to further develop the understanding of how breath analysis with SESI-HRMS can contribute to TDM for ASM, and how it could assist in the understanding and monitoring of type-1 diabetes. This to eventually assist clinicians with clinical decisions regarding type-1 diabetes and epilepsy. In what follows, we will outline the focus of each study in more detail.

2.1 Study I

Breath analysis has the potential to provide valuable information about an individual's health status through the detection of VOCs. However, there are significant gaps in our understanding of the identity of these VOCs, which hinders the translation of breath analysis into meaningful clinical interpretations[74] [12]. The identification of tentatively identified metabolites is crucial for advancing the field, but most untargeted metabolic studies utilizing real-time breath analysis have low levels of confidence[60]. The steps for communicating the confidence of tentatively identified molecules are outlined by Schymanski et al[75]. There are five levels of confidence, ranging from the lowest level (—the exact mass of interest)—to level 1 (confirmation of the structure of a molecule by a reference standard). The goal of tentatively reported metabolites is to be confirmed at level 1, however most untargeted metabolic studies utilizing real-time breath analysis remain on level 5 or level 4, typically due to the lack of suitable methods for compound identification. A study from our lab, which was reported to be able to perform TDM on ASM, can be used as an example[60]. This study showed that certain metabolites are up- or down-regulated when patients suffer from side effects or do not respond to medication. However, the identification was only level 4, an assigned molecular formula. As mentioned, for these studies to be translated into clinics and contribute to personalized medicine, compounds must be unequivocally confirmed.

Various methods have been proposed for the identification of VOCs in breath, each with its benefits and limitations. For example, methods utilizing sensors (such as electric noses) rely on pattern matching or complementary techniques to identify compounds[76]. Meanwhile, methods like GC-MS have the potential to confirm the structure of a molecule by a reference standard but have long analysis run times and sample integrity can be compromised due to the required sample preparation. Ideally, a real-time technique would be used, such as SIFT-MS, PTR-MS, or SESI-HRMS. One

drawback of utilizing SIFT-MS and PTR-MS is the low resolution associated with the technique, leading to overlapping peaks from isobars. This can be remedied by the simultaneous use of three different precursor ions available for chemical ionization. However, the overlying ions are compounded when isomers have to be identified. For this, the combination of GC and SIFT-MS has been proposed[77]. The drawback of this method is that compounds must be preselected, making it cumbersome for large-scale, untargeted metabolic studies.

One method that address these challenges is the analysis of EBC by the use of ultra-high-performance liquid chromatography (UHPLC) coupled to MS in combination with SESI-HRMS[78]. EBC analysis itself compromises sample integrity and has long analysis run times, however, when combined with a real-time technique such as SESI-HRMS, it can be a powerful confirmation tool. Therefore, in study I, the goal is to confirm the presence of metabolites previously reported[60] to be involved in response to and side effects from ASM in EBC by UHPLC-MS/MS to confidently infer information with a real-time approach.

2.2 Study II

In Study II we focus on ASM VPA and its therapeutic drug management. Drugs that need to be therapeutically managed have toxic effects when dosed too highly and are ineffective when dosed too low, therefore they need to be in a therapeutic range. Current TDM practices rely on blood analyses, but with a narrow window, large inter-subject variability, and analytical fluctuations it makes the TDM of these types of drugs challenging. In addition, the methods are invasive, expensive, and time intensive, requiring 1–3 days of analysis time[60]. Breath analysis has been seen as an attractive alternative because of its non-invasiveness and speed. VPA in particular has been extensively researched; as such, there is a lot of evidence that TDM for VPA is possible with the use of breath analysis. Current methods that have been published are SESI-HRMS[59, 79], low-temperature plasma mass spectrometry (LTP-MS) with[80] and without[81] the use of a substrate filter, and the analysis of EBC with a solid phase extraction method (SPE)[82, 83]. In addition, our lab recently published a paper in which we showed the possibility to perform TDM on VPA and in which we provided risk estimates for patients suffering from side effects or patients who were non-responding to medication[60]. It must be noted that real-time VPA analysis for children below the age of 4, the cognitively impaired, or the elderly is cumbersome, if not impossible. Techniques that do not have these limitations are off-line methods, nevertheless, the analysis of EBC and use of

substrate compromises sample integrity, require a long analysis time, and in addition, have no reported results on the ability to provide risk estimates for response to and side effects from ASM. Therefore, we propose a hybrid approach: breath samples from patients are collected off-line, but the mass spectrometric analysis is carried out with a real-time instrument. By doing so, we aim to explore the feasibility of off-line SESI-HRMS (collection with a bag system) for determining the blood concentration of ASM VPA, response to and side effects from ASMs in breath by comparing with the gold standard of real-time SESI-HRMS. The goal would thus be to transition off-line SESI-HRMS into clinics to be used with patients who are not capable of performing the required real-time exhalation maneuvers.

2.3 Study III

DKA, is a life-threatening and acute complication of diabetes. It is typically characterized by ketone body formation, hyperglycemia, and metabolic acidosis—an in-depth description will be presented in the next chapter. Currently, DKA diagnosis relies on blood analysis of blood glucose, blood pH, and blood base excess. These tests can be expensive, impractical, and even painful[68]. In addition, from 2009 to 2014 the annual DKA hospitalization rate increased by >6%, whereas the in-hospital mortality rate decreased. Although this trend is not fully understood, it is most likely because of a better understanding of the pathophysiology of DKA and the adoption of DKA treatment guidelines. Even though the continued decline of in-hospital DKA mortality is encouraging, a better understanding of the pathophysiology and evidence-based measures might help to slow or reverse hospitalization rates[84].

Breath analysis is an attractive tool for investigating the above-mentioned aspects. Previous reports have shown that breath analysis by SESI-HRMS captures large areas of the metabolism of the human body[85]. A portion of the metabolome measurable by SESI-HRMS breath analysis plays a significant role in DKA (examples of which include butanoate metabolism[86], tricarboxylic acid (TCA) cycle[87], amino acid metabolism[88, 89], and fatty acid metabolism[90]). The metabolites that play an important role have also been broadly confirmed (examples of which include amino acids[91], TCA cycle [72], ketones[92], and fatty acids[73]). This raises the question of which insights we can extract from breath analysis to aid the monitoring and pathophysiological understanding of DKA.

For this reason, we investigate the anabolic effect of insulin on metabolites captured in exhaled breath during acute DKA.

Chapter 3: Methods

In order to measure exhaled molecules, a specific device which operates sensitively, selectively, has high scan speeds, and is robust, is required. Such a device could be a mass spectrometer. For a mass spectrometer to work, exhaled molecules must be ionized. The ionization must also be highly efficient to reach high sensitivity without any pre-concentration step. A technique that meets this requirement is SESI, which is used together with HRMS to perform breath pharmacometabolomics within the realms of this thesis. After the analysis of exhaled compounds, one can start to infer information about the clinical state of a person with the use of statistical methods for personalized medicine. This section aims to give a broader understanding of the methodologies used within the thesis to reach the previously evaluated aims. For a more in-depth explanation, the reader is referred to the cited literature.

3.1 Ionization sources

In the next paragraph, we will summarize how a mass spectrometer works. Mass spectrometry is a technique to analyze matter. However, before, matter can be analyzed, it has to be ionized, i.e., charged. This ionization can happen in several ways. For the purpose of this thesis, we will only discuss the following ionization forms: ESI–which is used in combination with liquid chromatography, and SESI, which is used for breath analysis.

3.1.1 Electrospray ionization

ESI is a very common ionization technique because of its high sensitivity, its soft ionization (meaning that predominantly protonated or deprotonated ions are formed), and its compatibility with aqueous solutions. For the latter reason, it is often used in combination with liquid chromatography. Ionization occurs by applying a high voltage to a liquid sample which is introduced through a narrow capillary. The electric field causes a series of processes which leads to ions. The exact ionization mechanism is not yet fully understood, but the most widely accepted theory is that due to the electric field, fine aerosols of charged droplets are formed. As the charged droplets move towards the inlet of the mass spectrometer, the solvent around the aerosols start to evaporate and as a consequence, their size is reduced and the charge density enlarged. The droplets shrink in size until the surface tension can no longer sustain the Coloumb force and a “Coloumb

explosion" occurs, causing even smaller droplets to form. The solvent around the droplets evaporate further until individual ions are created. The ions then created enter the mass spectrometry to be subsequently detected. Through this process, ions ranging from small molecules (e.g., metabolites) to large biomolecules (e.g., proteins) can be created.

3.1.2 Secondary Electrospray Ionization

With conventional ESI the ionization takes place directly; with SESI it happens indirectly. A sample, such as breath, is directly mixed with the ion cloud coming from a nano-ESI source. A clean solvent—usually consisting of H_2O together with an additive (such as 0.1% formic acid) to increase conductivity—is electrosprayed to generate primary ions. Analytes, such as breath, are introduced to the plume where it was believed that the dominant ionization occurs through proton transfer reaction. Recent work by Španěl et al[58] [32] [33] however, aimed to further the understanding of the ionization mechanism, casting doubts on the believed ionization mechanism by proton transfer. Both experimental and theoretical approaches showed that the ion chemistry largely involves ligand switching which especially favors polar analytes, whereas it rarely occurs for non-polar molecules. This specific ionization mechanism results in the widely variable sensitivities observed and is thought to be dependent on the dipole moment and proton affinity of a molecule. In addition, an unknown factor must play a role in the ionization chemistry.

Therefore, the exact ionization mechanism, as with ESI, is yet to be fully understood, making absolute quantification cumbersome. The process, however, like ESI, is soft, resulting in predominantly primary ions. SESI is found to be the most sensitive to polar compounds and can reach up to sub-ppqv concentrations[93]. Moreover, SESI operates at ambient pressure, allowing compatibility with high-resolution instruments such as time-of-flight or Orbitrap.

3.2 Mass spectrometry

Mass spectrometry is a widely used technique in clinical chemistry that allows for the identification, quantification, and profiling of matter. Unlike spectroscopy, which separates light, mass spectrometry separates mass. The method is highly sensitive, meaning it can detect even trace amounts of a substance (up to parts per trillion, depending on the ionization source).

To separate mass, individual compounds must first be ionized, or charged, as we have seen previously seen with ESI and SESI. Mass spectrometry uses a combination of electric fields and/or

magnets to guide these ions, and these subsequently separate and detect them. It has proven to be a valuable tool for scientists and researchers in a wide range of fields, such as environmental monitoring, pharmaceutical development, and forensic analysis.

There are different types of mass spectrometers, such as quadrupoles, time-of-flight instruments and Orbitraps. We will discuss quadrupoles and Orbitraps as these will be relevant for the thesis.

3.2.1 Quadrupole

A quadrupole is a specific but common mass analyzer. Quadrupoles offer low resolving power and accuracy but are relatively cheap, simple to use, and robust. A quadrupole is mostly used as a mass filter in high-resolution instruments. It consists of four parallel rods with space in between to guide ions. To achieve this, direct current (DC) and radio frequency (RF) voltages are applied to the opposite polarities of the rods. This is illustrated in Figure 2. An alternating current is applied between the rods which creates an oscillating electric field. Subsequently, ions are passed through the oscillating electric field in which positive ions are attracted to negative rods and vice versa. The alternating current will result in a stable trajectory for ions, depending on their m/z ratio. Manipulating the current can therefore lead to different m/z being stable and therefore being analyzed. Which ions are stable within a specific alternating current is shown in a Mathieu diagram or can be explained by the Mathieu equation (Equation 1). The steeper the slope passing through the stability regions, the higher the resolution, but the lower the signal intensity as fewer ions are allowed for transmission. Usually, quadrupole systems are operated in a so-called triple quadrupole/tandem set-up. As the name implies, three sets of quadrupoles are used: two quadrupoles are used for mass filtering with the mechanism described above, and one is an RF voltage only for fragmentation by collision-induced dissociation (CID). Triple quadrupoles can operate in different modes with different identification or quantification goals. Generally, the goal of a triple quadrupole set-up is to achieve high confidence in identification together with quantification. This is done through multiple reaction monitoring. In this set-up, one specific m/z is selected in the first quadrupole, after which CID is applied in the collision cell. This will create an array of fragments, and the analyte of interest will create key fragments, i.e., fragments that are specific for that distinct analyte which can thereafter be analyzed by a detector such as a Faraday cup. In this thesis, we will only see a single quadrupole in action because in breath analysis mostly volatile, low molecular masses are seen, usually in the range of <200 Da. The quadrupole is used

as a mass filter to limit the range of masses measured. The fragmentation for study I is not done in a triple quadrupole but in a higher-energy C-trap dissociation (HCD) cell, which is a CID technique specific to an Orbitrap instrument.

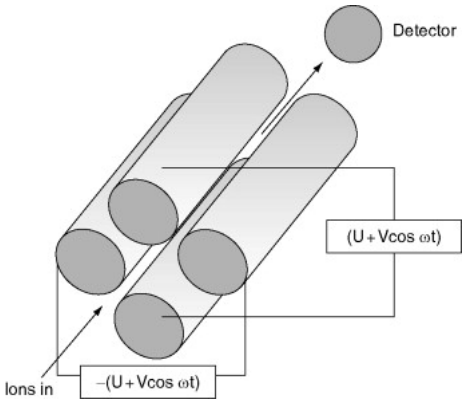


Figure 2 Schematic diagram of a quadrupole mass analyzer. Adopted from F.A. Mellon [94]

$$\frac{m}{z} = K\frac{V}{r^2 w^2}$$

Equation 1 Mathieu Equation, m = mass of ion, z = charge, K = constant, V = voltage, r = effective distance between electrodes, w = oscillation frequency

3.2.2 Orbitrap

The Orbitrap is a specific mass analyzer with high resolution and high mass accuracy. It has many applications ranging from metabolomics to environmental analysis. It was first introduced by Alexander Makarov in 2000, based on the principle of orbital trapping introduced by Kingdon. The Orbitrap is a variant of an ion trap and consists of a central rod-like electrode and two outer barrel-like electrodes that create a trap for the ions. Ions are confined in a so-called C-trap before being injected into the electric field of the Orbitrap. The ions are trapped in the electric field as the electrostatic attraction to the inner electrode is balanced by their inertia. The ions start to cycle around the inner electrode in elliptical trajectories. In addition to this circular motion, the ion also moves back- and forward in a helix-like motion. Each ion, depending on its m/z, will have a unique trajectory. This trajectory can be detected by measuring the image current on the outer electrodes. By applying a fast Fourier transform on the signals, the m/z ratio can be determined up to 500,000 mass resolution. Orbitraps are often accompanied by HCD cells, which allows for fragmentation

capabilities. The fragmentation of specific analytes can later be used as an identification tool, as we have used in Study I. A schematic overview of an Orbitrap is given in Figure 3.

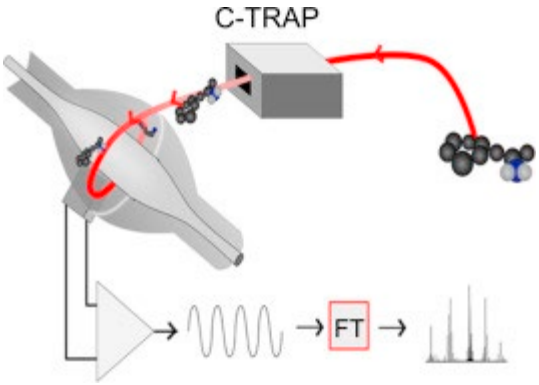


Figure 3 Schematic overview of the workings of an Orbitrap. Adopted from Beck et al [95]

3.3 Liquid chromatography

Liquid chromatography is a powerful analytical technique used to separate and identify different compounds in a sample. The separation is based on the principle of affinity, which refers to the specific binding interactions between an immobilized agent, known as the stationary phase, and its binding partner in the sample. Simply put, liquid chromatography involves pumping the sample or analyte in a specific liquid (mobile phase) through a column that contains a particular binding material or stationary phase packed into it. As the analyte moves through the column, it begins to interact with the stationary phase and temporarily binds to it. The strength of the interaction or bonding between the analyte and the stationary phase depends on the physical-chemical properties of the stationary phase and the analyte. Therefore, for each different application, a new method must be developed, or an existing one adapted. As the mobile phase continues to flow, the interaction between the analyte and the stationary phase weakens, and the analyte is released back into the mobile phase. This process of binding and release may occur multiple times during the time in the column—how often it occurs is related to the plate number. A higher plate number means higher separability between analytes and thus more interaction. The retention time or the time that the analyte spends in the column depends on the strength and frequency of its interaction with the stationary phase. When analytes are finally separated, they can be identified and quantified in an upstream detector such as a mass spectrometer. For this thesis, liquid chromatography has been used to separate and identify compounds in EBC.

3.4 Data processing

Most modern analytical chemistry techniques generate large amounts of data. The art of analyzing this data is called chemometrics or bioinformatics, if the context is chemical or biological respectively. Breath analysis, and more specifically real-time breath analysis by SESI-HRMS, produces two types of data structures, time traces, which is the sum of all ion currents, i.e., every signal a mass spectrum contains in one scan, and the underlying information in the total ion current, the mass spectrum. Usually, the acquisition of this data happens in both positive SESI+ and negative SESI-modes. Both the time traces and the mass spectra are used to get the dataset in the right format. After acquiring the data, the raw data is used to align and calibrate the mass spectra by using reference peaks. After alignment, initial peak picking is performed using a kernel density function. Breath fractions are identified and only the mass spectra coming from the breath fractions (i.e., exhalations) are used to average the signal intensities. Subsequently, one obtains an n*x matrix, in which n represents the number of observations (i.e., patients) and x represents the number of features. The data is often $\log_{10}$ transformed to approach a gaussian distribution allowing for statistical testing which assumes normality. A downstream analysis is then performed by using conventional statistics and machine learning algorithms. An overview of this process is shown in Figure 4. The exact pre-processing strategy for each study can be found in the corresponding manuscript.

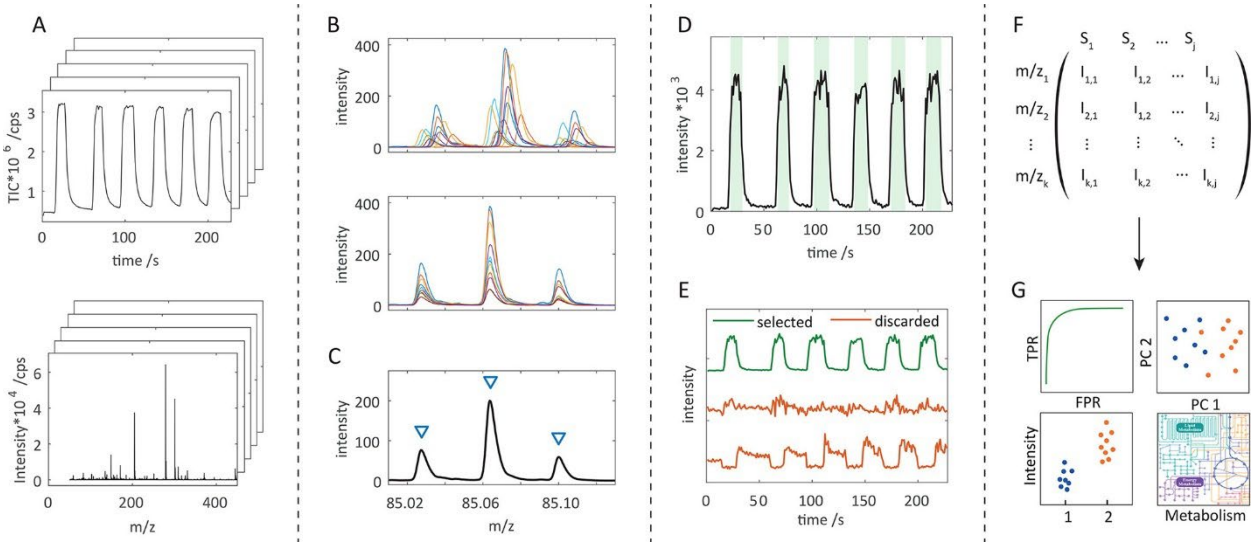


Figure 4 Overview of workflow of SESI-HRMS analysis. A) Type of data structures (time traces above, mass spectrum below). B) Alignment using reference peaks. C) Peak-picking. D) Extraction of regions of interest containing exhalations. E) Features selection. F) Generation of the final data structure. G) Analysis and interpretation of the generated data structure. Adopted from Bruderer et al. [78]

3.5 Statistical analysis

We have now seen, in a broad sense, how metabolic information can be detected out of breath and processed to infer clinically relevant information. However, statistical testing is required to be able to make sound statements about the quality and meaning of your data. In this section we will briefly touch upon statistical testing relevant to the thesis.

3.5.1 Hypothesis testing

Hypothesis testing lies at the core of scientific inquiry. This type of statistical analysis puts your assumptions about a population parameter to the test. Frequently undertaken in comparative studies such as healthy versus diseased, but has application in most scientific areas. The central question to be asked during hypothesis testing, is whether two populations, e.g., diseased versus healthy, have the same distribution. An example of such populations can be found in study III, in which we compare patients undergoing DKA versus patients who are in-control. The prerequisite for this test is that both populations should have the same type of distribution. The null hypothesis, H_0, is the base: it assumes that two groups have the same distribution. The alternative hypothesis, H_1, assumes that two distributions are not equal. A wide range of tests are available to test your hypothesis, with the student's t-test being the standard.

3.5.2 T-test

With the t-test you can assess whether two populations have different means, the prerequisite being however, that your data is equally distributed. It is therefore important to first test for normality by using, for example, the Kolmogorov-Smirnov test. The t-test is determined by calculating the difference between two samples' means and dividing it by the standard error of the difference. This value is then compared to a critical value from a t-distribution (Equation 2). If the test statistic is higher than the critical value, the null hypothesis can be rejected. Thereby, we can conclude that the data has two different means, and as such are two independent entities.

$$t = \frac{X_1 - X_2}{S_p \times \sqrt{\frac{1}{n_1}} + \sqrt{\frac{1}{n_2}}}$$

Equation 2 t-test with equal variances. X_1 and X_2 are sample means. S_p represent the estimate of the pooled standard deviation of the two samples.

3.5.3 ANOVA

T-tests will function properly when you have two distinguishable groups. However, in metabolomics, we often have several groups which need comparing at the same time. For example, in Study III, we not only investigate patients with diabetic ketoacidosis, but also diabetic patients who are in control, and non-diabetic patients. If we want to compare several groups at the same time, we can perform the analysis of variance (ANOVA), through which we can test multiple groups at the same time for their differences in means. ANOVA is similar to the previously mentioned t-test, but it allows for comparison among multiple different groups. The test statistic is calculated by dividing the variance between groups by the variance within groups and then compare it to the critical value from the F-distribution.

3.5.4 Machine learning

Machine learning is a tool to analyze data. It was first used in the 1940s but has since then, through the increase in computing power and methodological development, gained popularity. Machine learning models use algorithms and statistical models to analyze and draw inferences from patterns in data. This is popular with high dimensional data, and is hence attractive for metabolomics. Machine learning models can be separated into three categories: supervised, unsupervised, and semi-supervised learning. With supervised machine learning a labeled data set is used to train machine learning algorithms. These algorithms include the commonly used linear regression but extend to more complex models such as neural networks and support vector machines. Unlike supervised learning, no labeled data are being used with unsupervised machine learning. The algorithms discover hidden patterns in data without the need for human intervention. Probably the most used tool is principal component analysis (PCA), a tool we will use in Study III. With this method, the goal is to reduce the number of variables, and to detect patterns in the remaining components. Finally, semi-supervised learning lies between supervised and unsupervised learning. During the training of the model, a smaller labeled dataset is used to guide the classification and feature extraction from a larger unlabeled data set. It mostly sees its use when a dataset is too exhaustive to label or if one does not have enough data.

3.6 Clinical considerations

After analyzing breath by means of highly sophisticated instrumentation, computational processing, and statistical analysis, the data at hand needs to be explained from a clinical

perspective. To be able to do so, we need to understand how molecules end up in our breath and how diseases work. In this section, we will briefly touch upon this in the context of the thesis.

3.6.1 Physiology of blood–lung barrier

Different diseases cause different metabolites in different concentrations to be present in the blood. For compounds to be detectable in breath and to be able to infer the previously mentioned source of information, they have to diffuse through the air-blood barrier of the lung. The air-blood barrier is a relatively thin surface, which allows for molecular diffusion. Figure 5 shows a schematic representation of the air-blood barrier. Whether diffusion occurs depends solely on the physical chemical properties of a compound and at the moment, there is no consensus about the exact mechanism. One important property determining whether a compound will diffuse through the air-blood barrier is described by Henry's law constant. Henry's law constant describes the vapor pressure of a compound over an ideal solution. This law characterizes its lipophilicity, which will tell us how likely it is that the molecule will diffuse through the air-lung barrier. This has two reasons; firstly, the alveolar membrane is made of collagen, which is non-polar, and secondly, non-metabolized drugs or metabolites are usually non-polar by nature. Additionally, lipophilic compounds tend to accumulate at the surface of liquid solutions and vaporize more efficiently.

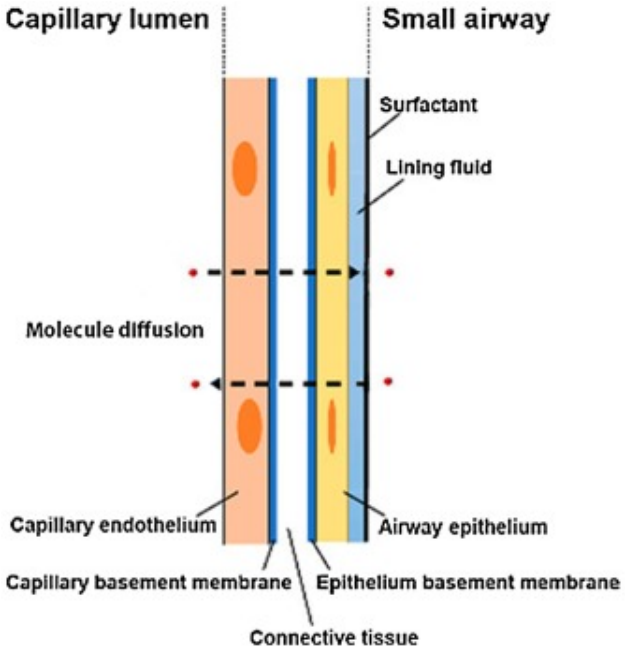


Figure 5 Representation of the air-blood barrier within the lung. The figure displays particle transport between the alveolar space. Adopted from Lopez-Lorente et al. [96]

3.6.2 Epilepsy

Epilepsy is one of the most common chronic neurological conditions in the world. It is estimated that at least 0.6% of Europe's population suffers from epilepsy[97]. It is characterized by neuronal disorder leading to epileptic seizures. There are many different types of seizures, and they vary in severity and duration. Which often occur because of a sudden and temporary imbalance in neuronal activity. The most important neurotransmitters in epilepsy are glutamate and gamma-aminobutyric acid (GABA); while glutamate acts as a main stimulant, GABA acts as a suppressor. ASMs can be used to modulate brain activity to prevent seizures, mainly affecting the neurotransmitter systems previously mentioned. These ASMs must usually be therapeutically managed, meaning they have a certain window in which their mode of action is optimal. However, non-responders to medication, and side-effects are commonly reported. Optimizing medication usually relies on trial and error. Moreover, epilepsy can have many causes, including genetics, head injury, and brain tumors. In some cases, the causes are unknown making it difficult to understand the exact physiology.

3.6.3 Type-1 Diabetes & diabetic ketoacidosis

Type-1 diabetes is characterized by insulin deficiency. It occurs when the body's immune system mistakenly destroys insulin-producing cells in the pancreas. It is usually diagnosed by measuring the patients' pH-levels, blood glucose levels, and base excess (BE) value. The metabolic process which governs type-1 diabetes is the inability to produce energy from the TCA cycle because the glycolysis and gluconeogenesis pathways are blocked; in other words, the body is unable to use sugars for energy production (due to insulin deficiency). In order to provide energy, it reverts to fatty acid breakdown, i.e., lipolysis. As a byproduct of lipolysis, ketone bodies are formed. In the case of total depletion of insulin, the patient comes into a clinical state called DKA. It usually includes extremely low blood pH-levels, high glucose levels, and BE values below -20 mmol/L. After receiving insulin, patients usually revert their metabolism to gluconeogenesis. A schematic overview of the general pathophysiology is given in Figure 6.

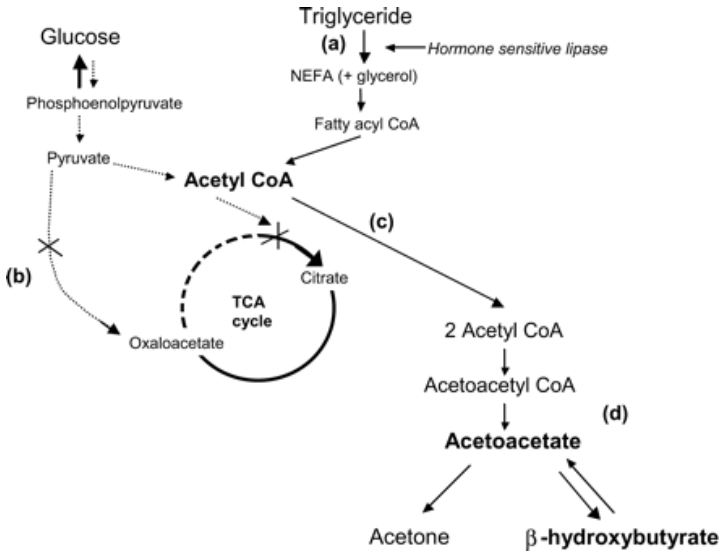


Figure 6 a) Increased lipolysis results in the production of acetyl CoA, b) The amount of available oxaloacetate is reduced, c) acetyl CoA is diverted from entering the TCA cycle, d) Acetyl CoA undergoes condensation to form acetoacetate which breaks down into acetone and beta-hydroxybutyrate. Adopted from Wallace et al[98]

References

1. Spikins, P.; Needham, A.; Tilley, L.; Hitchens, G., Calculated or caring? Neanderthal healthcare in social context. *World Archaeology* **2018,** *50* (3), 384-403.
2. Shampo, M. A.; Gibilisco, J. A., Dentists on stamps. *Journal of Clinical Pediatric Dentistry* **2007,** *31* (4), 222.
3. Cilliers, L.; Retief, F., Mesopotamian medicine: history of medicine: SAMJ forum. *South African Medical Journal* **2007,** *97* (1), 27-30.
4. Schwartz, G., Hippocrates revisited. *Einstein J. Biol. Med* **2004,** *21*, 33-34.
5. Jones, R., Leonardo da Vinci: anatomist. *British Journal of General Practice* **2012,** *62* (599), 319-319.
6. Greenstone, G., The history of bloodletting. *BC Medical Journal* **2010,** *52* (1), 12-14.
7. Sackett, D. L. In *Evidence-based medicine*, Seminars in perinatology, Elsevier: 1997; pp 3-5.
8. Schneider, M. V.; Orchard, S., Omics technologies, data and bioinformatics principles. *Bioinformatics for Omics Data: Methods and Protocols* **2011**, 3-30.
9. Newsholme, E. A.; Start, C., *Regulation in metabolism*. 1973.
10. Kaddurah‐Daouk, R.; Weinshilboum, R. M.; Network, P. R., Pharmacometabolomics: implications for clinical pharmacology and systems pharmacology. *Clinical Pharmacology & Therapeutics* **2014,** *95* (2), 154-167.
11. Kouba, E.; Wallen, E. M.; Pruthi, R. S., Uroscopy by Hippocrates and Theophilus: prognosis versus diagnosis. *The Journal of urology* **2007,** *177* (1), 50-52.
12. Singh, K. D.; Del Miguel, G. V.; Gaugg, M. T.; Ibañez, A. J.; Zenobi, R.; Kohler, M.; Frey, U.; Sinues, P. M., Translating secondary electrospray ionization–high-resolution mass spectrometry to the clinical environment. *Journal of breath research* **2018,** *12* (2), 027113.
13. Ibrahim, W.; Wilde, M. J.; Cordell, R. L.; Richardson, M.; Salman, D.; Free, R. C.; Zhao, B.; Singapuri, A.; Hargadon, B.; Gaillard, E. A., Visualization of exhaled breath metabolites reveals distinct diagnostic signatures for acute cardiorespiratory breathlessness. *Science translational medicine* **2022,** *14* (671), eabl5849.
14. Alahmadi, F. H.; Wilkinson, M.; Keevil, B.; Niven, R.; Fowler, S. J., Short-and medium-term effect of inhaled corticosteroids on exhaled breath biomarkers in severe asthma. *Journal of Breath Research* **2022,** *16* (4), 047101.
15. van Oort, P. M.; Povoa, P.; Schnabel, R.; Dark, P.; Artigas, A.; Bergmans, D. C.; Felton, T.; Coelho, L.; Schultz, M. J.; Fowler, S. J., The potential role of exhaled breath analysis in the diagnostic process of pneumonia—a systematic review. *Journal of breath research* **2018,** *12* (2), 024001.
16. Bos, L. D.; Sterk, P. J.; Fowler, S. J., Breathomics in the setting of asthma and chronic obstructive pulmonary disease. *Journal of Allergy and Clinical Immunology* **2016,** *138* (4), 970-976.
17. Miekisch, W.; Schubert, J. K.; Noeldge-Schomburg, G. F., Diagnostic potential of breath analysis—focus on volatile organic compounds. *Clinica chimica acta* **2004,** *347* (1-2), 25-39.
18. Woodfield, G.; Belluomo, I.; Laponogov, I.; Veselkov, K.; Lin, G.; Myridakis, A.; Ayrton, O.; Španěl, P.; Vidal-Diez, A.; Romano, A., Diagnostic Performance of a Noninvasive Breath Test for Colorectal Cancer: COBRA1 Study. *Gastroenterology* **2022,** *163* (5), 1447-1449. e8.
19. Beger, R. D.; Schmidt, M. A.; Kaddurah-Daouk, R., Current concepts in pharmacometabolomics, biomarker discovery, and precision medicine. *Metabolites* **2020,** *10* (4), 129.

20. Haukaas, T. H.; Euceda, L. R.; Giskeødegård, G. F.; Bathen, T. F., Metabolic portraits of breast cancer by HR MAS MR spectroscopy of intact tissue samples. *Metabolites* **2017,** *7* (2), 18.
21. Institute, B. History, religion and the breath. https://www.breathinstitute.co.uk/medical-library-bad-breath/bad-breath-through-the-ages/ (accessed 18-04-2023).
22. Dweik, R. A.; Amann, A., Exhaled breath analysis: the new frontier in medical testing. *Journal of breath research* **2008,** *2* (3), 030301.
23. Teranishi, R.; Mon, T.; Robinson, A. B.; Cary, P.; Pauling, L., Gas chromatography of volatiles from breath and urine. *Analytical chemistry* **1972,** *44* (1), 18-20.
24. Pereira, J.; Porto-Figueira, P.; Cavaco, C.; Taunk, K.; Rapole, S.; Dhakne, R.; Nagarajaram, H.; Câmara, J. S., Breath analysis as a potential and non-invasive frontier in disease diagnosis: an overview. *Metabolites* **2015,** *5* (1), 3-55.
25. Righettoni, M.; Tricoli, A.; Pratsinis, S. E., Si: WO3 sensors for highly selective detection of acetone for easy diagnosis of diabetes by breath analysis. *Analytical chemistry* **2010,** *82* (9), 3581-3587.
26. Sinues, P. M.-L.; Meier, L.; Berchtold, C.; Ivanov, M.; Sievi, N.; Camen, G.; Kohler, M.; Zenobi, R., Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease. *Respiration* **2014,** *87* (4), 301-310.
27. Smith, D.; Španěl, P., Application of ion chemistry and the SIFT technique to the quantitative analysis of trace gases in air and on breath. *International Reviews in Physical Chemistry* **1996,** *15* (1), 231-271.
28. Lindinger, W.; Jordan, A., Proton-transfer-reaction mass spectrometry (PTR–MS): on-line monitoring of volatile organic compounds at pptv levels. *Chemical Society Reviews* **1998,** *27* (5), 347-375.
29. Wu, C.; Siems, W. F.; Hill, H. H., Secondary electrospray ionization ion mobility spectrometry/mass spectrometry of illicit drugs. *Analytical chemistry* **2000,** *72* (2), 396-403.
30. Xu, M.; Tang, Z.; Duan, Y.; Liu, Y., GC-based techniques for breath analysis: current status, challenges, and prospects. *Critical Reviews in Analytical Chemistry* **2016,** *46* (4), 291-304.
31. Fens, N.; Van der Schee, M.; Brinkman, P.; Sterk, P., Exhaled breath analysis by electronic nose in airways disease. Established issues and key questions. *Clinical & Experimental Allergy* **2013,** *43* (7), 705-715.
32. Dryahina, K.; Polášek, M.; Smith, D.; Španěl, P., Sensitivity of secondary electrospray ionization mass spectrometry to a range of volatile organic compounds: Ligand switching ion chemistry and the influence of Zspray™ guiding electric fields. *Rapid Communications in Mass Spectrometry* **2021,** *35* (22), e9187.
33. Dryahina, K.; Som, S.; Smith, D.; Španěl, P., Reagent and analyte ion hydrates in secondary electrospray ionization mass spectrometry (SESI‐MS), their equilibrium distributions and dehydration in an ion transfer capillary: Modelling and experiments. *Rapid Communications in Mass Spectrometry* **2021,** *35* (7), e9047.
34. Smith, D.; Španěl, P.; Herbig, J.; Beauchamp, J., Mass spectrometry for real-time quantitative breath analysis. *Journal of breath research* **2014,** *8* (2), 027101.
35. Španěl, P.; Smith, D., Account: on the features, successes and challenges of selected ion flow tube mass spectrometry. *European Journal of Mass Spectrometry* **2013,** *19* (4), 225-246.
36. Beauchamp, J.; Herbig, J.; Dunkl, J.; Singer, W.; Hansel, A., On the performance of proton-transfer-reaction mass spectrometry for breath-relevant gas matrices. *Measurement Science and Technology* **2013,** *24* (12), 125003.

37. Beauchamp, J., Current sampling and analysis techniques in breath research—results of a task force poll. *Journal of Breath Research* **2015,** *9* (4), 047107.
38. Vidal, G., Super SESI-HRMS: Real-Time Analysis of Dynamic Metabolic Processes.
39. Liu, C.; Zeng, J.; Sinues, P.; Fang, M.; Zhou, Z.; Li, X., Quantification of volatile organic compounds by secondary electrospray ionization-high resolution mass spectrometry. *Analytica Chimica Acta* **2021,** *1180*, 338876.
40. Spanel, P.; Smith, D., Progress in SIFT-MS: breath analysis and other applications. *Mass Spectrom Rev* **2011,** *30* (2), 236-67.
41. Huang, J. Z.; Kumar, S.; Hanna, G. B., Investigation of C3-C10 aldehydes in the exhaled breath of healthy subjects using selected ion flow tube-mass spectrometry (SIFT-MS). *Journal of Breath Research* **2014,** *8* (3), 037104.
42. Aprea, E.; Cappellin, L.; Gasperi, F.; Morisco, F.; Lembo, V.; Rispo, A.; Tortora, R.; Vitaglione, P.; Caporaso, N.; Biasioli, F., Application of PTR-TOF-MS to investigate metabolites in exhaled breath of patients affected by coeliac disease under gluten free diet. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences* **2014,** *966*, 208-213.
43. Yan, Y. Y.; Wang, Q. H.; Li, W. W.; Zhao, Z. J.; Yuan, X.; Huang, Y. P.; Duan, Y. X., Discovery of potential biomarkers in exhaled breath for diagnosis of type 2 diabetes mellitus based on GC-MS with metabolomics. *Rsc Advances* **2014,** *4* (48), 25430-25439.
44. Guallar-Hoyas, C.; Turner, M. A.; Blackburn, G. J.; Wilson, I. D.; Thomas, C. L. P., A workflow for the metabolomic/metabonomic investigation of exhaled breath using thermal desorption GC-MS. *Bioanalysis* **2012,** *4* (18), 2227-2237.
45. Gong, X.; Shi, S.; Gamez, G., Real-Time Quantitative Analysis of Valproic Acid in Exhaled Breath by Low Temperature Plasma Ionization Mass Spectrometry. *J Am Soc Mass Spectrom* **2017,** *28* (4), 678-687.
46. Decrue, F.; Singh, K. D.; Gisler, A.; Awchi, M.; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of exhaled breath analysis with parallel lung function and FeNO measurements in infants. *Analytical Chemistry* **2021,** *93* (47), 15579-15583.
47. Slingers, G.; Vanden Eede, M.; Lindekens, J.; Spruyt, M.; Goelen, E.; Raes, M.; Koppen, G., Real‐time versus thermal desorption selected ion flow tube mass spectrometry for quantification of breath volatiles. *Rapid Communications in Mass Spectrometry* **2021,** *35* (4), e8994.
48. Španěl, P.; Dryahina, K.; Vicherková, P.; Smith, D., Increase of methanol in exhaled breath quantified by SIFT-MS following aspartame ingestion. *Journal of breath research* **2015,** *9* (4), 047104.
49. Smith, D.; Turner, C.; Španěl, P., Volatile metabolites in the exhaled breath of healthy volunteers: their levels and distributions. *Journal of breath research* **2007,** *1* (1), 014004.
50. Kari, E.; Miettinen, P.; Yli-Pirilä, P.; Virtanen, A.; Faiola, C. L., PTR-ToF-MS product ion distributions and humidity-dependence of biogenic volatile organic compounds. *International Journal of Mass Spectrometry* **2018,** *430*, 87-97.
51. Norman, M.; Hansel, A.; Wisthaler, A., O2+ as reagent ion in the PTR-MS instrument: Detection of gas-phase ammonia. *International Journal of Mass Spectrometry* **2007,** *265* (2-3), 382-387.
52. Takita, A.; Masui, K.; Kazama, T., On-line monitoring of end-tidal propofol concentration in anesthetized patients. *The Journal of the American Society of Anesthesiologists* **2007,** *106* (4), 659-664.

53. Bajtarevic, A.; Ager, C.; Pienz, M.; Klieber, M.; Schwarz, K.; Ligor, M.; Ligor, T.; Filipiak, W.; Denz, H.; Fiegl, M., Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC cancer* **2009,** *9* (1), 1-16.
54. Wittmer, D.; Chen, Y. H.; Luckenbill, B. K.; Hill, H. H., Electrospray ionization ion mobility spectrometry. *Analytical Chemistry* **1994,** *66* (14), 2348-2355.
55. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.
56. Sinues, P. M. L.; Kohler, M.; Zenobi, R., Human Breath Analysis May Support the Existence of Individual Metabolic Phenotypes. *Plos One* **2013,** *8* (4).
57. Martinez-Lozano Sinues, P.; Zenobi, R.; Kohler, M., Analysis of the exhalome: A diagnostic tool of the future. *Chest* **2013,** *144* (3), 746-749.
58. Španěl, P.; Dryahina, K.; Omezzine Gnioua, M.; Smith, D., Different reactivities of H3O+ (H2O) n with unsaturated and saturated aldehydes: ligand‐switching reactions govern the quantitative analytical sensitivity of SESI‐MS. *Rapid Communications in Mass Spectrometry* **2023,** *37* (9), e9496.
59. Li, X.; Martinez‐Lozano Sinues, P.; Dallmann, R.; Bregy, L.; Hollmén, M.; Proulx, S.; Brown, S. A.; Detmar, M.; Kohler, M.; Zenobi, R., Drug pharmacokinetics determined by real‐time analysis of mouse breath. *Angewandte Chemie International Edition* **2015,** *54* (27), 7815-7818.
60. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; Awchi, M.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U., Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications medicine* **2021,** *1* (1), 1-11.
61. Fang, M.; Ivanisevic, J.; Benton, H. P.; Johnson, C. H.; Patti, G. J.; Hoang, L. T.; Uritboonthai, W.; Kurczy, M. E.; Siuzdak, G., Thermal degradation of small molecules: a global metabolomic investigation. *Analytical chemistry* **2015,** *87* (21), 10935-10941.
62. Miekisch, W.; Fuchs, P.; Kamysek, S.; Neumann, C.; Schubert, J. K., Assessment of propofol concentrations in human breath and blood by means of HS-SPME–GC–MS. *Clinica chimica acta* **2008,** *395* (1-2), 32-37.
63. Poli, D.; Goldoni, M.; Corradi, M.; Acampa, O.; Carbognani, P.; Internullo, E.; Casalini, A.; Mutti, A., Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME–GC/MS. *Journal of Chromatography B* **2010,** *878* (27), 2643-2651.
64. Fitzgerald, J. E.; Bui, E. T.; Simon, N. M.; Fenniri, H., Artificial nose technology: status and prospects in diagnostics. *Trends in biotechnology* **2017,** *35* (1), 33-42.
65. Tisch, U.; Haick, H., Chemical sensors for breath gas analysis: the latest developments at the Breath Analysis Summit 2013. *Journal of Breath Research* **2014,** *8* (2), 027103.
66. Di Francesco, F.; Fuoco, R.; Trivella, M. G.; Ceccarini, A., Breath analysis: trends in techniques and clinical applications. *Microchemical journal* **2005,** *79* (1-2), 405-410.
67. Miller, G. F.; Coffield, E.; Leroy, Z.; Wallin, R., Prevalence and costs of five chronic conditions in children. *The Journal of School Nursing* **2016,** *32* (5), 357-364.
68. Minh, T. D. C.; Blake, D. R.; Galassetti, P. R., The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes research and clinical practice* **2012,** *97* (2), 195-205.
69. Chung, W. K.; Erion, K.; Florez, J. C.; Hattersley, A. T.; Hivert, M.-F.; Lee, C. G.; McCarthy, M. I.; Nolan, J. J.; Norris, J. M.; Pearson, E. R., Precision medicine in diabetes: a

consensus report from the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes care* **2020,** *43* (7), 1617-1635.
70. Sinha, S. R., Basic mechanisms of sleep and epilepsy. *Journal of Clinical Neurophysiology* **2011,** *28* (2), 103-110.
71. Bregy, L.; Nussbaumer-Ochsner, Y.; Sinues, P. M.-L.; García-Gómez, D.; Suter, Y.; Gaisl, T.; Stebler, N.; Gaugg, M. T.; Kohler, M.; Zenobi, R., Real-time mass spectrometric identification of metabolites characteristic of chronic obstructive pulmonary disease in exhaled breath. *Clinical Mass Spectrometry* **2018,** *7*, 29-35.
72. Tejero Rioseras, A.; Singh, K. D.; Nowak, N.; Gaugg, M. T.; Bruderer, T.; Zenobi, R.; Sinues, P. M., Real-Time Monitoring of Tricarboxylic Acid Metabolites in Exhaled Breath. *Anal Chem* **2018,** *90* (11), 6453-6460.
73. Gaugg, M. T.; Bruderer, T.; Nowak, N.; Eiffert, L.; Sinues, P. M. L.; Kohler, M.; Zenobi, R., Mass-Spectrometric Detection of Omega-Oxidation Products of Aliphatic Fatty Acids in Exhaled Breath. *Analytical Chemistry* **2017,** *89* (19), 10329-10334.
74. Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S., Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics* **2013,** *9*, 280-299.
75. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: communicating confidence. ACS Publications: 2014.
76. Güntner, A. T.; Pineau, N. J.; Mochalski, P.; Wiesenhofer, H.; Agapiou, A.; Mayhew, C. A.; Pratsinis, S. E., Sniffing entrapped humans with sensor arrays. *Analytical chemistry* **2018,** *90* (8), 4940-4945.
77. Kubišta, J.; Španěl, P.; Dryahina, K.; Workman, C.; Smith, D., Combined use of gas chromatography and selected ion flow tube mass spectrometry for absolute trace gas quantification. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up‐to‐the‐Minute Research in Mass Spectrometry* **2006,** *20* (4), 563-567.
78. Bruderer, T.; Gaisl, T.; Gaugg, M. T.; Nowak, N.; Streckenbach, B.; Müller, S.; Moeller, A.; Kohler, M.; Zenobi, R., On-line analysis of exhaled breath: focus review. *Chemical reviews* **2019,** *119* (19), 10803-10828.
79. Gamez, G.; Zhu, L.; Disko, A.; Chen, H.; Azov, V.; Chingin, K.; Krämer, G.; Zenobi, R., Real-time, in vivo monitoring and pharmacokinetics of valproic acid via a novel biomarker in exhaled breath. *Chemical Communications* **2011,** *47* (17), 4884-4886.
80. Gong, X.; Shi, S.; Zhang, D.; Gamez, G., Quantitative Analysis of Exhaled Breath Collected on Filter Substrates via Low-Temperature Plasma Desorption/Ionization Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2022,** *33* (8), 1518-1529.
81. Gong, X.; Shi, S.; Gamez, G., Real-time quantitative analysis of valproic acid in exhaled breath by low temperature plasma ionization mass spectrometry. *Journal of The American Society for Mass Spectrometry* **2016,** *28* (4), 678-687.
82. Jouyban, A.; Farajzadeh, M. A.; Mogaddam, M. R. A.; Nemati, M.; Khoubnasabjafari, M.; Jouyban-Gharamaleki, V., Molecularly imprinted polymer based-solid phase extraction combined with dispersive liquid–liquid microextraction using new deep eutectic solvent; selective extraction of valproic acid from exhaled breath condensate samples. *Microchemical Journal* **2021,** *161*, 105772.
83. Zhang, D.; Latif, M.; Gamez, G., Instantaneous differentiation of functional isomers via reactive flowing atmospheric pressure afterglow mass spectrometry. *Analytical Chemistry* **2021,** *93* (29), 9986-9994.

84. Benoit, S. R.; Zhang, Y.; Geiss, L. S.; Gregg, E. W.; Albright, A., Trends in diabetic ketoacidosis hospitalizations and in-hospital mortality—United States, 2000–2014. *Morbidity and Mortality Weekly Report* **2018,** *67* (12), 362.
85. Gisler, A.; Singh, K. D.; Zeng, J.; Osswald, M.; Awchi, M.; Decrue, F.; Schmidt, F.; Sievi, N. A.; Chen, X.; Usemann, J., An interoperability framework for multicentric breath metabolomic studies. *Iscience* **2022,** *25* (12), 105557.
86. Zhang, L.; Liu, C.; Jiang, Q.; Yin, Y., Butyrate in energy metabolism: there is still more to learn. *Trends in Endocrinology & Metabolism* **2021,** *32* (3), 159-169.
87. Forni, L. G.; McKinnon, W.; Lord, G. A.; Treacher, D. F.; Peron, J.-M. R.; Hilton, P. J., Circulating anions usually associated with the Krebs cycle in patients with metabolic acidosis. *Critical care* **2005,** *9* (5), 1-5.
88. Felig, P.; Marliss, E.; Ohman, J. L.; Cahill Jr, G. F., Plasma amino acid levels in diabetic ketoacidosis. *Diabetes* **1970,** *19* (10), 727-729.
89. Mathew, A. V.; Jaiswal, M.; Ang, L.; Michailidis, G.; Pennathur, S.; Pop-Busui, R., Impaired amino acid and TCA metabolism and cardiovascular autonomic neuropathy progression in type 1 diabetes. *Diabetes* **2019,** *68* (10), 2035-2044.
90. Decsi, T.; Szabó, É.; Kozári, A.; Erhardt, É.; Marosvölgyi, T.; Soltész, G., Polyunsaturated fatty acids in plasma lipids of diabetic children during and after diabetic ketoacidosis. *Acta Pædiatrica* **2005,** *94* (7), 850-855.
91. Gaugg, M. T.; Engler, A.; Bregy, L.; Nussbaumer-Ochsner, Y.; Eiffert, L.; Bruderer, T.; Zenobi, R.; Sinues, P.; Kohler, M., Molecular breath analysis supports altered amino acid metabolism in idiopathic pulmonary fibrosis. *Respirology* **2019,** *24* (5), 437-444.
92. Tassopoulos, C.; Barnett, D.; Fraser, T. R., Breath-acetone and blood-sugar measurements in diabetes. *The Lancet* **1969,** *293* (7609), 1282-1286.
93. Gould, O.; Drabińska, N.; Ratcliffe, N.; de Lacy Costello, B., Hyphenated mass spectrometry versus real-time mass spectrometry techniques for the detection of volatile compounds from the human body. *Molecules* **2021,** *26* (23), 7185.
94. Mellon, F., Mass spectrometry| principles and instrumentation. **2003**.
95. Beck, O.; Rylski, A.; Stephanson, N. N., Application of liquid chromatography combined with high resolution mass spectrometry for urine drug testing. In *Critical Issues in Alcohol and Drugs of Abuse Testing*, Elsevier: 2019; pp 321-332.
96. López-Lorente, C. I.; Awchi, M.; Sinues, P.; Gómez, D. G., Real-time pharmacokinetics via online analysis of exhaled breath. *Journal of Pharmaceutical and Biomedical Analysis* **2021**, 114311.
97. Forsgren, L.; Beghi, E.; Oun, A.; Sillanpää, M., The epidemiology of epilepsy in Europe–a systematic review. *European Journal of neurology* **2005,** *12* (4), 245-253.
98. Wallace, T.; Matthews, D., Recent advances in the monitoring and management of diabetic ketoacidosis. *Qjm* **2004,** *97* (12), 773-780.

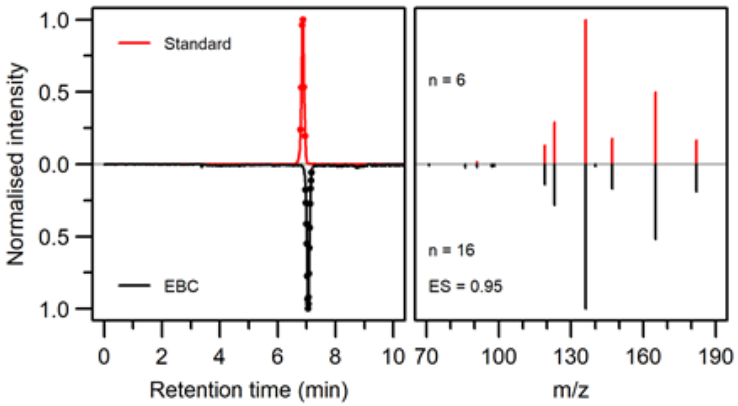
Chapter 4: Study I

UHPLC-MS/MS based identity confirmation of amino acids involved in response to and side effects from antiseizure medications

Mo Awchi[1,2], Pablo Sinues[1,2], Alexandre N. Datta[1], Diego García-Gómez[3], Kapil Dev Singh[1,2,*]

1. University Children's Hospital Basel, University of Basel, Spitalstrasse 33, 4056 Basel, Switzerland.
2. Department of Biomedical Engineering, University of Basel, Gewerbestrasse 14, 4123 Allschwil, Switzerland.
3. Department of Analytical Chemistry, University of Salamanca, Plaza de los Caídos s/n, 37008 Salamanca, Spain.

* Corresponding author: kapil.singh@ukbb.ch



Abstract

Real-time breath analysis using secondary electrospray ionization coupled to high-resolution mass spectrometry is a fast and non-invasive method to access the metabolic state of a person. However, it lacks the ability to unequivocally assign mass spectral features to compounds due to the absence of chromatographic separation. This can be easily overcome by using exhaled breath condensate and conventional LC-MS systems. In this study, to the best of our knowledge, we confirm for the first time the presence of six amino acids (GABA, Oxo-Pro, Asp, Gln, Glu and Tyr) previously reported to be involved in response to and side effects from antiseizure medications in exhaled breath condensate and by extension in exhaled human breath. Raw data are publicly available at MetaboLights with the accession number MTBLS6760.

Keywords

UHPLC, tandem mass spectrometry, compound identification, exhaled breath condensate, amino acids

Introduction

Within the last decade, real-time exhaled breath analysis has risen as a promising tool for personalised medicine. In addition to respiratory gases and water vapour, exhaled breath contains a minuscule amount of various volatile organic compounds (VOCs), of both endogenous and exogenous origin.[1] The composition and concentration of these VOCs give rise to the characteristic "breath-print" of an individual, which provides relevant metabolic information and makes exhaled breath analysis an attractive tool for clinical diagnosis. Moreover, in addition to being non-invasive and quick, the use of real-time analysis methods eliminates the sample handling and processing inconsistencies. All these advantages allow monitoring and quantification of various fatty acids,[2] amino acids,[3] TCA metabolites[4] among others in past using real-time breath analysis. Additionally, real-time breath analysis can be employed to perform drug pharmacokinetics[5] and phenotype lung diseases such as asthma[6] and chronic obstructive pulmonary disease.[7,8] Recently, we reported the successful implementation of real-time breath analysis using secondary electrospray ionization coupled to high-resolution mass spectrometry (SESI-HRMS), to provide reliable estimations of systemic drug concentrations along with risk estimates for drug response and side effects in epileptic patients.[9] Where risk estimates were based on features with differential abundance between either i) patients suffering from sides effects against no side effects or ii) patients not responding to their antiseizure medications (ASMs) against the responders. Additionally, features were assigned to compounds based accurate mass, revealing upregulation of serval amino acids (such as GABA, Asp, Gln, Glu, Pro, Lys, etc.) in patients suffering from side effects and downregulation of tyrosine metabolic pathway compounds (such as Tyr, Phe, Dopamine etc.) in non-responders.[9]

Despite a promising potential one of the main disadvantages of real-time breath analysis is its inability to unequivocally confirm mass spectral features to compound assignment, which is crucial to make it transition into clinics.[10] However, this can be overcome by using exhaled breath condensate (EBC)[11] and conventional LC-MS/MS to achieve "Level 1" compound identifications.[12] Here, we used EBC and LC-MS/MS to confirm six previously mentioned compound assignments based on accurate mass from our recent work.[9]

Experimental Section

Chemicals

Pure chemical standards of selected amino acids i.e., γ-Aminobutyric acid (GABA), 5-Oxo-proline (Oxo-Pro), Aspartic acid (Asp), Glutamine (Gln), Glutamic acid (Glu) and Tyrosine (Tyr) were purchased from Sigma (USA). All other LC solvents such as water, acetonitrile (ACN) and methanol with and without 0.1% formic acid (FA) were also purchased from Sigma. All solvents used were of at least LC-MS grade.

Exhaled breath condensate sampling and up-concentration

EBC was collected using a custom-built sampling device (containing dry ice and isopropanol[13]) based on the guidelines of ATS/ERS task force.[11] In total, 100 mL of EBC was collected after pooling samples from 6 healthy (3 males and 3 females) subjects. On average to collect 10 mL of EBC a subject had to exhale for 1.5 h. All subjects signed informed consent prior to sampling and the study was approved by the local ethics committee for Northwest and Central Switzerland (2018-01324).

Collected EBC was then up-concentrated using vendor-recommended solid phase extraction (SPE) procedure. Briefly, the Oasis hydrophilic-lipophilic balance (HLB) cartridge (6 cc, 150 mg, 30 µm) from Waters (Milford, USA) was first preconditioned using 2 mL methanol and 2 mL of water. Thereafter, 100 mL of sample was loaded on the cartridge and subsequently eluted using 2 mL of methanol without any intermediate washing. All solvents and sample were allowed to flow through the cartridge under gravity. The methanol from the eluent was evaporated under a gentle stream of nitrogen before reconstituting it in 500 µL of water (i.e., approx. 200 times theoretical up concentration).

Compound selection and standard solution

Extracted ion chromatograms (XICs) of features corresponding to compounds involved in response to and side effects from ASMs,[9] from the EBC sample, were screened to select initial set of compounds (Figure S1). Afterwards, only compounds with commercially available pure chemical standards were kept (Figure S1). All standards were first dissolved in water to create individual master stock solutions of concentration 0.1 mg/mL. Finally, all standard stock solutions were serially diluted 10000 times (except Asp, which was diluted to only 100 times) to achieve final signal between 1e6 to 1e8 orders of magnitude at the injection volume of 100 µL.

UHPLC-MS/MS

Data-dependent tandem MS (MS/MS) experiments were performed on a Q Exactive Plus mass spectrometer with heated electrospray ionization (HESI) source (Thermo Fisher Scientific, Germany) coupled to an ultra-high-performance liquid chromatography (UHPLC) system (Vanquish, Thermo Fisher Scientific, Germany). 100 µL of samples (either diluted standard solution or up-concentrated EBC) were separated on a reverse phase analytical column with embedded weak acidic ion-pairing groups (Primesep 200 from SIELC Technologies, 150 × 4.6 mm ID, 5 µm) at a flow rate of 0.7 mL/min maintained at 25 °C. Samples were eluted with a gradient between solvent A (water with 0.1% FA) and solvent B (ACN with 0.1% FA). The gradient profile was 15% solvent B between 0 and 2 min, 15-95% solvent B between 2 and 15 min, 95% solvent B between 15-20 min, followed by column re-equilibration to 15% solvent B in a total of 24 min run (including 2 mins of pre-injection equilibration).

Eluted samples were directly ionized in the HESI source (see Table S1 for tune settings). The mass spectrometer was operated in the positive polarity Full MS/dd-MS^2 (Top 5) mode with one full scan in Orbitrap (Scan range = 70-400 m/z, Resolution = 140000 (at 200 m/z), AGC target = 10^6, max. injection time = 200 ms), followed by higher-energy collisional dissociation (HCD) fragmentation of the five most intense ions (Intensity threshold = 10^4; isolation window = 1 m/z, NCE = 30, Dynamic exclusion = auto), and acquisition of MS/MS spectra in Orbitrap (Resolution = 70000 (at 200 m/z), AGC target = 1e6, max. injection time = 100 ms). Additionally, as previously recommended,[14] a static exclusion list consisting of features with average intensity higher than 1e4 from 3 blank runs was also used. The mass spectrometer was externally calibrated prior to the measurement using a commercially available calibration solution (product number 88340, Pierce™ Triple Quadrupole, extended mass range, Thermo Fisher Scientific, Germany) and internally calibrated using m/z 279.15909 (dibutyl phthalate, a common plasticizer[15,16]) as lock mass during measurements.

The raw data files are publicly available at MetaboLights (https://www.ebi.ac.uk/metabolights) repository,[17] with the accession number MTBLS6760.

Data analysis

The data for XICs and MS/MS spectra plotted in various figures were directly retrieved from the RAW files using an in-house C# console app based on RawFileReader (version 5.0.0.38) an open-

source .Net assembly from Thermo Fisher Scientific. Afterwards, all figures were plotted using R (version 4.1.0).[16]

Results and Discussion

Detection of amino acids in exhaled breath by SESI-HRMS

Recently, we showed that in epileptic patients, real-time breath analysis could be used to provide risk estimates for response to and side effects from antiseizure medication (ASM).[9] However, these estimates were based on mass-to-charge (m/z) ratios (i.e. features) assigned to compounds from various amino acid metabolic pathways based solely on measured accurate mass, which promoted us to caution readers until unambiguous chemical identification is provided (for more details see Figure 6a and Supplementary Data 6 from[9]). Here, we selected six amino acids (Table 1) previously reported to be involved in response to and side effects from ASMs, for further compound confirmation based on UHPLC-MS/MS. Figure 1A shows a typical real-time SESI-HRMS measurement setup, whereby a subject exhale 5-6 times directly into a SESI source coupled to an HRMS. The whole procedure usually lasts 5 minutes per polarity mode. Figure 1B shows the real-time traces for features assigned to the adduct of the six compounds of interest from one measurement of previous study as an example.

Compound		Identification		
ID	Name	m/z (positive polarity)	Adduct	Confidence code
C00334	Gamma aminobutyric acid (GABA)	104.07053	B	1
C01879	5-Oxoproline (Oxo-Pro)	147.07629 | 130.04994 | 113.02330 | 112.03930	A | B | C | D	NA | 1 | NA | NA
C00049	L-Aspartic acid (Asp)	134.04489 | 116.03422	B | D	1 | 2d
C00064	L-Glutamine (Gln)	147.07629 | 130.04994 | 129.06587	B | C | D	1 | 2d | NA
C00025	L-Glutamic acid (Glu)	148.06032 | 131.03395 | 130.04994	B | C | D	1 | NA | 2d
C00082	L-Tyrosine (Tyr)	182.08125 | 165.05467 | 164.07056	B | C | D	1 | 2d | NA

In Adduct: A = $[M + NH_4]^+$, B = $[M + H]^+$, C = $[M - NH_3 + H]^+$ and D = $[M - H_2O + H]^+$
In Confidence code: 1 = Level 1 (RT, MS/MS, reference standard) and 2d = Level 2d (RT, reference standard), NA = Not assigned

Table 1. Amino acids and different m/z features assigned to them, selected for further confirmation by UHPLC-MS/MS analysis.

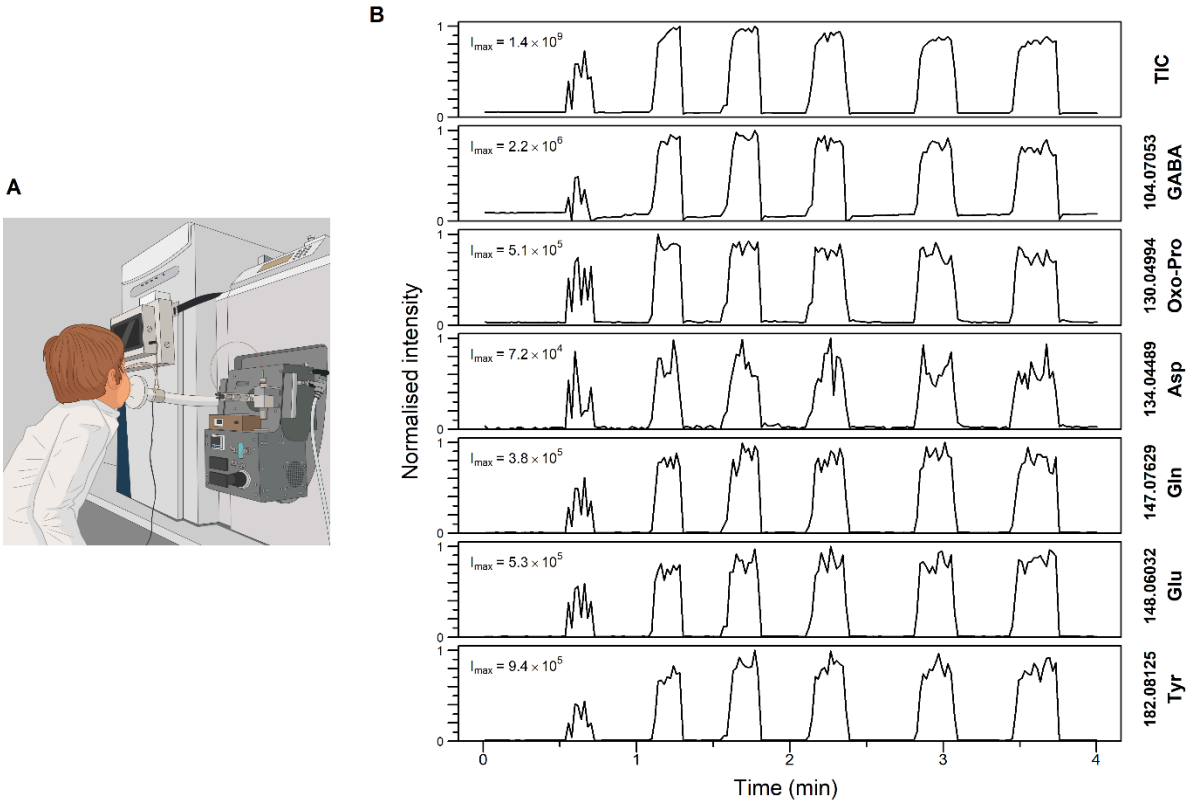


Figure 7. Detection of selected amino acids in exhaled breath. (A) Pictorial representation of a patient exhaling into SESI-HRMS analytical setup. (B) Total ion current and traces of features assigned to protonated molecule of six selected amino acids as measured in real-time exhaled breath.

Confirmation of amino acids presence in EBC

Different features were previously assigned to six selected amino acids owing to different adduct forms (Table 1). Initially, we focused on the protonated molecule, whereby matching retention time (RT) and MS/MS spectra between pure chemical standards and EBC in an UHPLC-MS/MS experiment unequivocally confirmed previous features to compound assignment (Figure 2). Spectrum similarity score, between the MS/MS spectra originating from standards and EBC was calculated for all compounds based on spectral entropy similarity (ES) method.[18] Additionally, we compared the retention time of other adduct forms between standards and EBC for all six compounds (Figure S2). For some adduct forms, due to low signal intensity, no clear peak was observed either in standards (Figure S2A,I) or in EBC (Figure S2K) or sometimes in both (Figure S2C,D,O). Suggesting the possibility that these ions are unstable in the conventional HESI settings as compared to the real-time environment of the SESI source. Table S1, shows the comparison of

tune settings between the HESI source used during UHPLC-MSMS analysis of this study and the SESI source used during real-time breath analysis from the previous study.[9] Although, most of the source-specific tune settings have little to no impact on the ion stability, recently it was shown that higher capillary temperature leads to "hard" ionization.[19] Furthermore, whenever a clear peak was observed for different adduct forms in the standard of the same compound, it had the same retention time, indicating that adduct formation happens after chromatography, perhaps in-source. In addition to the aforementioned confirmation of protonated forms, Figure S2F confirms m/z 116.03422 as dehydrated Asp, whereas Figure S2N confirms m/z 165.05467 as loss of ammonia adduct of Tyr.

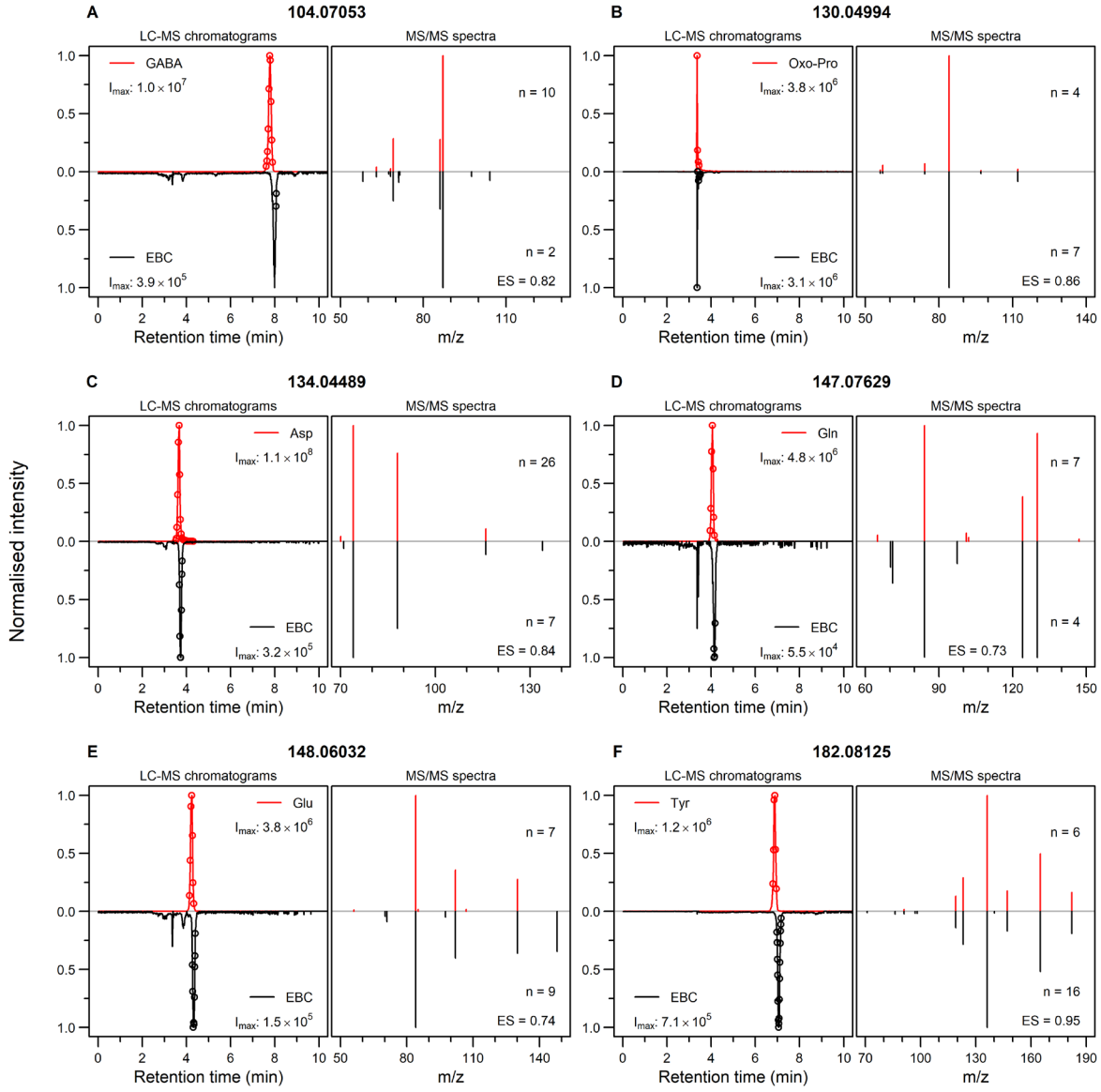


Figure 8. UHPLC-MS/MS confirmation of protonated molecule of selected amino acids. The figure shows the compound identification via UHPLC-MS, we confirmed that m/z 104.07053 belongs to protonated GABA (A), m/z 130.04994 belongs to protonated Oxo-Pro (B), m/z 134.04489 belongs to protonated Asp (C), m/z 147.07629 belongs to protonated Gln (D), m/z 148.06032 belongs to protonated Glu (E) and m/z 182.08125 belongs to protonated Tyr (F). Each panel show the compression of LC-MS chromatograms (i.e. XICs) and the average MS/MS spectra between the pure standards of suspected molecules and EBC. Number of MS/MS spectra used to obtain averaged MS/MS spectrum is denoted by n and the time at which those MS/MS spectra were triggered is denoted by the open circles in the corresponding LC–MS chromatograms. MS/MS spectrum similarity score based on spectral entropy similarity between standards and EBC is denoted by ES.

From Table 1, it is clear that m/z 130.04994 and 147.07629 were assigned to more than one compound based on different adduct forms. Furthermore, XIC for m/z 147.07629 in EBC has two resolved peaks (Figure 2D and Figure S3, same EBC trace). Where peak on the right have been established to corresponds to protonated Gln (Figure 2D). Unfortunately, as mentioned above, no clear peak was observed in the standard for the ammonium adduct of Oxo-Pro (Figure S2A and Figure S3A) but interestingly, the left peak has the same retention time as the standard of protonated Oxo-Pro. Hence, based on the fact that theoretically, ammonium adduct of Oxo-Pro will also have the same retention time as the protonated form of Oxo-Pro, we deduced that left peak in EBC probably corresponds to ammonium adduct of Oxo-Pro (Figure S3A). Similarly, in the measured EBC sample m/z 130.04994 mainly corresponds to protonated Oxo-Pro. However, a small fraction of this feature may also reflect dehydrated Glu and/or loss of ammonia adduct of Gln (Figure S3B).

This promoted us to look back into the real-time breath signal of these two features from the previous epileptic dataset,[9] and as now expected, we observed two distinct clusters in the data breath levels of m/z 147.07629 (Figure S3A inset in EBC trace). Furthermore, matching with the UHPLC data, breath levels for m/z 130.04994 show indistinct clusters (Figure S3B inset in EBC trace). These observations re- emphasize that by itself real-time SESI-HRMS is unable to resolve in-source fragments and adducts of different compounds that leads to same feature, hence one must remain cautious while working with uncharacterised features. Furthermore, this also necessitate the use of chromatographic separation for accurate feature characterization.[20]

Conclusion

As mentioned in Table 1, we concluded that for the selected compounds all of the protonated molecules can be confirmed at “Level 1” confidence (i.e. matching RT and MS/MS spectra with reference standards). Whereas, for other assignments, we could either reached to what we labelled in good agreement with Schymanski et al.,[12] “Level 2d” confidence (i.e., only matching RT with reference standards) or not assigned (NA) confidence. It must be noted that not assigned confidence does not mean a wrong assignment, it merely reflects the fact that we were unable to confirm the assignment. Additionally, one must not forget that during a real-time SESI measurement, single mass spectral feature could represent two or more adduct forms of different compounds as observed here in the case of m/z 147.07629 and 130.04994. Finally, simple

confirmatory studies like this are essential to develop and maintain a comprehensive breath features to compound database, which will ultimately help breath research to transition from bench to market.

Supporting Information

Table S1. Comparison of tune settings between the HESI source used during UHPLC-MSMS analysis of this study and the SESI source used during real-time breath analysis from the previous study.

Figure S1. Workflow used to select compounds for UHPLC-MSMS based confirmation.

Figure S2. Comparison of LC-MS chromatograms between standards of six selected compounds and EBC for different adduct forms, as defined in Table 1.

Figure S3. Features at m/z 147.07629 and 130.04994 appear to be associated with different adduct forms of more than one compound.

Author Contributions

P.S. and K.S. designed the study. M.A. performed the experiments. M.A. and K.S. analysed the data. A.N.D and D.G.G. provided guidance. K.S. wrote the manuscript with contributions of all authors. All authors have given approval to the final version of the manuscript.

Competing Interests

PS is a cofounder of Deep Breath Intelligence AG (Switzerland), which develops breath-based diagnostic tools. KS is a consultant for Deep Breath Intelligence AG (Switzerland). All other authors declare no competing interests.

Acknowledgments

We thank the participants who volunteered to take part in this study and the study nurses' team (Amelia Imolesi, Maya Weber and Isabel Gonzalez Novoa) for coordinating and maintaining study documents. Mélina Richard is gratefully acknowledged for her research coordination. This work is part of the Zurich Exhalomics project under the umbrella of University Medicine Zurich/Hochschulmedizin Zürich. PS received funding from Fondation Botnar (Switzerland) and the Swiss National Science Foundation (PCEGP3_181300).

References

1. Pereira, J.; Porto-Figueira, P.; Cavaco, C.; Taunk, K.; Rapole, S.; Dhakne, R.; Nagarajaram, H.; Câmara, J. S. Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview. Metabolites 2015, 5 (1), 3–55. https://doi.org/10.3390/metabo5010003.
2. Gaugg, M. T.; Bruderer, T.; Nowak, N.; Eiffert, L.; Martinez-Lozano Sinues, P.; Kohler, M.; Zenobi, R. Mass-Spectrometric Detection of Omega-Oxidation Products of Aliphatic Fatty Acids in Exhaled Breath. Anal. Chem. 2017, 89 (19), 10329–10334. https://doi.org/10.1021/acs.analchem.7b02092.
3. García-Gómez, D.; Gaisl, T.; Bregy, L.; Cremonesi, A.; Sinues, P. M.-L.; Kohler, M.; Zenobi, R. Real-Time Quantification of Amino Acids in the Exhalome by Secondary Electrospray Ionization-Mass Spectrometry: A Proof-of-Principle Study. Clin. Chem. 2016, 62 (9), 1230–1237. https://doi.org/10.1373/clinchem.2016.256909.
4. Tejero Rioseras, A.; Singh, K. D.; Nowak, N.; Gaugg, M. T.; Bruderer, T.; Zenobi, R.; Sinues, P. M.-L. Real-Time Monitoring of Tricarboxylic Acid Metabolites in Exhaled Breath. Anal. Chem. 2018, 90 (11), 6453–6460. https://doi.org/10.1021/acs.analchem.7b04600.
5. López-Lorente, C. I.; Awchi, M.; Sinues, P.; García-Gómez, D. Real-Time Pharmacokinetics via Online Analysis of Exhaled Breath. J. Pharm. Biomed. Anal. 2021, 205, 114311. https://doi.org/10.1016/j.jpba.2021.114311.
6. Weber, R.; Micic, S.; Streckenbach, B.; Welti, L.; Bruderer, T.; Perkins, N.; Inci, D.; Usemann, J.; Möller, A. Identification of Disease-Specific Molecular Breath Profiles in Patients with Allergic Asthma. In Paediatric asthma and allergy; European Respiratory Society, 2020; p 1213. https://doi.org/10.1183/13993003.congress-2020.1213.
7. Martinez-Lozano Sinues, P.; Meier, L.; Berchtold, C.; Ivanov, M.; Sievi, N.; Camen, G.; Kohler, M.; Zenobi, R. Breath Analysis in Real Time by Mass Spectrometry in Chronic Obstructive Pulmonary Disease. Respiration. 2014, 87 (4), 301–310. https://doi.org/10.1159/000357785.
8. Bregy, L.; Nussbaumer-Ochsner, Y.; Martinez-Lozano Sinues, P.; García-Gómez, D.; Suter, Y.; Gaisl, T.; Stebler, N.; Gaugg, M. T.; Kohler, M.; Zenobi, R. Real-Time Mass Spectrometric Identification of Metabolites Characteristic of Chronic Obstructive Pulmonary Disease in Exhaled Breath. Clin. Mass Spectrom. 2018, 7 (January), 29–35. https://doi.org/10.1016/j.clinms.2018.02.003.
9. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; Awchi, M.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U.; Datta, A. N.; Sinues, P. Personalised Therapeutic Management of Epileptic Patients Guided by Pathway-Driven Breath Metabolomics. Commun. Med. 2021, 1 (1), 21. https://doi.org/10.1038/s43856-021-00021-3.
10 Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial. Metabolomics 2013, 9 (2), 280–299. https://doi.org/10.1007/s11306-012-0482-9.
11.) Horváth, I.; Hunt, J.; Barnes, P. J.; Alving, K.; Antczak, A.; Baraldi, E.; Becher, G.; van Beurden, W. J. C.; Corradi, M.; Dekhuijzen, R.; Dweik, R. A.; Dwyer, T.; Effros, R.; Erzurum, S.; Gaston, B.; Gessner, C.; Greening, A.; Ho, L. P.; Hohlfeld, J.; Jöbsis, Q.; Laskowski, D.; Loukides, S.; Marlin, D.; Montuschi, P.; Olin, A. C.; Redington, A. E.; Reinhold, P.; van Rensen, E. L. J.; Rubinstein, I.; Silkoff, P.; Toren, K.; Vass, G.; Vogelberg, C.; Wirtz, H.; ATS/ERS Task Force on Exhaled Breath Condensate. Exhaled Breath Condensate: Methodological Recommendations and Unresolved Questions. Eur. Respir. J. 2005, 26 (3), 523–548. https://doi.org/10.1183/09031936.05.00029705.

12. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ. Sci. Technol. 2014, 48 (4), 2097–2098. https://doi.org/10.1021/es5002105.
13. Phipps, A. M.; Hume, D. N. General Purpose Low Temperature Dry-Ice Baths. J. Chem. Educ. 1968, 45 (10), 664. https://doi.org/10.1021/ed045p664.
14. Defossez, E.; Bourquin, J.; Reuss, S.; Rasmann, S.; Glauser, G. Eight Key Rules for Successful Data-dependent Acquisition in Mass Spectrometry-based Metabolomics. Mass Spectrom. Rev. 2021, No. April, 1–13. https://doi.org/10.1002/mas.21715.
15. Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M. Interferences and Contaminants Encountered in Modern Mass Spectrometry. Anal. Chim. Acta 2008, 627 (1), 71–81. https://doi.org/10.1016/j.aca.2008.04.043.
16. Weber, R. J. M.; Li, E.; Bruty, J.; He, S.; Viant, M. R. MaConDa: A Publicly Accessible Mass Spectrometry Contaminants Database. Bioinformatics 2012, 28 (21), 2856–2857. https://doi.org/10.1093/bioinformatics/bts527.
17. Haug, K.; Cochrane, K.; Nainala, V. C.; Williams, M.; Chang, J.; Jayaseelan, K. V.; O'Donovan, C. MetaboLights: A Resource Evolving in Response to the Needs of Its Scientific Community. Nucleic Acids Res. 2020, 48 (D1), D440–D444. https://doi.org/10.1093/nar/gkz1019.
18. Li, Y.; Kind, T.; Folz, J.; Vaniya, A.; Mehta, S. S.; Fiehn, O. Spectral Entropy Outperforms MS/MS Dot Product Similarity for Small-Molecule Compound Identification. Nat. Methods 2021, 18 (12), 1524–1531. https://doi.org/10.1038/s41592-021-01331-z.
19. Kaeslin, J.; Wüthrich, C.; Giannoukos, S.; Zenobi, R. How Soft Is Secondary Electrospray Ionization? J. Am. Soc. Mass Spectrom. 2022, 33 (10), 1967–1974. https://doi.org/10.1021/jasms.2c00201.
20. Xu, Y.-F.; Lu, W.; Rabinowitz, J. D. Avoiding Misannotation of In-Source Fragmentation Products as Cellular Metabolites in Liquid Chromatography-Mass Spectrometry-Based Metabolomics. Anal. Chem. 2015, 87 (4), 2273–2281. https://doi.org/10.1021/ac504118y.

Chapter 5: Study II

Prediction of systemic free and total valproic acid by off-line analysis of exhaled breath in epileptic children and adolescents

Mo Awchi[1,2], Kapil Dev Singh[1,2], Patricia E. Dill[1], Urs Frey[1], Alexandre N. Datta[1,*], Pablo Sinues[1,2,*]

1. University Children's Hospital Basel, Basel, Switzerland.

2. Department of Biomedical Engineering, University of Basel, Basel, Switzerland.

*pablo.sinues@unibas.ch

*alexandre.datta@ukbb.ch

Keywords: Breath analysis, Therapeutic Drug Monitoring, Valproic acid, Epilepsy, SESI-HRMS, Off-line analysis, Anti-Seizure Medication, Personalized Medicine

Abstract

Therapeutic drug monitoring (TDM) of medications with a narrow therapeutic window is a common clinical practice to minimize toxic effects and maximize clinical outcomes. Routine analyses rely on the quantification of systemic blood concentrations of drugs. Alternative matrices such as exhaled breath are appealing because of its inherent non-invasive nature. This is especially the case for pediatric patients. We have recently showcased the possibility of predicting systemic concentrations of valproic acid (VPA), an anti-seizure medication (ASM) by real-time breath analysis in two real clinical settings. This approach, however, comes with the limitation of the patients having to physically exhale into the mass spectrometer. This restricts the possibility of sampling from patients not capable or available to exhale into the mass spectrometer located on the hospital premises.

In this work, we developed an alternative method to overcome this limitation by collecting the breath samples in customized bags and subsequently analyzing them by secondary electrospray ionization coupled to high-resolution mass spectrometry (SESI-HRMS). A total of n = 40 patients (mean ± SD, 11.5 ± 3.5 y.o.) diagnosed with epilepsy and taking VPA were included in this study. The patients underwent three measurements: i) serum concentrations of total and free VPA, ii) real-time breath analysis and iii) off-line analysis of exhaled breath collected in bags. The agreement between the real-time and the off-line breath analysis methods was evaluated using Lin's Concordance Correlation Coefficient (CCC). CCC was computed for ten mass spectral predictors of VPA concentrations. Lin's CCC was > 0.6 for all VPA-associated features, except for two low-signal intensity isotopic peaks. Finally, free and total serum VPA concentrations were predicted using the off-line data set. Lin's CCC between actual and predicted serum concentration was 0.47 and 0.60 for total and free-VPA, respectively. As a secondary analysis, we explored whether exhaled metabolites previously associated with side-effects and response to medication could be rendered by the off-line analysis method. We found that five features associated with side effects showed a CCC > 0.6, whereas none of the drug response-associated peaks reached this cut-off. We conclude that the clinically relevant free fraction of VPA can be predicted by this combination of off-line breath collection with rapid SESI-HRMS analysis. This opens new possibilities for the breath-based TDM in clinical settings.

Introduction

Breath analysis holds promise as an attractive tool in a clinical context[1-18]. Therapeutic drug monitoring (TDM) is one of the applications whereby breath analysis has a large impact potential. Current TDM practice mainly relies on the analysis of blood levels of drugs with a narrow therapeutic range. Such a narrow window in addition to large inter-subject variability and analytical fluctuations makes therapeutic management challenging. Blood-based TDM assists clinicians to personalize drug dosage to avoid or minimize unwanted side effects and at the same time maximize clinical effects[19]. Alternative non-invasive matrices like breath are an appealing approach for TDM, especially for the pediatric population. Among the anti-seizure medications (ASMs) requiring TDM, valproic acid (VPA), in particular, has been extensively investigated using breath analysis[20-23].

Within the breath analysis by mass spectrometry realm, there exist two distinct approaches: Real-time[24-28] and off-line analysis[29, 30]. Real-time analysis requires patients to directly exhale into the mass spectrometer, whereas off-line techniques utilize a variety of gas collection devices to subsequently analyze the breath specimen by mass spectrometry. One main advantage of real-time analyses is that the biochemical information carried in the sample is less likely to be altered as a result of the sample collection, manipulation and analysis process[31]. However, this technique limits the possibility to screen a larger population. The subjects need to exhale directly into the mass spectrometer, which younger and intellectually impaired children are not capable to do. Children are normally capable to perform a controlled exhalation maneuver after the age of 4 years, whereas children with an intellectual disability are typically much later. In addition, patients being treated in clinical settings where such instrumentation is not available on site, could not benefit either from such clinical breath tests. Therefore, in this work, we propose a hybrid approach: breath samples from patients are collected off-line, but the mass spectrometric analysis is carried out in a real-time instrument[32-34], enabling prompt results and at the same time the flexibility of remote sampling. Here, we test this concept in a targeted approach by monitoring the stability of exhaled compounds previously associated with VPA, along with the response and side effects to ASMs[35].

Materials and Methods

Real-time and off-line mass spectrometric breath analysis

Real-time measurements were performed as previously described by Singh *et al.*[35]. In short, real-time breath was analyzed by exhaling into a SESI-HRMS. The breath interface consisted of an

Exhalion (FIT, Spain) to control the exhalation maneuvers by measuring the flow rate and exhaled CO_2, maximizing the reproducibility across individuals. The interface was connected to a SESI source (FIT, Spain) coupled to an HRMS instrument (Orbitrap, Thermo Fisher, USA). The mass spectrometer was operated by using a sheath gas flow rate of 60, auxiliary gas flow rate of 2, spray voltage at 3.5 kV in positive ion mode, the capillary temperature of 275 °C, and S-lens RF level of 55. A Nano-spray was created by utilizing 0.1% formic acid in water (Sigma Aldrich), whereby the current ranged between 70 and 120 nA. SESI sampling line temperature was set at 130 °C and the ion source at 90 °C. The MS was operated in full-scan mode ranging from 100-400 m/z, which was later extended to 70-1000 m/z for increased mass spectral coverage, with a maximum inject time of 500 ms, and automatic gain control target of 10^6. The mass spectral resolution was set at 140,000 at m/z 200. The mass spectrometer was calibrated weekly and a suitability test (see [36] for details) was performed before every day of measurements.

Offline breath measurements were performed as described previously[37]. In short, a custom-made Nalophan bag was inflated with approximately 2L of breath (i.e., dead volume + end-tidal fraction) after the online measurement. The bag samples were subsequently infused into the mass spectrometer within 30 minutes of collection. The mass spectrometer operated with the same settings and set-up as described above, except for the use of the Exhalion breath interface, which was removed for the deflation of the bags into the SESI-HRMS.

Safety precautions

Real-time measurements were performed using disposable spirometry mouthpieces incorporating bacterial and virus filter. Off-line collection of breath samples was carried out in single-use Nalophan bags.

Serum concentrations

Serum concentrations of ASMs were obtained from the clinical chemistry laboratory of University Hospital Basel. Total VPA was measured using enzyme-multiplied immunoassay technique, where the therapeutic range was 50-100 mg/L. Free VPA was measured by gas chromatography-mass spectrometry, with a therapeutic range of 5-10 mg/L. Free VPA required 4.7 mL of blood for analysis, whereas total VPA required 0-5-1 mL.

Study participants

In total, 40 patients (mean age 11.5 years, range 4.4-19.7 years, 50% male) diagnosed with epilepsy and taking at least valproic acid as medication were included in the study. For the complete

overview of patients, drugs intake and demographics see Table S1. Participants provided both, an online measurement and offline measurement, with at least one hour of no food intake, abstaining from chewing gum or brushing their teeth prior to the measurement, resulting in 40 pairs of measurement results. The Ethics Committee of North–western and Central Switzerland (ID 2017-01537 (Breath-TDM) and ID 2020-00778 (EBECA) approved the study and written informed consent was obtained.

Data Analysis

Data processing and statistical analysis were performed using MATLAB (version R2022a, MathWorks, USA). Raw centroid (intensity cut-off = 10^2 a.u.) and profile mass spectra were accessed using in-house C# console apps based on Thermo Fisher Scientific's RawFileReader (version 5.0.0.38). Centroid and profile mass spectra were recalibrated using reference peaks with formulae fulfilling the so-called "seven golden rules"[38] as well as common laboratory contaminants[39], present in at least 50% of the samples with an initial tolerance of at least 5 ppm. The in-house post-calibration algorithm uses a shape-preserving piecewise cubic interpolation to describe the experimental error across the entire m/z range. Outliers of the reference peaks, as assessed with a moving median algorithm, were excluded from the interpolation process. Centroids and profile peaks were subsequently shifted according to this described mass error. The process was iterated three times until achieving mass accuracies below 0.5 ppm across the entire mass range for all spectra evaluated. Histograms of the recalibrated centroids were subsequently binned using a Kernel density function. The bandwidth value controlling the smoothness of the resulting probability density curve was iterated to achieve Gaussian probability density functions of +/- 1 ppm at full width at half maximum. The recalibrated centroids falling within this window were used to construct the final data matrix of size 40 x 3,128 (samples x m/z). In this targeted study, we focused on 33 ions of interest: 10 associated with VPA, 15 with side effects, and 8 with drug response.

Zero imputation

Centroids below the intensity cut-off (i.e., 10^2 a.u.) were imputed according to the following criteria: ^{13}C and ^{18}O isotopes were calculated and imputed according to their expected natural abundance; Ammonium adducts were imputed by multiplying the corresponding $[M+H]^+$ signal intensity by the mean of all non-zero $[M+NH_4]^+$ values; zeros for the remaining features were

imputed using regression on order statistics (ROS) using the log method described in [40]. See Table S2 for further details.

Results and Discussion

One main advantage of SESI-HRMS is that it allows for real-time mapping of a broad range of metabolites[36]. However, as with any other real-time technique, the patients have to access the mass spectrometer physically, which limits its use. For this reason, we have recently evaluated in a systematic fashion the feasibility of collecting bedside breath samples in bags for subsequent rapid analysis[41]. This study revealed that around 55% of the ~1,200 mass spectral features detected in a typical breath fingerprint, can be recovered with an acceptable degree of comparability with the gold-standard counterpart real-time measurement. In this study, we use the same approach but target a small set of exhaled compounds relevant to ASMs.

After pre-processing the raw data, the final data matrix consisted of 80 samples (i.e., 40 pairs of real-time and offline measurements) x 3,128 mass spectral features. Figure 1a (top) shows a heatmap of the 40 x 3,128 submatrix corresponding to the real-time measurements in chronological order. Notably, it shows a relatively consistent signal intensity during the over four years of data collection period. The only exception is the bottom-left blue patch (indicating no detected compounds) in the region 70-100 m/z. This is simply explained by the changing of the settings of the tune file, which was adapted in early 2020, to increase the mass spectral coverage. Overall, this suggests that the technique and protocols used are suitable to monitor patients over extended periods of time. Figure 1a (bottom) shows the equivalent data matrix for the bag-collected breath samples. A superficial visual inspection indicates a fairly consistent picture across the time span of over four years. When comparing the real-time and off-line matrices, one can observe an overall relatively lower signal (bluish color) in the off-line matrix. This suggests that, as expected, some signals may be partially depleted during the collection and further analysis process. This is especially evident for heavier species (m/z > 200), which tend to be less volatile than lighter species, and hence more prone to losses during the collection process.

Systemic VPA prediction using breath collected in bags

We have recently shown that a combination of 11 ions detected in exhaled breath by real-time analysis of patients receiving VPA, could be used to predict with high accuracy systemic concentrations of free and total VPA[35]. The 11 features were assigned to four unique molecules:

VPA itself (detected in negative ion mode) and ten ions associated with three VPA-metabolites (i.e., 3-heptanone[22], 4-OH-γ-lactone[21] and heptanedione). For practical reasons, in this work we focused only on the ten ions detectable in positive ion mode. This clinical study has recently led to an IVD CE-marked breath test (DBI-EPIbreath®, Deep Breath Intelligence AG, Switzerland). In order to generalize the breath test, it would be crucial to be able to capture this information remotely for subsequent off-line analysis, which has been the main motivation of this work. Figure S1 zooms into the regions of interest (i.e., the 10 relevant ions detected in positive ion mode) of Figure 1a. Visual inspection of the heatmap reveals a relatively good consistency between the real-time and the off-line matrices. Figure 1b provides a more accurate representation by showing the raw mass spectra for the ten ions of interest (after preprocessing pipeline) for all measurements in real-time (upwards spectra) and paired off-line spectra (downwards spectra). Several messages emerge from this plot. Firstly, in the m/z dimension, it shows that the mass spectra are perfectly aligned, and the calibration error is well within 1 ppm even after four years of measurements. In addition, the high resolution (well in excess of 100,000), allows to obtain clearly resolved peaks without isobaric interfering species that might ultimately misrepresent the prediction of systemic concentrations. Finally, in the signal intensity dimension, it shows that the orders of magnitude are similar for both the real-time and off-line measurements. This provides an early indication that this off-line approach may potentially be used to capture the metabolic information carried by these 10 ions.

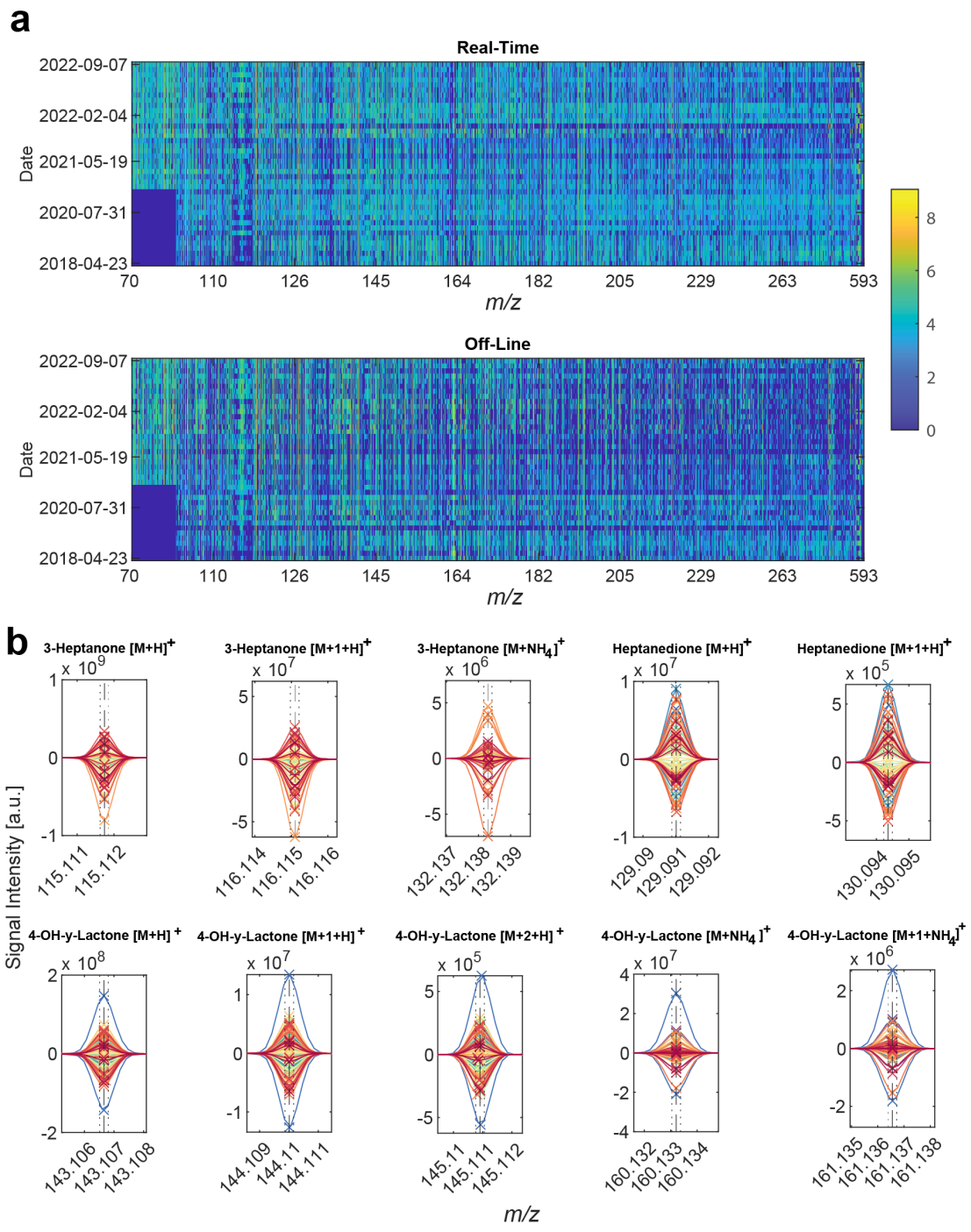


Figure 9 Breath collected in bags provides a similar metabolic signature as its counterpart by real-time analysis a) Signal intensities of all observed m/z ranging from April 2018 to September 2022. The color bar represents signal intensity (log 10 transformed). b) Raw mass spectra from identified biomarkers of VPA, the upper half are real-time measurements and the lower half are bag measurements. Dotted lines represent +- 1 ppm deviation. Dashed line represents theoretical mass. 'x' represents the centroid for each mass spectral feature. Each color represents a unique measurement (in pairs, on-line and off-line).

However, in order to ensure an accurate prediction, signal intensities between on-line (gold standard) and off-line approaches should correlate at an individual level. The agreement between the real-time and off-line measurements was evaluated using Lin's concordance correlation coefficient (CCC)[42], which is a widely accepted metric to benchmark newly developed methods against the current gold standard. CCC ranges from -1 to +1, representing the latter one a perfect agreement between both methods (i.e., how close the measurement pairs fall to the 45-degree line). Figure 2 illustrates this concept by representing the signal intensity for all pairs of measurements for all 10 ions and listing the computed CCC for each case. It shows that signal intensities typically expand over 2-3 orders of magnitude across all subjects. We found a CCC ~0.6 for the two ions corresponding to protonated heptanedione. For 3-Heptanone was higher with CCC ~0.7 for its protonated molecule and its ammonium adduct. Regarding the third metabolite, i.e. 4-OH-γ-lactone, the CCC for the protonated molecule was in the range of 0.6, whereas the agreement was found to be better (CCC ~0.75) for the ammonium adducts. Table 1 displays a more comprehensive overview of the Lin's CCC statistics, allowing for a more in-depth interpretation. For example, the lowest agreement (CCC = 0.45) was found for the low-abundance M+2 ^{18}O isotope of 4-OH-γ-lactone, whereby the two data points with the lowest signal intensity clearly depart from the identity line. A negative location shift of ~-0.3 was found for the protonated species of heptanedione, indicating a systematic depletion of this molecule during the collection process. In contrast, the protonated ions of 3-heptanone showed a positive location shift of ~0.65 (i.e. data points tend to fall underneath the identity line). This suggests that, at the particular m/z region, the bags are releasing some overlapping contaminant that adds to the signal of the VPA metabolite. This effect was less of an issue with the ammonium adduct (i.e., location shift 0.3). In contrast to the other two metabolites, 4-OH-γ-lactone shows no dramatic depletion, nor adulteration by interfering species in all five ions (i.e., location shift around 0). All in all, this provides an encouraging early indication that indeed the information for these VPA metabolites could be captured and measured with acceptable fidelity using the proposed method.

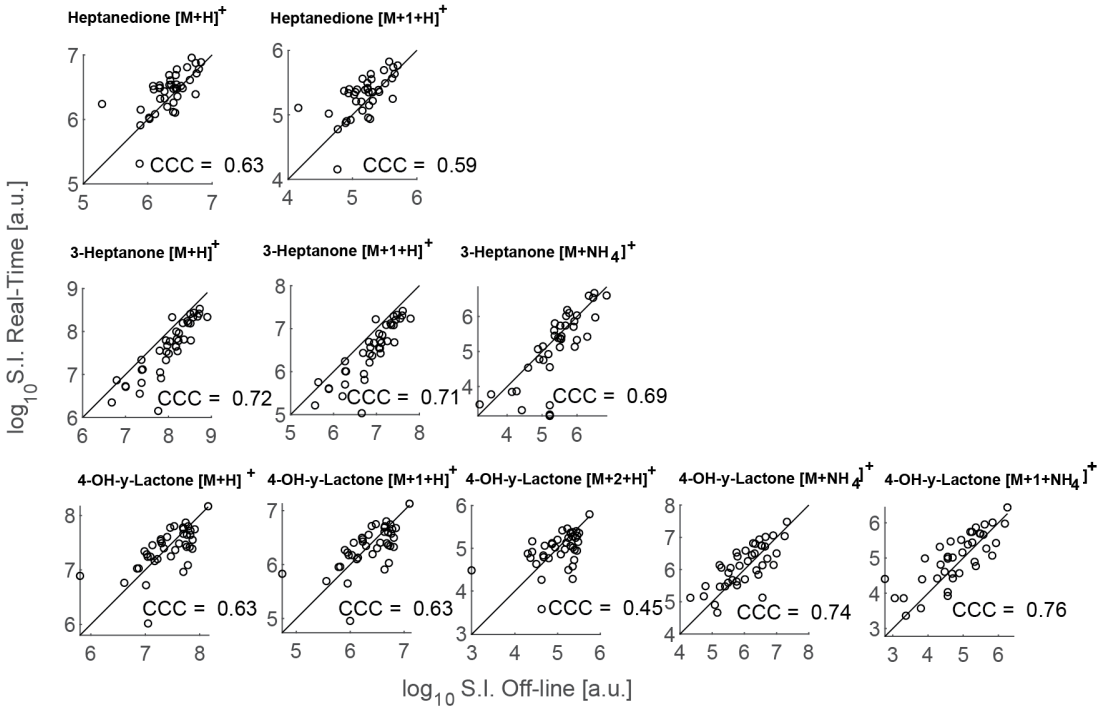


Figure 10 Off-line and the gold standard real-time analysis agree to a large extent | Pairs of measurements for all patients for each of the ions associated to VPA metabolites. CCC = concordance correlation coefficient. Identity line (i.e. 45-degree) added as reference line for visual inspection.

Molecule	m/z	ppm error	Lins CCC	Confidance Interval	Scale Shift	Location Shift	Bias Correction	Pearson Corr. Coeff.
3-Heptanone	115.1118	-0.0148	0.72	[0.57,0.82]	0.87	0.65	0.82	0.88
^{13}C3-Heptanone	116.1151	-0.0560	0.71	[0.57,0.82]	0.87	0.66	0.82	0.87
NH_4^+ 3-Heptanone	132.1383	-0.0053	0.69	[0.52,0.81]	0.74	0.33	0.91	0.76
Heptanedione	129.0910	0.0294	0.63	[0.41,0.78]	0.97	-0.31	0.95	0.66
^{13}C Heptanedione	130.0944	-0.0077	0.59	[0.36,0.75]	0.98	-0.34	0.94	0.63
4-OH-γ-Lactone	143.1066	0.0266	0.63	[0.40,0.78]	1.15	0.07	0.99	0.64
^{13}C 4-OH-γ-Lactone	144.1099	-0.0069	0.63	[0.40,0.78]	1.15	0.07	0.99	0.64
^{18}O 4-OH-γ-Lactone	145.1108	0.0551	0.45	[0.17,0.66]	1.25	0.09	0.97	0.46
NH_4^+4-OH-γ-Lactone	160.1332	-0.0331	0.74	[0.56,0.85]	1.08	-0.16	0.98	0.75
^{13}C NH_4^+ 4-OH-γ-Lactone	161.1366	-0.0006	0.76	[0.61,0.86]	1.15	-0.25	0.96	0.79

Table 2. Lin's CCC statistics (alpha = 0.05) analysis of the agreement between the mass spectrometric read-out between the real-time gold standard and the proposed off-line method for all 10 ions associated with three VPA metabolites. +1 denotes ^{13}C and +2 denotes ^{18}O isotopes.

However, the ultimate aim of such remote sampling development would be to enable the prediction of systemic concentrations of VPA. Total VPA and free VPA serum concentrations were predicted by first training and screening 19 regression models with leave-one-out cross-validation using the log10 transformed real-time data matrix. Based on the root mean square error (RMSE), linear support vector machine (Linear SVM) was chosen as the best-performing method. Thus, given the acceptable agreement between real-time and off-line methods, the real-time data matrix was used as the training set to predict total VPA blood levels (n=31) and free-fraction VPA blood levels (n=16), using the log10 transformed off-line data matrix as the test set.

Figure 3 displays the actual measured serum concentration versus the predicted one. The agreement for total VPAwas found to be modest (CCC = 0.47), which was mostly worsened by the four data points in the low-concentration region and the two data points in the highest-concentration region (i.e., below and above the therapeutic window). In contrast, despite the smaller sample size, the prediction of free VPA was substantially better (CCC = 0.60). This may be explained by the closer read-out that exhaled breath represents from the free fraction than the total concentration. The reason being that the metabolites used for the prediction stem from the β- and ω1- oxidation pathways of the free fraction. From a clinical perspective, this is an advantage because the free fraction is the one with a clinical effect[43, 44].

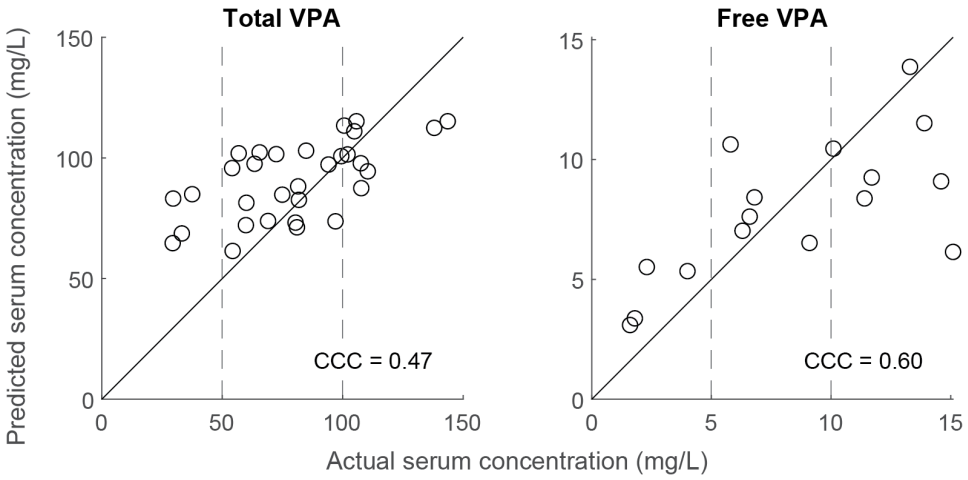


Figure 11 Collection of exhaled breath in bags allows for the prediction of systemic VPA | Actual versus predicted serum concentration for total and free VPA. Dashed lines represent therapeutic window. Identity line (i.e., 45-degree) added as reference line for visual inspection.

Side-effects and drug response

Along with the ability to detect drug metabolites, our previous study reported that the breath mass spectral fingerprints contained information on a set of endogenous metabolites that were associated with the actual clinical outcome (i.e., side effects and drug response). Whereby several amino acid metabolic pathways were found to be upregulated in patients suffering from side effects, whereas tyrosine metabolism was found to be downregulated in non-responders. During the last phase of this analysis, we investigated to what extent these endogenous compounds were preserved using the proposed off-line method. A total of 27 ions were previously associated with side effects and 10 ions with response to medication. Inclusion criteria for this study included the detection of the mass spectral feature within +/1 ppm of the theoretical mass and the detection of the ion in at least 60% of the patients. Thus, 15 ions for side effects and 8 for drug response fulfilled these criteria and were further evaluated. Figure 4a and 4b show the mass spectra for three representative examples of ions associated with side effects and drug response, respectively.

The spectra show again a very high mass accuracy and precision for all detected features. Similar to the ions used to predict systemic drug concentrations, the agreement between the off-line and the real-time analyses was evaluated using Lin's CCC. Figure 4c and 4d shows the real-time vs off-line scatter plot for the ions shown in Figure 4a and 4b, including the CCC. Profile peaks and

pairwise scatter plots of the remaining are shown in Figure S2 and S3, respectively. Lin's CCC statistics on all 23 ions are listed in Table S3.

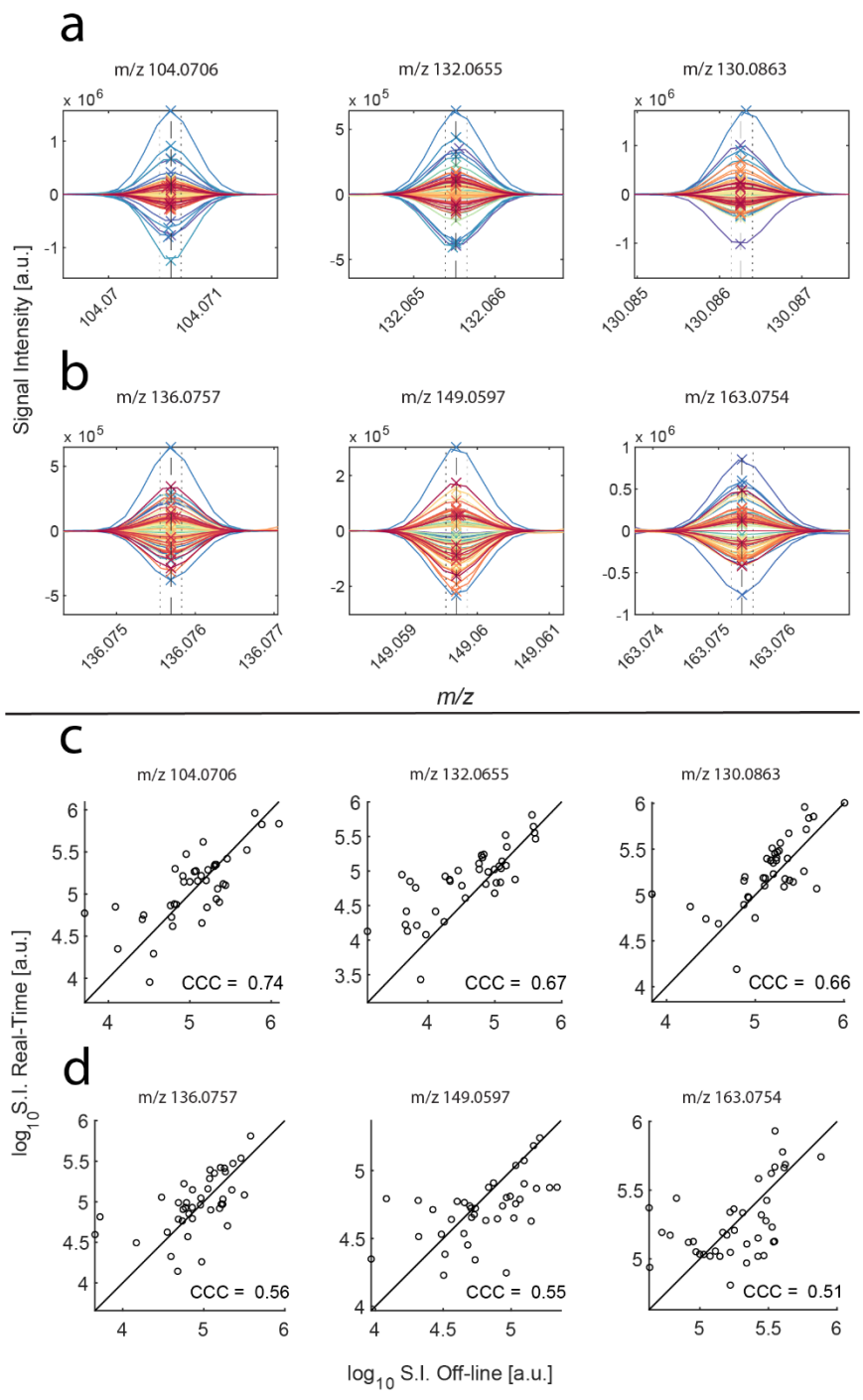


Figure 12 Collection of exhaled breath of metabolites associated with drug response and side effects | Profile peaks from selected side-effect (panel a) and drug response (panel b) -associated exhaled molecules., upper half are real-time measurements and lower half are bag measurements. Dotted lines represent +- 1 ppm deviation. Dashed line represents theoretical mass. 'x' represents identified mass for each

measurement. Each unique color is a unique measurement. Comparison between Real-Time and Off-line analysis for selected side-effect (panel c) and drug response (panel d) associated features. Identity line added as reference for visual aid.

We found a large disparity in the quality of the off-line read-out for these 23 ions. Out of 15 detected ions associated with side effects, four were found to have a CCC > 0.6, whereas none of the ions associated with drug response reached this quality threshold. While the complete compound ID of some these metabolites is ongoing (*Awchi et al.* under review), this data suggests that perhaps their endogenous nature and physicochemical properties make them less amenable for off-line analysis than the compounds used to predict systemic VPA. Further work will be required to determine whether this partial information will suffice to compute risk estimates for drug response and side effects.

Conclusion

In this work, we introduce the technique of remote sampling of exhaled breath and subsequent rapid mass spectrometric analysis of exhaled metabolites in association with the anti-seizure medication VPA and its clinical effect. We conclude that the proposed method allows for the prediction of systemic concentrations of VPA based on samples collected over a period of four years in a clinical setting. The method is particularly well suited to predict the free fraction, which ultimately has higher clinical relevance. In addition, we conclude that the method allows to capture endogenous compounds previously associated with drug response and side effects just partially. Since this technology can be used bed-side in a completely non invasive manner requiring no relevant cooperation of the patient, it has the potential for a wide clinical use in the context of the personalized treatment of patients with epilepsy. By reaching a wider range of patients, including young children and intellectually impaired individuals, we envision that the proposed method has the potential to enable large-scale screening of an already IVD CE-marked breath test (DBI-EPIbreath®, Deep Breath Intelligence AG, Switzerland) for ASMs TDM.

Acknowledgments

We thank the participants who volunteered to take part in this study and the study nurses' team (Amelia Imolesi, Maya Weber, and Isabel Gonzalez Novoa) for coordinating and maintaining

study documents. Mélina Richard is gratefully acknowledged for her full study coordination. Prof. Katharina Rentsch of University Hospital Basel is greatly acknowledged for analyzing the serum samples to determine the concentrations of ASMs. PS received funding from Fondation Botnar (Switzerland) and the Swiss National Science Foundation (PCEGP3_181300). This work is part of the Zurich Exhalomics project under the umbrella of University Medicine Zurich/Hochschulmedizin Zürich.

Author Contributions

AND, UF, and PS designed the study. AND and PED contributed to the screening and recruitment of patients. MA performed the mass spectral measurements. MA, KDS and PS developed the codes to analyze the data. MA, AND, UF and PS analyzed and interpreted the data. MA and PS wrote the manuscript. All authors reviewed and contributed to the final version of the manuscript.

Competing interests

PS is co-founder of Deep Breath Intelligence AG (Switzerland), which develops breath-based diagnostic tools and commercializes DBI-EPIbreath®. KS is consultant for Deep Breath Intelligence AG (Switzerland). All authors except PED are co-inventors of the European patent 20186274.5. The University Children's Hospital Basel is a shareholder of Deep Breath Intelligence AG (Switzerland).

References

1. Alahmadi, F. H.; Wilkinson, M.; Keevil, B.; Niven, R.; Fowler, S. J., Short- and medium-term effect of inhaled corticosteroids on exhaled breath biomarkers in severe asthma. *J. Breath Res.* **2022,** *16* (4).
2. Brinkman, P.; Ahmed, W. M.; Gómez, C.; Knobel, H. H.; Weda, H.; Vink, T. J.; Nijsen, T. M.; Wheelock, C. E.; Dahlen, S.-E.; Montuschi, P.; Knowles, R. G.; Vijverberg, S. J.; Maitland-van der Zee, A. H.; Sterk, P. J.; Fowler, S. J., Exhaled volatile organic compounds as markers for medication use in asthma. *Eur. Respir. J.* **2020,** *55* (2), 1900544.
3. van Oort, P. M.; Povoa, P.; Schnabel, R.; Dark, P.; Artigas, A.; Bergmans, D.; Felton, T.; Coelho, L.; Schultz, M. J.; Fowler, S. J.; Bos, L. D., The potential role of exhaled breath analysis in the diagnostic process of pneumonia-a systematic review. *J. Breath Res.* **2018,** *12* (2), 024001.
4. Bos, L. D.; Sterk, P. J.; Fowler, S. J., Breathomics in the setting of asthma and chronic obstructive pulmonary disease. *Journal of Allergy and Clinical Immunology* **2016,** *138* (4), 970-976.
5. Montuschi, P., Review: Analysis of exhaled breath condensate in respiratory medicine: methodological aspects and potential clinical applications. *Therapeutic Advances in Respiratory Disease* **2007,** *1* (1 %U http://tar.sagepub.com/content/1/1/5.abstract), 5-23.
6. Miekisch, W.; Schubert, J. K.; Noeldge-Schomburg, G. F. E., Diagnostic potential of breath analysis - focus on volatile organic compounds. *Clinica Chimica Acta* **2004,** *347* (1-2), 25-39.
7. Sukul, P.; Bartels, J.; Fuchs, P.; Trefz, P.; Remy, R.; Ruhrmund, L.; Kamysek, S.; Schubert, J. K.; Miekisch, W., Effects of COVID-19 protective face masks and wearing durations on respiratory haemodynamic physiology and exhaled breath constituents. *Eur. Respir. J.* **2022,** *60* (3).
8. Woodfield, G.; Belluomo, I.; Laponogov, I.; Veselkov, K.; Cobra, W. G.; Cross, A. J.; Hanna, G. B.; Boshier, P. R.; Lin, G. P.; Myridakis, A.; Ayrton, O.; panl, P.; Vidal-Diez, A.; Romano, A.; Martin, J.; Marelli, L.; Groves, C.; Monahan, K.; Kontovounisios, C.; Saunders, B. P., Diagnostic performance of a non-invasive breath test for colorectal cancer: COBRA1 study. *Gastroenterology* **2022**.
9. Belluomo, I.; Boshier, P. R.; Myridakis, A.; Vadhwana, B.; Markar, S. R.; Spanel, P.; Hanna, G. B., Selected ion flow tube mass spectrometry for targeted analysis of volatile organic compounds in human breath. *Nat. Protoc.* **2021,** *16*, 3419-3438.
10. Pesesse, R.; Stefanuto, P. H.; Schleich, F.; Louis, R.; Focant, J. F., Multimodal chemometric approach for the analysis of human exhaled breath in lung cancer patients by TD-GC x GC-TOFMS. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2019,** *1114*, 146-153.
11. Singh, K. D.; Del Miguel, G. V.; Gaugg, M. T.; Ibanez, A. J.; Zenobi, R.; Kohler, M.; Frey, U.; Sinues, P. M.-L., Translating secondary electrospray ionization-high-resolution mass spectrometry to the clinical environment. *J Breath Res* **2018,** *12* (2), 027113.
12. Beauchamp, J. D.; Pleil, J. D., Breath: An Often Overlooked Medium in Biomarker Discovery. In *Biomarker Validation*, Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, Germany: 2015; pp 75-93.
13. van Oort, P. M. P.; Nijsen, T.; Weda, H.; Knobel, H.; Dark, P.; Felton, T.; Rattray, N. J. W.; Lawal, O.; Ahmed, W.; Portsmouth, C.; Sterk, P. J.; Schultz, M. J.; Zakharkina, T.; Artigas, A.; Povoa, P.; Martin-Loeches, I.; Fowler, S. J.; Bos, L. D. J.; Consortium, B., BreathDx -

molecular analysis of exhaled breath as a diagnostic test for ventilator-associated pneumonia: protocol for a European multicentre observational study. *Bmc Pulmonary Medicine* **2017,** *17*.
14. Lawal, O.; Ahmed, W. M.; Nijsen, T. M.; Goodacre, R.; Fowler, S. J., Exhaled breath analysis: a review of 'breath-taking'methods for off-line analysis. *Metabolomics* **2017,** *13* (10), 1-16.
15. Mani-Varnosfaderani, A.; Gao, A.; Poch, K. R.; Caceres, S. M.; Nick, J. A.; Hill, J. E., Breath biomarkers associated withnontuberculosis mycobacteriadisease status in persons with cystic fibrosis: a pilot study. *J. Breath Res.* **2022,** *16* (3).
16. Stefanuto, P. H.; Romano, R.; Rees, C. A.; Nasir, M.; Thakuria, L.; Simon, A.; Reed, A. K.; Marczin, N.; Hill, J. E., Volatile organic compound profiling to explore primary graft dysfunction after lung transplantation. *Sci. Rep.* **2022,** *12* (1), 2053.
17. Ibrahim, W.; Wilde, M. J.; Cordell, R. L.; Richardson, M.; Salman, D.; Free, R. C.; Zhao, B.; Singapuri, A.; Hargadon, B.; Gaillard, E. A.; Suzuki, T.; Ng, L. L.; Coats, T.; Thomas, P.; Monks, P. S.; Brightling, C. E.; Greening, N. J.; Siddiqui, S.; Consortium, E.; Munton, R.; Le Quesne, J.; Goodall, A. H.; Pandya, H. C.; Reynolds, J. C.; Clokie, M. R. J.; Samani, N. J.; Barer, M. R.; Shaw, J. A., Visualization of exhaled breath metabolites reveals distinct diagnostic signatures for acute cardiorespiratory breathlessness. *Sci. Transl. Med.* **2022,** *14* (671), eabl5849.
18. McCartney, M. M.; Borras, E.; Rojas, D. E.; Hicks, T. L.; Hamera, K. L.; Tran, N. K.; Tham, T.; Juarez, M. M.; Lopez, E.; Kenyon, N. J.; Davis, C. E., Predominant SARS-CoV-2 variant impacts accuracy when screening for infection using exhaled breath vapor. *Communications Medicine* **2022,** *2* (1), 158.
19. Dasgupta, A., Chapter 1 - Introduction to Therapeutic Drug Monitoring: Frequently and Less Frequently Monitored Drugs. In *Therapeutic Drug Monitoring*, Academic Press: Boston, 2012; pp 1-29.
20. Gong, X.; Shi, S.; Gamez, G., Real-Time Quantitative Analysis of Valproic Acid in Exhaled Breath by Low Temperature Plasma Ionization Mass Spectrometry. *J Am Soc Mass Spectrom* **2017,** *28* (4), 678-687.
21. Gamez, G.; Zhu, L.; Disko, A.; Chen, H.; Azov, V.; Chingin, K.; Kramer, G.; Zenobi, R., Real-time, in vivo monitoring and pharmacokinetics of valproic acid via a novel biomarker in exhaled breath. *Chem. Commun.* **2011,** *47* (17), 4884-4886.
22. Erhart, S.; Amann, A.; Haberlandt, E.; Edlinger, G.; Schmid, A.; Filipiak, W.; Schwarz, K.; Mochalski, P.; Rostasy, K.; Karall, D.; Scholl-Burgi, S., 3-Heptanone as a potential new marker for valproic acid therapy. *J. Breath Res.* **2009,** *3* (1).
23. Gong, X.; Shi, S.; Zhang, D.; Gamez, G., Quantitative Analysis of Exhaled Breath Collected on Filter Substrates via Low-Temperature Plasma Desorption/Ionization Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2022,** *33* (8), 1518-1529.
24. Dryahina, K.; Polasek, M.; Smith, D.; Spanel, P., Sensitivity of secondary electrospray ionization mass spectrometry to a range of volatile organic compounds: Ligand switching ion chemistry and the influence of Zspray guiding electric fields. *Rapid Commun. Mass Spectrom.* **2021,** *35* (22), e9187.
25. Dryahina, K.; Som, S.; Smith, D.; Spanel, P., Reagent and analyte ion hydrates in secondary electrospray ionization mass spectrometry (SESI-MS), their equilibrium distributions and dehydration in an ion transfer capillary: Modelling and experiments. *Rapid Commun. Mass Spectrom.* **2021,** *35* (7), e9047.
26. Smith, D.; Španěl, P.; Herbig, J.; Beauchamp, J., Mass spectrometry for real-time quantitative breath analysis. *Journal of breath research* **2014,** *8*, 027101.

27. Španěl, P.; Smith, D., On the features, successes and challenges of selected ion flow tube mass spectrometry. *European Journal of Mass Spectrometry* **2013,** *19* (4), 225-246.
28. Beauchamp, J.; Herbig, J.; Dunkl, J.; Singer, W.; Hansel, A., On the performance of proton-transfer-reaction mass spectrometry for breath-relevant gas matrices. *Measurement Science and Technology* **2013,** *24* (12).
29. Xu, M.; Tang, Z.; Duan, Y.; Liu, Y., GC-based techniques for breath analysis: current status, challenges, and prospects. *Critical Reviews in Analytical Chemistry* **2016,** *46* (4), 291-304.
30. Beauchamp, J., Current sampling and analysis techniques in breath research-results of a task force poll. *Journal of Breath Research* **2015,** *9* (4).
31. Fang, M.; Ivanisevic, J.; Benton, H. P.; Johnson, C. H.; Patti, G. J.; Hoang, L. T.; Uritboonthai, W.; Kurczy, M. E.; Siuzdak, G., Thermal Degradation of Small Molecules: A Global Metabolomic Investigation. *Analytical Chemistry* **2015,** *87* (21), 10935-41.
32. Martinez-Lozano Sinues, P.; Landoni, E.; Miceli, R.; Dibari, V. F.; Dugo, M.; Agresti, R.; Tagliabue, E.; Cristoni, S.; Orlandi, R., Secondary electrospray ionization-mass spectrometry and a novel statistical bioinformatic approach identifies a cancer-related profile in exhaled breath of breast cancer patients: a pilot study. *J. Breath Res.* **2015,** *9* (3), 031001.
33. Bruhova Michalcikova, R.; Dryahina, K.; Smith, D.; Spanel, P., Volatile compounds released by Nalophan; implications for selected ion flow tube mass spectrometry and other chemical ionisation mass spectrometry analytical methods. *Rapid Commun. Mass Spectrom.* **2020,** *34* (5), e8602.
34. Spanel, P.; Smith, D., Quantification of volatile metabolites in exhaled breath by selected ion flow tube mass spectrometry, SIFT-MS. *Clin. Mass Spectrom.* **2020,** *16*, 18-24.
35. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; Awchi, M.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U.; Datta, A. N.; Sinues, P., Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications Medicine* **2021,** *1* (1), 21.
36. Gisler, A.; Singh, K. D.; Zeng, J.; Osswald, M.; Awchi, M.; Decrue, F.; Schmidt, F.; Sievi, N.; Chen, X.; Usemann, J.; Frey, U.; Kohler, M.; Li, X.; Sinues, P., An Interoperability Framework for Multicentric Breath Metabolomic Studies. *iScience* **2022,** *25* (12), 105557.
37. Decrue, F.; Singh, K. D.; Gisler, A.; Awchi, M.; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of Exhaled Breath Analysis with Parallel Lung Function and FeNO Measurements in Infants. *Analytical Chemistry* **2021,** *93* (47), 15579-15583.
38. Kind, T.; Fiehn, O., Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **2007,** *8*, 105.
39. Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M., Interferences and contaminants encountered in modern mass spectrometry. *Anal Chim Acta* **2008,** *627* (1), 71-81.
40. Lee, L.; Helsel, D., Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. *Computers & Geosciences* **2005,** *31* (10), 1241-1248.
41. Decrue, F.; Singh, K. D.; Gisler, A.; Awchi, M.; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of Exhaled Breath Analysis with Parallel Lung Function and FeNO Measurements in Infants. *Anal. Chem.* **2021**.
42. Lawrence, I. K. L., A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* **1989,** *45* (1), 255-268.

43. Patsalos, P. N.; Zugman, M.; Lake, C.; James, A.; Ratnaraj, N.; Sander, J. W., Serum protein binding of 25 antiepileptic drugs in a routine clinical setting : A comparison of free non – protein-bound concentrations. *Epilepsia* **2017,** *58* (7), 1234-1243.
44. Fisch, U.; Baumann, S. M.; Semmlack, S.; Marsch, S.; Ruegg, S.; Sutter, R., Accuracy of Calculated Free Valproate Levels in Adult Patients With Status Epilepticus. *Neurology* **2021,** *96* (1), e102-e110.

Chapter 6: Study III

Metabolic trajectories of diabetic ketoacidosis onset described by breath analysis

Mo Awchi[1,2], Kapil Dev Singh[1,2], Sara Bachmann Brenner[1], Marie-Anne Burckhardt[1], Melanie Hess[1], Jiafa Zeng[1,2], Alexandre N. Datta[1], Urs Frey[1], Urs Zumsteg[1], Gabor Szinnai[1,*], Pablo Sinues[1,2,*]

1. University Children's Hospital Basel, Basel, Switzerland.
2. Department of Biomedical Engineering, University of Basel, Basel, Switzerland.

*gabor.szinnai@ukbb.ch
*pablo.sinues@unibas.ch

Keywords**:** Breath analysis, Metabolomics, Diabetic Ketoacidosis, Pathophysiology, Acetone, Acetoacetate, Monitoring, SESI-HRMS

Summary

Objective

This feasibility study aimed to investigate the anabolic effect of insulin on metabolites captured in exhaled breath during acute diabetic ketoacidosis (DKA) for a better pathophysiological understanding, and monitoring.

Research Design and Methods

Breath analysis was conducted on 30 patients, 12 with type 1 diabetes (T1D) (5 with DKA, 2 with hyperglycemia (HG), 5 with euglycemia), 18 without. They inflated Nalophan bags, and its metabolic content was subsequently interrogated by secondary electrospray ionization high-resolution mass spectrometer (SESI-HRMS). The total amount of longitudinal individual breath measurements was 107 (88 from people with T1D). Clinical parameters, such as base excess and ketone bodies, were measured for patients admitted to the intensive care unit (ICU) and were used to identify associated metabolic markers in breath.

Results

The study showed that well-known significantly altered metabolites in DKA patients with base excess (BE) < -2 vs non-diabetic patients such as acetone ($p = 3x10^{-7}$), pyruvate ($p = 5x10^{-3}$) and acetoacetate ($p = 4x10^{-7}$) are readily detectable in breath and follow the expected dynamics during the switch from catabolic to anabolic state in DKA. In addition, a total of 665 mass spectral features were found to significantly correlate ($p < 0.05$; $q < 0.18$) with BE (219/446 features showed significant positive/negative correlation). These correlating features with BE prompt metabolic trajectories towards in-control state as they progress towards homeostasis, with clear inter-individual differences.

Conclusion

This study provides proof-of-principle for using exhaled breath analysis in a real ICU setting for DKA monitoring. We found that, apart from the well-known association between acetone and DKA state, a completely new panel of exhaled metabolites (e.g. acetoacetate), describe the transition between catabolic to anabolic state in DKA. These new insights provide a more comprehensive understanding on the complex dynamics of DKA, and suggest that such non-invasive new technology may help to differentiate from other metabolic disorders and complement the current standards used for DKA care.

Introduction

The annual hospitalization rate for diabetic ketoacidosis (DKA) has been increasing in recent years and a better understanding of the pathophysiology, evidence-based, and targeted prevention measures might help reversing the trend[1]. In addition, the exact pathophysiology of DKA on a metabolic level is still debated[2]. Furthermore, in the emergency care setting, diabetes is a one of the relevant differential diagnosis of unclear metabolic acidosis or unknown origin, rapid non-invasive diagnostic technologies may help to clarify the diagnosis and enable rapid specific treatment. The development of novel technologies with unprecedented capabilities are key to further expand our understanding of DKA pathophysiology[3]. In this pilot study we tested the hypothesis that exhaled breath analysis can capture novel insights of DKA patients switching from catabolic to anabolic state after start of insulin and rehydration therapy.

Research design and Methods

This observational, longitudinal study was conducted at the University Children's Hospital Basel, Basel, Switzerland. The Ethics Committee of North–Western and Central Switzerland ID 2020-00778 (EBECA) approved the study and written informed consent was obtained. The study has been registered in https://clinicaltrials.gov/ct2/show/NCT04461821 and the methods are described in detail in the supplementary information.

Briefly, this study included participants aged 5-20 years old with and without type 1 diabetes (T1D). Participants with T1D were categorized according to glycemia and pH as i) with DKA (pH<7.3, base excess (BE) >-2, blood glucose>10 mmol/L); ii) with hyperglycemia (pH>7.3, BE >-2, blood glucose>10 mmol/L), or iii) euglycemic . DKA participants admitted to the intensive care unit (ICU) were treated according to international guidelines.

Patients with epilepsy represented the control group without T1D.

Breath was analyzed using custom-made Nalophan bags inflated with 2L of breath[4]. Subsequently, the bag was infused into a secondary electrospray ionization high-resolution mass spectrometer (SESI-HRMS). Raw centroid and profile mass spectra were accessed using in-house C# console apps based on Thermo Fisher Scientific's RawFileReader (version 5.0.0.38). Subsequent data processing and statistical analysis were performed using MATLAB (version R2022a, Mathworks, USA). Detailed methods are described in the Supplementary Information.

Results

We recruited 30 participants, five of which were DKA participants admitted to the ICU. The biochemical characteristics of the DKA participants while treated at the ICU are shown in Table S1-S5. Longitudinal breath measurements (n = 32) of these 5 DKA participants were recorded during their stay at the hospital until stabilization. Similarly, two participants with hyperglycemia were monitored during hospitalization, providing a total of n = 6 breath measurements. In total, these 7 participants contributed n = 24 breath measurements while their BE was below -2 mmol/L and n = 14 with while their BE was stabilized in the range -2 to +2 mmol/L. In addition, five euglycemic patients (i.e., (-2 < BE < 2) mmol/L) provided a total of n = 50 measurements. Finally, 18 participants without T1D treated for epilepsy contributed n = 19 breath measurements. This resulted in a total of N = 107 individual breath measurements (Fig. S1).

In the first part of the study, we assessed the utility of breath analysis for monitoring the normalization of DKA after start of intravenous (IV) insulin and rehydration therapy. In the late 1960s, Tassopoulos *et al.*[5] were already able to show that breath-acetone correlates with plasma-beta-hydroxybutyrate (BHB) and to a larger extend the clinical state of people undergoing DKA treatment. Further studies have confirmed this association[6] [7]. We used these observations as an anchor point for our current study and to draw further conclusions. The current study used BE, a surrogate marker for acid/base disturbances in blood, to compare with breath acetone. In DKA, high levels of acidic blood ketones are produced due to insulin deficiency, resulting in a negative BE[8]. As expected, a similar trend is observed between sequentially measured breath acetone and BE during DKA treatment (Fig. 1a).

Our study also investigated the relationship between BE and other ketone bodies, specifically acetoacetate (AcAc) and BHB. AcAc, produced during lipolysis and the breakdown of ketogenic amino acids, can be converted to acetone through decarboxylation. Our findings demonstrated for the first time a similar trend between BE and the breath signal intensity of AcAc (Fig. 1b), including participants 054, for which acetone showed a lack of association at the beginning of the time series. This suggests that exhaled AcAc might be a more robust surrogate marker for BE. In contrast, no obvious association was found between the mass spectral feature corresponding to BHB and BE. Finally, we observed a significant correlation ($r = 0.71$; $p \sim 3x10^{-8}$) between breath acetone and AcAc (Fig. S2), which is consistent with their known biochemical interplay.

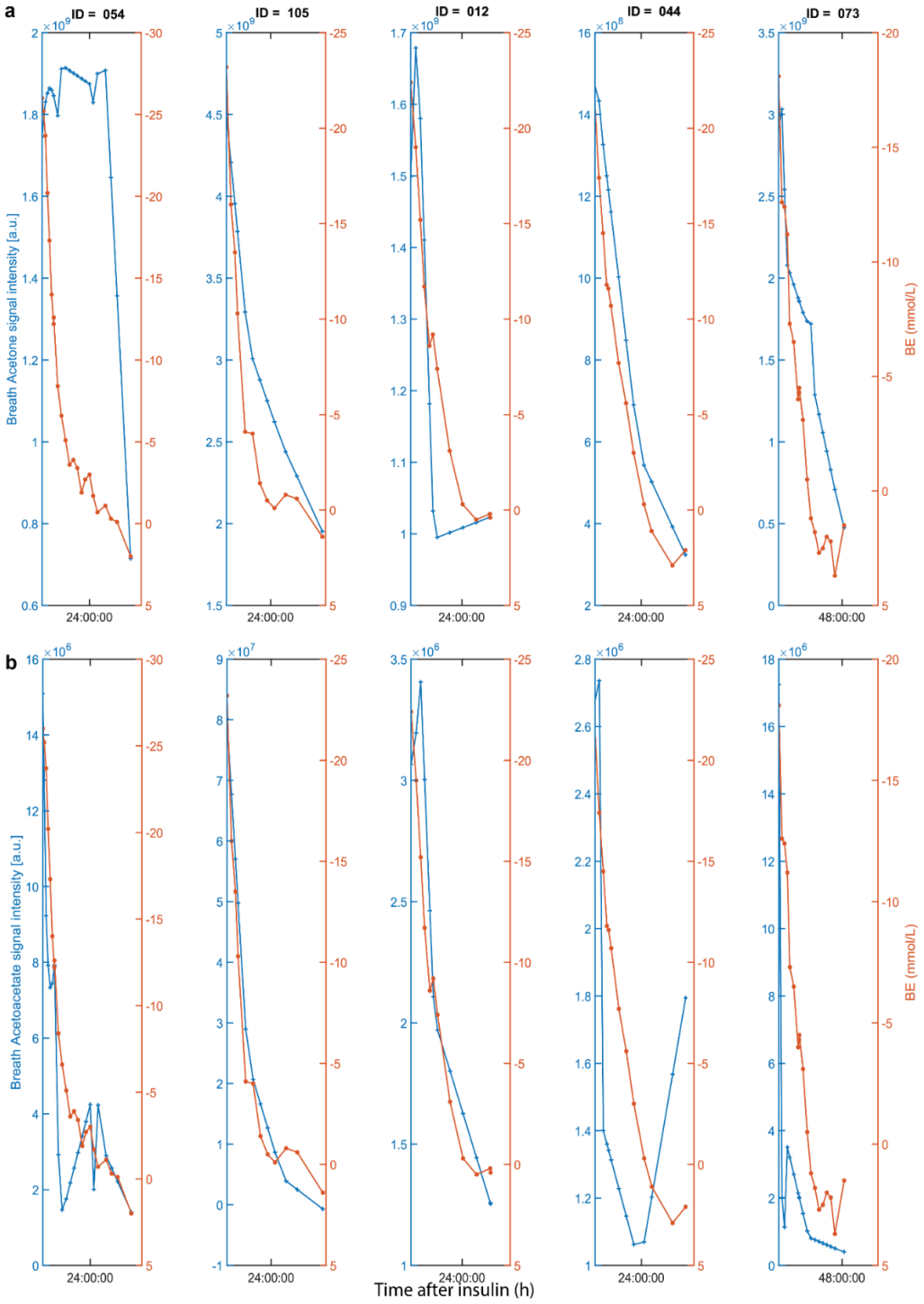


Figure 13 Time traces of a) breath acetone (left y-axis); b) and acetoacetate (left y-axis); in comparison with BE (right y-axis in a and b) for five DKA participantss admitted to ICU after start of insulin administration and rehydration therapy. X-axis represents the first 24 or 48 hours after start of iv insulin. Breath timepoints were linearly inter- and extrapolated to match with the clinical time-points.

Apart from acetone and AcAc, previous reports have shown the capability of breath analysis by SESI-HRMS to capture large areas of the human metabolism[9]. A portion of the metabolome covered by this breath analysis technique plays a significant role in DKA. For example, butanoate metabolism[10], tricarboxylic acid (TCA) cycle[11], amino acid metabolism[12] and fatty acid metabolism[13]. For this pilot study, we aimed to elucidate the generally accepted pathways, previously reported by Wallace *et al*[14]. Insulin deficiency (either in newly diagnosed T1D patients or in known T1D patients with insufficient dosing) is the cause of DKA. Insulin deficiency represents a catabolic state characterized by increased gluconeogenesis, glycogenolysis, and impaired peripheral glucose utilization resulting in hyperglycemia. Further, insulin deficiency causes increased lipolysis, ketogenesis and metabolic acidosis. As shown in Fig. 2 (please note Log10 scale), we observe downregulated levels of non-esterified fatty acids[15], in agreement with previously reported observations[13]. The increased lipolysis, results in formation of ketone bodies. As expected, and previously shown in Fig. 1, significant upregulation is observed for acetone and AcAc. Pyruvate shows an upregulated trend, which additionally is in agreement with previously reported observations[16]. The TCA cycle however, did not show significant alterations (with the exception of fumarate), which is in disagreement with previous reports[17]. Table S8 lists the multiple comparison statistical values for all these DKA-relevant exhaled metabolites.

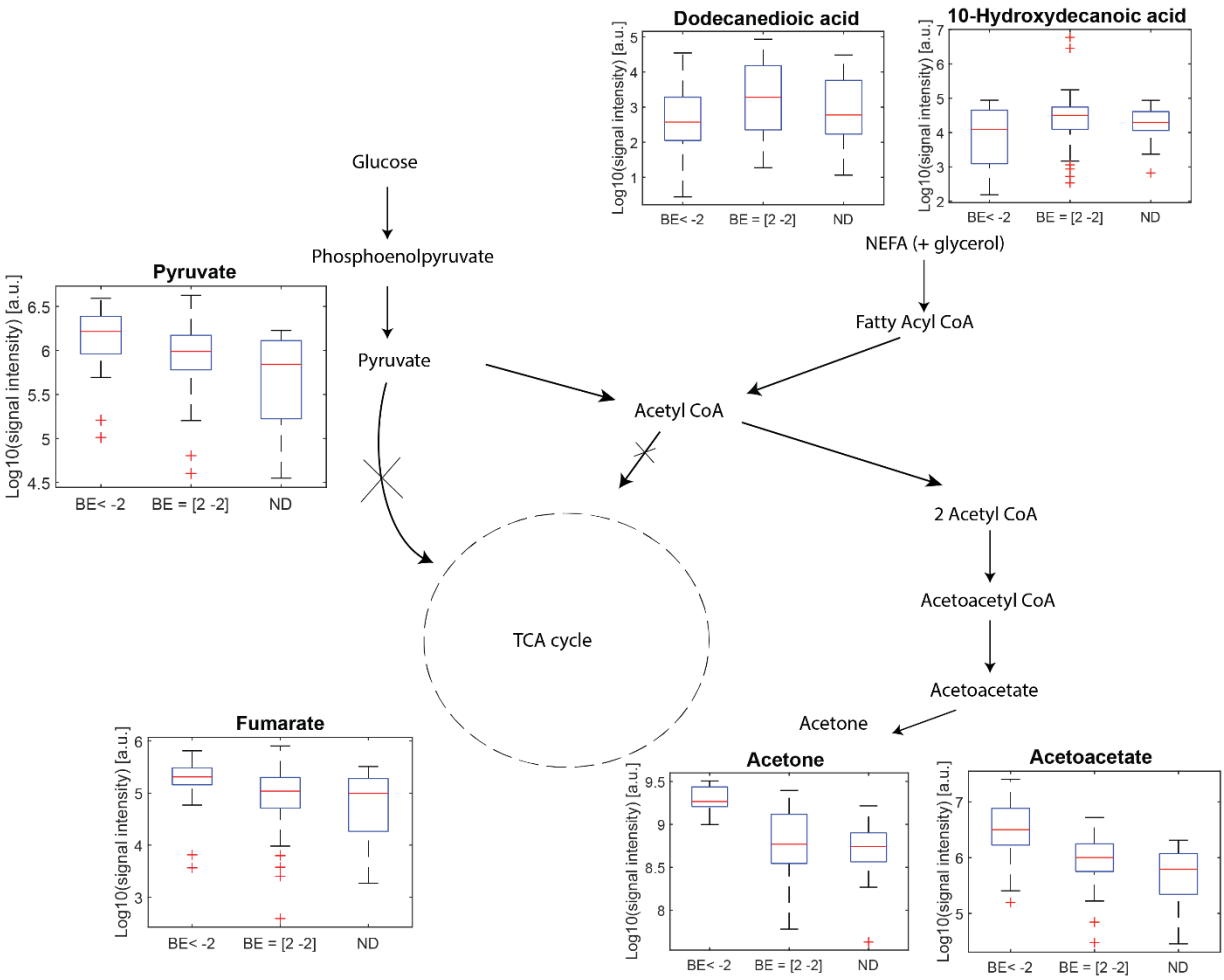


Figure 2 Breath analysis captures generally accepted DKA-altered pathways. Participants with T1D with -2 < BE show a significant up-regulation of pyruvate and fumarate, acetone and acetoacetate vs participants with T1D with normal BE values (i.e. +/- 2) and participants without T1D. In contrast, exhaled free fatty acids are downregulated.

Once we confirmed the soundness of the proposed breath analysis approach by confirming that AcAc and acetone mirror BE (Fig. 1), as well as an overall consistent picture of known altered metabolic pathways (Fig. 2), we further explored to what extent the comprehensive exhaled metabolic fingerprint may provide novel metabolic insights into DKA. This was pursued by investigating the correlation between BE values and the rest of the metabolic features detected in breath. A total of 665 features were found to significantly correlate ($p < 0.05$; $q < 0.18$) with BE (Table S9). 219 features showed significant positive correlation while 446 features showed a significant negative correlation. Two examples of such features showing a significant positive and negative correlation with BE levels are shown in Fig. S3. The feature with a positive correlation (r

= 0.67; q = $4x10^{-3}$) corresponds to a metabolite with molecular formula of $C_7H_{14}O$, while the one with a negative correlation with BE (r = -0.66; q = $5x10^{-3}$) could be mapped to $C_8H_6O_3$. To further visualize the overall contribution of the exhaled metabolites to the associated DKA stage, we subjected the dataset to principal component analysis (PCA). Fig. 3a shows the score plot of the first two PCs, explaining 43% of the variance. For visualization purposes, the data points were stratified according to their BE: i) participants with T1D with abnormal BE < -2, ii) participants with T1D without DKA (-2 < BE < 2), and iii) participants without T1D. Fig. 3a shows two distinct main clusters: participants with BE < -2 (lower quadrant; red data points) and the rest. Participants with and without T1D in euglycemia cluster together (i.e., overlapping 95% confidence interval (CI) ellipses). This indicates that, even though the participants are highly heterogeneous (i.e., participants with epilepsy and T1D in euglycemia), they are equally homeostatic from a metabolic acidosis point of view, and their breath fingerprints capture such state.

Fig. 3b shows how this technique can be used to map the individual metabolic trajectories of DKA participants after start of iv insulin and rehydration treatment. It clearly shows how the trajectory migrates from the lower half of the PCA space (catabolic state) when the BE values ranged between -15 and -12, to the upper half (anabolic state) when the BE value was -0.7. The remaining DKA patients' trajectories are shown in Fig. S4-S7. For patient 012, all measurements were recorded while BE was well below -2, and consistently his/her breath metabolic trajectory remained within the -2 < BE 95% CI ellipse. Patient 044 showed a similar behavior as 073, with a clear scattering around the lower half of the space while BE was well below -2, to then sharply transition to the in-control region once the BE was stabilized to -1.4. In contrast, patients 054 and 105 showed that they approached the in-control region as their BE stabilized, but they hardly reached the transition boundary between both regions. Somehow, these two patients seem to show a less "elastic" metabolic recovery. This is supported by the gradient of ketone bodies (Fig. S8), which shows a lagging behavior for patients 054 and 105, especially for the latter one, which takes up to 10h to reach the minimum in the gradient plot. Interestingly, these two patients were known T1D patients with insufficient metabolic control and a history of precedent DKA, while the other three patients experienced DKA at diagnosis of T1D.

Fig. S9-S10 show the metabolic trajectories of the two HG patients (053 and 023) for which BE values were available. Again, the metabolic picture falls in the in-control domain, however, rather in the interface between both regions, especially for the lowest BE data points of 053. Finally,

worth of note are the participants without diabetes 051 and 007, whose data points fall well within the BE < -2 space. Interestingly, these two patients were receiving anti-seizure medication Sultiame, which inhibits carboanhydrase, resulting potentially in metabolic acidosis[18], which may explain why they are mapped within the -2 < BE PCA region.

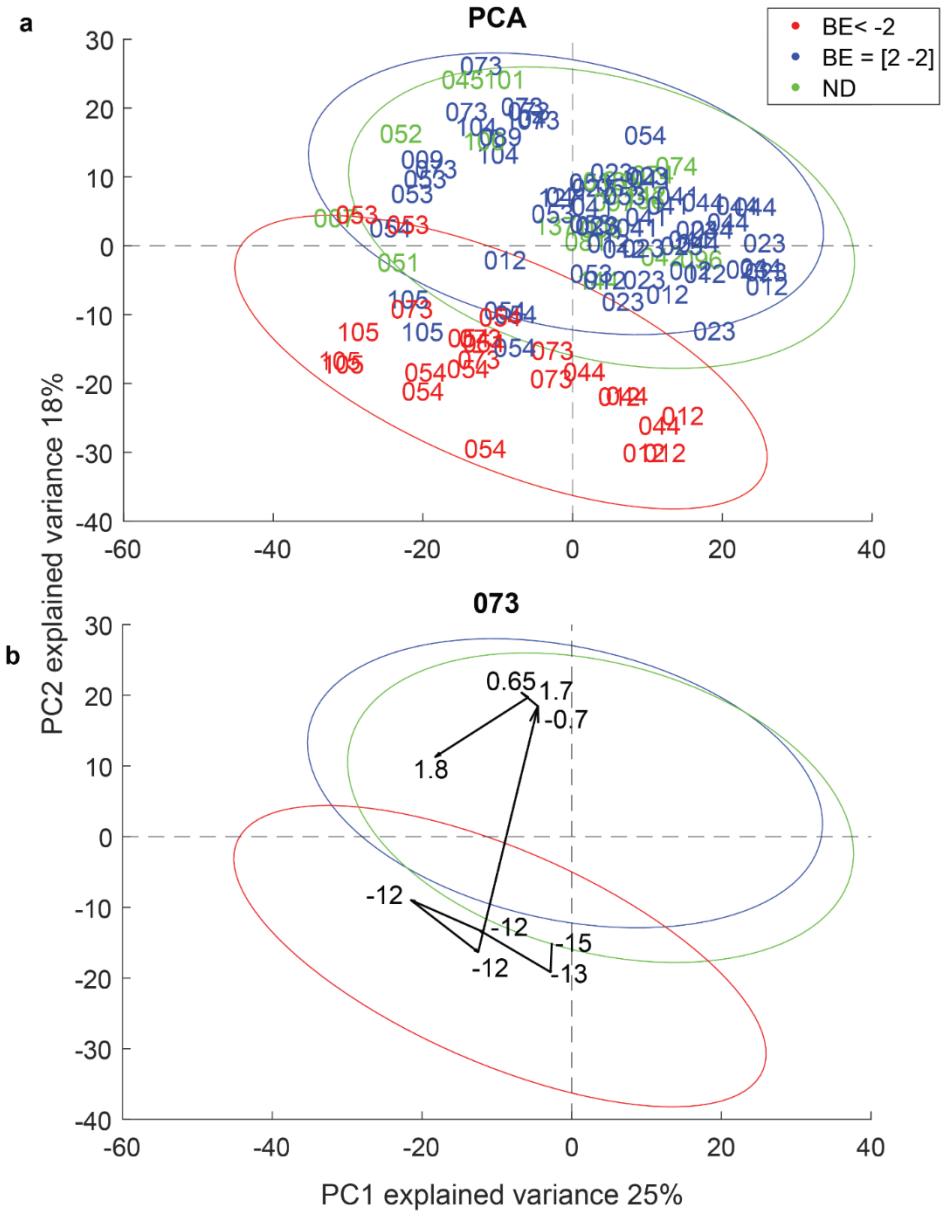


Figure 3 Panel a: PCA plot of all patients, Panel b: Metabolic trajectory of participant 073

Conclusion

We confirmed that the generally accepted DKA-associated perturbed pathways could be captured by exhaled metabolomics. A more untargeted approach further revealed metabolic trajectories of

DKA participants drifting apart during the course of their switch from catabolic to anabolic state, especially for those newly diagnosed. Taken all together, this pilot study suggests the exhaled profiles can track individualized metabolic trajectories mirroring the ketonic state.

However, our feasibility study has limitations that require an open discussion: i) The sample size is clearly limited. Thus, further independent validation studies will be required; ii) The identification of the majority of the metabolites reported here remains unknown, limiting the possibility of further interpreting our results from a mechanistic point of view. Further specialized measurements will be required to elucidate the structure of the most relevant metabolites[19]. Despite these noted limitations, it is fair to conclude that the method could be seamlessly integrated in a real ICU setting. It only requires inflating a breath sampling bag and the subsequent read-out is provided immediately after deflating the sample into the mass spectrometer, providing an attractive new approach to conveniently interrogate the metabolome of patients in a clinical setting. This opens new possibilities to integrate this technique for the monitoring, not only of DKA participants, but also to personalize the therapeutic management other patients with altered metabolism[20].

Acknowledgements

We thank Mélina Richard for coordinating the study. This work is part of the Zurich Exhalomics project under the umbrella of the University of Medicine Zurich/Hochschulmedizin Zürich. This study was funded by the Fondation Botnar, Switzerland (Professorship to Pablo Sinues), and the Swiss National Science Foundation (Grant No. 320030_173168).

Duality of Interests

PS is co-founder of Deep Breath Initiative AG (Switzerland), which develops breath-based diagnostic tools. KDS is a consultant for Deep Breath Initiative AG (Switzerland). The University Children's Hospital Basel is a shareholder of Deep Breath Intelligence AG (Switzerland). All other authors have no competing interest.

Author Contributions

MA, AND, UF, GS and PS designed the study. SBB, MAB, MH, AND, and GS contributed to the screening and recruitment of patients. MA and JZ performed the mass spectral measurements. MA, KDS and PS developed the codes to analyze the data. MA, KDS, AND, UF, GS and PS analyzed

and interpreted the data. MA and PS wrote the manuscript. All authors reviewed and contributed to the final version of the manuscript.

References

1. Vellanki, P.; Umpierrez, G. E., Increasing hospitalizations for DKA: a need for prevention programs. *Diabetes Care* **2018,** *41* (9), 1839-1841.
2. Chung, W. K.; Erion, K.; Florez, J. C.; Hattersley, A. T.; Hivert, M.-F.; Lee, C. G.; McCarthy, M. I.; Nolan, J. J.; Norris, J. M.; Pearson, E. R., Precision medicine in diabetes: a consensus report from the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes care* **2020,** *43* (7), 1617-1635.
3. Singh, K. D.; Del Miguel, G. V.; Gaugg, M. T.; Ibanez, A. J.; Zenobi, R.; Kohler, M.; Frey, U.; Sinues, P. M.-L., Translating secondary electrospray ionization-high-resolution mass spectrometry to the clinical environment. *J Breath Res* **2018,** *12* (2), 027113.
4. Decrue, F.; Singh, K. D.; Gisler, A.; Awchi, M.; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of Exhaled Breath Analysis with Parallel Lung Function and FeNO Measurements in Infants. *Anal. Chem.* **2021**.
5. Tassopoulos, C.; Barnett, D.; Fraser, T. R., Breath-acetone and blood-sugar measurements in diabetes. *The Lancet* **1969,** *293* (7609), 1282-1286.
6. Hancock, G.; Sharma, S.; Galpin, M.; Lunn, D.; Megson, C.; Peverall, R.; Richmond, G.; Ritchie, G. A.; Owen, K. R., The correlation between breath acetone and blood betahydroxybutyrate in individuals with type 1 diabetes. *Journal of Breath Research* **2020,** *15* (1), 017101.
7. Qiao, Y.; Gao, Z.; Liu, Y.; Cheng, Y.; Yu, M.; Zhao, L.; Duan, Y.; Liu, Y., Breath ketone testing: a new biomarker for diagnosis and therapeutic monitoring of diabetic ketosis. *BioMed research international* **2014,** *2014*.
8. Laffel, L., Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes/metabolism research and reviews* **1999,** *15* (6), 412-426.
9. Gisler, A.; Singh, K. D.; Zeng, J.; Osswald, M.; Awchi, M.; Decrue, F.; Schmidt, F.; Sievi, N. A.; Chen, X.; Usemann, J., An interoperability framework for multicentric breath metabolomic studies. *Iscience* **2022,** *25* (12), 105557.
10. Zhang, L.; Liu, C.; Jiang, Q.; Yin, Y., Butyrate in energy metabolism: there is still more to learn. *Trends in Endocrinology & Metabolism* **2021,** *32* (3), 159-169.
11. Mathew, A. V.; Jaiswal, M.; Ang, L.; Michailidis, G.; Pennathur, S.; Pop-Busui, R., Impaired amino acid and TCA metabolism and cardiovascular autonomic neuropathy progression in type 1 diabetes. *Diabetes* **2019,** *68* (10), 2035-2044.
12. Felig, P.; Marliss, E.; Ohman, J. L.; Cahill Jr, G. F., Plasma amino acid levels in diabetic ketoacidosis. *Diabetes* **1970,** *19* (10), 727-729.
13. Decsi, T.; Szabó, É.; Kozári, A.; Erhardt, É.; Marosvölgyi, T.; Soltész, G., Polyunsaturated fatty acids in plasma lipids of diabetic children during and after diabetic ketoacidosis. *Acta Pædiatrica* **2005,** *94* (7), 850-855.
14. Wallace, T.; Matthews, D., Recent advances in the monitoring and management of diabetic ketoacidosis. *Qjm* **2004,** *97* (12), 773-780.
15. Martinez-Lozano, P.; Fernandez de la Mora, J., Direct analysis of fatty acid vapors in breath by electrospray ionization and atmospheric pressure ionization-mass spectrometry. *Analytical Chemistry* **2008,** *80* (21), 8210-8215.
16. Ciechanowska, A.; Ladyzynski, P.; Wojcicki, J. M.; Sabalinska, S.; Krzymien, J.; Pulawska, E.; Karnafel, W.; Foltynski, P.; Kawiak, J., Microdialysis monitoring of glucose, lactate, glycerol, and pyruvate in patients with diabetic ketoacidosis. *The International journal of artificial organs* **2013,** *36* (12), 869-877.

17. Forni, L. G.; McKinnon, W.; Lord, G. A.; Treacher, D. F.; Peron, J.-M. R.; Hilton, P. J., Circulating anions usually associated with the Krebs cycle in patients with metabolic acidosis. *Critical care* **2005,** *9* (5), 1-5.
18. Leniger, T.; Wiemann, M.; Bingmann, D.; Widman, G.; Hufnagel, A.; Bonnet, U., Carbonic anhydrase inhibitor sulthiame reduces intracellular pH and epileptiform activity of hippocampal CA3 neurons. *Epilepsia* **2002,** *43* (5), 469-474.
19. Awchi, M.; Sinues, P.; Datta, A. N.; Garcia-Gomez, D.; Singh, K. D., UHPLC-MS/MS-Based Identity Confirmation of Amino Acids Involved in Response to and Side Effects from Antiseizure Medications. *J. Proteome Res.* **2023,** *22* (3), 990-995.
20. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; Awchi, M.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U.; Datta, A. N.; Sinues, P., Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications Medicine* **2021,** *1* (1), 21.

Chapter 7: General discussion and outlook

In this thesis, we have seen a few applications within the realm of pharmacometabolomics utilizing breath and SESI-HRMS. More specifically, we have discussed what breath analysis by SESI-HRMS can mean for TDM of ASMs and for the pathophysiological understanding and monitoring of diabetic ketoacidosis, all in the context of personalizing medication. In this chapter, the main findings are revisited and further elaborated together with their implications and we will make suggestions for future research.

7.1 Study I

In study I we saw the confirmation of 6 amino acids, namely GABA, Asp, Gln, Glu, Pro, and Lys. This was a follow-up work from a previous study[1], and, as compound identification is a bottle-neck in metabolomics, it comprised scientifically important work. The compounds were identified with an ES similarity of 0.95 (Tyr), 0.82 (GABA), 0.86 (Pro), 0.84 (Asp), 0.74 (Glu), and 0.73 (Gln), and could be level-1 identified according to the Szymanski diagram[2]. In addition, some adducts could be confirmed to be originating from the metabolites of origin at a "2d level''. To the best of our knowledge, this is the first time these metabolites have been confirmed in breath. Furthermore, we established that single mass features with an exact mass of, for example, m/z 147.07629 could originate from two or more forms of metabolites, leading to cautious interpretations of results from instrumentation for which no upstream separation is performed.

The work had some limitations; firstly, for some compounds, the ES similarity was below 0.75, (a threshold often used in large-scale metabolomics studies). This implies that the compound can be confirmed even below 0.75 of ES. However, validation work is needed to conclude with more certainty that these compounds were correctly annotated.

Secondly, another important limitation was that, in contrast to previous studies[3] [4], the MS/MS spectra could not be obtained in real-time. This raises the question whether these compounds are seen with the real-time technique, more sensitive instrumentation would be required to confirm this. Lastly, not all metabolites that were reported to be involved in response to and side effects from ASM could be confirmed. Future studies could potentially confirm these, if the annotation

was correct for these metabolites. In the meantime, the assigned metabolites must be interpreted cautiously. The following section will give an overview of further suggestions for future work.

7.1.1 Suggestions for future work

Origin confirmation

It must be noted that the confirmed amino acids have various functions within the human body, such as protein synthesis[5]. However, the main function of these metabolites is as (precursor) neurotransmitters[6]. Although this study focuses on epilepsy, a disease originating from the brain, no further conclusion could be drawn regarding the original source of the compounds. If these compounds truly originate from the brain, it could open many doors for future research utilizing breath, because neurotransmitters are a class of compounds used by the central nervous system (CNS) to transmit messages between neurons, or from neurons to muscles. They are synthesized by precursor neurotransmitters (p-neurotransmitters), metabolites that transform into key neurotransmitters, after which they can have three different effects on neurons: excitatory, inhibitory, or modulatory. The orchestration of this message transmission plays a key role in many aspects of our CNS as well as our bodily functions[6]. Consequently, many diseases originating from the brain, such as epilepsy[7], Alzheimer's disease[8], Parkinson's disease[9], depression[10], and even schizophrenia[11] are linked to imbalances of these neurotransmitters, emphasizing the importance of analytically identifying and quantifying these compounds. A possible method would be to use magnetic resonance spectroscopy (MRS)[12], to confirm that these detected compounds do originate from the brain. GABA values obtained from the MRS could be compared with GABA levels in breath, and—if significant correlation is observed—a link could be made.

Up-concentration

In this study only one up-concentration method was used: SPE with HLB sorbent to up-concentrate 100 mL of EBC down to 500 uL (~200 times theoretical up-concentration). However, different metabolites have different physical-chemical properties and therefore absorb at different rates to stationary phases. As breath compromises such a wide array of metabolites[13], different methodologies will yield a higher turnover of metabolites suggested to be present in breath. HLB is a very generic sorbent, and it may lead to specific compounds not being retained. Therefore, for future studies, a mix of sorbents could be used and combined afterwards for a final solution. The

proposed methods could be based on weak- and strong cation exchange polymers, weak and strong anion exchange polymers, and mix-mode polymers[14].

Metabolite annotation

An in-house algorithm was used to identify metabolites of interest. We took all the $M+H^+$ ions previously reported to be involved in response to and side effects from ASM and inspected the chromatogram. If there was no clear peak or no peak above 1e4, or if there was a mismatch between adducts and the $M+H^+$ peak, the corresponding metabolite was removed. Usually, in large-scale metabolomics studies, metabolite identification is done via spectral library searching or by comparing it to a molecular structure database in which fragmentation patterns are produced in-silico. These methods include models such as MetFrag[15], MAGMa[16], and CSI:fingerID[17]. The downside of spectral library searching is that when no standard exists, no annotation can be performed.

For future large-scale breath feature confirmation studies, an option such as CSI:fingerID should be considered, to be able to scan the entire range of possible metabolites present, after which a certain level of confidence about that compound could be obtained. Level 1 confirmation can then be performed on the metabolites that were a potential hit. Together with the up-concentration strategy previously mentioned, this ideally should yield high amounts of simultaneously confirmed metabolites.

7.2 Study II

In the second study, we looked at the ASM VPA, a drug that needs to be therapeutically managed because high doses can be toxic while low doses might not have any effect. In a previous study[1] we identified metabolic features that were mostly associated with the tyrosine metabolism being either up- or down-regulated by patients suffering from side effects or who were not responding to medication. The utility of breath analysis by SESI-HRMS for TDM of VPA had already been demonstrated in this previous study, even though it was done by real-time analysis. Currently, patients incapable of performing the required real-time exhalation cannot benefit from this technique. Therefore, we looked for an alternative option with the use of off-line SESI-HRMS analysis. We compared the results obtained from the golden standard (i.e., real-time SESI-HRMS) with the off-line technique to assess the transferability of the method. Off-line monitoring was

successfully introduced based on Lin's CCC (a method for comparing a new test with the golden standard) of 0.47 for total VPA and 0.60 for free VPA. The modest coherence for total VPA of 0.47 can be attributed to a non-linear relationship of the datapoints outside the therapeutic range. The same was observed for free VPA but to a lesser extent. We can assume that free VPA represents a closer read-out that exhaled breath represents from the free fraction in comparison to the total concentration, which includes protein bound VPA.

We also found a large disparity between the real-time and off-line analysis for the metabolites involved in the response to and side effects from anti-seizure medication. This suggests that we lose information with the off-line method: an observation previously reported[18]. A potential way to overcome this is to look firstly at alternative materials for off-line collection and secondly at preservation methods (a study currently being conducted). Ideally, the approach to performing TDM on ASMs by breath analysis should be extended to other ASMs beyond VPA. This will make SESI-HRMS an attractive approach for hospitals to perform TDM on a wide array of ASMs.

7.2.1 Suggestions for future work

Model selection and implementation

The selection of the model plausibly influenced the performance of Lin's CCC. Even though both linear (linear regression, linear SVM, etc.) and non-linear models (exponential GPR, cubic SVM) were trained, a modest fit was observed. The poor performance of the non-linear models can be attributed to the lack of the number of datapoints, thus, in future studies additional data points must be considered, to make a more accurate prediction. Another reason for the modest results could be the zero-imputation method. Zeroes were imputed by two different methods. For the ^{13}C and ^{18}O isotopes of VPA, the zeros were imputed by their natural abundance, whereas for the ammonium adducts, the zeros were imputed by multiplying the corresponding $[M+H]^+$ signal intensity by the mean of all non-zero $[M+NH4]^+$ values. Other zero imputation methods[19] such as k-mean's nearest neighbor could have been considered for potential model optimization. Finally, another method to boost performance is to engineer features. Some of the metabolites were directly associated with VPA itself whereas others were metabolized products. Boosting the importance (applying a weight on features) of metabolites directly coming from VPA might also improve the results.

7.3 Study III

In this study, we looked at the metabolic trajectories of metabolites in breath in patients with DKA after onset of insulin treatment. We observed that the metabolites which play a key role in DKA (Acetone, acetoacetate (AcAc), pyruvate, etc.) are readily detectable in breath and that their behavior mirrors previously reported results. This makes breath analysis an attractive approach to monitor these compounds during acute DKA. Additionally, both breath AcAc and breath acetone were identified as biomarkers for DKA monitoring. Breath AcAc showed a more comprehensive mirroring to BE, suggesting that it might be a more robust marker for DKA monitoring in comparison to acetone. It is worth mentioning that this was the first time that breath AcAc was proven to be a marker for DKA monitoring. A more untargeted approach revealed that 665 mass spectral features correlated significantly with BE: 219 showed positive correlations and 446 showed negative. PCA plots revealed that the DKA state is distinguishable between the -control group and non-diabetic patients. In addition, these correlating features with BE prompt metabolic trajectories toward the in-control state as they progress toward homeostasis. Overall, the study provided a proof-of-principle for using exhaled breath as a DKA monitoring tool in an intensive care setting. The study, however, also had some limitations, mainly with regards to sample size. This limitation requires additional (large-scale) validation studies to confirm the relationships observed. We must also note that most compounds, especially the correlating features, were only "level 4'' confirmed according to the Szymanski diagram[2]. Which requires further specialized EBC measurements for unequivocal confirmation is required, as we have previously seen in study I.

7.3.1 Suggestions for future work

Putting breath analysis to the test

As previously mentioned, we were able to confirm that Acetone makes a trajectory comparable with BE, a trend that is confirmed with previous studies which used ketones bodies rather than BE. Additionally, acetoacetate was identified in breath for the first time and also was found as correlating with BE. Besides these observations, we could dissect the generally accepted pathway of breath analysis, which was has only previously been observed in other bodily fluids like blood and cerebrospinal fluid (CSF). However, we found results which were not coherent with the current literature, for example the non-significant alteration of the TCA cycle metabolites. This can be interpreted in several ways. For one, the annotation of the metabolite might not have been correct.

For another, breath values might not mimic the exact metabolite concentrations as found in blood. Since there is a vast literature about the exact pathophysiology of DKA, one could use these observations as an ultimate test for breath. For future studies we therefore suggest to take blood samples simultaneously with breath. Specific questions could then be asked (e.g., are the amino acids previously observed to exist in breath up- or down-regulated, and how do these results correlate with the blood?), and results could then be compared. This will additionally provide insights into the mechanism of the air-blood barrier. If should the results between blood and breath, not be comparable this would mean that the air-blood barrier does not exactly mimic blood concentrations and is selective.

Multi-omics

For this study we only looked at the effect of insulin on metabolites captured in the breath of DKA patients. However, more processes contribute to the complex pathophysiology of DKA. An emerging field that focuses on all the aspects of the "-omic-fields" is multi-omics. This field takes genomic, transcriptomics, proteomics, and metabolomics into account by considering the whole pathophysiology of the disease. Previous studies have shown genomics[20], transcriptomics[21], proteomics[22], and—in our current study—metabolomics to be important parameters for DKA diagnosis, monitoring and pathophysiological understanding. Genomics, for example, can help identify genetic risk factors for DKA, while transcriptomics can provide insights into gene expression changes in response to DKA. Proteomics, on the other hand, can identify changes in protein expression. Undertaking such a study will lead to a greater understanding of key pathways and networks that are dysregulated in DKA, which in turn could provide novel insights into therapeutic targets.

7.4 Conclusion

In this thesis, we have seen the utility of breath analysis using SESI-HRMS in the context of pharmacometabolomics for type-1 diabetes and epilepsy. We have shown breath analysis to be a powerful technique that allows us to tap into a wide range of (pharmaco)metabolic information ranging from amino acids to drug metabolites. Samples are taken in a completely non-invasive manner, with no need for sample preparation, and results are obtained within 15 minutes. These attributes show the potential of breath analysis to become a widely adopted tool where it can aid

clinical assessment in diseases that have their origin in metabolites and are measurable with breath. Tremendous amounts of work are needed to uncover the full potential of breath analysis, as some bottlenecks remain which need to be addressed. One such bottleneck is the need for unequivocal identification of biomarkers. In this thesis, we showed an application of the confirmation of potential biomarkers via UHPLC-HRMS and made some suggestions for further studies on with regards to wide-range metabolite confirmation. In addition, reproducibility and quantification remain challenges to current-day breath analysis by SESI-HRMS. An effort to increase reproducibility has been performed by our lab and we observed that, over one month, the reproducibility between three sites could be decreased with ~20%[13], an observation comparable with other mass spectrometric analyses. However, to be fully reproducible, the ionization mechanism of secondary electrospray ionization must be better understood. This will aid quantification, as we have seen with SIFT-MS, for which the ionization mechanism is understood. Furthermore, validation studies must also be performed in order to make breath analysis reliable and more robust—an effort currently being pursued by our lab as well. Finally, as we saw in study I, the measurement of isomers with SESI-HRMS remain cumbersome, if not impossible. A possible strategy to overcome this is to combine SESI-HRMS with complementary techniques such as GC or ion mobility spectrometry.

In conclusion, this study demonstrated the potential of breath analysis for pharmacometabolomic studies in epilepsy and type-1 diabetes. Our findings provide valuable insight into the metabolic pathways and biomarkers associated with these disorders and offer new opportunities for non-invasive diagnosis and treatment monitoring. Overall, this study contributes to the growing body of literature on the use of metabolomics and breath analysis in biomedical research and highlights the importance of interdisciplinary approaches for advancing our understanding and aide of complex diseases.

References

1. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; Awchi, M.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U., Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications medicine* **2021,** *1* (1), 1-11.
2. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: communicating confidence. ACS Publications: 2014.
3. Gaugg, M. T.; Engler, A.; Bregy, L.; Nussbaumer-Ochsner, Y.; Eiffert, L.; Bruderer, T.; Zenobi, R.; Sinues, P.; Kohler, M., Molecular breath analysis supports altered amino acid metabolism in idiopathic pulmonary fibrosis. *Respirology* **2019,** *24* (5), 437-444.
4. Tejero Rioseras, A.; Singh, K. D.; Nowak, N.; Gaugg, M. T.; Bruderer, T.; Zenobi, R.; Sinues, P. M., Real-Time Monitoring of Tricarboxylic Acid Metabolites in Exhaled Breath. *Anal Chem* **2018,** *90* (11), 6453-6460.
5. Hamanaka, R. B.; O'Leary, E. M.; Witt, L. J.; Tian, Y.; Gökalp, G. A.; Meliton, A. Y.; Dulin, N. O.; Mutlu, G. M., Glutamine metabolism is required for collagen protein synthesis in lung fibroblasts. *American journal of respiratory cell and molecular biology* **2019,** *61* (5), 597-606.
6. Hyman, S. E., Neurotransmitters. *Current biology* **2005,** *15* (5), R154-R158.
7. Werner, F.-M.; Coveñas, R., Classical neurotransmitters and neuropeptides involved in generalized epilepsy in a multi-neurotransmitter system: How to improve the antiepileptic effect? *Epilepsy & Behavior* **2017,** *71*, 124-129.
8. Palmer, A. M.; Stratmann, G. C.; Procter, A. W.; Bowen, D. M., Possible neurotransmitter basis of behavioral changes in Alzheimer's disease. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **1988,** *23* (6), 616-620.
9. Hornykiewicz, O., Brain neurotransmitter changes in Parkinson's disease. In *Movement disorders*, Elsevier: 1981; pp 41-58.
10. Nutt, D. J., Relationship of neurotransmitters to the symptoms of major depressive disorder. *J Clin Psychiatry* **2008,** *69* (Suppl E1), 4-7.
11. Carlsson, A.; Waters, N.; Carlsson, M. L., Neurotransmitter interactions in schizophrenia—therapeutic implications. *Biological psychiatry* **1999,** *46* (10), 1388-1395.
12. Puts, N. A.; Edden, R. A., In vivo magnetic resonance spectroscopy of GABA: a methodological review. *Progress in nuclear magnetic resonance spectroscopy* **2012,** *60*, 29.
13. Gisler, A.; Singh, K. D.; Zeng, J.; Osswald, M.; Awchi, M.; Decrue, F.; Schmidt, F.; Sievi, N. A.; Chen, X.; Usemann, J., An interoperability framework for multicentric breath metabolomic studies. *Iscience* **2022,** *25* (12), 105557.
14. Rodríguez-Gonzalo, E.; Hernández-Prieto, R.; García-Gómez, D.; Carabias-Martínez, R., Development of a procedure for the isolation and enrichment of modified nucleosides and nucleobases from urine prior to their determination by capillary electrophoresis–mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis* **2014,** *88*, 489-496.
15. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *Journal of cheminformatics* **2016,** *8* (1), 1-16.
16. Ridder, L., MAGMa-Based Mass Spectrum Annotation in CASMI 2014. *Current Metabolomics* **2017,** *5* (1), 18-24.

17. Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S., Searching molecular structure databases with tandem mass spectra using CSI: FingerID. *Proceedings of the National Academy of Sciences* **2015,** *112* (41), 12580-12585.
18. Decrue, F.; Singh, K. D.; Gisler, A.; Awchi, M.; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of exhaled breath analysis with parallel lung function and FeNO measurements in infants. *Analytical Chemistry* **2021,** *93* (47), 15579-15583.
19. Armitage, E. G.; Godzien, J.; Alonso‐Herranz, V.; López‐Gonzálvez, Á.; Barbas, C., Missing value imputation strategies for metabolomics data. *Electrophoresis* **2015,** *36* (24), 3050-3060.
20. Robertson, C. C.; Rich, S. S., Genetics of type 1 diabetes. *Current opinion in genetics & development* **2018,** *50*, 7-16.
21. Qiu, H.; Novikov, A.; Vallon, V., Ketosis and diabetic ketoacidosis in response to SGLT2 inhibitors: basic mechanisms and therapeutic perspectives. *Diabetes/metabolism research and reviews* **2017,** *33* (5), e2886.
22. Bakinowska, L.; Vartak, T.; Phuthego, T.; Taylor, M.; Chandler, K.; Jerram, S. T.; Williams, S.; Feldmann, M.; Johnson, D. G.; Patel, K. A., Exocrine Proteins Including Trypsin (ogen) as a Key Biomarker in Type 1 Diabetes. *Diabetes Care* **2023,** *46* (4), 714-721.

Supporting Information

Study I

UHPLC-MS/MS based identity confirmation of amino acids involved in response to and side effects from antiseizure medications

Mo Awchi[1,2], Pablo Sinues[1,2], Alexandre N. Datta[1], Diego García-Gómez[3], Kapil Dev Singh[1,2,*]

1. University Children's Hospital Basel, University of Basel, Basel, Switzerland.
2. Department of Biomedical Engineering, University of Basel, Basel, Switzerland.
3. Department of Analytical Chemistry, University of Salamanca, Salamanca, Spain.

* Corresponding author: kapil.singh@ukbb.ch

Table of Contents

Table S1. Comparison of tune settings between the HESI source used during UHPLC-MSMS analysis of this study and the SESI source used during real-time breath analysis from the previous study.

	Settings	HESI source	SESI source
Source specific	Sheath gas flow rate	10	60
	Aux gas flow rate	0	2
	Sweep gas flow rate	0	0
	Spray voltage (|kV|)	4	3.5
	Probe heater temp. (°C)	30	-
	Sampling line temp. (°C)	-	130
	Core temp. (°C)	-	90
MS specific	Capillary temp. (°C)	320	275
	S-lens RF level	55	55

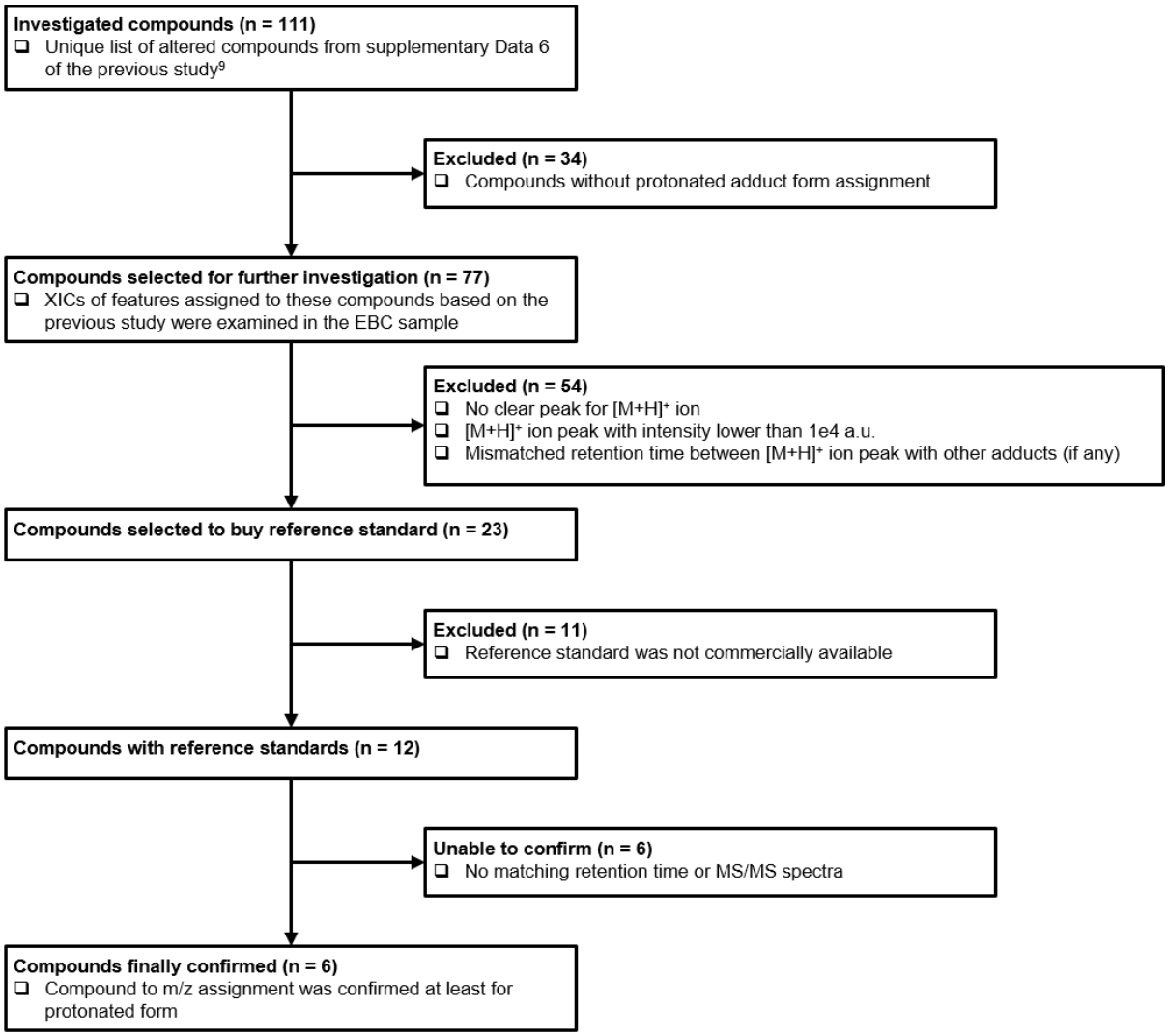


Figure S1. Workflow used to select compounds for UHPLC-MSMS based confirmation.

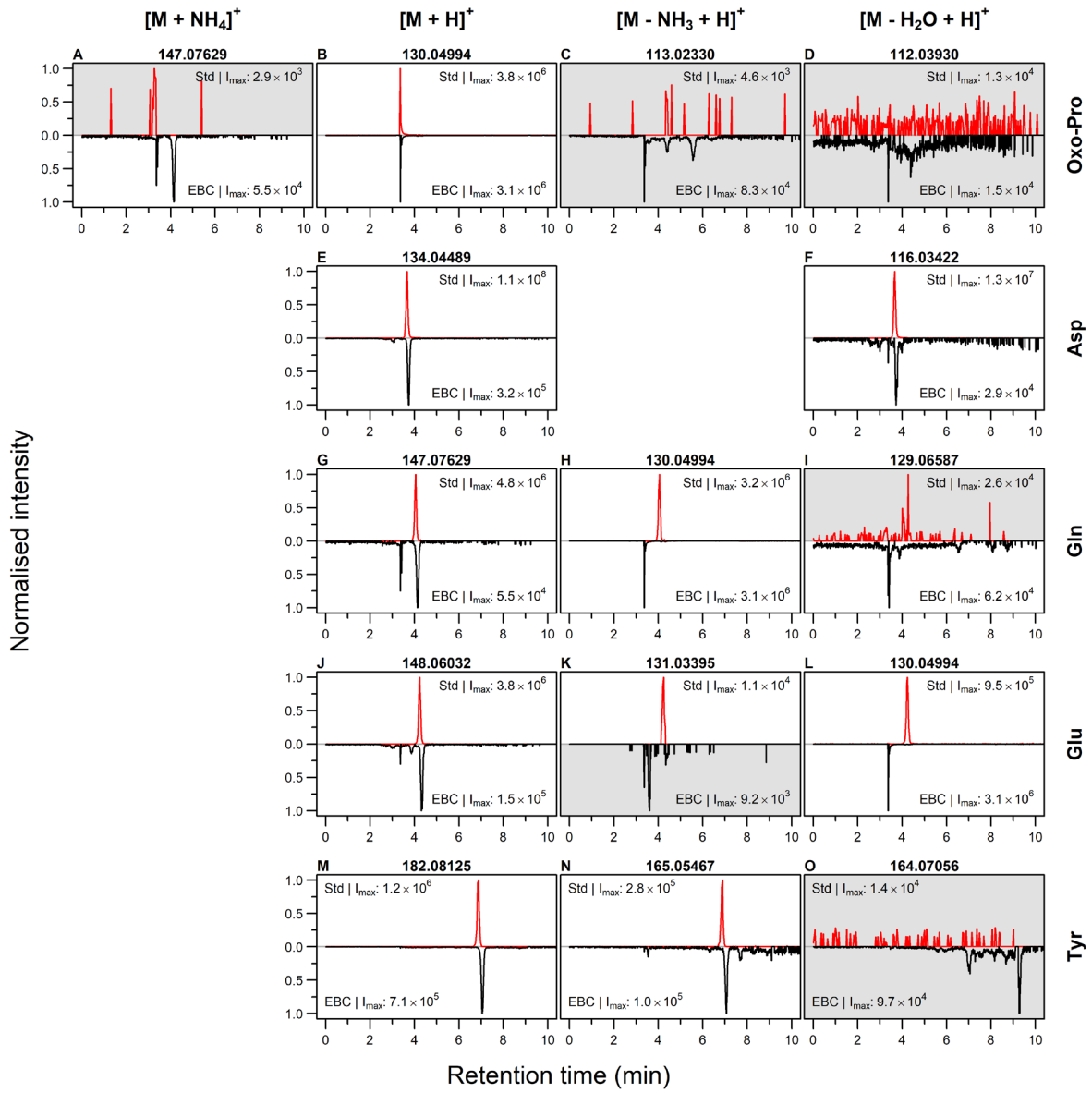


Figure S2. Comparison of LC-MS chromatograms between standards of six selected compounds and EBC for different adduct forms, as defined in Table 1. Grey background points to the LC-MS chromatogram (i.e. XIC) with no clear peak.

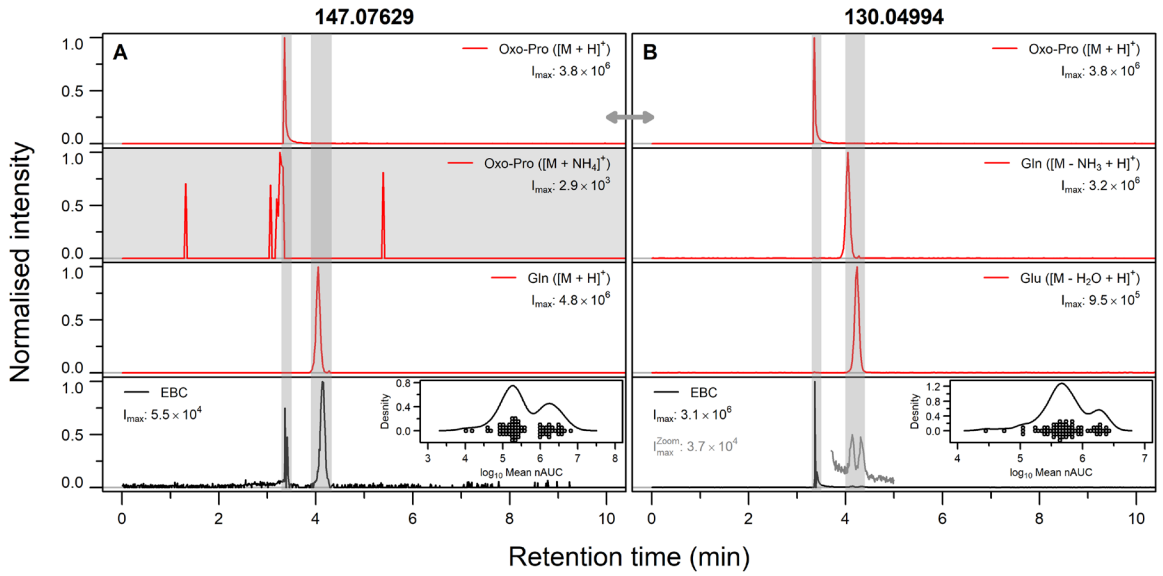


Figure S3. Features at m/z 147.07629 and 130.04994 appear to be associated with different adduct forms of more than one compound. The figure shows the comparison of LC-MS chromatograms between different standards (with different adduct forms) and EBC. Chromatograms shown here are same as the corresponding ones from Figure S2 (just rearranged here to ease comparison). Like Figure S2, chromatograms with no clear peak are shown with grey background. Furthermore, peaks in the XIC of EBC are highlighted with grey vertical bars to ease comparison. Finally, the grey trace in the EBC of panel B around 4 min represents ~100 times zoomed XIC in the y-axis, to reveal two smaller peaks. Inset in EBC panel shows the distribution of mean normalised area under the curve (nAUC) of selected feature from previous real-time breath analysis,[9] as density curves accompanied by actual data points, suggesting that m/z147.07629 possibility corresponds to two different compounds in the previous study. Note: The chromatogram for protonated Oxo-Pro in both panel is the same XIC for m/z 130.04994 as denoted by the bidirectional arrow, see text for more details.

Study II

Prediction of systemic free and total valproic acid by off-line analysis of exhaled breath in epileptic children and adolescents

Mo Awchi[1,2], Kapil Dev Singh[1,2], Patricia E. Dill[1], Urs Frey[1], Alexandre N. Datta[1,*], Pablo Sinues[1,2,*]

1. University Children's Hospital Basel, University of Basel, Basel, Switzerland.

2. Department of Biomedical Engineering, University of Basel, Basel, Switzerland.

*pablo.sinues@unibas.ch

*alexandre.datta@ukbb.ch

Table of Contents

Table S1 Subject identifier, demographics, measurement date and anti-seizure medication information

Subject Identifier	Measurement Date	Gender	Age	Anti-seizure Medication	Blood Concentration
BreathTDM_041-1	20180423	Female	13	VPA (FF)	107.5 (14.6)
BreathTDM_043-1	20180424	Female	6	LEV | VPA	24.04 | 54.08
BreathTDM_003-3	20180723	Female	9	LTG | VPA (FF)	13 | 65.5 (13.3)
BreathTDM_005-3	20180817	Male	14	LEV | VPA (FF)	63.69 | 59.8 (NA)
BreathTDM_028-2	20180817	Male	15	LTG | VPA	6.7 | 54.3
BreathTDM_056-1	20180821	Male	10	LTG | VPA (FF)	4.4 | 60.0 (6.3)
BreathTDM_064-2	20200225	Female	10	VPA (FF) | LEV	97 (9.1) | NA
BreathTDM_069-2	20200701	Female	7	VPA	NA
BreathTDM_064-3	20200706	Female	10	VPA (FF)	107.7 (15.1)
BreathTDM_005-5	20200717	Male	16	VPA (FF)	29.6 (1.6)
BreathTDM_056-3	20200731	Male	12	VPA (FF) | LTG	37.5 (2.3) | 5.8
EBECA_004-1	20200929	Female	11	VPA | LTG | STP | CLB	81 | 11.3 | 1.4 | 0.19
EBECA_005-1	20200929	Male	9	VPA	99.4
EBECA_006-1	20200929	Female	8	VPA (FF)	105.7 (13.9)
EBECA_004-2	20201216	Female	12	VPA (FF) | LTG | STP	69 (6.8) | 14.8 | 1.7
EBECA_061-1	20210330	Male	17	VPA (FF)	63.4 (4.0)
EBECA_062-1	20210331	Male	10	VPA (FF)	100.6 (11.7)
EBECA_064-1	20210423	Female	13	VPA	110.4
EBECA_070-1	20210512	Male	10	VPA	143.6
EBECA_005-2	20210519	Male	10	VPA	104.8
EBECA_006-2	20210519	Female	8	VPA	102
EBECA_075-1	20210527	Male	9	VPA	72.4
EBECA_081-1	20210618	Male	19	VPA	80.4
EBECA_082-1	20210706	Female	16	VPA	NA
EBECA_083-1	20210706	Female	4	VPA	94.1
EBECA_096-1	20210827	Female	14	LTG | VPA	11 | 29.4
EBECA_100-1	20210924	Female	12	VPA	NA
EBECA_109-1	20211011	Male	8	VPA (FF)	81.5 (6.6)
EBECA_102-1	20211109	Male	7	VPA	NA
EBECA_114-1	20220204	Male	10	VPA | LTG	NA | NA
EBECA_006-3	20220208	Female	9	VPA (FF)	81.8 (11.4)
EBECA_005-3	20220208	Male	11	VPA (FF)	74.9 (10.1)
EBECA_133-1	20220601	Female	12	VPA | ESX	NA | NA
EBECA_004-3	20220701	Female	13	VPA | LTG | CLB | STP	NA | NA | NA | NA
EBECA_141-1	20220713	Female	16	VPA	NA
EBECA_001-2	20220802	Male	20	VPA	56.8
EBECA_005-4	20220816	Male	11	VPA (FF)	84.8 (5.8)
EBECA_147-1	20220831	Male	17	VPA	NA
EBECA_148-1	20220906	Male	10	VPA | STP | CLB	138 | 0.9 | 0.145

EBECA_096-2	20220907	Female	15	VPA (FF) | LTG	33.2 (1.8) | 11.1

1) In antiseizure drug (ASD) information, names and blood concentration units are as follows:
   (a) Lamotrigine (LTG), Stiripentol (STP), and Valproic acid (VPA) are in mg/L.
   (b) Clobozam (CLB), Ethosuximide (ESX), Levetiracetam (LEV) and Sultiame (STM) are in µmol/L.
2) In blood concentration
   (a) Value inside parentheses represents concentration of free VPA.
   (b) NA = not available.

Table S2 Valproic acid biomarkers zero imputation details

Molecule	Imputation factor Real-Time	Imputation factor Offline	Number of zeros replaced Real-Time	Number of zeros replaced Off-line
3-Heptanone	NA	NA	0	0
^{13}C 3-Heptanone*	0.078	0.078	0	0
NH_4+ 3-Heptanone**	0.008	0.005	0	5
Heptanedione	NA	NA	0	0
^{13}C Heptanedione*	0.078	0.078	0	0
4-OH-γ-Lactone	NA	NA	0	0
^{13}C 4-OH-γ-Lactone*	0.089	0.089	0	0
^{18}O 4-OH-γ-Lactone*	0.008	0.008	1	3
NH_4+ 4-OH-γ-Lactone**	0.102	0.076	0	2
^{13}C NH_4+ 4-OH-γ-Lactone**	0.009	0.007	0	5

*denotes imputation computed by isotopic distribution
**denotes imputation computed by factor between $M+H^+$ and corresponding adduct

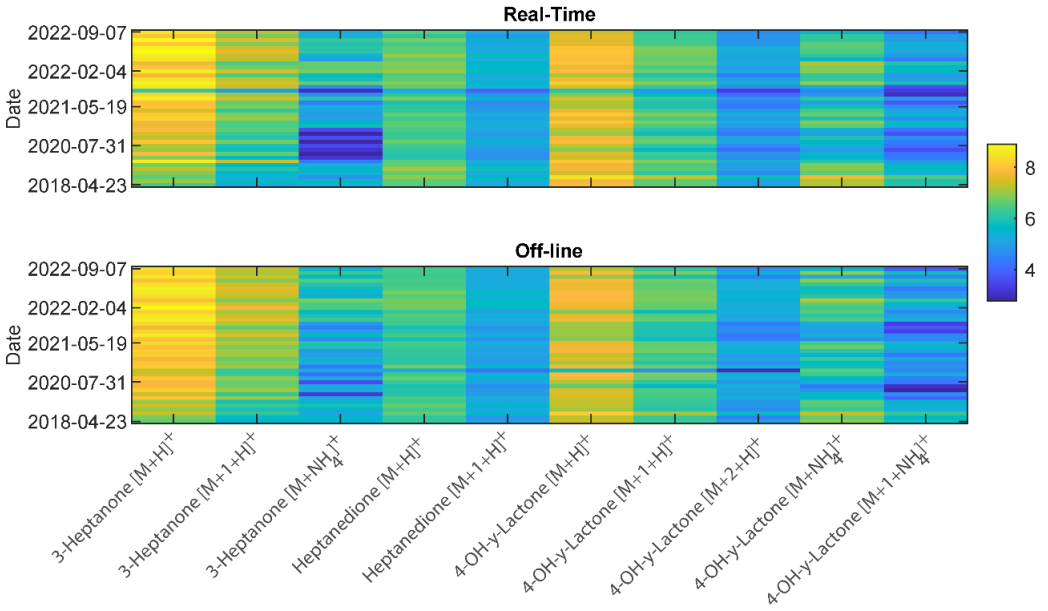


Figure S1 VPA biomarkers signal intensity over time

g

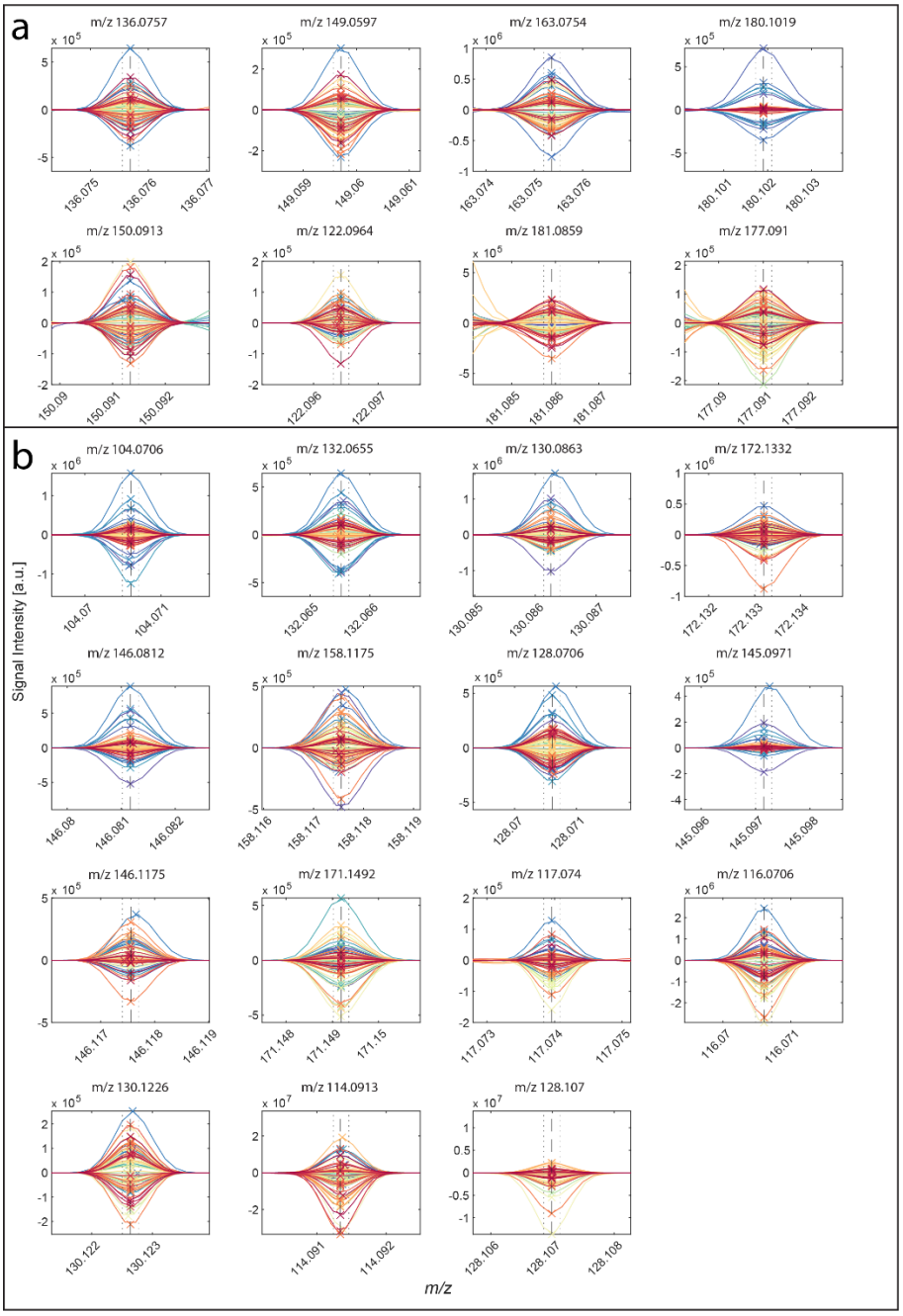


Figure S2 Drug response (panel a) and Side effect (panel b) metabolites profile mass spectral peaks

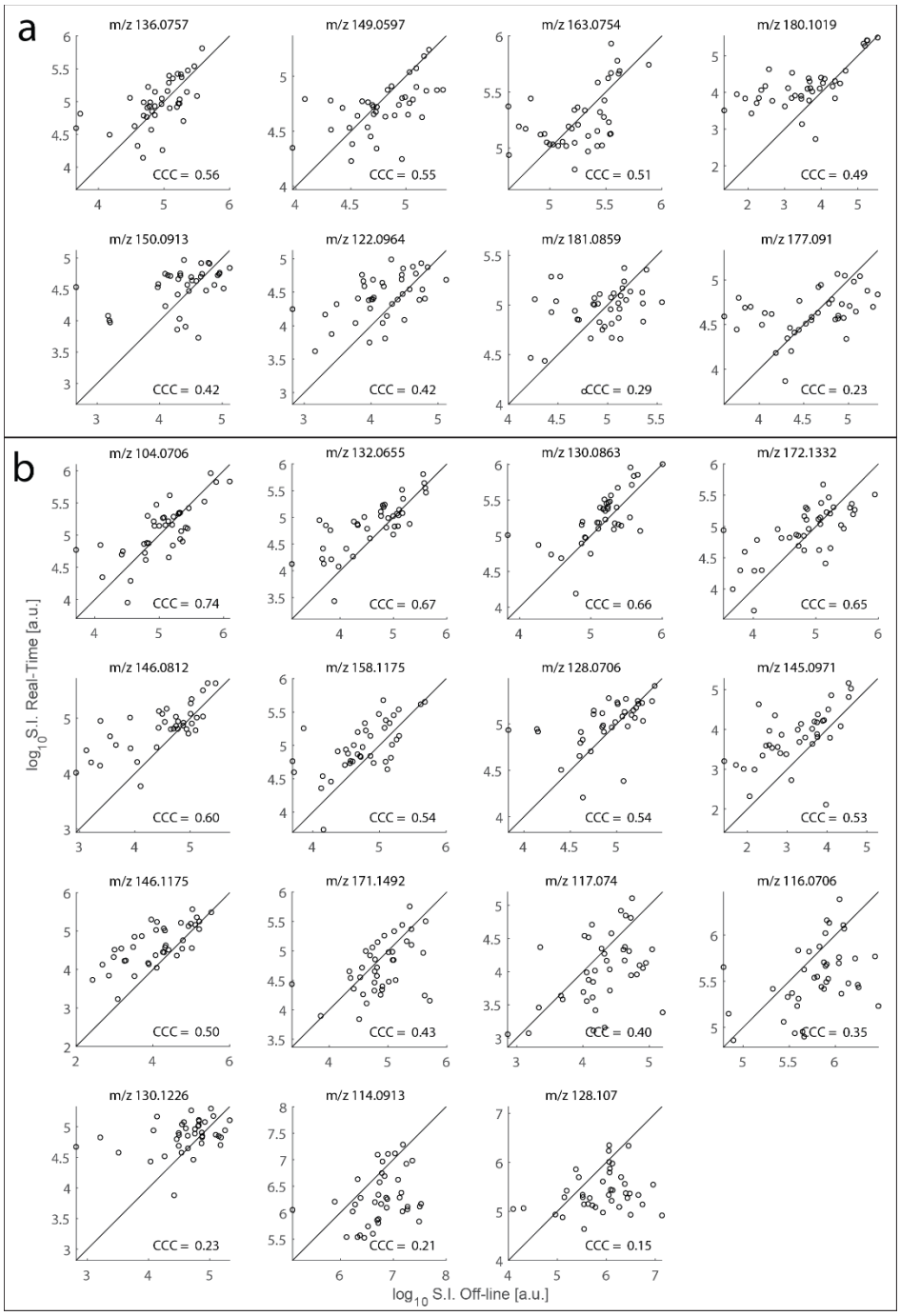


Figure S3 Lin's CCC correlation plots for drug response (panel a) and side effect metabolites (panel b)

Table S3 Lin's CCC statistics for all metabolites found in response to and side effects from antiseizure medications

m/z	ppm Error	Lins CCC	Confidance Interval	Scale Shift	Location Shift	Bias Correc-tion	Pearson Corr. Coeff.	SE or DR*
104.0706	-0.049	0.74	[0.57,0.85]	-0.10	0.99	0.75	1.12	SE
146.1176	0.002	0.67	[0.50,0.80]	-0.47	0.87	0.78	1.34	SE
128.0706	0.114	0.66	[0.45,0.80]	-0.30	0.96	0.69	1.03	SE
128.1070	-0.031	0.65	[0.45,0.78]	-0.19	0.94	0.69	1.37	SE
117.0740	0.001	0.60	[0.42,0.73]	-0.62	0.79	0.76	1.47	SE
130.1226	-0.034	0.54	[0.32,0.71]	-0.53	0.86	0.63	1.20	SE
146.0812	-0.040	0.54	[0.31,0.72]	-0.37	0.92	0.59	1.20	SE
171.1492	-0.029	0.53	[0.33,0.69]	-0.70	0.79	0.67	1.19	SE
116.0706	-0.036	0.50	[0.32,0.65]	-0.82	0.70	0.71	1.50	SE
114.0914	0.001	0.43	[0.16,0.65]	0.34	0.94	0.46	1.03	SE
172.1332	0.001	0.40	[0.14,0.61]	0.51	0.88	0.45	0.98	SE
130.0863	-0.044	0.35	[0.10,0.56]	0.66	0.82	0.42	1.04	SE
145.0972	-0.005	0.23	[0.02,0.43]	-0.67	0.70	0.33	1.89	SE
132.0655	0.082	0.21	[0.00,0.40]	1.04	0.65	0.32	0.99	SE
158.1176	-0.005	0.15	[-0.05,0.34]	0.98	0.63	0.24	1.56	SE
136.0757	-0.004	0.56	[0.32,0.74]	1.17	-0.16	0.98	0.58	DR
149.0597	0.026	0.55	[0.31,0.72]	1.28	0.27	0.94	0.59	DR
163.0754	0.024	0.51	[0.24,0.70]	1.12	-0.01	0.99	0.51	DR
180.1019	0.027	0.49	[0.30,0.64]	1.64	-0.74	0.71	0.68	DR
150.0913	-0.004	0.42	[0.18,0.62]	1.45	-0.44	0.86	0.49	DR
122.0964	0.033	0.42	[0.19,0.60]	1.38	-0.66	0.79	0.53	DR
181.0859	-0.004	0.29	[-0.01,0.54]	1.28	-0.06	0.97	0.30	DR
177.0910	0.021	0.23	[-0.06,0.49]	1.41	-0.03	0.94	0.24	DR

*SE= Side effects, DR = Drug response

Study III

Metabolic trajectories of diabetic ketoacidosis onset described by breath analysis

Mo Awchi1,2, Kapil Dev Singh1,2, Sara Bachmann Brenner1, Marie-Anne Burckhardt1, Melanie Hess1, Jiafa Zeng1,2, Alexandre N. Datta1, Urs Frey1, Urs Zumsteg1, Gabor Szinnai1,*, Pablo Sinues1,2,*

1. University Children's Hospital Basel, Basel, Switzerland.

2. Department of Biomedical Engineering, University of Basel, Basel, Switzerland.

*gabor.szinnai@ukbb.ch

*pablo.sinues@unibas.ch

Table of Contents

Extended Methods

Offline breath measurements were performed as described previously[1]. The bag samples were collected using custom made nalophan bags and subsequently infused into the mass spectrometer within 15 minutes of collection. The mass spectrometer operated using an interface which consisted of a SESI source (FIT, Spain) coupled to an HRMS instrument (Orbitrap, Thermo Fisher, USA). The instrumental setting for the mass spectrometer were as follows; sheath gas flow rate of 60, auxiliary gas flow rate of 2, spray voltage at 2.8 kV in positive ion mode, the capillary temperature of 275 °C, and S-lens RF level of 55. A nano-spray was created by utilizing 0.1% formic acid in water (Sigma Aldrich), whereby the current ranged between 70 and 140 nA. SESI sampling line temperature was set at 130 °C and the ion source at 90 °C. The MS was operated in full-scan mode ranging from 55-825 m/z, with a maximum inject time of 500 ms, and automatic gain control target of 10^6. The mass spectral resolution was set at 140,000 at m/z 200. The mass spectrometer was calibrated weekly and a suitability test (see Gisler et al[2] for details) was performed before every day of measurements.

Raw centroid and profile mass spectra were accessed using in-house C# console apps based on Thermo Fisher Scientific's RawFileReader (version 5.0.0.38). Centroid and profile mass spectra were recalibrated using reference peaks. Hereafter, time traces for each mass spectrometric feature was extracted and the area under the curve was integrated numerically. Zeros were imputed using the regression on order statistics method[3]. This procedure resulted in a 107 x 3505 data matrix. The time points between the clinical measurement and breath measurement differed, to obtain a BE value during the breath measurements, BE values were interpolated using linear interpolation on minute resolution for the whole-time frame (i.e. during their ICU stay). Time points overlapping with breath measurements were extracted and the corresponding BE value was used for the trajectory plots.

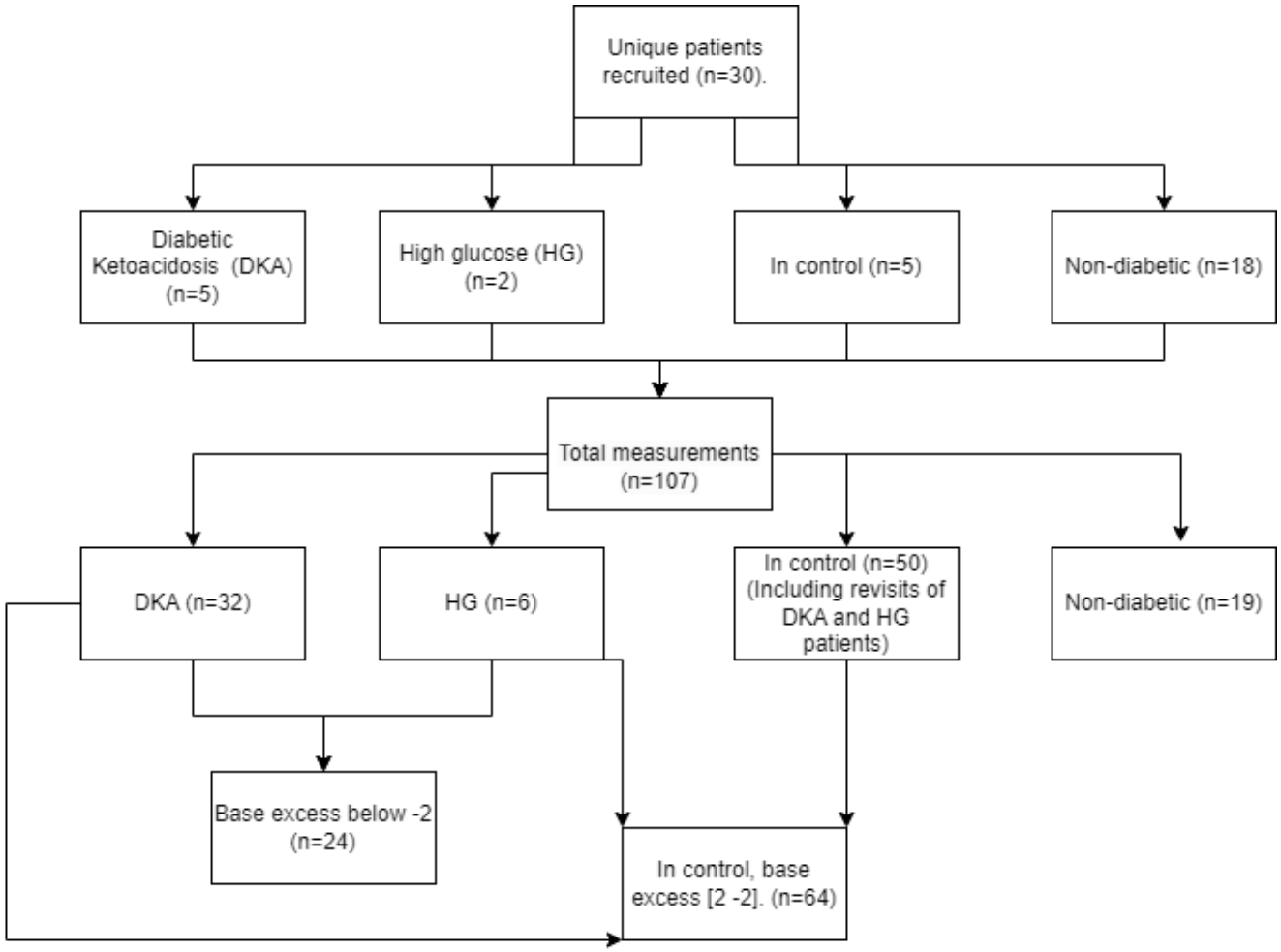


Figure S1. Study overview. Thirty unique patients were recruited of which five diagnosed with diabetic ketoacidosis, two with high glucose, five were in-control and 18 were non-diabetic. All patients contributed to a total of 107 measurements. Twenty-four came from the DKA patients, eight from HG patients, 56 measurements came from patients who were in control (these included patients who were initially recruited during DKA state but were re-measured later when they were in control). The eighteen non-diabetic patients resulted in 19 measurements. After patient stratification, the measurement were divided in BE below -2, which totaled twenty-four measurements, in control (BE between 2 and -2) with 64 measurements.

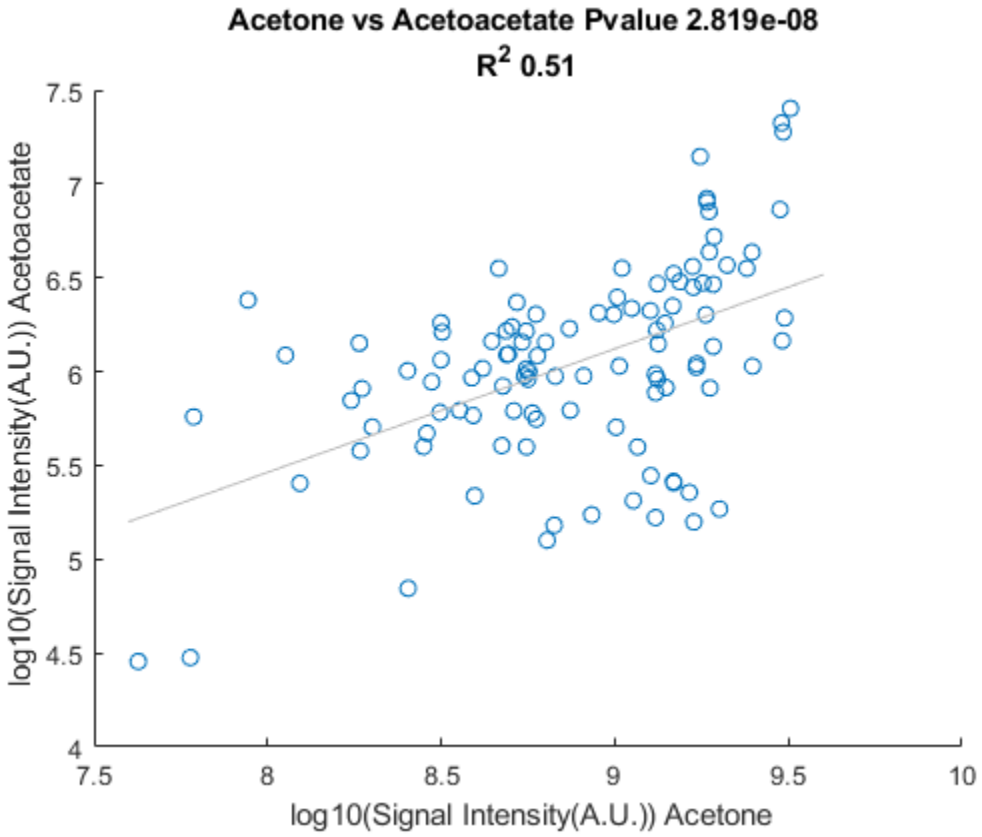


Figure S2 Acetone and acetoacetate show significant correlation among each other

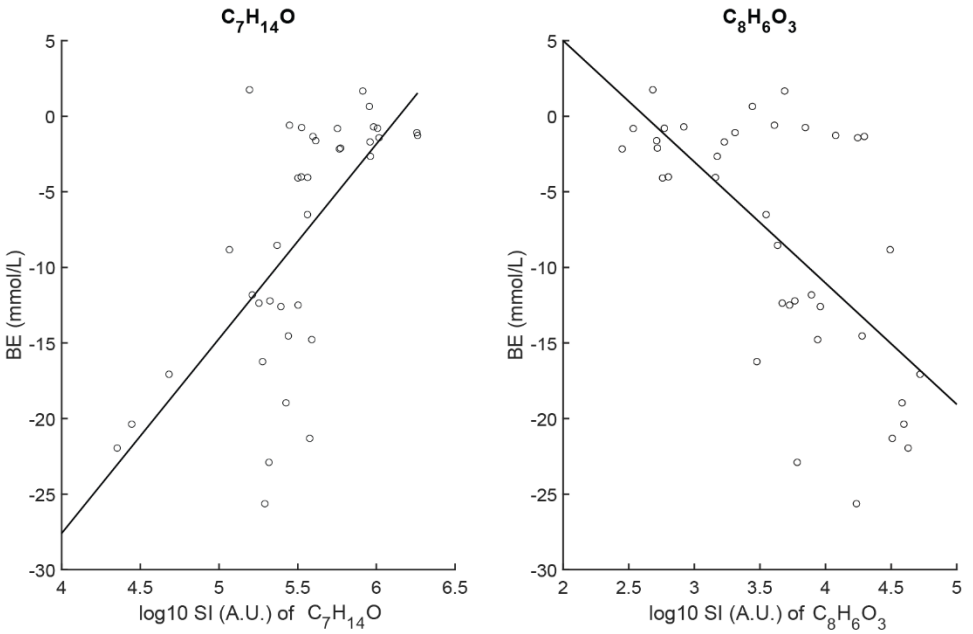


Figure S3 Two examples of features correlating positively and negatively with BE.

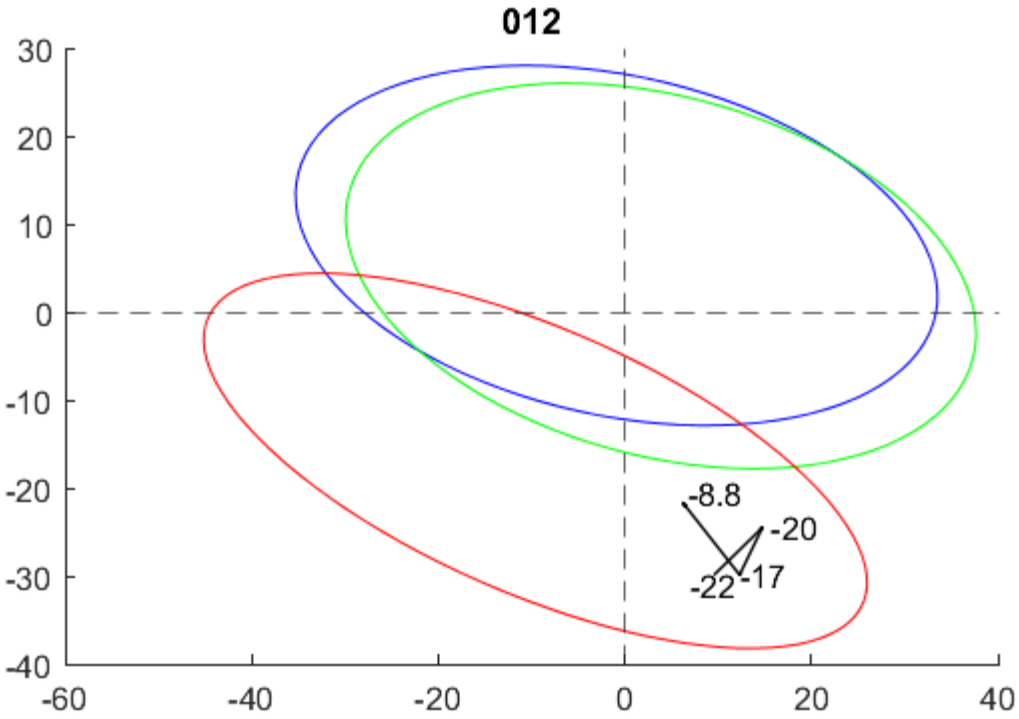


Figure S4 Trajectory of subject 012 during ICU stay (DKA)

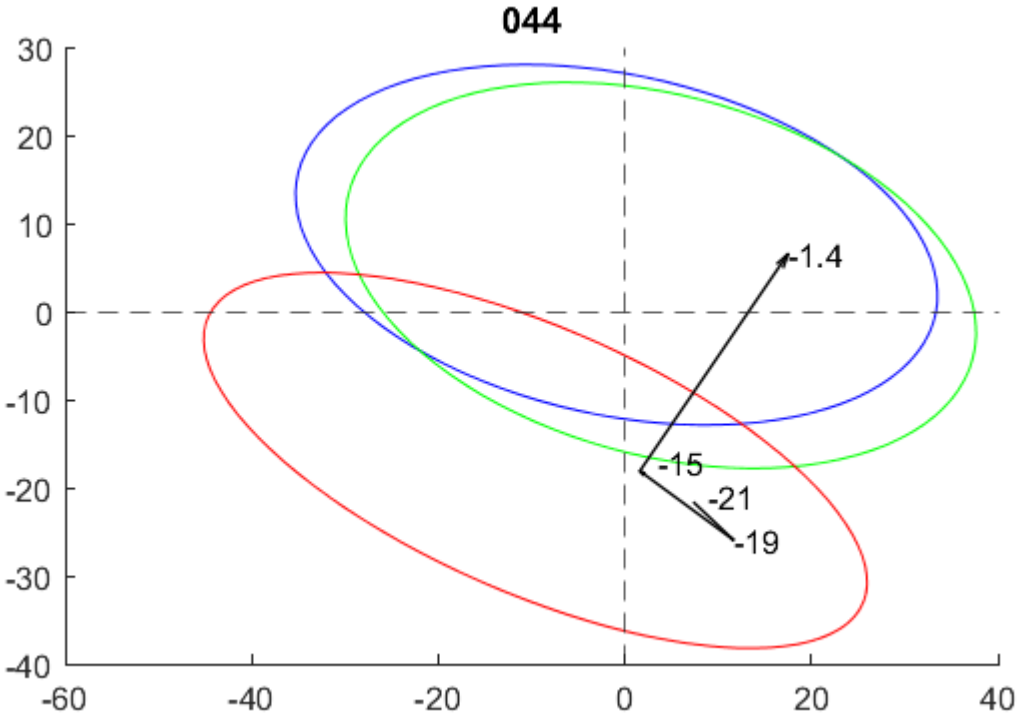


Figure S5 Trajectory of subject 044 during ICU stay (DKA)

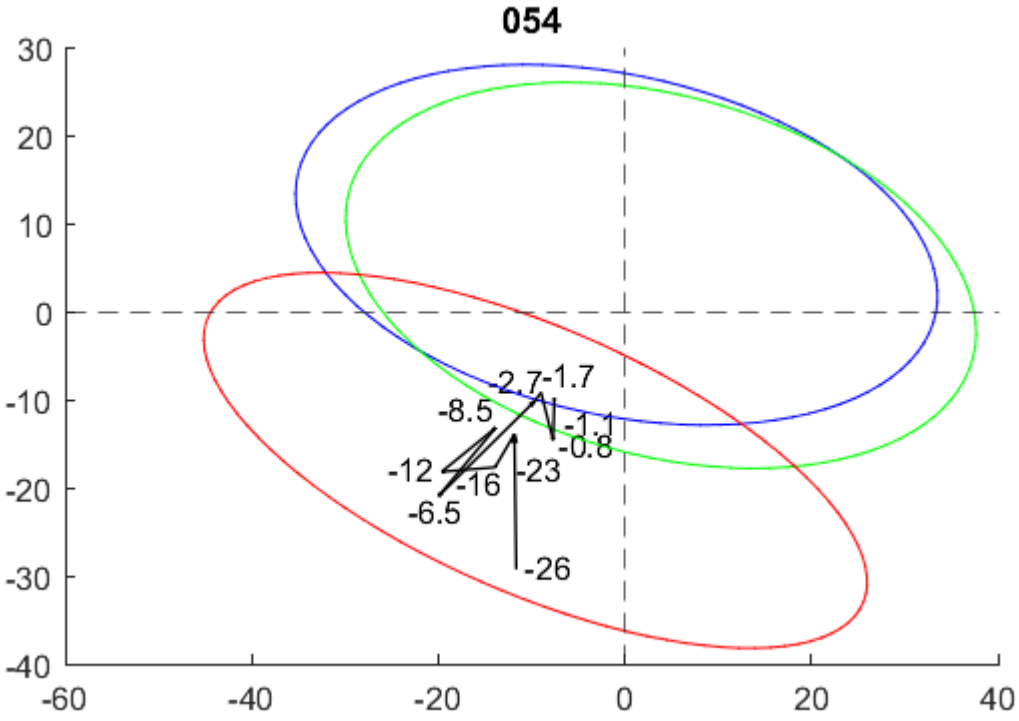


Figure S6 Trajectory of subject 054 during ICU stay (DKA)

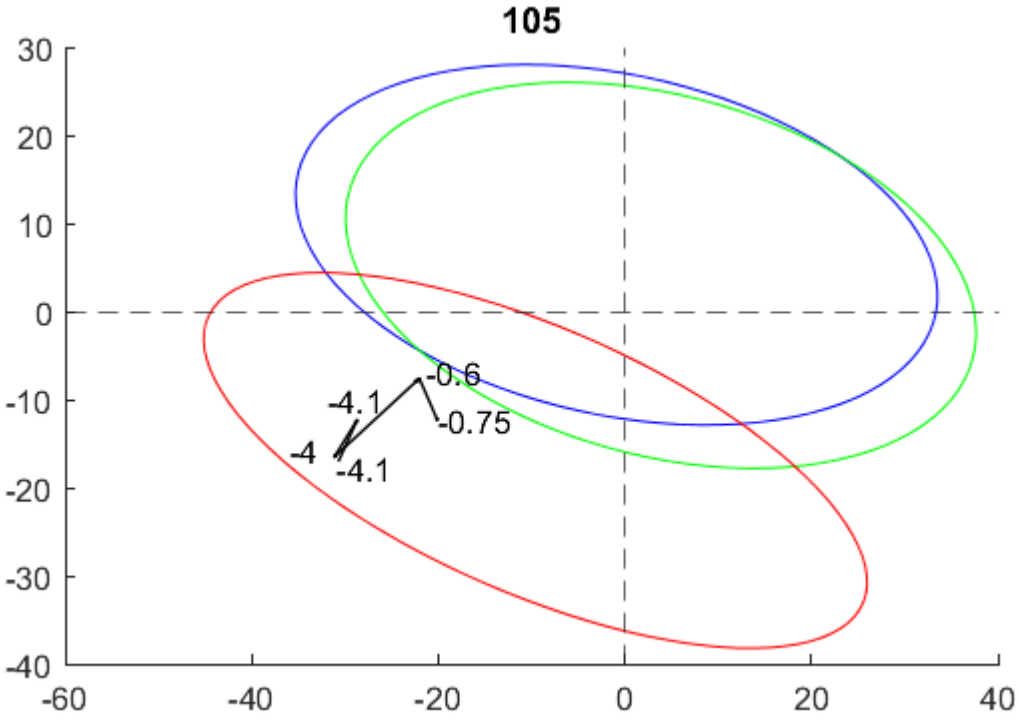


Figure S7 Trajectory of subject 105 during ICU stay (DKA)

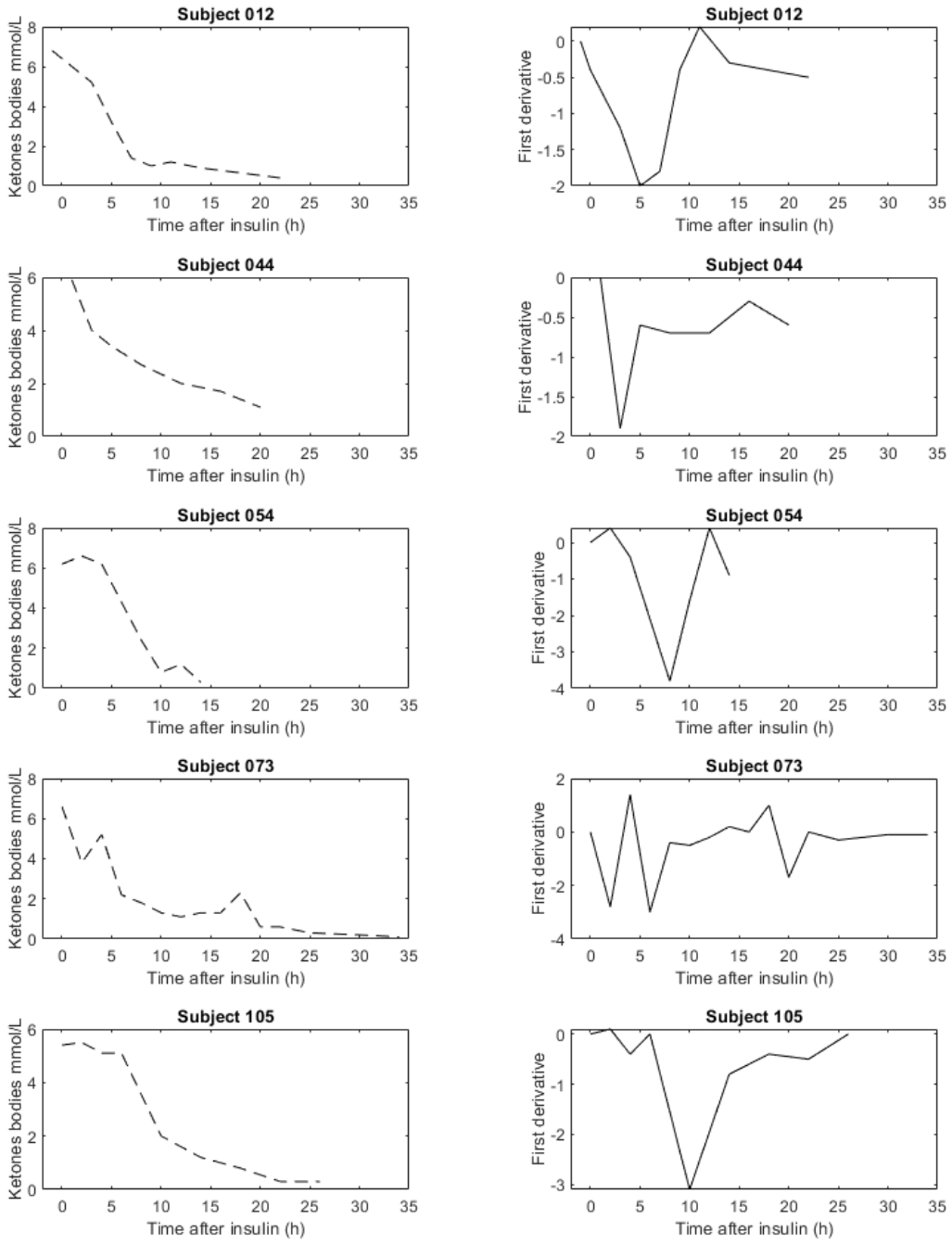


Figure S8 Ketones bodies and its first derivative

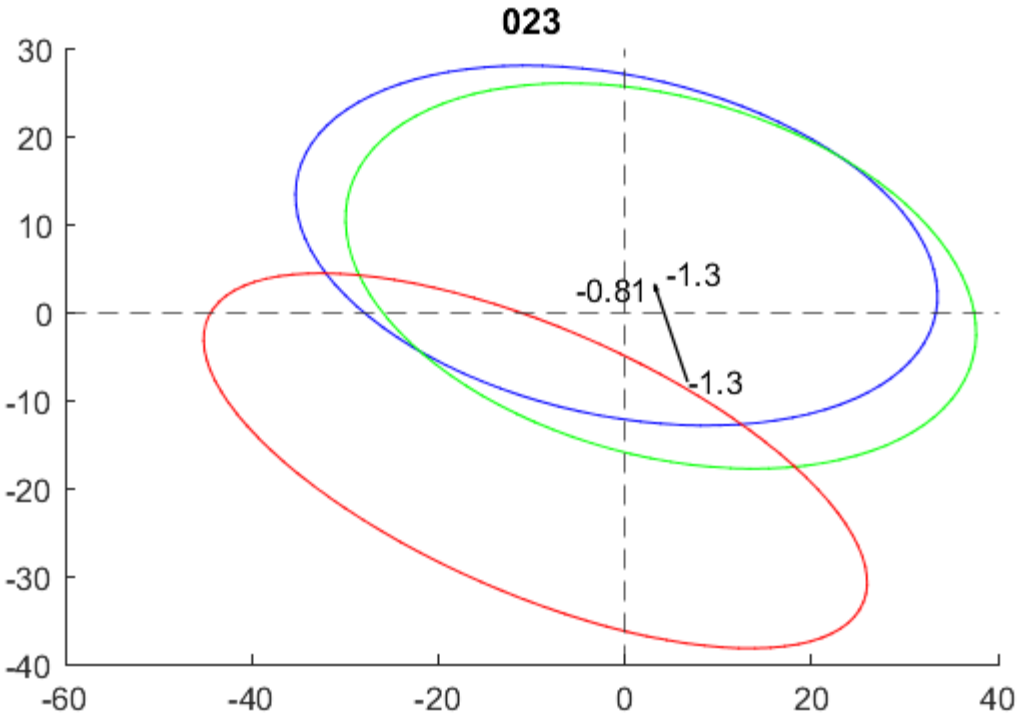


Figure S9 Trajectory of subject 023 during hospital stay (HG)

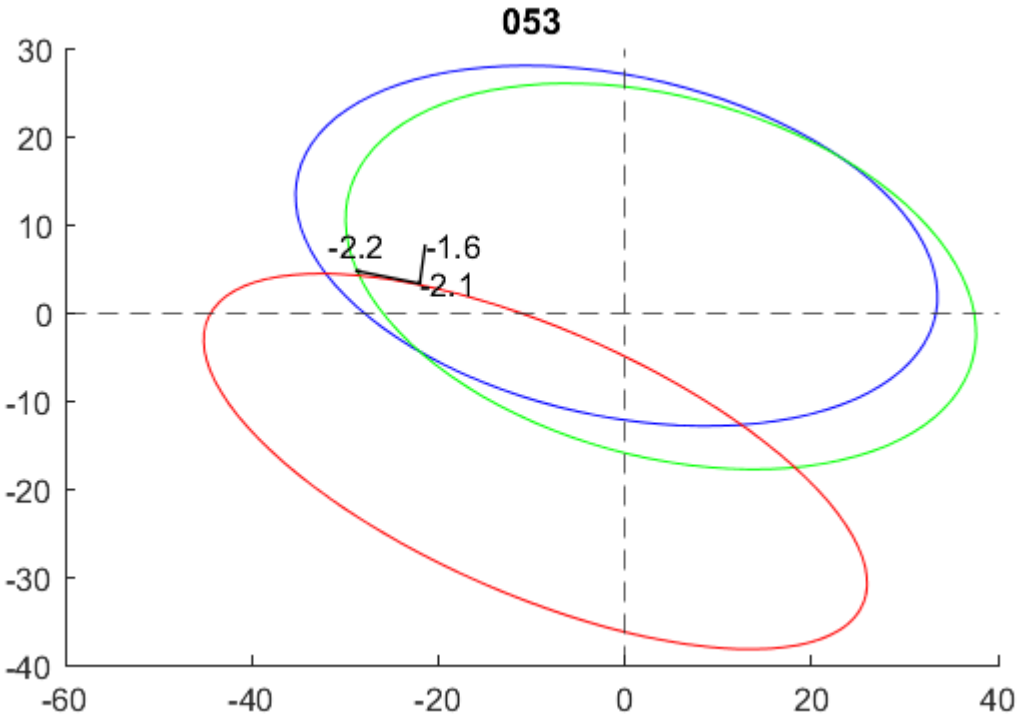


Figure S10 Trajectory of subject 053 during hospital stay (HG)

References

1. Decrue F, Singh KD, Gisler A, Awchi M, Zeng J, Usemann J, Frey U, Sinues P. Combination of Exhaled Breath Analysis with Parallel Lung Function and FeNO Measurements in Infants. Analytical Chemistry 2021;93:15579-15583
2. Gisler A, Singh KD, Zeng J, Osswald M, Awchi M, Decrue F, Schmidt F, Sievi NA, Chen X, Usemann J. An interoperability framework for multicentric breath metabolomic studies. Iscience 2022;25:105557
3. Lee L, Helsel D. Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. Computers & Geosciences 2005;31:1241-1248

Table S1 (DKA)

Subject 012																					
Date	05.10.2020	05.10.2020	05.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	06.10.2020	07.10.2020	07.10.2020	07.10.2020
Time	18:52	20:39	21:40	00:04	01.00	02:12	03.00	04:02	05.00	06:30	07.00	08:07	09.00	10:08	12.00	15:58	20.00	22:04	04:26	10:52	11:01
pH	7.172	7.156	7.161	7.231		7.242		7.266		7.299		7.315		7.324		7.396		7.444	7.422	7.453	7.422
BE (mmol/L)	-19.6	-22.3	-22.4	-19		-15.2		-11.7		-8.6		-9.2		-7.4		-3.1		-0.3	0.5	0.2	0.4
Glucose (mmol/L)	26	19.9	21.2	13.6		15.9		15.2		14.1		12.7		18.2		12.4		15.6	9.4	15.9	17.7
Lactate (mmol/L)	1.6	1.7	1.7	1.5		1		1.1		1.4		1.2		1		1.1		1.3	1.3	1.3	1.2
PCO2 (mmHg)	22	17	16	18		26		32		35		31		34		34		34	38	33	38
HCO3 (mmol/L)	7.8	5.8	5.5	7.2		10.7		13.9		16.6		15.4		17.3		20.5		22.7	24.2	23	24.1
Sodium (mmol/L)	134	138	140	143		144		144		144		146		144		145		141	142	140	138
Anion gap K+ (mmol/L)	NA	23.8	25.8	23.4		19		16.5		13.9		16.3		15.3		14.1		13.2	11.4	12.2	11.3
Blood ketones		6.8	6.4		5.2		3.2		1.4		1		1.2		0.9		0.4				
Time form Start insulin		-1	0		3		5		7		9		11		14		22				

Table S2 (DKA)

Subject 044																		
Date	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	29.01.2021	30.01.2021	30.01.2021	30.01.2021	30.01.2021	30.01.2021	30.01.2021	31.01.2021	31.01.2021
Time	12:32	14:22	15:52	17:00	18:06	19:00	20:12	21:00	22:06	22:57	00:18	04:16	08:07	12:03	17:23	21:20	08:09	14:53
pH	7.16	7.17	7.19		7.26		7.32		7.36	7.33	7.34	7.36	7.40	7.43	7.448	7.460	7.472	7.462
BE (mmol/L)	-19.6	-21.2	-21.5		-17.4		-14.5		-11.8	-11.6	-10.7	-7.7	-5.6	-3	-0.3	1.1	2.9	2.1
Glucose (mmol/L)	23.8	20.2	18		11.4		8.5		10.2	9.6	8.3	6.7	8.7	15.6	4.4	8.4	8.2	10.4
Lactate (mmol/L)	2.3	1.3	1.4		1.5		1.1		1.2	1	1	1.1	1.1	1.7	1.4	1.3	1.7	2.4
PCO2 (mmHg)	25	18	16		18		20		21	24	25	29	29	31	33	34	36	36
HCO3 (mmol/L)	8.4	6.3	5.7		8		9.7		11.2	12.4	13.2	16.1	17.6	19.8	22.7	24.1	25.9	25.3
Sodium (mmol/L)	134	134	138		141		142		141	141	143	142	141	144	144	143	143	145
Anion gap K+ (mmol/L)	NA	NA	19.5		17.9		16.8		14.5	14.6	15.3	13.9	14.6	12.7	NA	NA	NA	NA
Blood ketones				5.9		4		3.4			2.7	2	1.7	1.1				
Start from insulin therapy				1		3		5			8	12	16	20				

Table S3 (DKA)

Subject 054																									
Date	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	12.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	13.03.2021	14.03.2021	14.03.2021
Time	09:21	10:16	10:55	11:56	12:57	13:59	15:04	16:05	16:08	18:08	20:01	22:07	00:09	02:04	04:03	06:12	08:04	10:14	12:03	14:12	18:15	21:03	00:08	07:00	
pH	6.922	7.054	7.075	7.110	7.189	7.242	7.258	7.318	7.343	7.341	7.332	7.359	7.392	7.382	7.382	7.446	7.401	7.408	7.416	7.415	7.435	7.441	7.418	7.444	
BE (mmol/L)	-27.9	-26	-25.2	-23.7	-20.2	-17.3	-14	-12.2	-12.6	-8.4	-6.6	-5.1	-3.6	-3.9	-3.4	-1.9	-2.7	-3	-1.7	-0.7	-1.1	-0.3	-0.1	2	
Glucose (mmol/L)	55	48	44	42	39	35	31	27	27	18.8	14.3	14.8	14.2	13.1	12.4	10.4	10.8	7.7	7.1	10.7	8.2	12.1	3.2	7.1	
Lactate (mmol/L)	6.8	3.4	2.8	2.4	2.1	1.9	1.6	1.7	1.9	1.4	1.3	1.8	1.5	1.6	1.3	1.3	1	1.2	1.1	1.1	2.1	1.9	2.4	1.9	
PCO2 (mmHg)	32	16	17	18	19	21	27	24.2	21	30	35	35	34	34	35	31	34	33	34	36	33	34	37	38	
HCO3 (mmol/L)	6.3	4.3	4.8	5.4	6.9	8.6	11.6	12	11	15.7	18.1	19.1	20	20	20.5	20.7	21	20.3	21.7	22.9	21.9	22.9	23.6	25.6	
Sodium (mmol/L)	125	126	130	133	134	136	137	137	136	142	143	144	145	145	146	147	147	150	149	149	150	146	149	144	
Anion gap K+ (mmol/L)	NA	31.1	30.6	27	27.9	25.3	23.4	20.5	18.9	19.2	15.8	14.9	13.7	14.2	13.1	10.2	12.4	16.7	12.5	11.5	13.1	11.4	12.6	11.5	
Blood ketones		6.2		6.6		6.2				2.4	0.8	1.2	0.3												
Time form Start insulin		0		2		4				8	10	12	14												

Table S4 (DKA)

Subject 073																													
Date	24.05.2021	24.05.2021	24.05.2021	24.05.2021	24.05.2021	24.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	25.05.2021	26.05.2021	26.05.2021	26.05.2021	26.05.2021	26.05.2021
Time	11:42	13:40	16:05	18:01	20:11	22:01	00:01	01:02	02:00	04:22	05:09	05:16	06:00	08:02	10:00	11:14	12:00	14:04	15:00	17:03	20:10	23:04	00:01	02:03	05:03	08:12	15:12		
pH	7.167	7.222	7.320	7.286	7.377	7.356		7.326		7.384	7.380	7.361		7.410		7.442		7.457		7.459	7.501	7.449		7.439	7.420	7.453	7.483		
BE (mmol/L)	-16.7	-18.1	-12.6	-12.4	-11.2	-7.3		-6.5		-4	-4.3	-4.5		-3.1		-0.5		1.2		1.8	2.7	2.5		2	2.2	3.7	1.5		
Glucose (mmol/L)	19.4	19.3	14.2	14.1	20.7	16		14.3		13.1	12.7	13.3		13.2		13.7		16.7		10.4	14	9.4		11.4	8.9	9	14.3		
Lactate (mmol/L)	1.3	1	1.2	1.3	1.1	1		0.8		1.2	0.8	0.8		1.1		1.5		1.6		1.5	3.2	2.8		2.4	1.5	1.8	3.1		
PCO2 (mmHg)	33	21	2.3	27	20.6	31		36		34	34	36		32		33		35		36	32	38		39	41	40	32		
HCO3 (mmol/L)	11.3	8.2	11.5	12.6	11.8	16.7		18.4		19.8	19.6	19.7		20		22.4		24.1		24.9	25	26		25.7	26.4	27.4	23.7		
Sodium (mmol/L)	140	139	141	141	141	143		144		141	140	142		141		142		143		142	144	142		142	141	141	143		
Anion gap K+ (mmol/L)	NA	21	15.6	16.6	12.9	12.4		12.8		10	8.5	12.5		12.7		11.6		12.1		11.7	12	11.1		11.5	9.8	8.3	14.5		
Blood ketones		6.6	3.8	5.2	2.2	1.8	1.3		1.1	1.3			1.3	2.3	0.6		0.6		0.3		0.2		0.1						
Time form Start insulin		0	2	4	6	8	10		12	14			16	18	20		22		25		27		31						

Table S5 (DKA)

Subject 105														
Date	14.11.2021	14.11.2021	14.11.2021	14.11.2021	14.11.2021	14.11.2021	14.11.2021	14.11.2021	15.11.2021	15.11.2021	15.11.2021	15.11.2021	16.11.2021	
Time	02:16	04:02	06:31	08:38	10:04	14:12	18:13	22:12	02:06	06:04	12:06	18:09	08:01	
pH	7.061	7.094	7.222	7.241	7.281	7.333	7.355	7.376	7.379	7.359	7.343	7.397	7.370	
BE (mmol/L)	-22.6	-23.2	-16	-13.5	-10.3	-4.1	-4	-1.4	-0.5	-0.1	-0.8	-0.6	1.4	
Glucose (mmol/L)	23.4	23.9	17.1	16.9	16.7	11.9	15.2	11.3	9.2	8.6	8.8	6.7	5.3	
Lactate (mmol/L)	5.4	4.5	2.4	2.3	1.8	1.4	1.1	1	0.8	0.8	1.1	0.8	0.8	
PCO2 (mmHg)	32	23	26	31	33	41	38	40	42	46	47	39	48	
HCO3 (mmol/L)	8.5	6.7	10.4	12.6	15.1	21.1	20.5	23.1	24.2	25.2	24.9	23.5	26.8	
Sodium (mmol/L)	139	136	138	139	138	140	139	140	139	140	142	141	141	
Anion gap K+ (mmol/L)	NA	29.4	24.2	23	19.7	14.7	15.9	13.4	8.5	7.2	NA	NA	NA	
Blood ketones		5.4	5.5	5.1	5.1	2	1.2	0.8	0.3	0.3				
Time from start insulin		0	2	4	6	10	14	18	22	26				

Table S6 (HG)

Subject 023				
Date	25. 11.2020	25. 11.2020	25. 11.2020	26. 11.2020
Time	11:40	12:51	18:05	08:00
pH	7.350	7.380	7.428	7.395
BE (mmol/L)	-1.9	-1.9	-0.8	-2.9
Glucose (mmol/L)	30	22.2	7.9	12.3
Lactate (mmol/L)	1.6	1.6	2.3	1.4
PCO2 (mmHg)	43	39	34	35
HCO3 (mmol/L)	23.3	22.4	22.3	20.8
Sodium (mmol/L)	135	137	144	140
Anion gap K+ (mmol/L)	NA	NA	NA	NA

Table S7 (HG)

Subject 053			
Date	10.03.2021	10.03.2021	11.03.2021
Time	14:37	18:19	08:19
pH	7.400	7.452	7.402
BE (mmol/L)	-1.2	-2.1	-2.6
Glucose (mmol/L)	19.4	7	6.3
Lactate (mmol/L)	1.3	2.9	1.4
PCO2 (mmHg)	38	30	34
HCO3 (mmol/L)	22.7	20.4	21
Sodium (mmol/L)	136	140	140
Anion gap K+ (mmol/L)	13.7	NA	NA

Table S8 Metabolites and their p-values and Log2foldchange

Molecule	m/z	Assigned Molecular	Pvalue BE<-2 vs [-2 2]	Log2FC BE<-2 vs [-2 2]	Pvalue BE<-2 vs ND	Log2FC BE<-2 vs ND	Pavlue [-2 2] vs ND	Log2FC [-2 2] vs ND
Acetone	59.04914	C3H6O	3.5E-08	1.29	3.44E-07	1.74	0.57	0.46
Acetoacetate	-101.0244	C4H6O3	6.0E-06	2.21	3.95E-07	2.92	0.11	0.71
Pyruvate	-87.00875	C3H4O3	0.20	0.53	4.65E-03	1.15	0.07	0.61
Fumarate	-115.0037	C4H4O4	1.6E-01	0.56	0.06	0.99	0.58	0.43
Oxoglutarate	-145.01425	C5H6O5	3.4E-01	-0.83	0.74	-0.40	0.90	0.43
(Iso)citrate	-191.0197	C6H8O7	9.9E-01	-0.95	0.91	1.01	0.81	1.96
Aconitate	-173.00915	C6H6O6	7.7E-01	-0.89	0.46	1.41	0.11	2.30
Succinate	-117.01935	C4H6O4	8.8E-01	0.07	0.61	-0.80	0.27	-0.87
Dodecanedioic acid	-229.14455	C12H22O4	4.0E-02	-1.13	0.55	-0.09	0.56	1.03
10-Hydroxydecanoic acid	-187.13395	C10H20O3	2.2E-02	-2.90	0.37	-0.19	0.66	2.71

Table S9 m/z and their assigned molecular formula, p-value, q-value and r-value

m/z	Assigned Molecular Formula	p-value	q-value	r
-165.01935	C8H6O4	1.10E-06	2.54E-03	-0.70
115.11174	C7H14O	3.68E-06	4.25E-03	0.67
-73.9965	unassignable	1.07E-05	4.95E-03	-0.65
-149.02445	C8H6O3	7.97E-06	4.95E-03	-0.66
-331.1342	C22H20O3	1.03E-05	4.95E-03	-0.65
-91.0286	unassignable	1.89E-05	7.29E-03	-0.63
116.11512	unassignable	3.19E-05	9.14E-03	0.62
-96.9931	C4H2O3	3.08E-05	9.14E-03	-0.62
-101.01555	unassignable	3.56E-05	9.14E-03	-0.62
-254.22065	unassignable	3.95E-05	9.14E-03	-0.62
-89.0131	unassignable	7.73E-05	1.58E-02	-0.60
-163.0037	C8H4O4	9.57E-05	1.58E-02	-0.59
-190.9986	C9H4O5	9.12E-05	1.58E-02	-0.59
-279.23295	C18H32O2	9.30E-05	1.58E-02	-0.59
-161.0244	C9H6O3	1.07E-04	1.65E-02	-0.59
91.05885	unassignable	1.45E-04	2.09E-02	0.58
91.05837	unassignable	1.55E-04	2.10E-02	0.58
92.98593	unassignable	2.01E-04	2.10E-02	0.57
-85.0659	C5H10O	1.82E-04	2.10E-02	-0.57
-87.0143	unassignable	2.44E-04	2.10E-02	-0.56
-89.0171	unassignable	1.88E-04	2.10E-02	-0.57
-89.0187	unassignable	1.93E-04	2.10E-02	-0.57
-98.0373	unassignable	2.46E-04	2.10E-02	0.56
-214.1894	unassignable	2.20E-04	2.10E-02	-0.56
-214.9949	C6H4O7N2	2.45E-04	2.10E-02	-0.56
-223.0976	C12H16O4	2.17E-04	2.10E-02	-0.57
-281.1385	C15H22O5	1.79E-04	2.10E-02	-0.57
-122.0329	unassignable	2.61E-04	2.15E-02	-0.56
91.05737	unassignable	3.54E-04	2.17E-02	0.55
92.0607	unassignable	3.09E-04	2.17E-02	0.55
92.77384	unassignable	3.15E-04	2.17E-02	0.55
-156.0747	unassignable	3.53E-04	2.17E-02	-0.55
-161.0357	C8H6O2N2	3.23E-04	2.17E-02	-0.55
-193.0506	C10H10O4	3.56E-04	2.17E-02	-0.55
-201.14965	C11H22O3	3.06E-04	2.17E-02	-0.55
-213.186	C13H26O2	3.34E-04	2.17E-02	-0.55
-214.15325	unassignable	2.83E-04	2.17E-02	-0.56

m/z	Assigned Molecular Formula	p-value	q-value	r
-313.07715	C23H10N2	3.55E-04	2.17E-02	-0.55
-200.10095	unassignable	3.83E-04	2.21E-02	-0.55
-205.08705	C12H14O3	3.75E-04	2.21E-02	-0.55
92.05691	H5ON5	4.09E-04	2.30E-02	0.54
-146.09045	unassignable	4.33E-04	2.38E-02	-0.54
-87.00875	C3H4O3	4.50E-04	2.42E-02	-0.54
-103.0286	unassignable	4.95E-04	2.44E-02	-0.54
-111.01995	C4H4O2N2	4.76E-04	2.44E-02	-0.54
-198.0405	C8H9O5N	4.95E-04	2.44E-02	-0.54
-329.26395	C11H30N12	4.80E-04	2.44E-02	-0.54
-148.0163	unassignable	5.09E-04	2.45E-02	-0.54
-239.05955	unassignable	5.54E-04	2.61E-02	-0.53
-97.0124	unassignable	5.76E-04	2.66E-02	-0.53
-101.01745	unassignable	6.23E-04	2.83E-02	-0.53
-168.07425	unassignable	6.40E-04	2.85E-02	-0.53
-156.1111	unassignable	6.69E-04	2.86E-02	-0.53
-283.06705	C9H16O10	6.62E-04	2.86E-02	-0.53
-71.05015	C4H8O	6.99E-04	2.94E-02	-0.53
217.10705	C10H16O5	8.22E-04	2.98E-02	-0.52
-89.0302	unassignable	7.92E-04	2.98E-02	-0.52
-101.03135	unassignable	7.92E-04	2.98E-02	-0.52
-156.99645	unassignable	8.15E-04	2.98E-02	-0.52
-185.0819	C9H14O4	8.25E-04	2.98E-02	-0.52
-239.20165	C15H28O2	8.12E-04	2.98E-02	-0.52
-242.22065	unassignable	8.15E-04	2.98E-02	-0.52
-255.0878	C12H16O6	8.15E-04	2.98E-02	-0.52
-284.2676	unassignable	7.37E-04	2.98E-02	-0.52
-162.01965	C8H5O3N	8.65E-04	3.03E-02	-0.52
-177.0557	C10H10O3	8.53E-04	3.03E-02	-0.52
157.1587	C10H20O	9.34E-04	3.05E-02	0.52
-58.0379	unassignable	9.61E-04	3.05E-02	-0.51
-89.0318	unassignable	9.78E-04	3.05E-02	-0.51
-147.02	C7H4O2N2	9.20E-04	3.05E-02	-0.52
-147.0815	C10H12O	9.66E-04	3.05E-02	-0.51
-150.0557	C8H9O2N	9.66E-04	3.05E-02	-0.51
-257.23915	unassignable	9.07E-04	3.05E-02	-0.52
-282.25195	unassignable	8.96E-04	3.05E-02	-0.52
-74.99735	unassignable	1.04E-03	3.18E-02	-0.51
-143.9937	unassignable	1.06E-03	3.18E-02	-0.51
-209.1547	C13H22O2	1.06E-03	3.18E-02	-0.51

m/z	Assigned Molecular Formula	p-value	q-value	r
-281.12375	C11H22O8	1.08E-03	3.20E-02	-0.51
96.99517	unassignable	1.14E-03	3.22E-02	0.51
143.14305	C9H18O	1.17E-03	3.22E-02	0.51
206.19844	C8H23ON5	1.16E-03	3.22E-02	0.51
-68.99805	C3H2O2	1.18E-03	3.22E-02	-0.51
-179.0921	C7H16O5	1.18E-03	3.22E-02	-0.51
-223.0612	C11H12O5	1.18E-03	3.22E-02	-0.51
-297.27805	C15H34N6	1.18E-03	3.22E-02	-0.51
-283.1191	C14H20O6	1.22E-03	3.28E-02	-0.51
-172.06965	unassignable	1.28E-03	3.31E-02	-0.50
-257.21255	C15H30O3	1.25E-03	3.31E-02	-0.50
-314.24315	C12H29ON9	1.27E-03	3.31E-02	-0.50
-72.9901	unassignable	1.38E-03	3.45E-02	-0.50
-255.26055	unassignable	1.36E-03	3.45E-02	-0.50
-269.23155	unassignable	1.37E-03	3.45E-02	-0.50
-109.0407	C5H6ON2	1.42E-03	3.47E-02	-0.50
-158.11315	unassignable	1.42E-03	3.47E-02	-0.50
-177.0193	C9H6O4	1.41E-03	3.47E-02	-0.50
-148.0334	unassignable	1.48E-03	3.57E-02	-0.50
-231.01465	C8H8O8	1.51E-03	3.60E-02	-0.50
209.11722	C12H16O3	1.56E-03	3.63E-02	0.50
-281.2486	C18H34O2	1.55E-03	3.63E-02	-0.50
92.05757	unassignable	1.59E-03	3.67E-02	0.49
-227.05625	C10H12O6	1.61E-03	3.67E-02	-0.49
-94.03795	unassignable	1.63E-03	3.70E-02	0.49
222.22163	C15H27N	1.69E-03	3.73E-02	0.49
-69.00095	unassignable	1.66E-03	3.73E-02	-0.49
-225.2223	C15H30O	1.68E-03	3.73E-02	-0.49
71.04986	unassignable	1.72E-03	3.74E-02	0.49
71.05018	unassignable	1.75E-03	3.76E-02	0.49
75.06904	unassignable	1.79E-03	3.76E-02	0.49
-172.106	unassignable	1.79E-03	3.76E-02	-0.49
-301.23665	C13H30O2N6	1.78E-03	3.76E-02	-0.49
91.05588	unassignable	1.82E-03	3.79E-02	0.49
-93.0346	C6H6O	1.89E-03	3.90E-02	0.49
-146.054	unassignable	1.92E-03	3.91E-02	-0.49
-253.2173	C16H30O2	1.93E-03	3.91E-02	-0.49
75.0715	unassignable	2.07E-03	3.91E-02	0.48
151.13109	unassignable	2.03E-03	3.91E-02	0.48
-114.06415	unassignable	2.09E-03	3.91E-02	-0.48

m/z	Assigned Molecular Formula	p-value	q-value	r
-178.9832	unassignable	2.12E-03	3.91E-02	0.48
-179.0714	C10H12O3	2.06E-03	3.91E-02	-0.48
-184.02505	C7H7O5N	2.12E-03	3.91E-02	-0.48
-185.0092	C7H6O6	2.00E-03	3.91E-02	-0.49
-191.03465	C10H8O4	2.08E-03	3.91E-02	-0.48
-229.18095	C13H26O3	2.02E-03	3.91E-02	-0.49
-251.1289	C14H20O4	2.00E-03	3.91E-02	-0.49
-267.216	C9H28O3N6	1.96E-03	3.91E-02	-0.49
-102.0197	C3H5O3N	2.17E-03	3.94E-02	-0.48
-283.23065	unassignable	2.16E-03	3.94E-02	-0.48
171.17433	C11H22O	2.25E-03	4.02E-02	0.48
-225.186	C14H26O2	2.25E-03	4.02E-02	-0.48
-123.02	C5H4O2N2	2.29E-03	4.04E-02	-0.48
-123.0815	C8H12O	2.28E-03	4.04E-02	-0.48
-172.06205	C7H11O4N	2.36E-03	4.12E-02	-0.48
73.04679	unassignable	2.44E-03	4.21E-02	0.48
120.11002	unassignable	2.47E-03	4.21E-02	0.48
-72.9931	unassignable	2.48E-03	4.21E-02	-0.48
-170.0095	C6H5O5N	2.45E-03	4.21E-02	-0.48
117.06988	C9H8	2.57E-03	4.21E-02	0.48
217.11996	C10H12N6	2.56E-03	4.21E-02	-0.48
-123.08575	unassignable	2.52E-03	4.21E-02	-0.48
-124.00405	C5H3O3N	2.59E-03	4.21E-02	-0.47
-186.08525	unassignable	2.54E-03	4.21E-02	-0.48
-283.2966	unassignable	2.58E-03	4.21E-02	-0.47
109.06479	C7H8O	2.73E-03	4.38E-02	0.47
-209.11895	C12H18O3	2.72E-03	4.38E-02	-0.47
-89.0244	C3H6O3	2.78E-03	4.42E-02	-0.47
-90.02775	unassignable	2.81E-03	4.42E-02	-0.47
-166.09535	unassignable	2.83E-03	4.42E-02	-0.47
-191.05595	C7H12O6	2.82E-03	4.42E-02	-0.47
-164.998	C11H2O2	2.91E-03	4.51E-02	-0.47
-237.18605	C15H26O2	2.95E-03	4.54E-02	-0.47
-240.0608	unassignable	2.97E-03	4.54E-02	-0.47
131.08867	unassignable	3.01E-03	4.55E-02	0.47
219.19546	C12H26O3	3.00E-03	4.55E-02	-0.47
-144.03835	unassignable	3.07E-03	4.61E-02	-0.47
-175.04035	C10H8O3	3.28E-03	4.89E-02	-0.46
147.06521	C6H10O4	3.37E-03	4.99E-02	0.46
-158.09035	unassignable	3.42E-03	5.03E-02	-0.46

m/z	Assigned Molecular Formula	p-value	q-value	r
-151.9981	unassignable	3.51E-03	5.06E-02	-0.46
-203.09245	C9H16O5	3.49E-03	5.06E-02	-0.46
-205.1234	C13H18O2	3.48E-03	5.06E-02	-0.46
-183.1027	C10H16O3	3.57E-03	5.12E-02	-0.46
-136.0166	unassignable	3.63E-03	5.15E-02	-0.46
-165.97915	unassignable	3.62E-03	5.15E-02	-0.46
109.09925	unassignable	3.72E-03	5.18E-02	0.46
-142.05905	unassignable	3.71E-03	5.18E-02	-0.46
-162.05615	C9H9O2N	3.68E-03	5.18E-02	-0.46
-241.2173	C15H30O2	3.76E-03	5.18E-02	-0.46
-314.24395	C16H33O3N3	3.77E-03	5.18E-02	-0.46
-257.17585	C14H26O4	3.80E-03	5.19E-02	-0.46
-72.99615	unassignable	3.89E-03	5.29E-02	-0.46
-116.03965	unassignable	4.00E-03	5.41E-02	-0.46
-203.01975	C7H8O7	4.07E-03	5.47E-02	-0.46
109.06793	unassignable	4.13E-03	5.52E-02	0.45
-179.035	C9H8O4	4.17E-03	5.54E-02	-0.45
-155.99375	C5H3O5N	4.20E-03	5.54E-02	-0.45
-147.04515	C9H8O2	4.39E-03	5.77E-02	-0.45
-140.00685	unassignable	4.45E-03	5.81E-02	-0.45
71.04914	C4H6O	4.48E-03	5.82E-02	0.45
-152.01155	unassignable	4.53E-03	5.82E-02	-0.45
-153.0386	unassignable	4.53E-03	5.82E-02	-0.45
-180.07825	C8H11O2N3	4.79E-03	6.09E-02	-0.45
-231.17545	C16H24O	4.80E-03	6.09E-02	-0.45
117.07305	unassignable	4.90E-03	6.19E-02	0.45
-171.0663	C8H12O4	4.99E-03	6.27E-02	-0.45
-139.999	C5H3O4N	5.10E-03	6.37E-02	-0.45
-57.0346	C3H6O	5.23E-03	6.47E-02	-0.44
-162.95365	unassignable	5.21E-03	6.47E-02	-0.44
-123.00875	C6H4O3	5.33E-03	6.52E-02	-0.44
-240.05945	unassignable	5.32E-03	6.52E-02	-0.44
-211.0612	C10H12O5	5.43E-03	6.61E-02	-0.44
-247.17035	C16H24O2	5.52E-03	6.68E-02	-0.44
120.0889	unassignable	5.62E-03	6.74E-02	0.44
-211.17035	C13H24O2	5.63E-03	6.74E-02	-0.44
-213.0407	C9H10O6	5.68E-03	6.77E-02	-0.44
-227.1289	C12H20O4	5.78E-03	6.85E-02	-0.44
-117.0443	unassignable	5.83E-03	6.86E-02	-0.44
-167.02105	C5H4O3N4	5.90E-03	6.86E-02	-0.44

m/z	Assigned Molecular Formula	p-value	q-value	r
-184.98435	C5H2O6N2	5.86E-03	6.86E-02	-0.44
-255.23295	C16H32O2	5.88E-03	6.86E-02	-0.44
-112.98835	C4H2O4	6.00E-03	6.93E-02	-0.44
88.04741	unassignable	6.28E-03	7.01E-02	0.44
89.07093	C3H8ON2	6.26E-03	7.01E-02	-0.44
127.02123	unassignable	6.24E-03	7.01E-02	0.44
-59.98345	unassignable	6.22E-03	7.01E-02	0.44
-174.1217	unassignable	6.27E-03	7.01E-02	-0.44
-191.07135	C11H12O3	6.18E-03	7.01E-02	-0.44
-231.051	C9H12O7	6.27E-03	7.01E-02	-0.44
-183.06635	C9H12O4	6.31E-03	7.01E-02	-0.44
-146.98755	C11O	6.43E-03	7.07E-02	-0.43
-186.12165	unassignable	6.41E-03	7.07E-02	-0.43
-181.0356	unassignable	6.47E-03	7.08E-02	-0.43
-59.02505	unassignable	6.69E-03	7.18E-02	-0.43
-114.0026	unassignable	6.73E-03	7.18E-02	-0.43
-115.0705	unassignable	6.76E-03	7.18E-02	-0.43
-116.0353	C4H7O3N	6.77E-03	7.18E-02	-0.43
-124.0163	unassignable	6.68E-03	7.18E-02	-0.43
-166.01465	C7H5O4N	6.75E-03	7.18E-02	-0.43
-233.10305	C10H18O6	6.64E-03	7.18E-02	-0.43
124.08383	unassignable	6.91E-03	7.29E-02	0.43
-149.00945	unassignable	6.98E-03	7.33E-02	-0.43
-103.0037	unassignable	7.02E-03	7.34E-02	-0.43
-128.0717	C6H11O2N	7.16E-03	7.45E-02	-0.43
-255.10285	C16H16O3	7.29E-03	7.55E-02	-0.43
-143.0961	unassignable	7.37E-03	7.61E-02	-0.43
91.05316	unassignable	7.47E-03	7.62E-02	0.43
-213.11325	C11H18O4	7.47E-03	7.62E-02	-0.43
-227.16525	C13H24O3	7.49E-03	7.62E-02	-0.43
-113.99595	unassignable	7.56E-03	7.64E-02	-0.43
-184.14275	unassignable	7.64E-03	7.64E-02	-0.43
-191.0197	C6H8O7	7.60E-03	7.64E-02	-0.43
-247.04585	C9H12O8	7.64E-03	7.64E-02	-0.43
-123.961	unassignable	7.70E-03	7.67E-02	-0.43
-279.15955	C16H24O4	7.82E-03	7.76E-02	-0.42
95.01611	unassignable	7.88E-03	7.77E-02	0.42
-241.1809	C14H26O3	7.90E-03	7.77E-02	-0.42
185.18999	C12H24O	8.02E-03	7.83E-02	0.42
-140.0718	C7H11O2N	8.06E-03	7.83E-02	-0.42

m/z	Assigned Molecular Formula	p-value	q-value	r
-160.106	unassignable	8.06E-03	7.83E-02	-0.42
107.04914	C7H6O	8.22E-03	7.85E-02	0.42
119.08866	unassignable	8.24E-03	7.85E-02	0.42
125.13248	C9H16	8.21E-03	7.85E-02	0.42
-93.0395	unassignable	8.29E-03	7.85E-02	0.42
-140.0116	unassignable	8.28E-03	7.85E-02	-0.42
-256.2363	unassignable	8.27E-03	7.85E-02	-0.42
-59.0012	unassignable	8.38E-03	7.90E-02	-0.42
139.01546	unassignable	8.46E-03	7.90E-02	-0.42
-166.02725	unassignable	8.48E-03	7.90E-02	-0.42
-245.1395	C12H22O5	8.43E-03	7.90E-02	-0.42
-297.24175	C14H30ON6	8.53E-03	7.91E-02	-0.42
109.08592	C4H12O3	8.63E-03	7.98E-02	0.42
-235.09745	C13H16O4	8.70E-03	8.01E-02	-0.42
-299.20005	C16H24N6	8.90E-03	8.16E-02	-0.42
119.06849	unassignable	8.95E-03	8.17E-02	0.42
93.03664	unassignable	9.07E-03	8.22E-02	0.42
-255.1238	C13H20O5	9.07E-03	8.22E-02	-0.42
79.02121	unassignable	9.15E-03	8.26E-02	-0.42
-158.05395	unassignable	9.37E-03	8.43E-02	-0.42
128.10874	unassignable	9.50E-03	8.43E-02	-0.42
139.14814	C10H18	9.52E-03	8.43E-02	0.42
-113.98325	unassignable	9.47E-03	8.43E-02	-0.42
-257.1031	C12H18O6	9.47E-03	8.43E-02	-0.42
109.09982	unassignable	9.57E-03	8.44E-02	0.42
-156.03025	C6H7O4N	9.60E-03	8.44E-02	-0.41
-207.0658	C11H12O4	9.64E-03	8.44E-02	-0.41
126.06307	unassignable	9.73E-03	8.45E-02	0.41
-153.03105	C6H6O3N2	9.73E-03	8.45E-02	-0.41
135.06517	C5H10O4	9.79E-03	8.47E-02	0.41
217.12235	C14H16O2	9.87E-03	8.51E-02	-0.41
97.00417	unassignable	9.97E-03	8.51E-02	0.41
121.06781	unassignable	9.99E-03	8.51E-02	0.41
191.14301	C13H18O	1.00E-02	8.51E-02	0.41
239.09496	unassignable	9.99E-03	8.51E-02	0.41
-213.1496	C12H22O3	1.01E-02	8.51E-02	-0.41
106.07324	unassignable	1.01E-02	8.52E-02	0.41
-162.98245	C11O2	1.01E-02	8.52E-02	-0.41
97.00462	unassignable	1.05E-02	8.75E-02	0.41
-124.0121	unassignable	1.05E-02	8.76E-02	-0.41

m/z	Assigned Molecular Formula	p-value	q-value	r
105.07396	unassignable	1.05E-02	8.76E-02	0.41
-279.12235	C11H16O3N6	1.07E-02	8.84E-02	-0.41
-283.1915	C16H28O4	1.07E-02	8.84E-02	-0.41
94.99054	unassignable	1.09E-02	8.99E-02	0.41
-234.1943	C9H25O2N5	1.10E-02	9.05E-02	-0.41
-154.01455	C6H5O4N	1.11E-02	9.08E-02	-0.41
-223.0284	unassignable	1.12E-02	9.14E-02	-0.41
-61.996	unassignable	1.13E-02	9.16E-02	-0.41
-280.11635	C10H15O3N7	1.13E-02	9.16E-02	-0.41
-115.9989	unassignable	1.14E-02	9.17E-02	-0.41
99.08212	unassignable	1.15E-02	9.18E-02	0.41
-168.103	C9H15O2N	1.15E-02	9.18E-02	-0.41
-240.13225	C7H15ON9	1.15E-02	9.18E-02	-0.41
-113.0972	C7H14O	1.16E-02	9.23E-02	-0.41
121.06841	unassignable	1.18E-02	9.32E-02	0.40
-217.10815	C10H18O5	1.19E-02	9.32E-02	-0.40
-285.2707	unassignable	1.18E-02	9.32E-02	-0.40
94.08626	C3H11O2N	1.19E-02	9.35E-02	0.40
-100.0121	unassignable	1.21E-02	9.37E-02	-0.40
-145.087	C7H14O3	1.21E-02	9.37E-02	-0.40
-243.1238	C12H20O5	1.20E-02	9.37E-02	-0.40
81.00027	unassignable	1.25E-02	9.63E-02	0.40
121.06903	unassignable	1.25E-02	9.63E-02	0.40
-211.04605	unassignable	1.26E-02	9.66E-02	-0.40
-59.01695	unassignable	1.27E-02	9.72E-02	-0.40
-178.10005	unassignable	1.27E-02	9.72E-02	0.40
121.06969	unassignable	1.30E-02	9.83E-02	0.40
-247.0823	C10H16O7	1.29E-02	9.83E-02	-0.40
105.05462	C4H8O3	1.31E-02	9.86E-02	0.40
121.07037	unassignable	1.32E-02	9.86E-02	0.40
149.09609	C10H12O	1.32E-02	9.86E-02	0.40
-158.01735	unassignable	1.32E-02	9.86E-02	-0.40
-126.02775	unassignable	1.33E-02	9.87E-02	-0.40
-166.04035	C10H5N3	1.33E-02	9.87E-02	-0.40
85.10117	C6H12	1.37E-02	9.89E-02	0.40
93.09101	C4H12O2	1.34E-02	9.89E-02	0.40
95.01274	C5H2O2	1.38E-02	9.89E-02	0.40
95.05028	unassignable	1.41E-02	9.89E-02	0.40
95.05118	unassignable	1.36E-02	9.89E-02	0.40
95.05163	unassignable	1.35E-02	9.89E-02	0.40

m/z	Assigned Molecular Formula	p-value	q-value	r
97.00555	unassignable	1.42E-02	9.89E-02	0.39
97.006	unassignable	1.39E-02	9.89E-02	0.40
97.00634	unassignable	1.41E-02	9.89E-02	0.40
110.07936	unassignable	1.42E-02	9.89E-02	0.39
221.08438	unassignable	1.35E-02	9.89E-02	0.40
222.19333	C8H23O2N5	1.41E-02	9.89E-02	0.39
-76.0121	unassignable	1.39E-02	9.89E-02	-0.40
-84.9931	C3H2O3	1.37E-02	9.89E-02	-0.40
-99.00405	unassignable	1.40E-02	9.89E-02	-0.40
-118.02985	C7H5ON	1.40E-02	9.89E-02	-0.40
-180.03025	C8H7O4N	1.38E-02	9.89E-02	-0.40
-184.106	unassignable	1.35E-02	9.89E-02	-0.40
-215.0197	C8H8O7	1.42E-02	9.89E-02	-0.39
-219.06665	C12H12O4	1.37E-02	9.89E-02	-0.40
-285.20715	C16H30O4	1.35E-02	9.89E-02	-0.40
-156.03785	unassignable	1.43E-02	9.92E-02	-0.39
119.05232	unassignable	1.44E-02	9.95E-02	0.39
-245.1758	C13H26O4	1.45E-02	9.97E-02	-0.39
169.05224	unassignable	1.45E-02	1.00E-01	0.39
-250.14485	C14H21O3N	1.46E-02	1.00E-01	-0.39
117.07176	unassignable	1.47E-02	1.00E-01	0.39
-200.1737	unassignable	1.47E-02	1.00E-01	-0.39
119.08149	C4H10O2N2	1.48E-02	1.00E-01	0.39
-160.06975	unassignable	1.48E-02	1.00E-01	-0.39
110.06816	unassignable	1.49E-02	1.01E-01	0.39
-115.0037	C4H4O4	1.50E-02	1.01E-01	-0.39
108.08077	C7H9N	1.52E-02	1.01E-01	0.39
-104.04345	unassignable	1.52E-02	1.01E-01	-0.39
-112.004	C4H3O3N	1.52E-02	1.01E-01	-0.39
-113.0608	C6H10O2	1.51E-02	1.01E-01	-0.39
-118.04105	C6H5N3	1.51E-02	1.01E-01	-0.39
95.05073	unassignable	1.52E-02	1.01E-01	0.39
-113.00745	unassignable	1.54E-02	1.02E-01	-0.39
-74.02475	C2H5O2N	1.55E-02	1.02E-01	-0.39
-146.02475	C8H5O2N	1.56E-02	1.02E-01	-0.39
-153.00665	unassignable	1.56E-02	1.02E-01	-0.39
91.05527	unassignable	1.61E-02	1.03E-01	0.39
117.05463	C5H8O3	1.60E-02	1.03E-01	0.39
121.07605	C7H8N2	1.60E-02	1.03E-01	0.39
-112.01215	unassignable	1.60E-02	1.03E-01	-0.39

m/z	Assigned Molecular Formula	p-value	q-value	r
-213.07685	C10H14O5	1.61E-02	1.03E-01	-0.39
-215.05615	C9H12O6	1.60E-02	1.03E-01	-0.39
-249.09755	C10H18O7	1.60E-02	1.03E-01	-0.39
-257.13945	C13H22O5	1.61E-02	1.03E-01	-0.39
-313.2366	C14H30O2N6	1.58E-02	1.03E-01	-0.39
123.08044	C8H10O	1.62E-02	1.03E-01	0.39
235.24212	C17H30	1.63E-02	1.03E-01	0.39
-154.08735	C8H13O2N	1.63E-02	1.03E-01	-0.39
159.14041	unassignable	1.65E-02	1.04E-01	0.39
169.0681	unassignable	1.65E-02	1.04E-01	0.39
-150.01965	C7H5O3N	1.65E-02	1.04E-01	-0.39
97.44194	unassignable	1.67E-02	1.04E-01	-0.39
-253.1078	C13H18O5	1.67E-02	1.04E-01	-0.39
99.08261	unassignable	1.69E-02	1.05E-01	0.39
-61.00465	H2O2N2	1.68E-02	1.05E-01	0.39
121.07201	unassignable	1.71E-02	1.05E-01	0.38
218.1104	unassignable	1.73E-02	1.05E-01	-0.38
219.21072	C16H26	1.72E-02	1.05E-01	-0.38
-154.0274	unassignable	1.73E-02	1.05E-01	-0.38
-169.01425	C7H6O5	1.73E-02	1.05E-01	-0.38
-195.06625	C10H12O4	1.71E-02	1.05E-01	-0.38
-200.1373	unassignable	1.72E-02	1.05E-01	-0.38
-299.22075	C13H28O2N6	1.73E-02	1.05E-01	-0.38
-172.00145	unassignable	1.76E-02	1.07E-01	-0.38
-139.00365	C6H4O4	1.77E-02	1.07E-01	-0.38
-215.12885	C11H20O4	1.77E-02	1.07E-01	-0.38
85.02936	unassignable	1.79E-02	1.08E-01	0.38
95.04914	C6H6O	1.80E-02	1.08E-01	0.38
-75.0043	unassignable	1.81E-02	1.08E-01	-0.38
-157.087	C8H14O3	1.81E-02	1.08E-01	-0.38
-287.221	C12H28O2N6	1.80E-02	1.08E-01	-0.38
121.07438	unassignable	1.82E-02	1.08E-01	0.38
-197.04555	C9H10O5	1.83E-02	1.09E-01	-0.38
199.20563	C13H26O	1.84E-02	1.09E-01	0.38
73.04548	unassignable	1.85E-02	1.09E-01	0.38
-102.0278	unassignable	1.86E-02	1.09E-01	-0.38
-199.99175	unassignable	1.86E-02	1.09E-01	0.38
-111.04515	C6H8O2	1.88E-02	1.10E-01	-0.38
-168.1108	unassignable	1.89E-02	1.11E-01	-0.38
108.0525	unassignable	1.90E-02	1.11E-01	0.38

m/z	Assigned Molecular Formula	p-value	q-value	r
109.06619	unassignable	1.91E-02	1.11E-01	0.38
94.02634	unassignable	1.95E-02	1.12E-01	0.38
110.07262	unassignable	1.96E-02	1.12E-01	0.38
114.01856	C4H3O3N	1.95E-02	1.12E-01	0.38
201.18489	C12H24O2	1.96E-02	1.12E-01	0.38
-84.01725	unassignable	1.99E-02	1.12E-01	-0.38
-87.0127	unassignable	1.99E-02	1.12E-01	-0.38
-144.07475	unassignable	1.98E-02	1.12E-01	-0.38
-146.96245	unassignable	1.95E-02	1.12E-01	-0.38
-158.0461	C6H9O4N	1.99E-02	1.12E-01	-0.38
-170.09035	unassignable	1.99E-02	1.12E-01	-0.38
-195.10265	C11H16O3	1.95E-02	1.12E-01	-0.38
-209.08195	C11H14O4	1.98E-02	1.12E-01	-0.38
-283.26425	C18H36O2	1.94E-02	1.12E-01	-0.38
-124.0037	C5H3O3N	2.00E-02	1.12E-01	-0.38
104.04231	unassignable	2.02E-02	1.13E-01	0.38
143.03389	C6H6O4	2.03E-02	1.14E-01	0.38
94.07323	unassignable	2.04E-02	1.14E-01	0.37
-251.04085	C8H12O9	2.07E-02	1.15E-01	-0.37
87.04504	unassignable	2.08E-02	1.15E-01	0.37
100.05186	unassignable	2.09E-02	1.16E-01	0.37
-223.0459	C7H12O8	2.11E-02	1.16E-01	-0.37
-255.1966	C15H28O3	2.11E-02	1.16E-01	-0.37
108.05695	unassignable	2.12E-02	1.16E-01	0.37
114.09945	unassignable	2.15E-02	1.16E-01	0.37
121.0671	unassignable	2.13E-02	1.16E-01	0.37
-61.99315	unassignable	2.14E-02	1.16E-01	-0.37
-116.98295	unassignable	2.14E-02	1.16E-01	-0.37
-145.0294	C9H6O2	2.14E-02	1.16E-01	-0.37
-228.205	unassignable	2.15E-02	1.16E-01	-0.37
60.04916	unassignable	2.20E-02	1.17E-01	-0.37
132.09204	unassignable	2.18E-02	1.17E-01	0.37
196.16957	C12H21ON	2.18E-02	1.17E-01	0.37
222.16998	C10H23O4N	2.19E-02	1.17E-01	0.37
222.18521	C14H23ON	2.17E-02	1.17E-01	0.37
-171.10265	C9H16O3	2.20E-02	1.17E-01	-0.37
-184.06165	C8H11O4N	2.19E-02	1.17E-01	-0.37
221.1536	C14H20O2	2.21E-02	1.17E-01	0.37
-109.0659	C7H10O	2.21E-02	1.17E-01	-0.37
115.96998	unassignable	2.25E-02	1.17E-01	0.37

m/z	Assigned Molecular Formula	p-value	q-value	r
116.01647	unassignable	2.26E-02	1.17E-01	-0.37
209.13245	C16H16	2.23E-02	1.17E-01	0.37
223.13287	C13H18O3	2.23E-02	1.17E-01	0.37
-115.0149	C3H4O3N2	2.25E-02	1.17E-01	-0.37
-155.035	C7H8O4	2.26E-02	1.17E-01	-0.37
-158.1172	unassignable	2.24E-02	1.17E-01	-0.37
-217.1445	C11H22O4	2.26E-02	1.17E-01	-0.37
-225.1496	C13H22O3	2.25E-02	1.17E-01	-0.37
145.10434	unassignable	2.28E-02	1.18E-01	0.37
109.02841	C6H4O2	2.29E-02	1.18E-01	0.37
-96.9766	unassignable	2.29E-02	1.18E-01	-0.37
-61.9941	unassignable	2.31E-02	1.19E-01	-0.37
142.13077	unassignable	2.35E-02	1.20E-01	0.37
-101.0244	C4H6O3	2.34E-02	1.20E-01	-0.37
-152.07175	C8H11O2N	2.34E-02	1.20E-01	-0.37
-174.0489	unassignable	2.34E-02	1.20E-01	-0.37
88.05187	unassignable	2.37E-02	1.20E-01	0.37
94.0281	unassignable	2.37E-02	1.20E-01	0.37
-123.02825	unassignable	2.36E-02	1.20E-01	-0.37
-195.13905	C12H20O2	2.38E-02	1.20E-01	-0.37
103.11173	C6H14O	2.39E-02	1.20E-01	0.37
94.0777	unassignable	2.40E-02	1.21E-01	0.37
93.05465	C3H8O3	2.41E-02	1.21E-01	0.37
-280.115	C9H19O7N3	2.43E-02	1.22E-01	-0.36
97.97299	unassignable	2.44E-02	1.22E-01	-0.36
139.12297	C8H14N2	2.46E-02	1.23E-01	-0.36
91.03897	C3H6O3	2.48E-02	1.23E-01	0.36
121.0648	C8H8O	2.49E-02	1.23E-01	0.36
122.08119	C4H11O3N	2.50E-02	1.23E-01	0.36
151.12292	C9H14N2	2.50E-02	1.23E-01	0.36
236.17262	C8H21O3N5	2.49E-02	1.23E-01	-0.36
-148.0404	C8H7O2N	2.47E-02	1.23E-01	-0.36
-271.1012	unassignable	2.50E-02	1.23E-01	-0.36
297.27779	C19H36O2	2.52E-02	1.24E-01	-0.36
97.00507	unassignable	2.55E-02	1.24E-01	0.36
101.08469	unassignable	2.57E-02	1.24E-01	0.36
132.08824	unassignable	2.53E-02	1.24E-01	0.36
137.1344	unassignable	2.54E-02	1.24E-01	-0.36
-137.06085	C8H10O2	2.55E-02	1.24E-01	-0.36
-143.07135	C7H12O3	2.57E-02	1.24E-01	-0.36

m/z	Assigned Molecular Formula	p-value	q-value	r
-229.10815	C11H18O5	2.56E-02	1.24E-01	-0.36
-241.10815	C12H18O5	2.56E-02	1.24E-01	-0.36
-198.0853	unassignable	2.58E-02	1.24E-01	-0.36
135.02188	unassignable	2.59E-02	1.24E-01	0.36
94.04132	unassignable	2.61E-02	1.25E-01	0.36
119.07027	C5H10O3	2.63E-02	1.25E-01	0.36
-76.9702	unassignable	2.62E-02	1.25E-01	-0.36
-142.06385	unassignable	2.63E-02	1.25E-01	-0.36
-156.0065	unassignable	2.64E-02	1.26E-01	-0.36
-170.99345	C6H4O6	2.66E-02	1.26E-01	-0.36
122.06814	unassignable	2.68E-02	1.27E-01	0.36
-61.99505	unassignable	2.69E-02	1.27E-01	-0.36
-161.00915	C5H6O6	2.69E-02	1.27E-01	-0.36
114.01052	unassignable	2.73E-02	1.28E-01	0.36
159.13796	C9H18O2	2.74E-02	1.28E-01	0.36
171.08046	C12H10O	2.74E-02	1.28E-01	0.36
-139.0765	C8H12O2	2.74E-02	1.28E-01	-0.36
-144.0666	C6H11O3N	2.73E-02	1.28E-01	-0.36
81.06863	unassignable	2.78E-02	1.29E-01	0.36
84.05698	unassignable	2.77E-02	1.29E-01	0.36
169.04958	C8H8O4	2.78E-02	1.29E-01	0.36
-96.9601	unassignable	2.78E-02	1.29E-01	-0.36
-247.1187	C11H20O6	2.77E-02	1.29E-01	-0.36
-193.0354	C6H10O7	2.83E-02	1.30E-01	-0.36
152.01647	unassignable	2.87E-02	1.32E-01	-0.36
136.01738	unassignable	2.89E-02	1.32E-01	0.35
-136.05305	unassignable	2.89E-02	1.32E-01	-0.35
-145.05065	C6H10O4	2.91E-02	1.33E-01	-0.35
81.00128	unassignable	2.93E-02	1.34E-01	0.35
-249.13435	C11H22O6	2.93E-02	1.34E-01	-0.35
147.13791	C8H18O2	2.94E-02	1.34E-01	0.35
-329.15495	C23H22O2	2.95E-02	1.34E-01	-0.35
133.04955	C5H8O4	2.98E-02	1.34E-01	0.35
165.11606	unassignable	2.98E-02	1.34E-01	-0.35
-59.01785	unassignable	2.98E-02	1.34E-01	-0.35
-115.03165	unassignable	2.97E-02	1.34E-01	-0.35
116.48525	unassignable	2.99E-02	1.34E-01	0.35
119.03389	C4H6O4	3.00E-02	1.34E-01	0.35
-89.00735	unassignable	3.01E-02	1.35E-01	-0.35
-251.1483	C8H16N10	3.01E-02	1.35E-01	-0.35

m/z	Assigned Molecular Formula	p-value	q-value	r
-221.18125	unassignable	3.05E-02	1.36E-01	-0.35
-115.04605	unassignable	3.07E-02	1.37E-01	-0.35
-98.95585	unassignable	3.08E-02	1.37E-01	-0.35
136.01815	C10HN	3.10E-02	1.38E-01	0.35
-112.0485	unassignable	3.11E-02	1.38E-01	-0.35
116.03064	unassignable	3.13E-02	1.38E-01	-0.35
-99.08155	C6H12O	3.15E-02	1.39E-01	-0.35
-190.0146	C9H5O4N	3.18E-02	1.40E-01	-0.35
60.0505	unassignable	3.21E-02	1.41E-01	-0.35
119.06037	C7H6N2	3.21E-02	1.41E-01	0.35
-156.0797	unassignable	3.22E-02	1.41E-01	-0.35
121.03178	unassignable	3.24E-02	1.41E-01	0.35
-221.1824	unassignable	3.23E-02	1.41E-01	0.35
-237.1496	C14H22O3	3.24E-02	1.41E-01	-0.35
-158.0094	C5H5O5N	3.28E-02	1.42E-01	-0.35
-165.0557	C9H10O3	3.28E-02	1.42E-01	-0.35
201.16376	C15H20	3.29E-02	1.42E-01	0.35
217.1435	C11H20O4	3.29E-02	1.42E-01	-0.35
98.03624	unassignable	3.33E-02	1.44E-01	0.35
-140.0433	unassignable	3.35E-02	1.44E-01	-0.35
-166.06365	unassignable	3.36E-02	1.44E-01	-0.35
-142.0874	C7H13O2N	3.38E-02	1.45E-01	-0.35
-186.1587	unassignable	3.42E-02	1.46E-01	-0.34
-95.95225	unassignable	3.43E-02	1.46E-01	-0.34
87.52958	unassignable	3.47E-02	1.48E-01	-0.34
165.11215	C7H16O4	3.52E-02	1.48E-01	-0.34
202.97983	unassignable	3.52E-02	1.48E-01	0.34
217.19508	C16H24	3.50E-02	1.48E-01	0.34
297.3141	C20H40O	3.53E-02	1.48E-01	-0.34
-77.9959	unassignable	3.53E-02	1.48E-01	0.34
-112.99925	C3H2O3N2	3.53E-02	1.48E-01	-0.34
-121.0295	C7H6O2	3.50E-02	1.48E-01	-0.34
-128.0114	unassignable	3.50E-02	1.48E-01	-0.34
-231.1238	C11H20O5	3.53E-02	1.48E-01	-0.34
-185.1547	C11H22O2	3.56E-02	1.49E-01	-0.34
91.05273	unassignable	3.60E-02	1.50E-01	0.34
107.10665	C5H14O2	3.60E-02	1.50E-01	0.34
-75.0132	unassignable	3.59E-02	1.50E-01	-0.34
-215.16525	C12H24O3	3.61E-02	1.50E-01	-0.34
-234.15805	C8H21O3N5	3.60E-02	1.50E-01	-0.34

m/z	Assigned Molecular Formula	p-value	q-value	r
96.05552	C4H5N3	3.65E-02	1.51E-01	0.34
135.01367	unassignable	3.67E-02	1.52E-01	0.34
-142.0146	C5H5O4N	3.68E-02	1.52E-01	-0.34
-182.0459	C8H9O4N	3.68E-02	1.52E-01	-0.34
-181.087	C10H14O3	3.73E-02	1.54E-01	-0.34
-61.99645	unassignable	3.76E-02	1.54E-01	-0.34
-142.9986	C5H4O5	3.76E-02	1.54E-01	-0.34
-170.0469	C8H5N5	3.75E-02	1.54E-01	-0.34
156.14645	unassignable	3.78E-02	1.54E-01	0.34
189.12737	C13H16O	3.78E-02	1.54E-01	0.34
96.05249	unassignable	3.80E-02	1.54E-01	0.34
180.13829	C11H17ON	3.80E-02	1.54E-01	0.34
109.08414	unassignable	3.82E-02	1.55E-01	0.34
-118.0227	unassignable	3.84E-02	1.55E-01	-0.34
-139.0401	C7H8O3	3.84E-02	1.55E-01	-0.34
-197.11825	C11H18O3	3.83E-02	1.55E-01	-0.34
105.06988	C8H8	3.89E-02	1.56E-01	0.34
227.16419	C13H22O3	3.89E-02	1.56E-01	-0.34
-93.0407	unassignable	3.91E-02	1.57E-01	0.34
-101.9832	unassignable	3.94E-02	1.57E-01	-0.34
-155.00995	C5H4O4N2	3.94E-02	1.57E-01	-0.34
-235.11875	C10H20O6	3.94E-02	1.57E-01	-0.34
-241.1445	C13H22O4	3.94E-02	1.57E-01	-0.34
166.12268	C10H15ON	3.97E-02	1.57E-01	0.34
-127.97515	unassignable	3.97E-02	1.57E-01	-0.34
-143.02345	unassignable	3.96E-02	1.57E-01	-0.34
75.04406	C3H6O2	4.01E-02	1.58E-01	0.33
137.56998	unassignable	4.01E-02	1.58E-01	-0.33
121.04957	C4H8O4	4.05E-02	1.59E-01	0.33
137.11723	C6H16O3	4.05E-02	1.59E-01	0.33
151.11174	C10H14O	4.06E-02	1.59E-01	0.33
207.12022	C5H14O3N6	4.03E-02	1.59E-01	0.33
235.20978	unassignable	4.08E-02	1.59E-01	0.33
-123.04885	unassignable	4.04E-02	1.59E-01	0.33
-144.0302	C5H7O4N	4.07E-02	1.59E-01	-0.33
-165.0044	unassignable	4.07E-02	1.59E-01	-0.33
-271.17425	C8H20ON10	4.06E-02	1.59E-01	-0.33
-164.11615	unassignable	4.13E-02	1.60E-01	-0.33
-297.04585	C22H6N2	4.13E-02	1.60E-01	-0.33
162.1277	C11H15N	4.14E-02	1.60E-01	0.33

m/z	Assigned Molecular Formula	p-value	q-value	r
-115.0848	unassignable	4.15E-02	1.60E-01	-0.33
-269.19515	C9H22N10	4.15E-02	1.60E-01	-0.33
91.05195	unassignable	4.17E-02	1.60E-01	0.33
133.06078	C4H8O3N2	4.17E-02	1.60E-01	0.33
-171.0552	unassignable	4.18E-02	1.60E-01	-0.33
101.08716	unassignable	4.20E-02	1.61E-01	0.33
-247.13145	C11H16ON6	4.21E-02	1.61E-01	-0.33
110.08117	C3H11O3N	4.26E-02	1.61E-01	0.33
137.12977	unassignable	4.25E-02	1.61E-01	0.33
157.97394	unassignable	4.27E-02	1.61E-01	0.33
171.13796	C10H18O2	4.26E-02	1.61E-01	-0.33
222.15696	C7H19O3N5	4.24E-02	1.61E-01	0.33
-144.00655	unassignable	4.29E-02	1.61E-01	-0.33
-160.95665	unassignable	4.27E-02	1.61E-01	-0.33
-168.03025	C7H7O4N	4.28E-02	1.61E-01	-0.33
-199.17035	C12H24O2	4.26E-02	1.61E-01	-0.33
-223.13395	C13H20O3	4.29E-02	1.61E-01	-0.33
108.06553	C3H9O3N	4.30E-02	1.61E-01	0.33
-194.09515	unassignable	4.30E-02	1.61E-01	0.33
104.97322	unassignable	4.37E-02	1.62E-01	-0.33
109.03966	C5H4ON2	4.40E-02	1.62E-01	0.33
117.0895	unassignable	4.35E-02	1.62E-01	-0.33
121.0284	C7H4O2	4.37E-02	1.62E-01	0.33
137.98356	unassignable	4.40E-02	1.62E-01	-0.33
143.08872	unassignable	4.40E-02	1.62E-01	0.33
-61.0037	unassignable	4.38E-02	1.62E-01	0.33
-77.01305	unassignable	4.42E-02	1.62E-01	-0.33
-95.01385	C5H4O2	4.39E-02	1.62E-01	-0.33
-111.08155	C7H12O	4.35E-02	1.62E-01	-0.33
-137.9958	unassignable	4.36E-02	1.62E-01	-0.33
-167.07135	C9H12O3	4.36E-02	1.62E-01	-0.33
-235.17035	C15H24O2	4.35E-02	1.62E-01	0.33
-299.2574	C14H32ON6	4.41E-02	1.62E-01	-0.33
-301.2001	C12H26O3N6	4.35E-02	1.62E-01	-0.33
86.03175	unassignable	4.45E-02	1.63E-01	0.33
82.06512	C5H7N	4.51E-02	1.65E-01	0.33
-83.01385	C4H4O2	4.53E-02	1.65E-01	-0.33
-128.08795	unassignable	4.54E-02	1.65E-01	-0.33
97.00972	unassignable	4.59E-02	1.66E-01	0.33
97.0102	unassignable	4.59E-02	1.66E-01	0.33

m/z	Assigned Molecular Formula	p-value	q-value	r
97.01066	unassignable	4.59E-02	1.66E-01	0.33
97.01113	unassignable	4.59E-02	1.66E-01	0.33
91.04056	unassignable	4.60E-02	1.66E-01	0.33
97.10118	C7H12	4.63E-02	1.67E-01	0.33
142.12266	C8H15ON	4.63E-02	1.67E-01	-0.33
-158.0826	C7H13O3N	4.65E-02	1.67E-01	-0.32
143.06374	unassignable	4.69E-02	1.68E-01	0.32
121.05867	unassignable	4.76E-02	1.70E-01	0.32
97.00926	unassignable	4.83E-02	1.71E-01	0.32
103.07536	C5H10O2	4.82E-02	1.71E-01	0.32
122.07107	C6H7N3	4.81E-02	1.71E-01	0.32
197.19	C13H24O	4.80E-02	1.71E-01	0.32
-59.01075	unassignable	4.80E-02	1.71E-01	-0.32
-138.0197	C6H5O3N	4.80E-02	1.71E-01	-0.32
-151.07645	C9H12O2	4.85E-02	1.71E-01	-0.32
-163.0248	C5H8O6	4.84E-02	1.71E-01	-0.32
-167.035	C8H8O4	4.83E-02	1.71E-01	-0.32
-217.07175	C9H14O6	4.84E-02	1.71E-01	-0.32
130.9385	unassignable	4.90E-02	1.72E-01	-0.32
139.01818	C10H2O	4.90E-02	1.72E-01	-0.32
-95.0251	C4H4ON2	4.88E-02	1.72E-01	-0.32
-171.0299	C7H8O5	4.89E-02	1.72E-01	-0.32
-172.02495	C6H7O5N	4.92E-02	1.72E-01	-0.32
-151.11305	C10H16O	4.94E-02	1.73E-01	-0.32
98.01389	unassignable	4.95E-02	1.73E-01	-0.32
224.0562	C11H5ON5	4.96E-02	1.73E-01	-0.32
-243.19655	C14H28O3	4.98E-02	1.73E-01	-0.32
-100.98805	unassignable	4.99E-02	1.73E-01	-0.32

CV

MO AWCHI

28th February 1996

Address: Mulhausserstrasse 78
4056, Basel,
Switzerland

Mobile: +41 76 481 4851

E-mail: mohamad.awchi@gmail.com

EDUCTION

Ph.D. Biomedical Engineering Supervisor: Prof. Pablo Sinues	University of Basel	Sep 2019 to Sep 2023
Master Analytical Chemistry	Stockholm University	Sep 2017 to Mar 2019
Bachelor of Chemistry	University of Applied Science Utrecht	Sep 2013 to Jun 2017

WORKING EXPERIENCE

Master thesis at RIKILT, Wageningen University and Research — 1st Jun18 to 30th Mar19

- Method development of an LC-HRMS (Orbitrap) and LC-MS/MS approach for the identification and detection of novel perfluorinated alkylated substances

Bachelor's thesis at AkzoNobel, Specialty Chemicals — 1st Mar17 to 1st Jul17

- Data generation and development of a software tool for multi-capillary column-ion mobility spectrometry data

Intern at Radboud University (chemometrics department) — 1st Sep16 to 30th Feb17

- Development of an outlier detection technique and pre-process strategies for multi-dimensional flow cytometry data

Intern at RIKILT, Wageningen University and Research — 1st Sep15 to 1st Jul16

- Method development for the analysis of Bisphenol A and analogues with LC-MS/MS

PUBLICATIONS

1. **Awchi, M**.; Sinues, P.; Datta, A. N.; García-Gómez, D.; Singh, K. D., UHPLC-MS/MS-Based Identity Confirmation of Amino Acids Involved in Response to and Side Effects from Antiseizure Medications. *Journal of Proteome Research* **2023,** *22* (3), 990-995.
2. Gisler, A.; Singh, K. D.; Zeng, J.; Osswald, M.; **Awchi, M**.; Decrue, F.; Schmidt, F.; Sievi, N. A.; Chen, X.; Usemann, J., An interoperability framework for multicentric breath metabolomic studies. *Iscience* **2022,** *25* (12), 105557.
3. **Awchi, M.**; Gebbink, W. A.; Berendsen, B. J.; Benskin, J. P.; van Leeuwen, S. P., Development, validation, and application of a new method for the quantitative determination of monohydrogen-substituted perfluoroalkyl carboxylic acids (H–PFCAs) in surface water. *Chemosphere* **2022,** *287*, 132143.
4. Singh, K. D.; Osswald, M.; Ziesenitz, V. C.; **Awchi, M**.; Usemann, J.; Imbach, L. L.; Kohler, M.; García-Gómez, D.; van den Anker, J.; Frey, U., Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications Medicine* **2021,** *1* (1), 21.
5. López-Lorente, C. I.; **Awchi, M.**; Sinues, P.; García-Gómez, D., Real-time pharmacokinetics via online analysis of exhaled breath. *Journal of Pharmaceutical and Biomedical Analysis* **2021,** *205*, 114311.
6. Decrue, F.; Singh, K. D.; Gisler, A.; **Awchi, M.**; Zeng, J.; Usemann, J.; Frey, U.; Sinues, P., Combination of exhaled breath analysis with parallel lung function and FeNO measurements in infants. *Analytical Chemistry* **2021,** *93* (47), 15579-15583.
7. van Leeuwen, S. P.; Bovee, T. F.; **Awchi, M.**; Klijnstra, M. D.; Hamers, A. R.; Hoogenboom, R. L.; Portier, L.; Gerssen, A., BPA, BADGE and analogues: A new multi-analyte LC-ESI-MS/MS method for their determination and their in vitro (anti) estrogenic and (anti) androgenic properties. *Chemosphere* **2019,** *221*, 246-253.