Identification and risk assessment of unknown contaminants migrating from Food Contact Materials

Pieke, Eelco Nicolaas

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Humans are exposed to possibly thousands of chemical compounds through food alone, yet the nature or effect of these chemicals is often poorly understood. Food contact materials (FCM), designed to store, preserve, and protect the food are also one of the major sources of chemicals in food. Currently, focus on FCM-borne chemicals is for a small number of high profile, well-studied chemical compounds, yet little focus for a much larger group of unknown chemical compounds. We demonstrate the possibility to provide preliminary data on unknown chemical compounds with novel analytical strategies. These explorative strategies can be utilized in risk prioritization and act as one of the forefront tools to improve managing chemical safety in food, especially with regards to the many unknown and unaddressed compounds.

Research Group for Analytical Food Chemistry
National Food Institute
Technical University of Denmark

Kemitorvet
Building 202
DK-2800 Kgs. Lyngby, Denmark
Tel. +45 35 88 70 00
Fax +45 35 88 70 01

http://www.food.dtu.dk/english
ISBN 978-87-93565-17-3
ISSN

E.N. Pieke

Identification & Risk Prioritization of Unknown Contaminants from Paper & Board FCMs

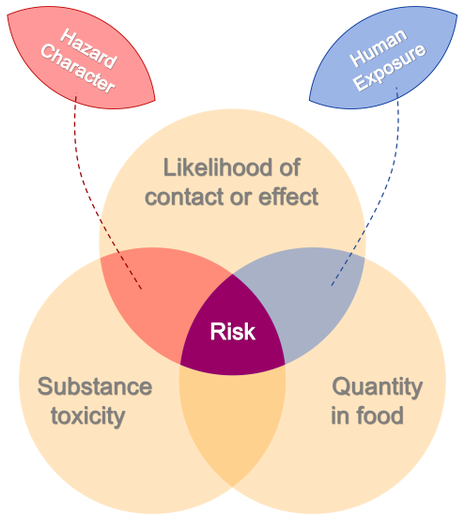
DTU

Identification and risk prioritization of unknown contaminants migrating from paper and board food contact materials

Ph.D. thesis

Eelco Nicolaas Pieke
May 2018

DTU Food
National Food Institute



Technical University of Denmark

Identification and risk prioritization of unknown contaminants migrating from paper and board food contact materials

PhD thesis

Eelco Nicolaas Pieke

DTU Food
National Food Institute

http://www.food.dtu.dk/english

1st Edition

May, 2018

Title	Identification and risk prioritization of unknown contaminants migrating from paper and board food contact materials
Author	Eelco Nicolaas Pieke
Supervision	Dr. Kit Granby *Associate Professor, Technical University of Denmark* Professor Jørn Smedsgaard *Professor Analytical Chemistry, Technical University of Denmark*
Evaluation Committee	Professor Anne Marie Vinggaard (Chair) *Professor Molecular Toxicology, Technical University of Denmark* Dr. Koni Grob *Former head Official Food Control Authority of the Canton of Zurich, Switzerland* Dr. Nikoline Juul Nielsen *Assistant Professor Environmental Chemistry and Physics, Copenhagen University, Denmark*
Funding	This project was funded by the Technical University of Denmark.
Copyright	
Cover	Eelco N. Pieke
ISBN	978-87-93565-17-3
Availability	This thesis is available at: National Food Institute Technical University of Denmark Kemitorvet Building 202 2800 Kgs. Lyngby Denmark Web www.food.dtu.dk/english E-mail food@food.dtu.dk Tel nr. +45 35 88 70 00 Fax nr. +45 35 88 70 01

Preface

This thesis disserts the primary results of three years research at the National Food Institute at the Technical University of Denmark as a PhD student in order to obtain the PhD degree. This research was performed at the Research Group for Analytical Food Chemistry, formerly the Division of Food Chemistry, in the National Food Institute at the Technical University of Denmark from 4 December 2014 until 31 January 2018. This research was funded internally by DTU.

The project was carried out under supervision of Associate Professor Dr. Kit Granby and co-supervision by Researcher Dr. Xenia Trier from 4 December 2014 until 24 February 2016. From March 2016 until January 2018, co-supervision responsibility was transferred to Professor Dr. Jørn Smedsgaard. Supervision during the external research stay was by Dr. Gilles Rivière, Deputy Head of Unit “Food safety risk assessment” and Dr. Bruno Teste, Scientific and Technical Project Manager, at the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) in Maisons-Alfort, Paris region (Oct 2016 – Apr 2017).

We are supposed to learn from other's mistakes.

But what if there are no mistakes, what if *we* are the first?

Then it is up to us to leave the right message …

And up to us to make the mistakes others can learn from

Acknowledgements

This section is most likely the most-read section and, based on the complexity of the topic, also the best-understood. Needless to say, with work spanning over three years and countless of interactions, there are a lot of people to thank for helping establish this thesis.

First, I would like to thank the supervisors and people that made this project possible. Kit Granby is thanked extensively for being a great and caring supervisor, her tireless devotion, guidance, assistance, discussions, hospitality, and as I like to say: being a little bit my replacement mother in Denmark. Jørn Smedsgaard is thanked for his visionary contribution in later parts of the project, in-depth and great discussions on the strategy, and providing solid feedback. Xenia Trier is thanked for her invaluable assistance in setting up the project, defining the project scope, and all the great scientific discussions we had. Lastly, I thank Elena Boriani for our cooperation on especially the risk assessment topics, and providing some of the last critical puzzle pieces in the study.

I thank all my colleagues at DTU Food, especially those in Research Group for Analytical Food Chemistry. These are too many to list them all, and I can sadly only mention a few specific. Amelie, Philipp, Tingting, Petra, Line, and Demi: thank you for being there as fellow PhD's and all the fun we had with parties, dinners, and in the office(s). I have a lot of fond memories of everything we did together. Faranak, they say that common enemies make good allies, and I think our alliance against the Q-TOF has shown this. I greatly enjoyed working with you and sharing the love-hate relationship with our equipment. Lisbeth K., thank you for all your help with the migration testing and data analysis, and a few unforgettable parties. Liljana, thanks for tirelessly helping me out with all the DTU systems and, of course, the nice conversations and jokes. Just stop calling me Ashton, please? Gitte A., thank you for all the help on risk assessment and regulation, and all the scientific discussions we were able to have.

There’s also the odd number of people to thank for my enjoyable stay in ANSES, France. First and foremost: large thanks to Gilles Rivière and Bruno Teste for supervising me in this period. Gilles, thank you for all the time and effort you’ve invested in making my stay at ANSES productive. I enjoyed all the discussions we frequently had. Bruno, it was a pleasure to discuss with you on very many aspects, including interesting political figures, and thank you for devoting so much time to discuss my project. Sebastien, thanks for all the remarkable exchanges on the most impossible of topics, and teaching me that yes, there is a “too far” area for jokes. Finally, all the persons at ANSES that made my stay more enjoyable and scientifically sound: Julien, Fernando, Geraldine and the various Risk Assessment panel members at ANSES.

I greatly want to thank my family and friends for their support and visits during this period, and making my return journeys “home” much more enjoyable. However, please don’t make me see that mermaid ever again. Mom, thank you for many and frequent fly-by visits, your unwavering support, the food supplies, the comfort visits, and all the fun experiences here. Dad, thank you for support, visit during Danish summer™, for proofreading, and sticking up with some of my early-Christmas presents gone bad. Lisette, cheers for the fun visits (all the way from Australia) and your faith in me! I also want to thank my family in-law, Arnold, Monique, Sabrina, and Marvin for being supportive and having resolute faith in me, and of course taking care of Sylvana when I was out doing science stuff. Tomas and Merlijn, thanks for all the fun and games we’ve managed to do at a distance. Raji, thanks for being an equally good enemy as friend.

And then there’s the one person to thank the most. Dear Sylvana, thank you for the support, love, and strength you’ve given me to continue going for the last three — but perhaps even the last ten — years. Yet even more important than that, you have made sure that every day I have had a home to return to, rather than just a house. It hasn’t always been an easy road, and while the product of it all is written here by me, the adventure and experience along the road is ours to keep and cherish. Our adventure doesn’t need to be written down, because it will be in my heart always. Thank you, Syl.

Executive Summary

The exposure of humans to possibly thousands of chemical compounds through food poses a health risk that is questioning our ability to ensure high standards for food safety. Food contact materials (FCM) are a major source of extraneous chemical compounds in food, yet not much knowledge is available on all compounds present due to FCM. The practicability of comprehensive studies, like risk assessment (RA), is questionable for an increasingly large number of chemical compounds. As a consequence, most research is focused on a small number of well-studied chemical compounds, but little is dedicated to the much larger number of unknown compounds. How are FCM safe for use if the greater part of it is unidentified, unassessed, and possibly completely unknown?

Here, development is shown for two new analytical strategies: semi-quantification and tentative identification; along with possible application for FCM RA. Paper and board FCM were extracted for migratable content, followed by analysis by liquid chromatography (LC) high resolution quadrupole-time of flight (Q-TOF) mass spectrometry (MS). Compounds were semi-quantified by comparing to non-identical reference standards after dedicated system optimisation. For identification, Q-TOF MS/MS utilizing automated precursor selection was used to actively collect non-target fragmentation spectra of compounds in the chromatogram. A risk prioritization approach that classified chemical compounds according to expected risk was developed based on applied tentative data and subsequent data interpretation by expert assessors.

Semi-quantification was demonstrated to work for almost any compound detected by LC-QTOF-MS analysis (Manuscript A). The errors in the predicted concentrations were at maximum up to 3-fold error with average around up to 2-fold error. These errors were attainable after dedicated optimisation of the LC-MS system to produce uniform responses that favour improving response of weak-response compounds. Semi-quantification did not

require chemical identification or standard-matching. For a single sample, more than 300 compounds were simultaneously quantified. Consequently, semi-quantification is a valuable strategy for prioritization based on concentration and for acquiring quantitative data without prior identification or available reference standards.

The tentative identification of compounds was demonstrated by non-targeted structural data acquisition (Manuscript B). Fragmentation spectra collected by Q-TOF MS/MS were correlated with *in silico* generated spectra using chemical structure databases to find the best-matching chemical compound to the spectrum, thereby removing the need for a reference standard. A total of five structure databases were used, resulting in structure prediction for over 130 compounds discovered in a recycled paper and board pizza box. For most of the 130 compounds, structure predictions were successful with good correlation scores, resulting in an impression of the chemical structure. The tentative identity of some compounds was evaluated for possible risk based on concentration and existing hazard data. Tentative identification is a promising strategy used to obtain significant chemical information about compounds in complex samples.

Tentative data was used to prioritize risk of identified compounds, differentiating between high-risk and low-risk compounds based on predicted exposure and predicted hazards (Manuscript C). This approach mimics RA procedures by converting tentative data to hazard- and exposure estimates, followed by a combined assessment based on expertise judgement. The expertise of several trained risk assessors was used to assign risk profiles to compounds in order to achieve a risk ranking. Although tentative data contains uncertainty, interpretation by experts produces a viable risk ranking of known and unknown chemical compounds by implementing a consensus model of expert interpretations. Risk prioritization is successful in classifying estimated risk based on predicted exposure and predicted hazard, and is valuable for to preliminary RA studies.

The overarching strategy in this study shows that explorative techniques are valuable tools to help ensure food safety in the future. Tentative data and risk prioritization are key concepts that, when developed further in combination with predictive tools like structure-activity modeling or migration modeling, could be at the forefront of identifying present and future risks. The need for these strategies is clear: tentative and explorative data is needed because the current alternative is often no data.

Dansk Resumé (Danish Summary)

Menneskers eksponering for muligvis tusindvis af kemiske forbindelser via fødevarer udgør en sundhedsrisiko, der sætter spørgsmålstegn ved vores evne til at sikre høje standarder for fødevaresikkerhed. Fødevarekontaktmaterialer (FCM) er en kilde til kemisk forurening af fødevarer, og det til trods er der ikke meget viden tilgængelig om alle de forbindelser, der kan migrere over fra FCM. Det er tvivlsomt om der i praksis kan udføres omfattende studier, som risikovurdering, for et stadigt større antal kemiske forbindelser. Som en følge heraf er de fleste undersøgelser fokuseret på et lille antal velundersøgte kemiske forbindelser, mens kun få undersøgelser er dedikeret til det meget større antal ukendte forbindelser. Hvordan kan FCM være sikkert at bruge, hvis de indeholder mange uidentificerede stoffer, der ikke er risikovurderede og muligvis er helt ukendte?

Her vises udvikling af to nye analytiske strategier: semikvantifikation og tentativ identifikation, med deres mulige anvendelser inden for risikovurdering af FCM. Papir og pap FCM blev ekstraheret for migrerbart indhold efterfulgt af analyse ved væskekromatografi (LC) med højt opløseligt *'quadrupole-time of flight'* (Q-TOF) massespektrometri (MS). De kemiske forbindelser blev efter dedikeret system optimering semikvantificeret ved at sammenligne respons med ikke-identiske referencestandarder. Til identifikation blev der anvendt Q-TOF MS/MS med automatisk precursor selektion, for aktivt at opsamle non-target fragmenteringsspektre af forbindelser i kromatogrammet. For at klassificerede kemiske forbindelser efter formodet risiko blev der udviklet en risikoprioriterings metode baseret på tentative data og datafortolkning ved hjælp af ekspertvurderinger.

Det blev demonstreret at semikvantifikation kunne anvendes for næsten aller forbindelser detekteret ved LC-QTOF-MS (Manuskript A). Fejlene i de prædikterede koncentrationer var maksimalt en faktor tre, med et gennemsnit på en faktor to. Disse fejlestimater kunne

opnås efter dedikeret optimering af LC-MS-systemet til at producere ensartede respons, der favoriserer en forbedring af respons for forbindelser med et i forvejen svagt respons. Semikvantificering krævede ikke kemisk identifikation eller et match med standarder. Mere end 300 forbindelser blev kvantificeret i en enkelt prøve, dette gør semikvantifikation til en værdifuld strategi for prioritering baseret på koncentrationsindhold og til at få kvantitative data uden forudgående identifikation eller tilgængelighed af referencestandarder.

Den tentative identifikation af forbindelser blev demonstreret ved *'non-targeted'* strukturel dataopsamling (Manuskript B). Fragmentationsspektre opsamlet på Q-TOF MS/MS blev korreleret med *in silico* genererede spektre fra kemiske strukturdatabaser, for at finde de bedst matchende kemiske forbindelser, hvorved man undgår behovet for referencestandarder. I alt fem strukturdatabaser blev anvendt, hvilket resulterede i struktur forudsigelser for over 130 forbindelser fundet i en pizzabakke af genbrugspap. For de fleste af de 130 forbindelser var strukturforudsigelserne vellykkede med gode korrelationsscorer, hvilket gav et indtryk af den kemiske struktur. Den tentative identitet af de kemiske forbindelser blev evalueret for mulig fødevaresikkerheds-risici baseret på koncentrationsdata og eksisterende toksikologisk viden. Tentativ identifikation er en lovende strategi, der kan bruges til at opnå væsentlig information om kemiske forbindelser i komplekse prøver.

Tentative data blev brugt til at prioritere risici for identificerede forbindelser, ved differentiering mellem høj risiko og lav risiko baseret på prædiktiv eksponering og prædiktiv farlighed (Manuskript C). Denne fremgangsmåde efterligner risikovurderings-procedurer ved at konvertere tentative data til fare- og eksponeringsestimater efterfulgt af en samlet vurdering baseret på ekspert skøn. Risikovurderingseksperter blev bedt om at vurdere risikoprofiler for de kemiske forbindelser for at skabe en risikoprioritering. Selvom tentative data indeholder en usikkerhed, kan man ved implementering af en konsensusmodel for ekspertvurderingerne få en brugbar risikoprioritering for kendte og ukendte kemiske forbindelser. Risikoprioritering er velegnet til klassificering af estimeret risiko baseret på prædikteret eksponering og prædikteret farlighed, og er værdifuld i indledende risikovurderinger.

Den overordnede strategi i dette studie viser, at eksplorative teknikker er værdifulde værktøjer for at sikre fødevaresikkerheden i fremtiden. Tentative data og risikoprioritering

er nøglebegreber der, når de udvikles yderligere i kombination med prædiktive værktøjer som *in silico* strukturaktivitetsmodellering eller migrationsmodellering, kan være på forkant for at identificere nuværende og fremtidige risici. Behovet for disse strategier er evident: Der er behov for tentative og eksplorative data fordi det nuværende alternativ ofte er fravær af data.

List of Manuscripts

Manuscripts published:

Manuscript A:

Pieke, Eelco N., Kit Granby, Xenia Trier, and Jørn Smedsgaard. 2017. “A Framework to Estimate Concentrations of Potentially Unknown Substances by Semi-Quantification in Liquid Chromatography Electrospray Ionisation Mass Spectrometry.” *Analytica Chimica Acta* 975 (July). Elsevier Ltd: 30–41. doi:10.1016/j.aca.2017.03.054.

Manuscript B:

Pieke, Eelco N., Jørn Smedsgaard, and Kit Granby. 2017. “Exploring the Chemistry of Complex Samples by Tentative Identification and Semi-Quantification: A Food Contact Material Case.” *Journal of Mass Spectrometry*, December. doi:10.1002/jms.4052.

Manuscripts submitted for publication:

Manuscript C:

Pieke, Eelco N., Kit Granby, Bruno Teste, Jørn Smedsgaard, and Gilles Rivière. “Prioritization before Risk Assessment: the viability of uncertain data for food contact materials.” *Submitted to Regulatory Toxicology and Pharmacology.*

List of Dissemination Activities

Dissemination Activity I:

Date:	5 May 2015
Conference:	Extractables and Leachables in packaging materials
Organiser:	Waters Corporation
Location:	Copenhagen, Denmark
Type:	Presentation
Title:	Strategy to determine response factors for "unknowns" and NIAS in paper and board by LC-ESI-MS
Authors:	Eelco Nicolaas Pieke

Dissemination Activity II:

Date:	16–18 November 2016
Conference:	6th International Symposium on Food Packaging - Scientific Developments Supporting Safety and Innovation
Organiser:	ILSI Europe Packaging Materials Task Force
Location:	Barcelona, Spain
Types:	Posters
(1) Title:	Chemicals in paper and board food contact material: towards more knowledge, analytical and prioritization analysis
(1) Authors:	Eelco Nicolaas Pieke, Kit Granby, Elena Boriani
(2) Title:	Classification and prioritization of chemicals present in food contact material
(2) Authors:	Elena Boriani, Eelco Nicolaas Pieke, Emilio Benfenati

Dissemination Activity III:

Date:	31 May–2 June 2017
Conference:	GoFood 2017: The 2nd Global Food Safety & Technology Forum
Organiser:	Joint Centre of Excellence in Food Safety
Location:	Lund, Sweden; Copenhagen, Denmark
Type:	Presentation
Title:	Bridging the gap between food packaging safety and the world of unknown chemicals
Authors:	Eelco Nicolaas Pieke

Dissemination Activity IV:

Date:	2–3 November 2017
Conference:	Food Contact Material Safety Symposium (FCS)
Organiser:	Guangdong Inspection and Quarantine Technology Center (IQTC)
Location:	Xi'an, China
Type:	Keynote presentation
Title:	Food contact paper & board chemicals: setting priority based on screening methodology
Authors:	Eelco Nicolaas Pieke

Dissemination Activity V:

Date:	5 November 2017
Conference:	Challenges and Management of Non-intentionally Added Substances (NIAS)
Organiser:	ILSI Focal Point China; China National Centre for Food Safety Risk Assessment (CFSA); Danone China
Location:	Beijing, China
Type:	Keynote presentation
Title:	How does chemical exploration aid in the risk assessment of NIAS in FCM?
Authors:	Eelco Nicolaas Pieke

Dissemination Activity VI:

Date:	31 November–1 December 2017
Conference:	Joint International Symposium: Global Past, Present and Future Challenges in Risk Assessment – Strengthening Consumer Health Protection
Organiser:	German Federal Institute for risk assessment (BfR)
Location:	Berlin, Germany
(1) Type:	Presentation
(1) Title:	Fighting unknown chemicals: analytical strategies for risk prioritization
(1) Authors:	Eelco Nicolaas Pieke
(2) Type:	Poster
(2) Title:	Burden of disease of exposure to endocrine-disrupting chemicals in recycled paper used for food packaging in Denmark: a case-study
(2) Authors:	Elena Boriani, Eelco Pieke, Tine Hald, Sara Pires, Julie Boberg, Lea Sletting Jakobsen

Dissemination Activity VII:

Date:	8 February 2018
Conference:	Webinar
Organiser:	Food Packaging Forum
Location:	Available on: http://www.foodpackagingforum.org/events/prioritization-of-unknown-substances-in-fcms (date accessed: 04 April 2018) Watch directly: https://youtu.be/ny9nl7-x_yI (date accessed: 05 April 2018)
Type:	Webinar presentation
Title:	Prioritization of unknown substances in FCMs
Authors:	Eelco Nicolaas Pieke

List of Abbreviations

AOI	Analyte of Interest
APCI	Atmospheric Pressure Chemical Ionisation
APPI	Atmospheric Pressure Photoionisation
CMR	Carcinogenicity, Mutagenicity, Reproductive Toxicity
CRM	Charge Residue Model
Da	Dalton
DB	Database
DDA	Data-Dependent Acquisition
DM	Decision Module
DS	Data Source
DU	Decision Unit
EC	European Commission
EFSA	European Food Safety Authority
EI	Electron Impact ionisation
ESI	Electrospray Ionization
EU	European Union
FCM	Food Contact Material
FDA	United States Food and Drug Administration
FDA	U.S. Food and Drug Administration

FID	Flame Ionisation Detector
FPF	Food Packaging Forum
GC	Gas Chromatography
GPT	Gas-phase Proton Transfer
HILIC	Hydrophobic Interaction Chromatography
IARC	International Agency for Research on Cancer
IAS	Intentionally Added Substance
IEM	Ion Evaporation Model
ILSI	International Life Sciences Institute
LC	Liquid Chromatography
m/z	Mass to charge ratio
MS	Mass Spectrometry
MS/MS	Dual-stage mass spectrometry
NIAS	Non-Intentionally Added Substance
NMR	Nuclear Magnetic Resonance
OML	Overall Migration Limit
ORPI	Oligomers, Reaction Products, and Impurities
PEG	Polyethylene Glycol
Q	Quadrupole mass analyser
q	Quadrupole collision cell
QM	Quantification Marker
QSAR	Quantitative Structure-Activity Relationship

Table of Contents

Chapter 1

Introduction

1.1. Food contact materials and food safety: two different worlds?

Chemical innovations are one of the man-made developments that changed daily life, driving other innovations in many industries by providing tailored solutions. However, a consequence of increased chemical compound usage is that these chemical compounds may end up in unintended and often undesirable situations.

Food is a particularly undesirable location where chemical compounds may end; yet, the reality is that food contains a cocktail of chemical compounds including a large number of manmade compounds which were never intended for consumption. Some of these chemical compounds are present by nature, but a very large number originate from external influences, like contamination from processing, exposure to the environment, or migration from other sources into food (Figure 1). Most of the human exposure to these chemical compounds occurs at levels that are below thresholds to incur smell or off-flavour, and thereby raise little awareness. Therefore, sensitive analytical equipment is needed to detect these chemical compounds. While the concentration of chemical contamination in food is rarely at doses where acute effects can be observed, it is known that chemical compounds can cause adverse health effects even at low doses providing the exposure is during sufficient time. Consequently, food currently presents an almost unavoidable risk to human health, while it is simultaneously critical to survival.

While there is uncertainty on a fraction of the chemical compounds regarding identity or origin, it is certain that the number of chemical compounds present as contaminations in food is substantial (Koster et al. 2014; Summerfield & Cooper 2001). Tracing the origin of some chemical compounds is complex because the production chain of food is highly complex. For example, the process of transferring and processing food for consumers involves a great number of steps, e.g., applying production aids (pesticides, veterinary drugs, additives, etc.), transport, processing, preservation, and packaging. Each possible step involves the probability of introducing unwanted chemical agents into food, so the more food is processed and manipulated, the more likely chemical contamination becomes. The type, (geographic) location, and specifications of the process in each step of the production chain also influence which and how many chemical compounds may end up in the food, which makes it very hard to predict possible contaminations.

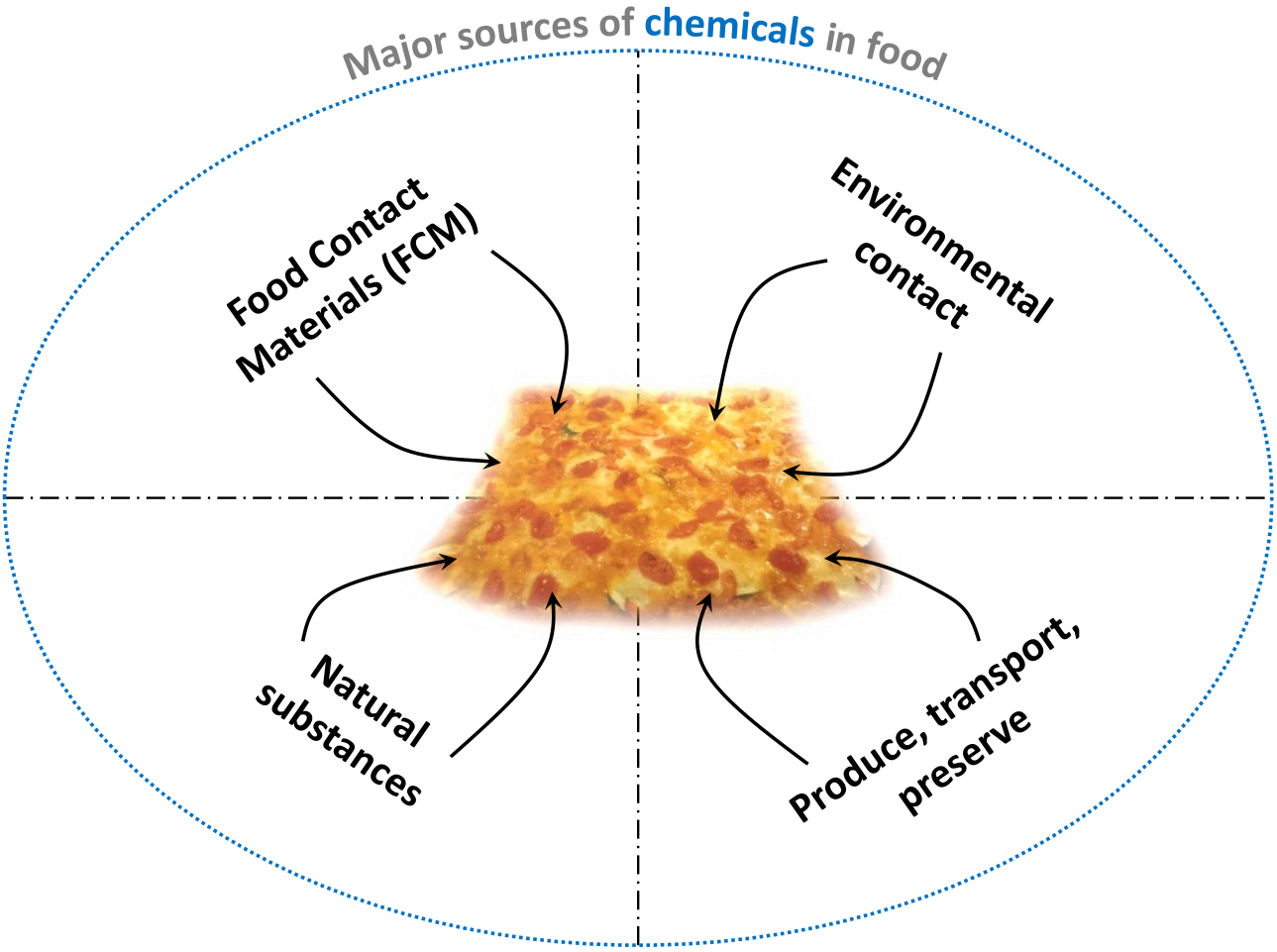


Figure 1: Food contains chemical compounds from a variety of sources. Numerous chemical compounds in food have the potential to induce adverse effects on humans. Many of these chemicals are man-made, but also some chemical compounds are natively present in the food. Man-made chemical compounds are not hazardous per definition, as many natural compounds are potent. However, man-made compounds are of particular interest since they may be avoided.

The influence of food packaging materials, also named the Food Contact Materials (FCM) such as plastic, paper, and board, is both an interesting and alarming because the variety of chemical compounds present in food due to FCM is substantial. For example, a printing ink regulation was adopted by the Swiss government in 2005 for printing inks contained well over 5,000 authorized chemical compounds (Food Safety and Veterinary Affairs (FSVO) 2017). The actual number of chemical compounds is likely higher, since this list does not include the unknown by-products or contaminations. Solely from packaging material, a few examples of possible types of chemical compounds to find in food are

packaging chemicals, production chemicals, printing inks, and degradation or combination products of all the previous (Castle, Offen, et al. 1997; Castle, Damant, et al. 1997; Nerín & Asensio 2007; Triantafyllou et al. 2007; Zülch & Piringer 2010). In addition, multiple food types are packaged differently and even within the same food type the packaging can differ greatly. An interesting group of food packaging materials are paper and board, which along with wood constitutes the majority of the growing market of packaging materials (Eurostat 2017). Therefore, paper and board are expected to be used increasingly more in the future especially so for luxury foods (Smithers Pira 2014), but are currently non-regulated on the harmonised European Union (EU) level.

1.2. Understanding the risk of unintended food chemicals

The risk posed by known or unknown chemical compounds migrating from FCM to food requires specific data to be identified. Risk can be split into two components (Figure 2): the average intake of the chemical, also known as human exposure, and the hazardous dose-effects relationship. The assessment of human exposure requires information on the concentration in the packaging materials and extent of migration from FCM into food. The combination of this information results in the estimated concentration of a chemical compound in food, although the concentration can also be measured directly. Comparing the concentration in food with the average human consumption of that particular food gives an indication of the actual human exposure. Hazards and adverse effects require extensive studies with a purified version of the chemical. These studies can either be performed by animal testing (*in vivo*) following extrapolation to possible human health effects, or can be obtained by testing the chemical in microbiological or human cellular essays (*in vitro*) and extrapolating the results to human health. More recent, the assessment of chemical hazards by computer-based structure assessment models (*in silico*) has become available, but this has not yet gained widespread acceptance as substitute for *in vivo* or *in vitro* testing.

However, it becomes clear that the strict requirement on information prior to assessing the risk greatly limits the number of compounds that can be assessed. The hazard of chemical compounds is difficult to study: the exposure occurs at levels that are insufficient to cause acute adverse effects, but chemical compounds can exert adverse effects on humans at

low doses and continuous exposure over longer time (Goodson et al. 2015; Lee et al. 2016). The major problem associated with this chronic, low dose exposure is that the visible effects, e.g., disease or health deterioration, are not instantly correlated to chemical exposure as there may be a relatively long time between cause and effect. In addition, the concurrent presence of other chemical compounds the correlation between the responsible chemical and the effect can be extremely difficult to assess. In addition, *in vivo* testing is subject to ethical considerations due to the need for animal tests, while it is costly for a large number of chemical compounds as many animals and tests will be required (Scholz et al. 2013).

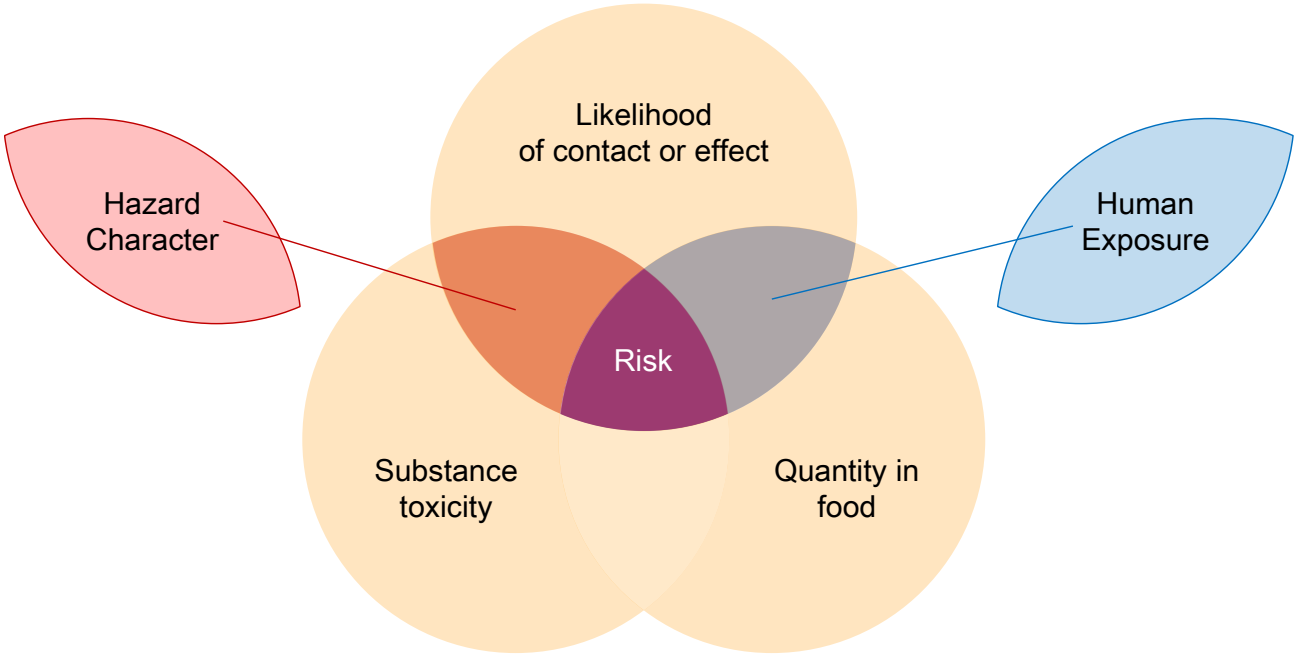


Figure 2: Risk can be represented as the combination of a hazard character and human exposure, which originate from toxicity testing, likelihood of human contact or likelihood of human effect, and the total amount that may be present in food. If for one of these parameters data is not available, then assessing risk is more complex and requires estimates.

The exposure to chemical compounds is equally difficult to monitor. Analytical methodologies that can determine the concentration of chemical compounds in foodstuffs are often highly specific to a certain chemical or a group of chemical compounds. When faced with a very large number of chemical compounds possibly present in food, the workload for determining the concentration of all chemical compounds is large. The

analytical challenges are discussed in more detail in 1.3.1. A further problem in exposure determination is consumption statistics, which may not be available for a specific type or food or for the population of interest. In addition, it may not be known how much the contact area is between food and packaging, or how large the migration potential of a chemical is given certain circumstances. Due to large number of unknown variables, assumptions are often used. However, the validity of assumptions can always be questioned and may not be applicable to specific cases. Hence, these assumptions are often constructed conservatively to ensure safety.

Perhaps the largest bottleneck in data demand of both hazard and exposure assessment is that the large number of possible chemical compounds demands a tremendous workload for all possible known and unknown compounds in food. For most of the unknown chemical compounds there is very little information available, and often the compound has neither been identified nor quantified, or has never been detected. Hence, without any available data or methods to obtain data, food safety control ends up in a deadlock (Grob 2014). The major question is: how is it possible to understand or assess the risk of a chemical compound that is unknown? As it turns out, the deficiency in information can be self-sustaining: the capacity to obtain information on unknown chemical compounds is limited by a deficit in knowledge on the chemical compound.

1.3. Limitations in ensuring safety in food contact materials

1.3.1. Analytical methods and principles are too specific

The way that analytical methods are designed is part of the reason that the lack of information is self-sustaining. Targeted methodology is common, so the compounds of interest are selected before the analysis and the method is optimized in a very limited scope. These selective methods offer the advantage of being sensitive, specific, and standardized; however, targeted methods are often specifically optimized, therefore unable to determine chemical compounds different than the selected targeted compounds. As a result, for every different group of chemical compounds different targeted methods are required. Furthermore, a wide range of chemical compounds cannot be simultaneously analysed because each chemical compound may have different optimal analysis settings,

which can be conflicting. While targeted methodology has been the backbone to safeguard food against a very large number of chemical compounds and are largely responsible for establishing safe food, they currently fail to discover new problems in an ever increasing problem in a rapid changing global market.

The limitations of targeted methodology do not just apply to existing analytical methods, but also to some fundamental analytical principles. As an example, quantification is inherently a targeted technique, because it is performed by comparing a measured signal in the sample to the measured signal of a known amount of the compound. Consequently, a compound needs to be known before it can be quantified, and the actual quantification is subject to availability or attainability of a reference standard. If the exact structure of the chemical compound is not known, quantification is usually not possible. Similarly, the identification of an unknown compound can be time-consuming and require skilled analysts if there is no prior information. In general, identification is rapid and relatively straightforward if the recorded experimental data can be compared to a library of existing data. When this is not possible, by mass spectrometry (MS) and/or nuclear magnetic resonance (NMR) spectroscopy it is possible to manually reconstruct a part of the molecular structure, but this is a time- and labour-intensive process that requires skilled analysts. Hence, if no recorded data exists for a chemical compound then reconstructing the molecular structure is an intensive task, especially when a large number of chemical compounds have to be analysed. In addition, the confirmation of the proposed molecular structure requires a reference standard to compare against.

The combination of highly-specific analytical methods and limitations of analytical principles to properly deal with unknowns causes a stalemate in analytical science. However, there has been an increase in analytical methods that deviate from the conventional strategy using non-target or untargeted approaches. These developments are needed, as they permit a much wider scope of selected compounds to be dealt with simultaneously, thereby increasing the coverage of analytical methods and reducing the workload. However, these methods eventually are limited by the fact that the discovered analytes can hardly be quantified or identified without reference standards and help of databases. In practice these methods are unable to substitute the targeted methodology as they do not generate the accurate and unambiguous data needed to draw final

conclusions. Hence, non-targeted methods are often only used as screening methods. In essence, the lack of information around unknown analytes can exist for longer periods because there are few methods or practices able to break this stalemate. Consequently, a significant portion of scientific work is dedicated to a small number of high profile, well-studied chemical compounds whereas a very small portion is directed towards the much larger group of unknown compounds.

1.3.2. Risk Assessment and Risk Management is reactive

There is a clear difference in the data that Risk Assessment (RA) requires and the data which can realistically be generated on large-scale by the existing analytical methods and hazard-assessment methods. In order for RA to be meaningful, a good exposure estimate and a proper hazard character is needed beforehand. Yet, as discussed, a requirement for the exposure and hazard data to be available is that there is information already available on the chemical compound: the chemical structure and methods to identify and quantify the compound. The challenge to estimate the real exposure and real hazard of a range of chemical compounds significantly hampers the ability to perform RA on chemical compounds. A priori knowledge is almost always required to perform a proper assessment, but there is a great shortage of methods capable of acquiring this information.

If these methods are absent — as is the case for unknown chemical compounds — then providing data is often a slow and labour-intensive process, if possible at all. Consequently, RA most often focuses on existing problems, rather than being proactive against emerging problems. Because of this, new “problem” chemical compounds are difficult to regulate or mitigate, as they require knowledge that is often only present at the stage where the problem is well-known and research has been carried out. In the current world with fast-paced changes and global trade, a reactive tool like the current RA is too slow and complex to completely ensure the food safety.

1.4. Hypotheses

It is possible to develop analytical strategies that combine elements of quantification, chemical identification, and risk prioritization that can be used to minimize the knowledge gap of unknown chemical compounds potentially migrating from food contact materials. Ultimately, this assessment strategy can be used in scientific risk assessment and prioritization.

I. The relative or absolute response factor (RF) for mass spectrometry (MS) can be optimized by carefully choosing the MS source parameters and liquid chromatography (LC) parameters for simultaneous analysis and semi-quantification of numerous unknown compounds. The normalized response factor can consequently be used in a semi-quantitative way based on:
    a. Suggested chemical structure based on standards with similar properties;
    b. Retention time;
    c. Some of the molecular properties.

II. Concentration estimates, with error margins, can be made through semi-quantitative determination based on comparison to non-ideal standards. Through careful selected assumptions, these estimates can be converted to a possible exposure and thereby be useful for risk assessment. The determinations can give concentration estimates of either:
    a. Groups of compounds defined by similar molecular properties of relevance;
    b. Individual detected molecules.

III. For an unknown compound eluting from a chromatographic separation it is possible to obtain a partial of full elucidated chemical structure based on accurate mass, isotopic information, and fragmentation analysis by Time of Flight (TOF) mass spectrometric analysis.

IV. Unknown compounds can be elucidated by strategies and accompanying methods proposed in I – III, and can subsequently be used in risk assessment via risk prioritization. By using risk prioritization with tentative data, an early exploration tool for discovering potentially high-risk compounds is attained.

Chapter 2

Background

2.1. Food Contact Materials

2.1.1. The role of paper and board packaging

The most commonly used packaging materials are paper and paperboard, plastics, glass, and metals. Of these four main packaging categories, the two most commonly used are plastic and paper/paperboard, accounting for somewhat over 70% of all food packaging (ALL4PACK 2016). Applications of packaging include food packaging materials and other types of food contact materials. Within each material group, there are large variations in types of packaging. For example, many different varieties of paper and paperboard exist, from carton boxes to cupcake cups. Similarly the material groups of plastics, glass, and metals also have numerous types and varieties, so the overall denominator of “food packaging” is very broad. Each specific application of packaging has different properties, which can consequently vary between the same packaging materials in a different application.

One of the reasons packaging is used frequently and extensively for food products is because if offers a multitude of advantages, and some are highlighted in Figure 3. Firstly, packaging will increase shelf-life and protect the product quality, and this is arguably one of the most influential effects of food packaging (Robertson 2010). By packaging the product, it is protected from environmental effects such as oxygen, microorganisms, light, and heat. Secondly, packing positively affects product hygiene (Dallyn & Shorten 1988). Thirdly, packaging permits for branding, handling, and consumer information. Companies use packaging as a way to brand and share information about their product, of which the latter is mandatory in the EU (European Parliament and Council of the European Union 2011b). Finally, packaging can provide physical protection of the products. Fragile products, such as eggs, could break easily if not closely surrounded by shock-absorbing pulp carton.

However, packaging is not exclusively advantageous. There are a number of considerations to take into account for packaging, see Figure 3. Firstly, product costs increase significant as a result of the added costs due to packaging. In 1997, the costs of packaging were estimated around 8.5% (USDA 1997), but more recent estimates range up from 25% to 40% extra costs (Robertson 2010; KnowThis.com 2017). The increase in cost

may also reduce the accessibility to food. Secondly, packaging has large detrimental environmental impact, most notably plastics ending up in the ocean (Ellen MacArthur Foundation 2016), or groundwater contamination by leaching from landfill of packaging waste (Renou et al. 2008). The continued use of disposable packaging products will increase the burden of waste and cause pollution globally (Derraik 2002). Thirdly, food packaging can alter flavour perception (Sajilata et al. 2007). The presence of contaminants even at low concentrations can contribute to significant off-flavours in food. Finally, food packaging will introduce chemical contaminants to the food, which in turn may be of risk (Suciu et al. 2013; Seltenrich 2015). The safety of food products should be improved by packaging, because packaging protects food against external influences and contaminants; however, the added protection may be mitigated if there are potentially hazardous chemical compounds migrating from the packaging to the food.

Figure 3: The use of food packaging has considerable advantages and disadvantages. Among these, improving shelf-life is considered a major benefit, whereas effects on the environment are considered a major drawback. Chemical safety, although not widely known as a problem, is discussed extensively in this thesis, whereas other factors are only discussed briefly.

Finally, there is the matter of food waste (Figure 3). Improving the shelf-life of food, either via protective atmosphere, improved hygiene, or protection will mitigate food waste at both retail and consumers (Halloran et al. 2014; Coma 2008; Robertson 2010). However, packaging materials can also contribute to food waste: the predetermined portion size leads consumers to choose more of the product than needed, which can result in thrown-away food, but waste can also occur when the packaging can be too difficult to empty completely (Williams et al. 2012; Silvenius et al. 2014). Hence, whether packaging reduces or increases food waste is subject to discussion, and depends strongly on the application.

2.1.2. Use and problems of paper packaging materials

The total use of food packaging materials has been steadily increasing over the years, but paper and wood constitutes the majority of the growing market of food-related packaging materials (Eurostat 2017). Hence, advancements in paper and wood materials provide interesting opportunities and challenges for the food market. In addition, by weight paper constitutes the largest waste contribution of all the packaging materials at 41%, whereas plastic contributes only 19% (Eurostat 2017). Paper is especially interesting for food packaging, and it is being increasingly used from 2012 onward while plastic use is stagnating; therefore, it is expected that a larger share of packaging will be paper as opposed to plastic in the coming decade, and especially so for luxury foods (Smithers Pira 2014). Therefore, it makes sense to take a closer look at paper and board products with regards to food safety, also because there is already significant research and knowledge specifically targeted at plastic packaging products.

While paper is a product from a natural resource, a large number of modifications are needed before paper has desirable packaging properties. The expanding and widened usage of paper requires more and different modification to improve the natural properties of paper, and also require controlled and consistent properties from batch-to-batch. For example, paper is not water or oil-repellent, is difficult to print upon, and has very inconsistent quality regarding tensile strength and density (Roberts 1996). Chemical compounds are added at various stages in the production process to overcome some of these limitations e.g., fillers, wet-strength additives, sizing agents, dyes, and brightening agents, or to assist in the production, e.g., retention aids, biocides, dispersing agents, and

defoamers. A part of the chemical compounds added in the production process are intended to remain within the paper to ensure permanent properties, while others are production aids and should not remain in paper. Aside from the chemical modifications in the paper production process, paper for food contact material is often made fit for application through printing, gluing, coating, whitening, or other modifications that introduce more chemical compounds into the material. Here, the details on the paper production process and subsequent modifications are not discussed, but the reader is invited to relevant work by Roberts (1996) and Biermann (1996).

Paper is often praised as sustainable packaging because it is biodegradable and easily recyclable. Paper is constructed predominantly from wood or plant fibres and is poorly resistant to environmental factors such as water, micro-organisms or mechanical force, which breaks up paper and board almost completely and makes it rapidly biodegradable (Venelampi et al. 2003; Yabannavar & Bartha 1993; Pivnenko et al. 2015). Once subjected to the environment, paper degrades into modified wood or plant fibres. Paper products are therefore especially suitable for recycling, as these fibres can be reused in recycling processes to turn old paper fibres into new paper products or as fuel (Hubbe et al. 2007). Currently, more than 71% of paper and board is being recycled in Europe, whereas the target for 2020 is set at 74% (European Paper Recycling Council 2017). On average, a cellulose fibre can be recycled approximate five to seven times before it becoming too short for any use (U.S. Environmental Protection Agency 2016). For plastic, only a limited number of types are recyclable, and often it is not economically feasible, as the costs of recycling may exceed that of using raw materials (O'Leary & Manavalan 2015), which is currently the case with relatively low oil prices. For glass, recycling is possible and often employed, but it requires significant energy to regain glass from recycled materials, albeit less than producing it from raw materials (Gaines et al. 1994; Vellini & Savioli 2009).

Recycling of paper and board to be used as food packaging material (FCM) has a profound negative impact on food safety due to the chemical compounds that are not sufficiently removed in the recycling process. Because paper and board are treated extensively with chemical compounds, these have a probability of re-entering the paper during the recycling process. Normally, care is taken to ensure chemical compounds are not present on the contact surface of virgin materials, but this is not the case with recycled

fibres. It has been discovered that large portions of the ink curatives and additives occur due to recycling rather than from migrating from the printed side through the paper (Castle, Damant, et al. 1997). The more recent case of presence off mineral oil also seems to be predominantly caused by recycling of printed board (Biedermann et al. 2011). Phthalates can also be traced back to recycling practices (Gärtner et al. 2009), and a number of other chemical compounds in paper and board can also be strongly related to recycling (Rosenmai et al. 2017; Bengtström 2014; Pivnenko et al. 2015). Consequently, while recycling is potentially valuable for a sustainable economy, the use of recycled materials for food packaging requires critical review.

2.1.3. Migration from paper and board materials into food

The ability of packaging to contribute a range of extraneous chemical compounds to the food has a profound research interest. In the public perception, packaging is seen as something inert: it protects the food from anything coming in and prevents anything going out. However, from a physical-chemical point of view the exact opposite happens: chemical compounds in food and chemical compounds in packaging are exchanged. The physical displacement of chemical compounds by migration is especially impactful if chemical compounds from the packaging transfer to food, whereas the opposite might instigate a mostly economic risk. Migration of chemical compounds from the packaging to food poses a serious concern. After all, the food is consumed and the packaging is waste, so which chemical compounds are migrating to the food and how can these impact human health once consumed? The presence of chemical compounds as a result of the production process and recycling is already discussed, but the process of migration into food has not been discussed.

Whether a chemical compound is capable to migrate to food is not immediately apparent, and depends on a number of molecular parameters and environmental parameters. The presence of a chemical compound in a FCM does not present a risk if it is not capable to migrate into food. Small molecules are for example much more likely to migrate to food than larger molecules (Jickells et al. 2005). There are many factors that affect the migration of a chemical compound, e.g., the total surface area, contact time, and contact temperature (Triantafyllou et al. 2007; Poças et al. 2011; Boccacci Mariani et al. 1999).

While some parameters are relatively easy to determine, some more intricate parameters like diffusion constants often require migration experiments (Triantafyllou et al. 2007). Overall, migration is a very complex process that involves a large number of parameters and it is difficult to predict accurately. In the end, a combination of all these parameters decides whether and how much a chemical can migrate to the food.

When considering the migration from FCM to food, there are significant differences between the different types of FCM. Different physical-chemical properties and production processes create very different migration profiles consisting of different chemical compounds, so each case needs to be individually treated. Besides, the chemical profile of potential contaminants in a material, the material itself plays a critical role in the type of chemical compounds observed in migration. Here, we distinguish between three different main groups of materials: impermeable, permeable, and porous, see Figure 4. In general, the more solid a material is, the lower the diffusivity and thereby a lower migration is expected (Castle 2006). In addition, solid materials typically only have surface migration, whereas more porous materials can have migration from the non-surface areas of the material. Hence, impermeable materials like metals and glass, have a lower migration than permeable materials, like plastics and rubbers, whereas the migration is highest for porous materials like paper and board (Poças et al. 2011).

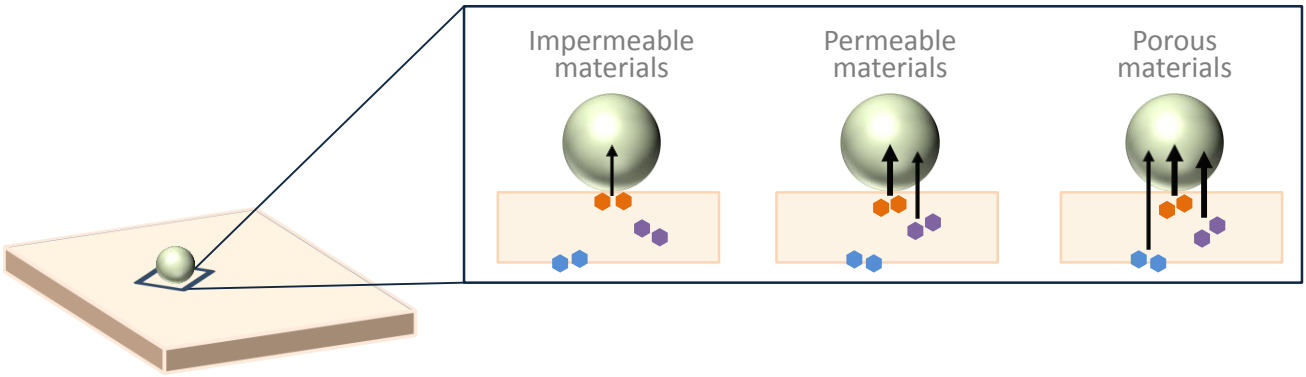


Figure 4: Migration within and from different porosity materials types. Impermeable materials, like glass and metals, permit migration only from the direct surface. Permeable materials, like plastics or rubbers, permit migration from surface and from layers close to the surface. Porous materials, like paper and board, permit migration from the surface, from the internal layers, and from the outer layers given the right circumstances.

The migration from a FCM into food follows to a great extent the laws of diffusion. The transfer of chemical compounds from a plastic FCM into a food is often well-described by Fick's laws of diffusion (Brandsch et al. 2002; Arvanitoyannis & Bosnea 2004). This has allowed for development of modelling tools for plastic materials which can be used to estimate the migration of chemical compounds to food in given circumstances, thereby reducing the need for time-consuming and expensive migration testing (Brandsch et al. 2002; Piringer 2007; Poças et al. 2011). However, for porous or multilayer plastic or paper materials, migration modelling becomes increasingly complex and no longer strictly follows Fick's Laws on diffusion, therefore requiring more detailed parameters and descriptions (Nerín & Asensio 2007; Barnkob & Petersen 2013; Zülch & Piringer 2010). Consequently, paper and board materials are notoriously complex to model accurate, and predicting the migration from these materials is not always possible, although there have been significant advances in these tools (Zülch & Piringer 2010; Hauder et al. 2013).

As paper and board materials are porous, the amount and diversity of migration chemical compounds can expected to be higher. The porosity of the paper and board structure enables displacement of chemical compounds within the packaging material towards the contact area. As a result, chemical compounds in areas not in contact with the food can migrate internally towards the food contact area, increasing the number of chemical capable of direct migration. Moreover, diffusion may also involve chemical compounds deeper inside the paper, or from the non-contact side containing printing, towards the contact surface (Castle, Offen, et al. 1997; M. Biedermann et al. 2013). This means that chemical compounds normally not expected to migrate, e.g., printing or adhesive chemical compounds on the non-contact side, are in fact capable of migration to food through paper and board (Bengtström 2014). Consequently, paper and board materials are often poor barriers for chemical migration.

2.1.4. Migration via direct contact by liquid or solid material

Direct migration occurs only when there is a contact area between the food and the packaging material, sometimes called "touching". In contact migration as shown in Figure 5, chemical compounds diffuse from an initial point of high concentration (i.e., the packaging material) to a point that has initial low concentration (i.e., the food) through

direct surface contact. Direct migration involves the direct transfer of a molecule from the packaging material contact area into the food, which confines the migrating chemical compounds to the contact area in case the FCM is not porous (Castle 2006; Gärtner et al. 2009). Contact between food and packaging can be either solid or liquid, and the contact area is often largest for liquids, followed by fine dispersed small solids (e.g., flour), and least for large solids (e.g., rice) (Eicher et al. 2015). Apart from a larger contact area, liquid contact also causes a faster migration due to fluid dynamics which permit higher concentrations of chemical compounds to be transferred to the food (Han et al. 2016).

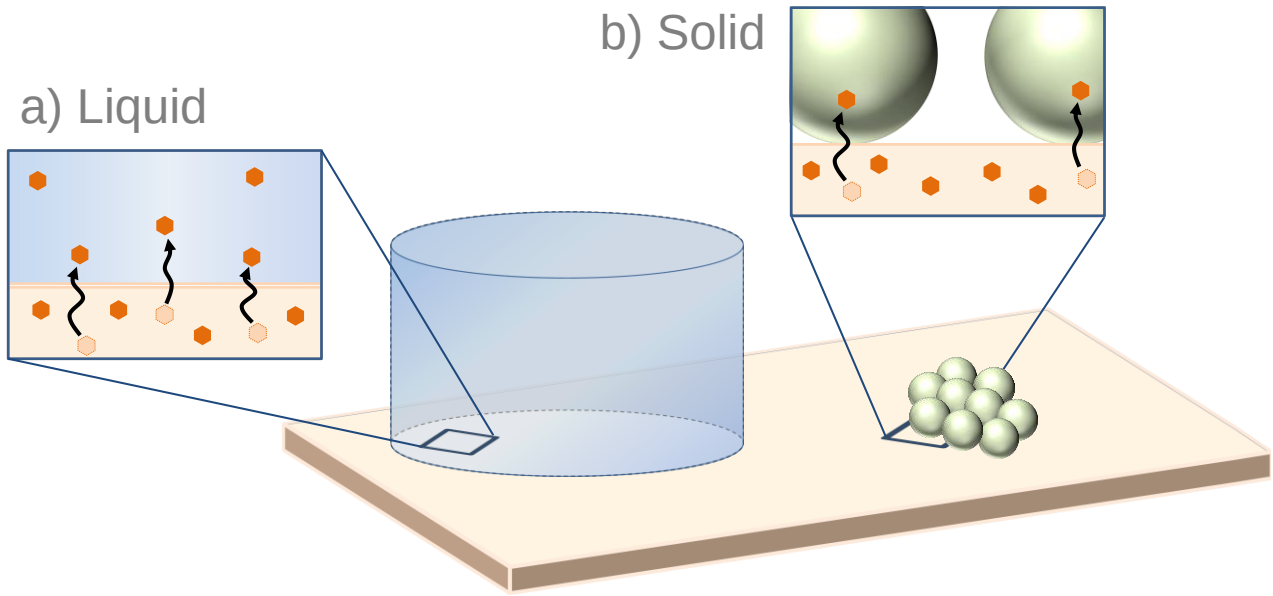


Figure 5: Migration by direct contact occurs either from liquid or from solid at the surface contact area. Migration into liquids is faster than that for solids, whereas finely dispersed solids have an effective larger area than large particles. The porosity affects the migration from larger depths of the material.

It has been shown that migration from direct contact is not negligible for paper and board: migration can occur even for non-volatile compounds (Triantafyllou et al. 2007; Biedermann-Brem et al. 2012). In addition, there are examples of food contact by paperboard that question the assumption of exclusively dry indirect migration, e.g., pizza boxes, snacks, fast food, or fruits (Binderup et al. 2002; Bradley 2006). Paper and board pose an interesting scenario in direct contact migration. Because paper is porous and permits internal displacement of especially small molecular compounds, the migration is

not confined to solely the contact area. As a consequence, a large range of chemical compounds that are not expected at the contact side of food have been found to migrate to food, e.g., photoinitiators or curing chemical compounds used in printing inks, recycled cardboard retained chemical compounds, and adhesives (Anderson & Castle 2003; Castle, Offen, et al. 1997; Aurela et al. 1999; Triantafyllou et al. 2007; Lago & Ackerman 2016).

The porosity of paper also permits liquid contact, for example from a fatty surface of food, to penetrate the paper material. The migration in these cases is better characterized by active extraction, which means chemical compounds migrate faster and at higher levels than diffusion would permit (Triantafyllou et al. 2007). In addition, the speed and depth of penetration into the packaging material will alter the type of chemical compounds extracted as the inside of the paper material is considered as a different phase, containing different chemical compounds than the surface of the paper (Zülch & Piringer 2010). Ultimately, penetration of liquid into the fibres of paper will weaken the cohesion between the fibres, which may release chemical compounds that were previously tightly-bound into the fibre network. For all of these reasons, uncoated paper and board is seldom used for packaging with direct contact with liquids.

2.1.5. Migration via indirect contact

Indirect migration can refer to migration from a secondary layer of packaging (e.g., the carton box around a plastic bag) (Jickells 2007), or as shown in Figure 6 by a phase transfer to the gas phase prior to migration into the food (Boccacci Mariani et al. 1999; Jickells et al. 2005). Since gas is abundantly present in almost all packaging designs — and definitely in those of paper and board — indirect migration is very common in all types of packaging and foods, but the magnitude of migration varies greatly. Migration via the gas phase occurs when chemical compounds in the packaging volatilize into the gas phase inside the packaging, after which they can migrate into the food. This does not require a contact surface, so migration can occur from any place in the packaging. Migration from a secondary packaging can occur in multiple ways, like rolled-up sheets of paper can off-set external printing inks to internal contact layers (Lago & Ackerman 2016; Bentayeb et al. 2013), or due to insufficient barrier properties of the primary packaging (M.

Biedermann et al. 2013; Fiselier & Grob 2012; Johns et al. 2000). Indirect migration is often suggested as the primary route of migration for paper and board materials because of frequent use for packaging of dry foods, e.g., rice, pasta, which have a relatively large particle size implying small surface contact with paper (Eicher et al. 2015; Bouma et al. 2003).

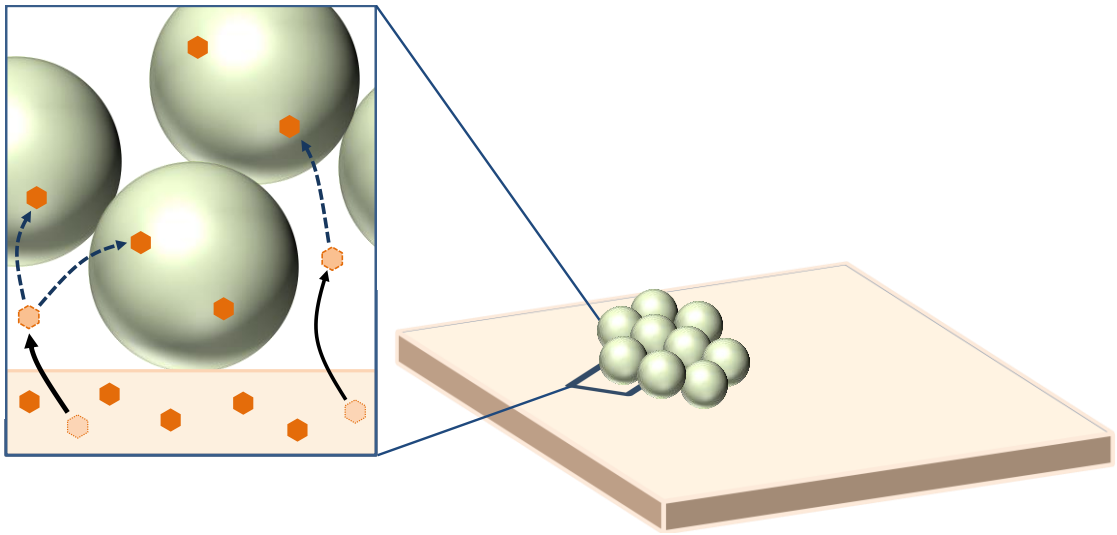


Figure 6: Migration by indirect contact occurs irrespective of the contact at the surface. Chemical compounds require a phase change from the material into the gas phase, after which they can diffuse and further distribute within the food.

Indirect migration via the gas phase requires that chemical compounds in the packaging are able to volatilize prior to migrating into the food. Because of gas-phase dependency, the process is strongly limited to chemical compounds able to enter into the gas phase with relative ease (i.e., volatile compounds) which excludes a great deal of possible compounds. The cut-off has been reported as a boiling point at around 400°C, meanwhile stating vapour pressure is likely to be more important (Jickells et al. 2005). In addition, the transfer rate to the food may also vary significantly with temperature and fat content (Triantafyllou et al. 2007), as the chemical compounds are required to adsorb to the food surface in order to enter the food. Lastly, the environment at which the food is stored greatly affects the type and level of compounds that are capable of migrating (Tehrany & Desobry 2004; Barnkob & Petersen 2013).

2.1.6. Compounds in paper and board: intentional and non-intentional

We can classify the chemical compounds in paper and board into two major groups as shown in Figure 7: the Intentionally Added Substances (IAS) and the Non-Intentionally Added Substances (NIAS). The latter has also been referred to as Oligomers, Reaction Products, and Impurities (ORPI) (Grob 2014), which covers the actual composition of this group better. Irrespective of the name, the differentiation between IAS and NIAS is predominantly due to available knowledge. Compounds that are added in the production process, are used as starting materials, or are added to the material to improve certain properties are generally considered as IAS. These compounds are known to be present, known to be (in)capable of migration, generally well-studied, have existing testing methodology, and are nowadays rarely a problem except in recycled materials (Binderup et al. 2002), where they may have become NIAS. Compounds that do not classify as IAS are then NIAS as long as they do not exceed 1000 Daltons (Da) (Silano et al. 2008). NIAS are reaction by-products of existing chemical compounds, impurities, contaminations within the process, or otherwise chemical contaminations not added (Koster et al. 2010; Wagner 2014). However, the difference between IAS and NIAS is not always clear from onset (Koster et al. 2010), and a chemical compound may be an IAS or NIAS depending on where it is discovered.

An important statement on the differentiation between IAS and NIAS is that the classification is not necessarily based on perceived safety. An IAS is not necessarily classified as safe, and NIAS are not necessarily a risk. For example, benzophenone is a widely used chemical in printing inks, and is considered an IAS for its use in printing inks (Koster et al. 2010). However, studies in paper and board have shown it migrates to food in significant amounts (Anderson & Castle 2003) and this may be a cause for health concerns (Snedeker 2014). Consequently, chemical compounds marked as IAS are only considered acceptable if these remain below the defined concentration thresholds (Specific Migration Limit: SML) set out for these compounds. The problem for NIAS is that few defined limits exist, like an SML or a *quantum satis* (as much as needed, but not more), because most NIAS are completely unknown. Instead, NIAS are covered under the Overall Migration Limit (OML), which is a maximum total migration of 10 mg dm^{-2} for plastic packaging material, or 60 mg kg^{-1} food for infants and young children. The OML is

relatively high considering some NIAS are possibly chemical compounds with high potency, or perhaps show carcinogenic properties. Yet, sometimes the relatively high OML can be exceeded through contact materials (S. Biedermann et al. 2013). The OML as safety margin for NIAS is highly questionable, and it is likely not an adequate safety approach for chemical compounds that have unknown toxicological effects and/or those that are relatively potent.

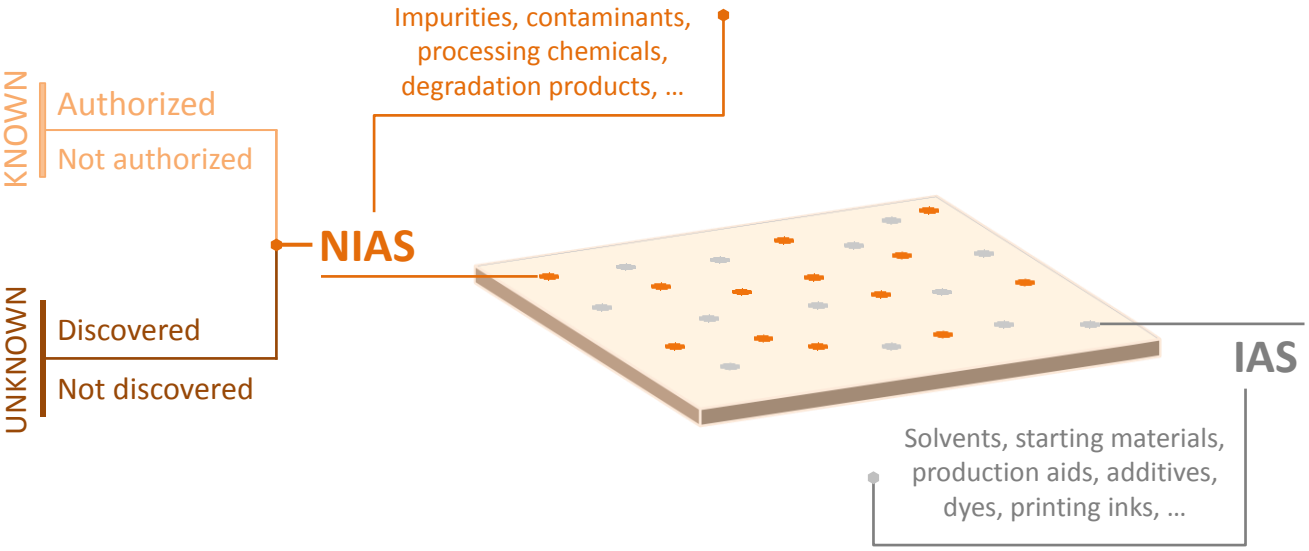


Figure 7: Differentiation between IAS and NIAS. The number of chemical compounds that are defined as NIAS is much larger than the number of IAS, and there are different levels of knowledge on the NIAS. Especially the unknown NIAS are of high concern, as no adequate safety assessment is available. Note that the differentiation between IAS and NIAS is not based on safety but mostly on origin and knowledge.

Because of the difference in knowledge between IAS and NIAS, it can be stated that risk associated with IAS is often understood and easier to control, whereas the risk attributed to NIAS may not be known or is not monitored, therefore has the possibility to be a high risk. Therefore, when problematic chemical compounds from food contact materials are discussed, the discussion almost exclusively deals with NIAS. The term NIAS is non-descriptive for the scale of the problem: this is a large and diverse group of chemical compounds where estimates in numbers go from tens- to hundreds of thousands (Grob et al. 2006). Due to the size of this group, finding and identifying each individual chemical compound is not feasible and perhaps impossible (Biedermann & Grob 2013), and here

we will highlight only a few sources of NIAS with some examples in paper and board materials that could be expected, but this is by great length not an exhaustive list. Because these are known NIAS, there are often well-investigated, whereas the unknown NIAS are poorly covered. Some examples:

I. phthalates due to recycled paper fibres (Pivnenko et al. 2015; Gärtner et al. 2009);
II. biocides used in production or recycling (Castle, Offen, et al. 1997; Lin et al. 2011);
III. printing inks or photoinitiators as a result of set-off migration, e.g., stacked or rolled FCM (Bentayeb et al. 2013), direct use on the material (Snedeker 2014), or recycling of materials (Boccacci Mariani et al. 1999; Anderson & Castle 2003);
IV. mineral oils from recycling (M. Biedermann et al. 2013; Biedermann et al. 2011).

2.2. The Atmospheric Pressure Ionisation process: Electrospray Ionisation

In mass spectrometry, analysis requires that chemical compounds are present as or are converted to gaseous ions. These ions can then be separated according to mass to charge (*m/z*) using electric or magnetic fields by a mass spectrometer. In practice, only a limited fraction of chemical compound is actually detected due to transmission losses in the mass spectrometer (Figure 8). A significant part of transmission loss can be attributed to ionisation and ion focus: only a small fraction of the analyte molecules are transmitted to the mass analyser (Hawkridge 2014; Leito et al. 2008). The transfer of compounds from the liquid or dissolved phase into gaseous ions requires an ion source capable of controlled phase transfer. The most well-known liquid sources are the electrospray ionisation (ESI), atmospheric pressure chemical ionisation (APCI), and atmospheric pressure photoionisation (APPI). The by far most common ionisation used in LC-MS is *electrospray ionisation* (ESI) illustrated simplified in Figure 8, and in more detail in Figure 9. The advantage of ESI is that it is simple to operate, generates ions from a wide range of compounds, has relatively high efficacy, is capable of dealing with large molecules, and is a soft ionisation technique that retains the original structure of many chemical compounds (Wilm 2011).

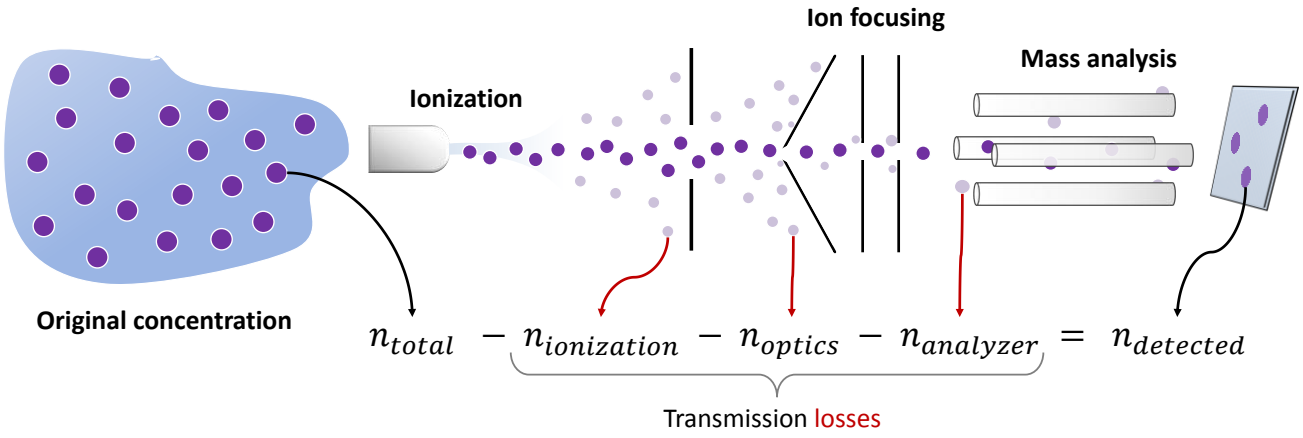


Figure 8: Ion loss within a mass spectrometry system. The magnitude of the loss depends on many factors including cleanliness, optimisation, and mass analysis. In practice, only a small fraction of the originally present molecules are detected as ions.

2.2.1. Mechanism of ion formation

In practice, ESI is achieved by applying a high voltage (typically ±2–5 kV) to the tip of a metal capillary outlet containing a liquid, which could come from a liquid chromatography instrument. A strong electric field (typically $E \approx 10^6$ Vm^{-1}) is created on the solution, which induces a charge separation at the surface of the liquid at the tip. In a positive electric field the negatively charged ions are drawn to the tip whereas the positively charged ions are repelled. If the field strength is sufficiently high, a Taylor cone will form at the tip that becomes unstable (Blades et al. 1991), releasing a jet of droplets from the tip that is similar in composition to the surface, i.e., has an excess charge (De La Mora 1992). Due to evaporation of liquid from the droplets, they are shrunk and the excess charge increases. At a certain point, the repelling Coulombic forces between excess ions will exceed the surface tension of the solvent: the Rayleigh limit is reached (Rayleigh 1882). When the Rayleigh limit is reached, the internal Coulombic forces cause fission of the droplet. The droplet deforms into a tail or cone, and a jet of offspring microdroplets are ejected parallel to the flow (Gomez & Tang 1994). Offspring droplets contain only a minor fraction of the mass of the parent droplet, a parent mass loss of 1.0 to 2.3%, while 10 to 18% of the charge is lost from the parent (Taflin et al. 1989). Because the offspring

droplets quickly reach the Rayleigh limit, the process is repeated until the n^{th} generation contains only nano-sized droplets (Kebarle 2000).

The exact mechanism of ion formation by ESI is complex and will not be considered here. Instead, some of the fundamental principles are discussed briefly, which are also shown in Figure 9. The formation of gas-phase ions is understood to occur through two simultaneous, yet competitive, mechanisms: the Charge Residue Model (CRM), as originally proposed by Dole et al. (1968); and, the Ion Evaporation Model (IEM) as proposed by Iribarne (1976). The CRM is not limited by the mass of the ion, so larger molecular ionise predominantly through CRM (Wilm 2011). In the IEM, however, formation of ions happens prior to the CRM through field evaporation via a solvated state (Loscertales & Fernández de la Mora 1995). IEM is most prone to happen at low ion concentrations ($<10^{-4}$ M) and exclusively for smaller, singly charged ions (Iribarne 1976; Iribarne et al. 1983).

2.2.2. Consequences of ion formation mechanisms

The ionisation processes depends on *surface-chemistry* in the droplets and of the compounds. Factors that influence the surface, compound concentration on the surface, or desorption from the surface will affect the total ion response. The surface of droplets can become saturated with ions when ion concentrations are too large, making ESI a concentration-sensitive technique. (Kebarle & Tang 1993; Tang & Kebarle 1991; Bruins 1998) The highest concentration or highest surface-affinity ions will be most likely to be found in offspring droplets, hence have the highest number of ions formed. Furthermore, the ionisation is surface associated: for any given set of molecules, those that have highest surface affinity will be those that are most likely to be found in offspring droplets (Cech & Enke 2000; Osaka & Takayama 2014; Chalcraft et al. 2009; Enke 1997; Iribarne et al. 1983; Iribarne 1976). In addition, *Gas-phase Proton Transfer* (GPT) happens when proton affinities between two gaseous molecules are sufficiently large that a proton can be abstracted, which causes two ionized species to compete for charge past the liquid stage (Amad et al. 2000).

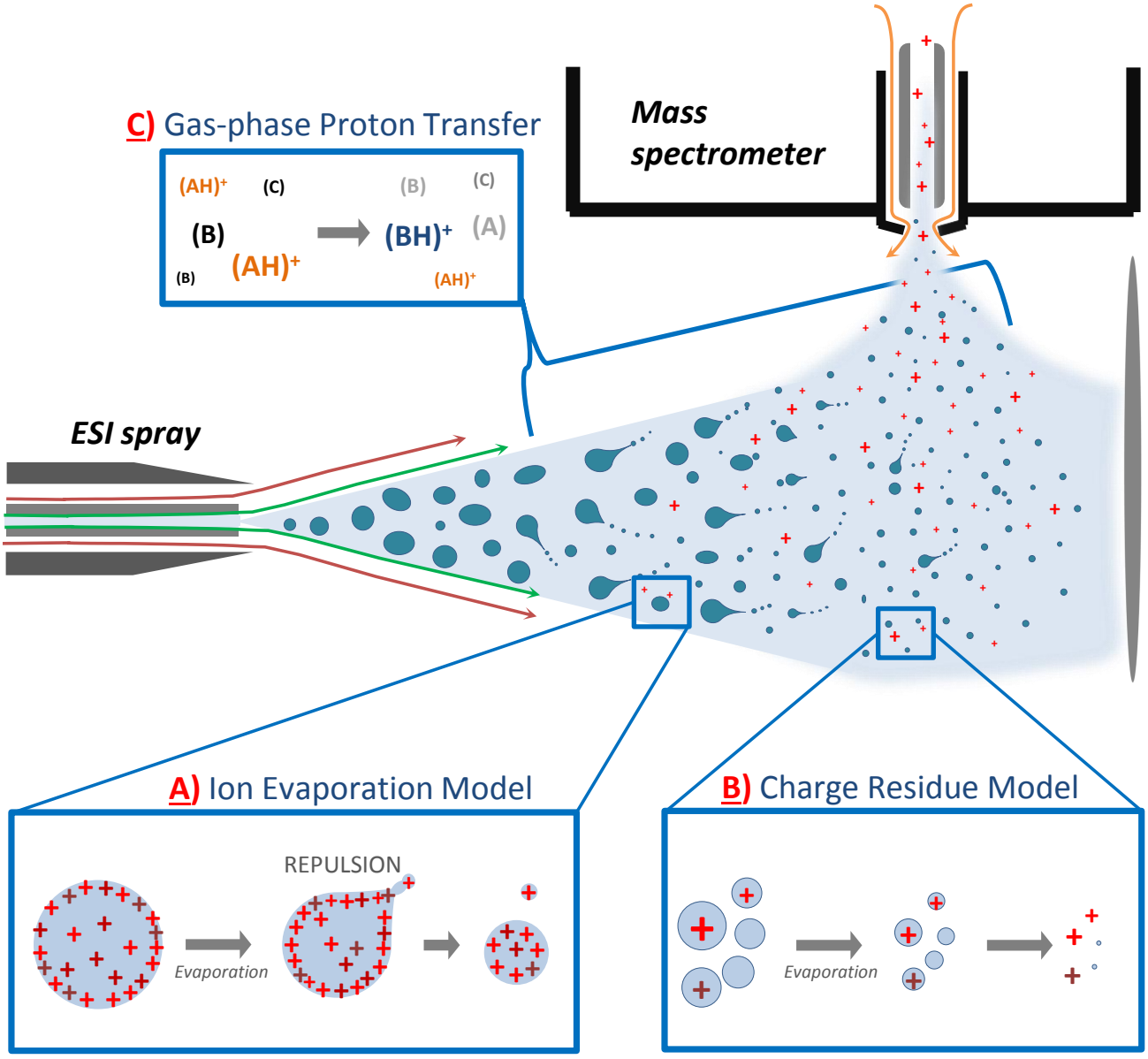


Figure 9: Schematic and simplistic visualization of the ESI process. A) The Ion Evaporation Model (IEM) forms solvated ion-clusters expelled from the primary droplet, which evaporate remaining solvent to become free ions. B) The Charge Residue Model (CRM) occurs after repeated process of fission and shrinking up to nano-sized droplets, where only a single ion is present in a droplet. C) Gas-phase Proton Transfer (GPT) happens as tertiary process, where charge is transferred from one gas-phase ion to another with different proton affinities.

The primary consequence of surface association is that when concentrations and differences in concentration, surfactancy, and proton affinity between ions become large enough, ESI becomes a competitive process (Bonfiglio et al. 1999). Effectively, this implies that the measured response factor (RF) of an analyte does not only depend on the concentration in the sample, but is also affected by competitive processes. Therefore, the RF is strongly related to how well a compound ionizes and is capable to compete in the LC-MS interface, and this may possibly be at the cost of RF of co-eluting ions.

The competitive and complex ionisation mechanisms observed in ESI affect the way quantification and identification are performed. Firstly, ESI is capable of producing several types of adducts with available salts, like solvents, water, formate, or sodium clusters. Ions will not predominantly ionize as the protonated adducts if other adducts are favourable in energy, e.g., ammonium ions or sodium ions. Hence, when employing targeted methods on ESI ions, the intensity of the protonated adduct may be lower than expected. Different adducts of the same ion behave differently, have different *m/z*, different stability, and have different fragmentation patterns. For the quantification of unknown compounds, the interface between LC and MS should ideally not discriminate between chemical compounds. However, it has been frequently reported that the nature of the analyte is strongly determinant of the response obtained (Bonfiglio et al. 1999; Cech et al. 2001; Cech & Enke 2000; P. J R Sjöberg et al. 2001; Leito et al. 2008; Kruve et al. 2014)

2.3. Detection and quantification of unknown chemical compounds

2.3.1. The current state of analytical methodology

The investigation of migration from paper and board FCM is often by adsorption to poly(2,6-diphenyl-p-phenylene oxide) — TENAX — from the gas phase (van Den Houwe et al. 2017). The TENAX is then desorbed (thermally or by solvent) and analysed by gas chromatography (GC) analysis, because the main route of migration are expected to be via the gas phase. Hence, the use of GC-MS has many advantages. Firstly, the ionisation technique most used in GC-MS, electron impact (EI), causes extensive molecular fragmentation, where the fragments can be used to deconvolve overlapping peaks, and

the fragments can contain useful information on the molecule structure. Secondly, identification of unknown compounds from GC-MS is simplified due to the reproducible and descriptive spectral qualities of EI, which permit database compilation and subsequent searching in these structure-spectrum databases. Moreover, the chromatographic resolving power of GC is high, and method optimisation is relatively simple, so the chromatograms are easy to interpret as the main chromatographic driving force is boiling point difference.

However, it is questionable to almost exclusively use GC-MS methods for analysis of paper and board FCM. Migration from direct contact is not negligible and can occur even for non-volatile compounds (Triantafyllou et al. 2007; Biedermann-Brem et al. 2012), like food contact by paperboard with non-dry indirect migration, e.g., pizza boxes, snacks, fast food, or fruits (Binderup et al. 2002; Bradley 2006). Finally, paper is a hydrophilic medium due to the presence of negative charges and carbohydrate-based fibres, so the retention of hydrophilic compounds is large. As gas chromatography and adsorption by TENAX requires compounds to be volatile, these methods deal poorly with highly polar chemical compounds, and are inherently limited in the analysis of potential migrants that occur via direct contact which do not need to be volatile. Hence, while gas chromatography is suitable for the well-studied migrants like diisopropylnaphthalenes or benzophenones, it may not be suitable for other chemical compounds in paper and board FCM.

As a supplement to GC, it may be valuable to equally investigate paper and board samples by liquid chromatography (LC). Many of LC applications in food safety are purely targeted methods designed to determine specific compounds like pesticides (Malik et al. 2010). However, there is an increase in LC-based screening approaches, as they are proving to be highly valuable to investigate the compounds not suitable to GC analysis. As an example, it has been shown that LC-MS is capable of covering a wide range of pesticide screening using an untargeted approach (García-Reyes, Hernando, Ferrer, et al. 2007). Recent work has also shown LC-MS is useful for screening for non-volatile contaminants from adhesives (Vera et al. 2013). In addition, surface analysis of food contact materials by LC-MS techniques are useful for rapid identification of unknowns, possibly for detecting previously unassessed substances (Ackerman et al. 2009). These advances illustrate that LC-MS is becoming an important screening methodology for food

samples, perhaps in a way similar to how environmental or metabolomics samples are screened (Viant & Sommer 2013; Krauss et al. 2010). However, LC-MS is still limited by a number of factors, for example the lack of existing spectra databases like GC-MS, the use of primarily soft ionisation techniques that require dedicated fragmentation, lack of compound-specific fragmentation, and ionisation that is strongly influenced by source parameters.

2.3.2. The role of chromatographic separation

Chromatography is the tool of choice when faced with mixtures, as it has high separation power for complex mixtures; however, optimising the separation is often a time-consuming process. In recent decades, chromatographic requirements have become less strict because of developments in selective, high-resolution, and multi-trace detectors like mass spectrometers. Multi-trace detectors are able to record several traces within a single chromatogram, which allows separation in two dimension (time * channel), whereas single-trace detectors only have a single dimension (time) available for separation. The added separation information of a multi-trace detector is needed for complex samples with many chemical substances, because adequate peak separation (resolution ≥ 1.5) cannot be obtained by chromatography alone. The large number of chromatographic peak resulting from complex samples imply that the retention window of the chromatogram needs to be sufficiently large or the peaks extremely narrow. As optimization of chromatographic separation is challenging and time-consuming, the availability of detectors like mass spectrometers that can separate chromatographically overlapping peaks by selectivity of the detector facilitate much faster method development.

The advantage of using selective multi-trace detectors in chromatography is schematically shown in Figure 10. In cases where the chromatographic resolution is sufficient, a selective multi-trace detector offers more detailed information per peak, but does not provide a significant chromatographic benefit. When the resolution is lower, multiple traces allow the deconvolution of two chromatographic peaks, whereas this is more complicated for a single-trace detector. If the resolution drops to unacceptable levels a single-trace detector is unable to differentiate between two peaks, appearing as a single peak, and in these cases the need for a multi-trace detector is obvious.

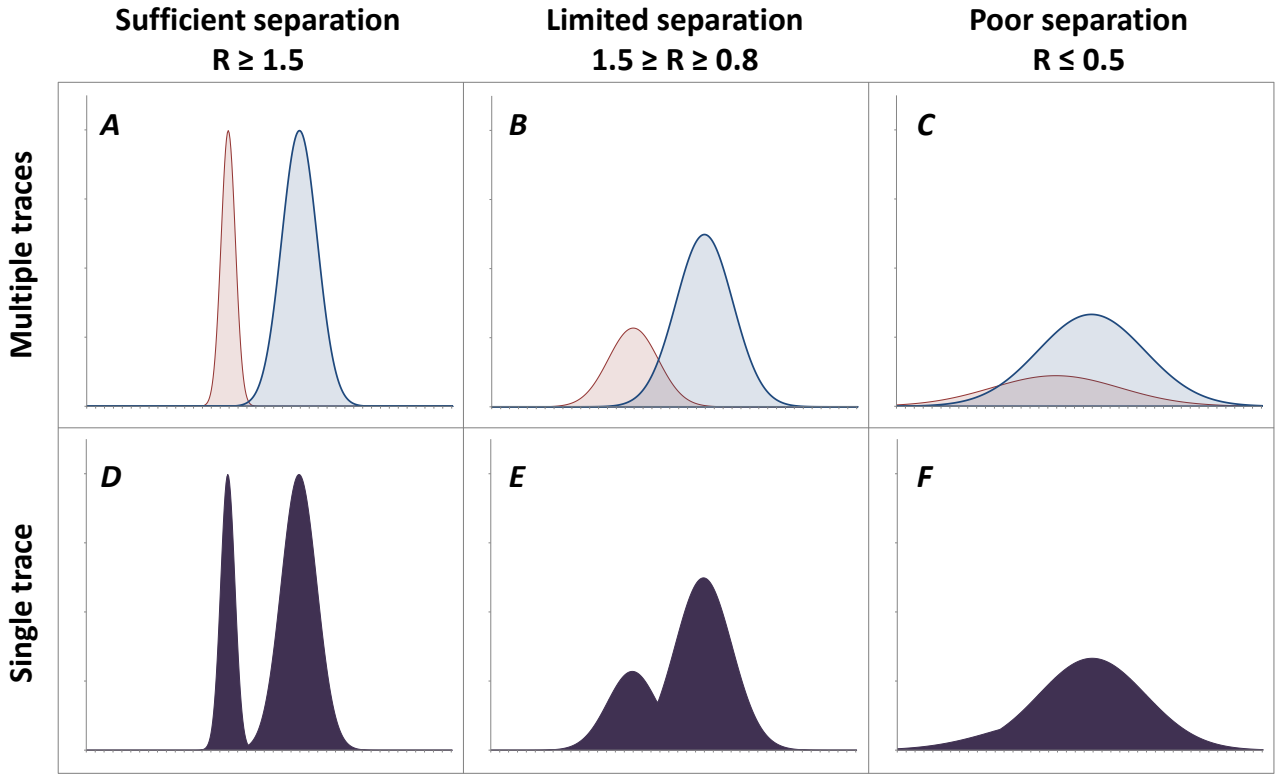


Figure 10: Resolution is less important with specific multi-trace detectors like mass spectrometers. In cases where the resolution is insufficient (C and F), detector specificity is crucial. For higher resolutions (A–E), the difference is less impactful, but specific detectors can provide additional information.

Despite the common availability of multi-trace detectors, the need to perform chromatography is present still. Too many simultaneously eluting analytes can cause suppression and interference effects in the detector, which can eliminate analyte signals or make these indistinguishable from others, and this is especially a problem in low resolution LC-MS (Furey et al. 2013; Berendsen et al. 2013). Chromatography also offers an important clean-up of relevant analytes versus a background signal. In addition, one should not underestimate the importance of chromatography in providing substantial information about the analyte., since the partitioning in chromatography is directly related to physical-chemical properties in the analyte, which in turn are related to chemical structure and this can be utilized for prediction of a number of useful properties, e.g., by Bökman et al. (2006) or Cech et al. (2001). Finally, by spacing out analytes over a time dimension the detector can use a larger amount of time collecting data per analyte peak, which can improve the quality or quantity or information gathered, and avoids that analyte peaks are missed due to co-elution with a much larger chromatographic peak.

2.3.3. Quantitative analysis

The primary objective of quantitative analysis is to obtain data on the occurrence of a specific chemical compound in a sample. Concentration data is valuable in prioritization: compounds in higher concentrations are likely to be of higher importance than those in lower concentrations given unknown toxicity. To estimate the concentration of an analyte of interest (AOI) in a given sample, the relationship between concentration and detector response should be known, the response factor (RF), which can vary greatly between different chemical compounds (Figure 11).

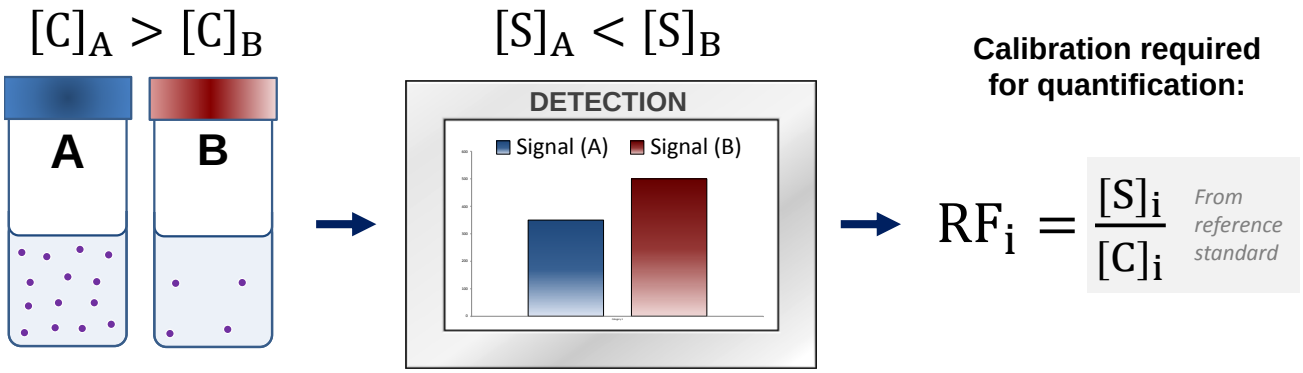


Figure 11: An idealized detector has a linear and unified signal to response factor (RF). In practice and as portrayed, almost all detectors behave non-ideal and show significant differences in response ratios. Hence, the ratios between signals may not be indicative of the relative concentrations in the original sample. Consequently, an observed signal cannot easily be converted to an actual concentration.

However, predicting the RF is not straightforward especially in LC-MS where differences in RF are large (Espinosa et al. 2015). The RF is predominantly determined experimentally by the analysis of a known concentration of the AOI under the same conditions as the regular analysis. Yet, experimental determination of RF is complicated and unrealistic for complex samples which may contain thousands of chemical compounds. In addition, the chemical structure of some compounds may be unknown, so obtaining a relevant standard for these compounds is impossible without further identification studies. The requirement for quantification to have accessible reference standards is severely limiting the application range and the practicality of quantification on complex samples (Chalcraft et al. 2009; Pieke, Granby, et al. 2017).

Despite the necessity to perform quantification in many cases it requires considerable resources to do it properly, therefore limiting the number of compounds that can included in analysis. The main reason reference standards are needed is because RF values are very variable for different compounds, mass spectrometry instruments, and analysis conditions. The measured response depends on a large number of variables like the molecular structure, matrix effects, sample purity, ion generation methods, ion optics, mass analyser type, and even detectors (Straub & Voyksner 1993; Per J R Sjöberg et al. 2001; Raji & Schug 2009; Tang et al. 2004). Additionally, there are a large number of environmental factors that affect the RF, which may influence day-to-day variation, e.g., instrument condition and maintenance, sampling performance, and temperature. Electrospray Ionisation is an important source of RF variation, which is the cause that LC-MS is often considered difficult as quantitative method (Taylor 2005).

When faced with complex samples containing many unknown chemical compounds, alternative quantification methods may be needed. One of these alternative strategies is to perform semi-quantification which relies on estimation of the concentration of an unknown chemical compounds by using a *different* compound. In a way semi-quantification is similar to the procedure of the additional of an internal standard to a sample to correct for errors, but in semi-quantification the internal standard acts as the quantification endpoint. Semi-quantification is a potentially interest technique because of its potential to avoid the more common quantification methods that require reference standards, which can be costly or impossible to acquire. The concept of applying semi-quantification has seen increased scientific attention in recent years, e.g., by Zou et al. (2017), Broecker et al. (2011), and Bu et al. (Bu et al. 2014). The implementation of semi-quantitative procedures is substantially more straightforward if all of the AOI are chemically similar, so that the variation in RF is minimal (Kovalchik et al. 2017). However, the application of semi-quantification in routine analytical methodology is a novel field, and still requires considerable research.

The premise of semi-quantification strongly contrasts the concept of traditional quantification, where the response of the investigated analyte (the AOI) and the quantification analyte (the QM) are often dissimilar. Especially LC-MS, where the response factor (RF) variance is large due to the mechanisms of ESI, seems to lend itself poorly to comparisons involving different chemical compounds. In order for semi-quantification to be viable in these applications, one of the key aspects is to ensure that the RF across the entire measurement range shows minimal variation. In those conditions, the worst-case quantification comparison will still give a controlled prediction error. That implies that even in the case where the RF between QM and AOI show a large difference, semi-quantification will give an *indication range* of the concentration. Despite semi-quantification being relatively novel field, it has been shown that there is merit in performing semi-quantification on complex samples and thereby avoiding the need for authentic standards, while also allowing prioritization (Pieke, Smedsgaard, et al. 2017; Bu et al. 2014).

The process of optimising for a minimal RF variance across the chromatogram is called normalisation, whereas conventional optimisation is regularly maximisation (Figure 12). Normalisation is applied to ensure that even a suboptimal combination of AOI and QM will result in a manageable measurement error. Traditionally, the system optimisation in LC-MS is performed in targeted analysis, i.e., *response maximisation* of a limited number of AOI to improves detection limits, e.g., as demonstrated by Gros et al. (2006). Maximisation of signals for specific compounds is not suitable for semi-quantification, as it is likely to increase the variation in response by optimising for a limited number of targeted analytes. Since maximisation favours specific signal increase, it may actually penalize the response of analytes with low responses. While those analytes may currently not be of interest to the particular method, they may become relevant in the near future. Hence, the optimisation in untargeted methods must ensure acceptable response for the low-RF compounds. In essence, since low-response analytes are already difficult to detect, they should have improved responses. Thus, response optimisation for the purpose of untargeted evaluation should consider the possibility of normalising responses and avoid penalising the difficult to detect analytes.

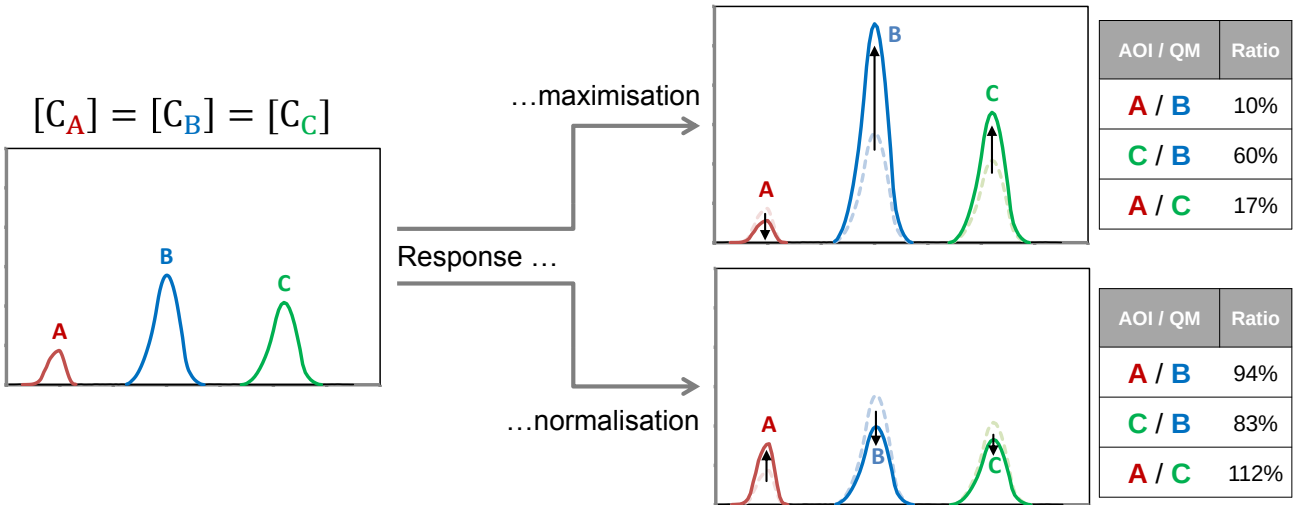


Figure 12: Method optimisation often is directed towards maximisation of a limited number of analyte signals. Since the responses of non-relevant analytes are ignored, the differences between signals can greatly increase as a result. In normalisation, all analytes and the ratio between analytes are considered as optimisation parameters, so that it favours increasing low-response analytes possibly at the cost of high-response analytes.

A further consideration is to select pairs of QM and AOI. Here, a limitation is that in semi-quantification of the AOI may be completely unknown which is opposed to an often known AOI in normal quantification. Because little information may be available on the AOI, selecting a proper QM can be challenging. Ideally, the QM and AOI can be selected on a large structural similarity so that the responses are possibly similar. However, to select AOI-QM based on structural similarity, the structure of AOI must be known. Instead, the selection of AOI-QM should be based on information that is sufficient for proper selection but also readily available. For example, the analyte retention time and the analyte mass could be used, since these are available from the MS measurements and do not require any knowledge of the structure. If this is possible the identification and quantification are completely decoupled, and would provide a major advancement to better understand unknown compounds.

2.4. Identification of unknown chemical compounds

2.4.1. Chemical identification

The objective of chemical identification (structure elucidation or compound identification) is to obtain as much structural information as possible of a given chemical compound. Chemical identification can be performed on different levels that result in basic information about the chemical compound, e.g., the elements present in a molecule, up to specific chemical structural information, e.g., the whole molecular structure. In general, when detailed information is required about a chemical compound, for example to do risk assessment and to mitigate risk, more detailed experiments are needed. These may require state of the art instruments, complex data processing and comparing to reference standards. The process of chemical identification can be visualized as a pyramid in Figure 13, where each layer implies that preceding layers are known or can be known from the current layer. Because the effort needed to reach the higher layers in the pyramid, there are fewer molecules for which detailed structural information is known. Essentially, the purpose of identification is attaining a level within the pyramid that is possible with current knowledge and resources, or otherwise sufficient for the current purpose. Consequently, many unknown NIAS will be only be identified to the base level of the pyramid or simply absent from it, whereas IAS generally are identified to a higher level.

The first three base levels of the pyramid contain fairly non-specific information about the structure of the chemical in question, but the information is not without relevance. First, the presence (or absence) of certain elements can prove to be valuable information. For example, isotope-based screening methods can be used to identify the presence of halogenated chemical compounds (Cincinelli et al. 2012), e.g., polychlorinated compounds (Focant et al. 2004), brominated flame retardants by looking for bromine ions (Stapleton et al. 2011), or chloride-containing pesticides (García-Reyes, Hernando, Molina-Díaz, et al. 2007). Similarly, indicate aliphatic or aromatic structures can be specifically detected and quantified by some detectors (Driscoll et al. 1978). Second, the elemental composition of a molecule can reveal the exact number of atoms, as well as the molecular weight and sometimes structural features. Often, one of the first steps in identification is to obtain the elemental composition. The use of high resolution mass spectrometry to obtain exact

mass and isotope patterns can be used to find the chemical composition of a compound, e.g., as shown by Schymanski et al. (2014) or Jaeger et al. (2017). Finally, functional groups are the first level that begins to reveal how the chemical structure is arranged and bonded. Examples of functional groups are alcohols (C-OH), ketones or aldehydes (C=O), aromatic rings, acids, bases, and so forth. Functional groups often reveal very practical information about the molecule, and methods can be optimized towards specific functional groups (Eberlin 2006).

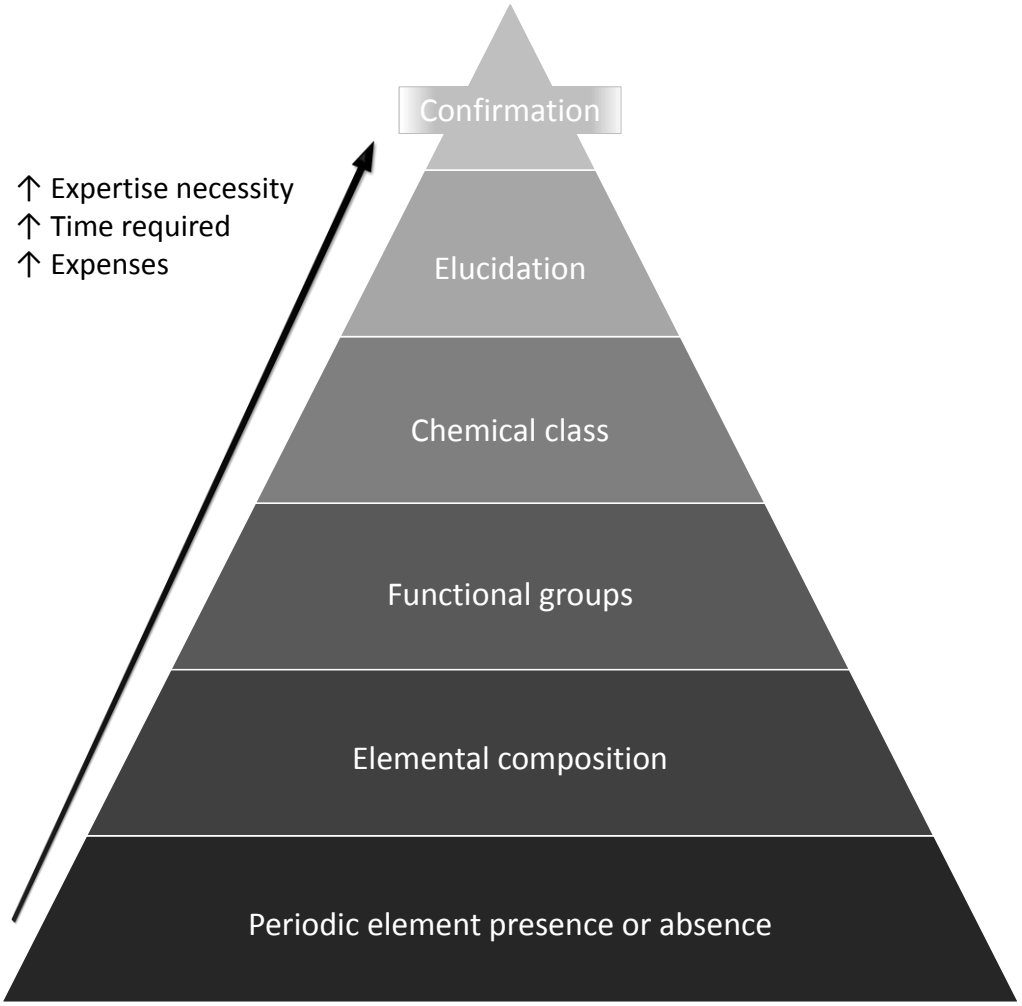


Figure 13: The identification of a chemical structure shows several levels of information. Generally, the higher in the pyramid the required knowledge, the more research is required, and often the requirements for instruments are higher. The highest level, confirmation, requires references to an authentic standard.

The top levels of the pyramid are significantly more informative about the exact configuration of the molecule, but are also far more difficult to achieve. More complex

techniques, like MS and Nuclear Magnetic Resonance (NMR) spectroscopy play a dominant role compared to simpler techniques like ultraviolet detection (UV) or flame ionisation detection (FID) analysis, and require more expertise. Firstly, the chemical class provides information on the arrangement of functional groups within a molecule, e.g., phthalates, or the functionality of a molecule, e.g., pesticides. The chemical class level is the first level where highly specific techniques are developed to discover and characterize different variations, e.g., bisphenols (Rosenmai et al. 2014) or phthalates (Heudorf et al. 2007; Aurela et al. 1999). Because similar structures can imply similar effects knowing the class of a chemical compound is highly informative on subsequent analysis or identification techniques. The following level of the pyramid, elucidation, is where information about the structure is becoming abundant and requires selective evaluation of data in order to avoid false positive results (Schymanski et al. 2015). Confirmation is often achieved by comparing the measurement results with that of a known standard, which will not be discussed here.

2.4.2. Dual-stage mass spectrometry: MS/MS

The retention of intact ions inside the mass spectrometer, especially in liquid chromatography – mass spectrometry, allows for more detailed investigation of ions. The detected *m/z* of an ion can be used to determine the molecular mass, isotopes, accurate mass (if sufficient resolution), and possibly a molecular formulation. However, the *m/z* of the molecule contains little to no information about the chemical structure of the molecule. Consequently, while high resolution greatly increases the specificity in molecular formulations that can be assigned to a detected ion, it is alone not sufficient for more in-depth analysis of molecular structure (Kind & Fiehn 2006). In order to obtain more chemical structural information on intact ions from a soft ionisation source, the ions can be collided with inert molecules under controlled conditions, which cause the original ion to break up into two or more fragments (Figure 15). Bond-breaking and the resulting fragments can be used to elucidate (part of) the molecular structure.

The combination of controlled in-source fragmentation and using multiple mass analysers (MS/MS) provides controlled fragmentation and thereby structural information, which can be used for structural elucidation. The first mass analyser selects ions of a specific *m/z*,

called the precursor, which is fragmented using different collision energies followed by determination of precursor and fragment ions by a second mass analyser, see Figure 14. Fragmentation data is valuable because all fragments are related to the precursor ion, and fragmentation happens in a systematic and somewhat predictable way (Hill & Mortishire-Smith 2005). Hence, fragments contain information about the original molecule structure much like a fingerprint, and certain fragments are indicative for a certain group of compounds. Identification via fragmentation becomes especially interesting if the second mass analyser is capable of accurate mass determination, as this improves the prediction accuracy (Gallart-Ayala et al. 2011). This means that fragmentation spectra can be used to predict the molecular structure, e.g., (Zhu et al. 2014; Jaeger et al. 2016; Bilbao et al. 2015), much like a puzzle being put together.

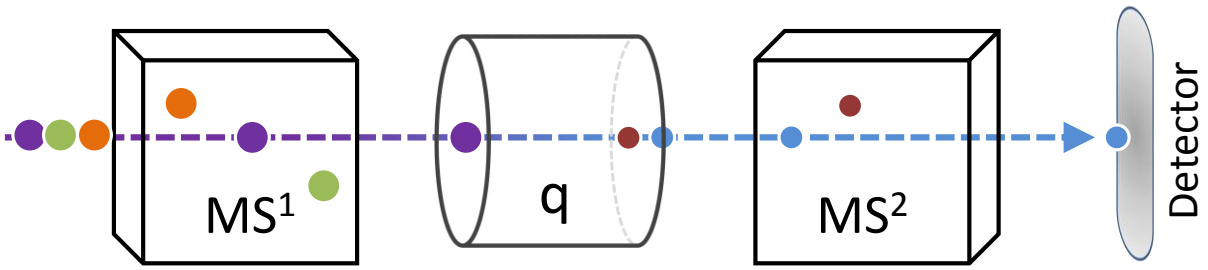


Figure 14: Simple schematic of a typical dual stage MS/MS setup. MS^1 often acts as ion filter for precursor selection, q is used to fragment the precursor ion, and MS^2 separates the resulting fragments either by selective mass analyser, like quadrupole, or a non-selective analyser, like TOF.

We can differentiate between three levels of identification by MS/MS in order of difficulty:

1. Confirming by available authentic standard;
   Requires the compound its reference standard to be available or obtainable.
2. Comparing to existing databases of spectra;
   Requires high-quality spectra, and the compound spectra must be available in the database.
3. By manual and/or automatic spectral interpretation.
   Requires high-quality spectra, a trained MS analyst, and possibly additional data.

Most identification studies are performed by the first two levels: using standards and/or databases. For example, Lago & Ackerman (2016) have recently demonstrated

identification by standard-matching and database searches of contaminants in food packaging using accurate mass and standard matching in LC-MS and GC-MS. A well-known problem with LC-MS is that the second option is rarely viable: no sizable database of spectra exists for compounds. This is in contrast to the availability of databases for GC-MS, for example the NIST database (NIST Mass Spectrometry Data Center 2017). As a result, LC-MS is often performed with available standards or, more rarely, by interpretation for spectra. For example, Ibáñez et al. (2005) showed third-level identification of unknown chemical compounds using mass spectra, but also indicated the need for mass spectrometric expertise and that it is time-consuming for even a small set of compounds. Often, identification on the third level is possible only when MS is assisted by other methods, e.g., as demonstrated by Blok-Tip et al. (2004) or Reepmeyer & Woodruff (2006).

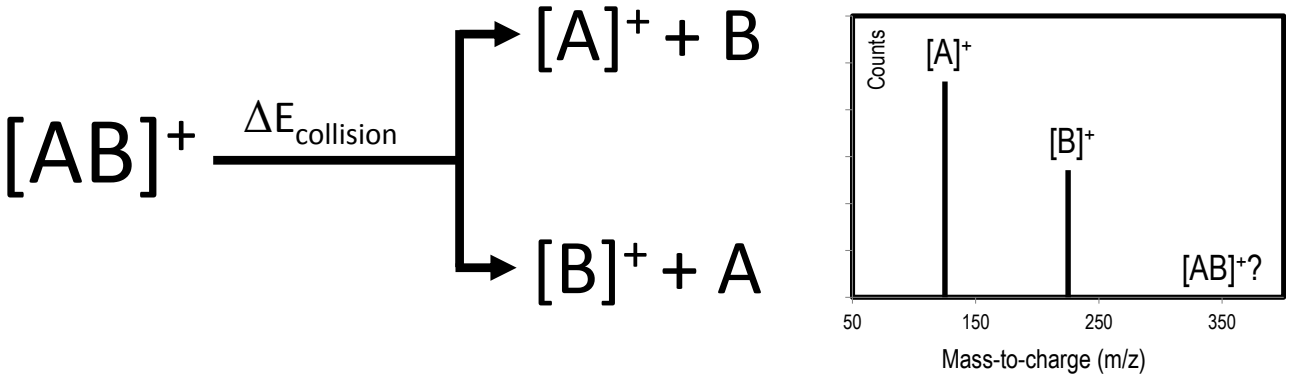


Figure 15: Fragmentation contains information about the molecular structure of a chemical compound. A molecular ion $[AB]^+$ consisting of groups A and B can fragment into each group. A complete fragmentation of $[AB]^+$ results that the molecular ion is not detected in the spectrum. In reality, fragmentation is more complex, and further chemical reactions or fragmentation can lead to unpredictable fragments.

While MS/MS proves to be a useful tool to tentatively identify chemical structures, it is often not sufficient to confirm the structure of a compound. The information obtained from the fragments in (high resolution) mass spectrometry experiments is rarely sufficient for a stand-alone identification (Krauss et al. 2010). First, the information from the fragments may be limited: it can be the result of a too high collision energy resulting in small, less informative fragments, or due to the formation of adducts that fragment poorly (Pieke, Smedsgaard, et al. 2017). Furthermore, the mass spectra obtained from MS/MS are not

sufficiently information-dense or predictable to elucidate the structure without additional information, like comparing to an existing database, so there is a greater dependency on the quality and actualisation of available information (van der Hooft et al. 2013; Kind & Fiehn 2010; Krauss et al. 2010). Finally, there is the potential for false-positive results because mass spectra are not always unique or sufficiently informative (van der Hooft et al. 2013; Kind & Fiehn 2006; Pieke, Smedsgaard, et al. 2017).

2.4.3. Non-targeted data acquisition in MS/MS

Multi-stage mass spectrometry (MS/MS), e.g., triple quadrupole QqQ, Q-TOF, or ion traps, are often associated with targeted analyses: the *m/z* of AOI is preselected, and fragmentation spectra are collected only for preselected fragments. Therefore, MS/MS generally requires compound-specific knowledge. Targeted MS/MS analysis is suitable when prior information is available on AOI: certain analytes can be expected and/or measured consistently. Due to the selective nature, targeted MS and MS/MS methods excel with a very low detection limit, and are highly specific, which minimises the probability of false-positive results. Because the methodology is standardized, results are comparable across instruments and across labs, These specific methods are based on information like the retention time, expected mass, and impurities, and specific fragmentation of chemical compounds. Because of prior information it is also possible to look for specific fragmentation patterns in MS/MS of closely retarded compounds. Hence, the identity of a chemical compound and closely related compounds can often be confirmed with high reliability by means of a reference standard or by existing spectra.

However, targeted methods can be subject to discussion in fields that do not have sufficient prior information, and thus require non-targeted approaches. For most samples, targeted methods are blind to chemical compounds that may be of interest but which are not expected or are unknown to be present, e.g., the presence of plasticizers in paper packaging or unknown NIAS. Food scandals, notably the melamine case (Chen 2009; Handford et al. 2016), have clearly delivered the message that exclusively using methodology with greatly limited scope is not sufficient to ensure food safety, and a similar message applies to food contact materials. Non-targeted analyses strategies can be used in cases where targeted analysis strategies are insufficient. The method design in non-

targeted analysis is different: instead of specific optimisation, these methods are developed to include as many analytes as possible. Hence, non-target methods consider all detected traces instead of a limited number of selected traces. However, non-targeted strategies are capable to include various levels of targeting, e.g., inquiry of a chemical database of suspected compounds. Generally, the sensitivity and specificity of non-targeted analysis is substantially lower; therefore, performing quantification and confirmation is often not directly possible in non-targeted methods.

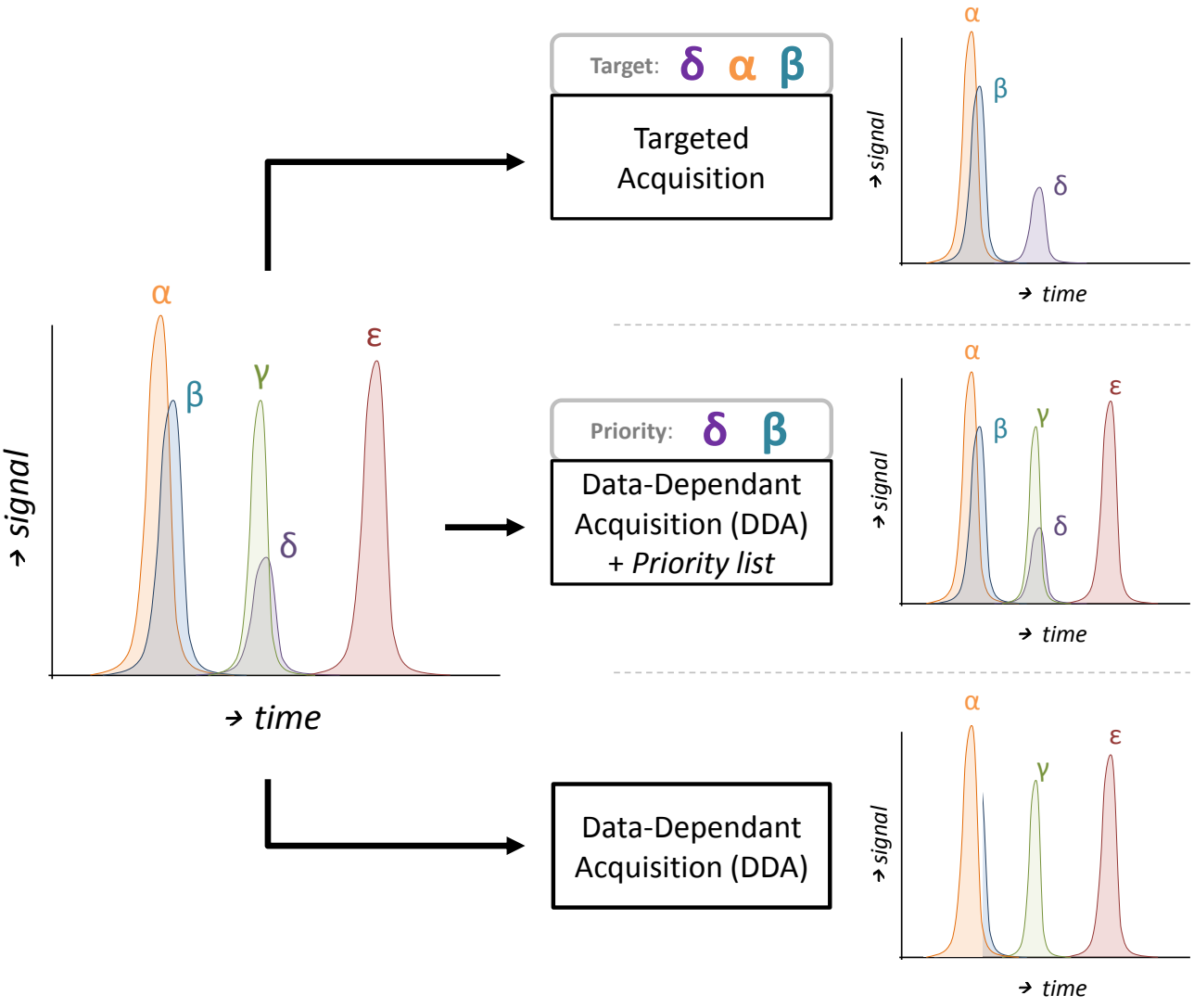


Figure 16: Different types of acquisition modes in MS/MS will result in different parts of the chromatogram covered. Targeted analysis (top) only covers the preselected ion in the method. Data-dependent acquisition (DDA, bottom) dynamically selects ions of interest, but deals poorly with overlapping ions with different intensities. A combination of DDA and targeting (middle) provides an in-between method capable of analysing both target and non-target compounds, but requires some prior knowledge on the sample.

The targeted nature of MS/MS appears incompatible with the non-targeted approach needed for samples containing unknown chemical compounds. Ideally, to use MS/MS in non-targeted analysis, the way precursor ions are selected needs a revision to be more dynamic. An alternative data acquisition method is data-dependent acquisition (DDA) (Thomas et al. 2012; Schwudke et al. 2007). In DDA, real-time MS data is evaluated for relevant precursor ions, and an algorithm selects precursors for fragmentation, see Figure 16. This permits the investigation of unknown analytes, as no prior definition is needed. DDA results in a repeating cycle of MS analysis, ion selection, and MS/MS analysis to collect fragmentation spectra. Due to the limited time spent in MS analysis, quantification is rarely possible when operating in DDA, as this is optimal on the MS signal. Another limitation associated with DDA is that complex samples may contain many unknown compounds which may not all be selected by DDA (Hopfgartner et al. 2012). The number of precursor ions to be analysed by fragmentation is limited by time required per scan cycle, so only a limited number of precursors can be included within a given time range (Zhu et al. 2014), and in these cases a type of precursor priority system may be needed.

2.4.4. Hazard identification after structure prediction

The reconstruction of a chemical image from a mass spectrum will often contain gaps of knowledge, as the reconstructed chemical structure may not be definite or could be only partly elucidated. This gap of knowledge can also be defined by stating the reconstruction provides a *tentative identification*. However, an important consequence of tentative identification is that subsequent acquisition of hazard data is hindered because this requires a chemically pure reference standard to test with, which can be complicated to obtain when the identified chemical structure has uncertainties. Similarly, literature studies on the chemical compound can take considerable more time if there is no clearly defined chemical structure, as several possibilities will need to be explored. Hence, while a tentative identification provides information on unknown chemical compounds, the extent of this information may be not sufficient to test or find hazard data on the actual chemical compound, which may differ from the predicted structure. Consequently, the well-established methods for hazard identification of a chemical compound are ultimately unsuitable for predicted structures.

Instead of acquiring hazard identification via traditional means, e.g., animal testing, advancements in alternative predictive toxicology might provide an outcome (Scholz et al. 2013). One of these is the modelling processes of structure-to-hazard, commonly referred to as Quantitative Structure-Activity Relationship (QSAR) modelling. Briefly, QSAR establishes a correlation between an endpoint, e.g., a hazardous effect, and properties of the chemical compound based on the chemical structure. The basis for this modelling is made on the assumption that if two molecules have similar properties and features, the effects they can exert on the human body are also likely similar. Like most modelling tools, the outcome usually contains a larger uncertainty than testing by experiments. Instead of elaborating upon the deeper and complex characteristics of QSAR, primarily the possible application for hazard identification for predicted structures is discussed here. Recent work on QSAR, e.g., Cherkasov et al. (2014) or Rosenberg (2017) provides a better insight into the mechanisms and fundaments of QSAR.

QSAR has seen significant advances for FCM in the recent decade and is considered one of the key methods for future hazard identification (Mays et al. 2012). The acceptance of QSAR models instead of regular testing is increasing, as QSAR is currently included as alternative method in EU regulation (European Parliament and Council of the European Union 2006). QSAR may also be used, among other things, to fill data gaps for hazard and risk assessment, as well as in priority setting (European Chemicals Agency 2008). Hence, the use of QSAR on explorative data fits well and is compatible with current regulation standards. Recent work by van Bossuyt et al. (2017) clearly demonstrated the added value QSAR can have in early-stage identification of hazards.

Consequently, QSAR is extremely useful when combined with tentative identification, as the input requirements of QSAR are molecular structures, so conveniently similar to the output of tentative identification. However, tentative identification provides the best-matching structure which may not be similar to the definite chemical structure. As a result, the QSAR prediction output may be based on a different molecule than truly present, so the result can deviate. However, because the assumption in QSAR is based on the hypothesis that similar molecules have similar activities the predicted activity may be adequately descriptive if the predicted chemical structure is similar to the true chemical structure.

2.5. The role of tentative data in risk studies

2.5.1. Existing regulation on food contact materials

Food contact materials as a source of chemical compounds in food are not necessarily a new phenomenon. There are regulations that, although expressed in general terms, seek to reduce and control the amount of chemical compounds that can be present in food as result of migration from packaging. Currently, the most important non-specific regulation in the European Union (EU) to ensure consumer health is the General Food Law EC directive 178/2002 (European Parliament and Council of the European Union 2002). The General Food Law prohibits food being placed on the market if it is unsafe: injurious to health or unfit for consumption (see Figure 17). The Food Law is there to ensure consumer health protection, but the definition of unsafe is ambiguous. For example, it is not clear what the food safety requirements of packaging are in the General Food Law. Controversially, the WHO recently reported that red meat and processed meat are probably carcinogenic to humans (World Health Organization 2015), so it *could* be argued that this may conflict with the General Food Law principle. Hence, to ensure that the food is safe all possible risks to human health need to be identified.

In light of those shortcomings, frameworks exist that permit specific regulations for FCM. The most important framework regulation is the Regulation EC 1935/2004 on materials and articles intended to come into contact with food (European Parliament and Council of the European Union 2004). This regulation defines that food packaging used under normal conditions must not transfer constituents that can endanger human health or adversely affect the quality of the food (see Figure 17). It includes all packaging materials that are intended to be brought into contact with food, are already in contact with food and were intended for that purpose, or can reasonably be expected to be brought into contact with food. One of the critical points in this regulation is that it paves the way for specific regulations for selected contact materials, which may include lists of permitted or banned compounds, migration limits for chemical compounds, additional rules, or analytical testing methods to ensure compliance. Essentially, this regulation permits further specific regulations that can ensure the safety of selected materials and thereby facilitates the acquisition of knowledge needed to build new regulations.

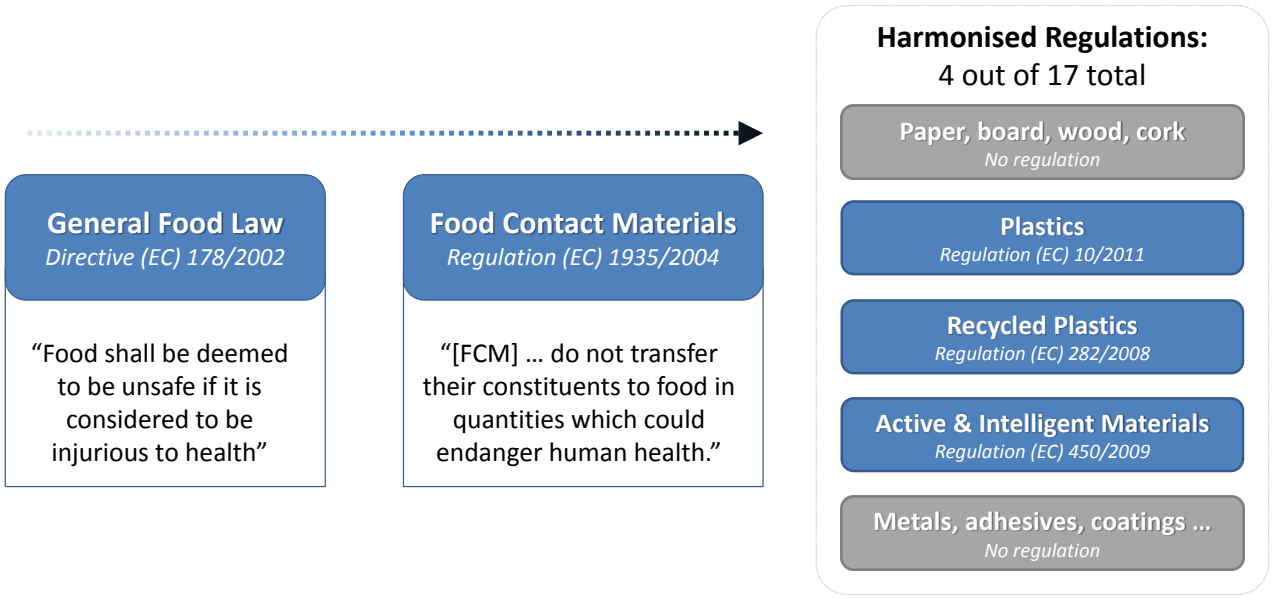


Figure 17: The progression of regulation is becoming specific for certain FCM materials. Currently, only 4 out of 17 materials are harmonized on EU levels, whereas the remaining 13 materials are regulated on national level or are unregulated. Paper and board, as well as printing inks and adhesives commonly used therein, are currently not harmonized.

The most important specific harmonised regulation on FCM as a result of the call for specific regulations by (EC) 1935/2004 is Commission Regulation (EU) No 10/2011 on plastic food contact materials (European Parliament and Council of the European Union 2011a). Regulation (EC) 10/2011 defines a positive list of chemical compounds permitted to be exclusively used in production of plastic materials, and these are often IAS. As a result, substances found in the material not present on the positive list will need to be authorized for use beforehand. For some of the listed substances, a migration limit into food simulants is defined by an SML. Secondly, the regulation dictates that the use of modelling tools is permitted to substitute migration testing given that these are at least as severe as the migration testing, and this paves the way for *in silico* models and exploration methods. Thirdly, the regulation defines a set of migration tests and food simulants to be used for testing, which can provide the basis for specific regulation of other materials.

Here, the Regulation (EC) 10/2011 that harmonises the specifications of plastic contact materials across the EU Member States will not be discussed extensively; an in-depth report as guidance through the document is published by the European Commission's Directorate General for Health and Consumers (DG Sanco 2014). However, a few interesting points of concern can be identified from Regulation (EC) 10/2011:

I. printing inks, adhesives and coatings do not have to comply with the compositional requirements of the Plastics Regulation;

II. NIAS are not subject to authorisation and listing in the Union list, but have to be absent of risk for human health;

III. FCM not covered by this legislation (non-harmonised, e.g., paper/board, rubbers, glass, etc.) are subject to, at best, national legislation.

These points demonstrate that risk assessment of FCM is a very complex issue. A recent report by the EFSA Scientific Cooperation lists 2800 entries used in the manufacturing on non-harmonised materials (not covered by EU legislation) like paper and board, printing inks, coatings, rubber, colorants, wood and cork (European Food Safety Authority 2012). Non-harmonised materials are more common than harmonised materials representing 13 out of 17 defined materials (see Figure 17). These materials are — in the best case — covered by national legislation; however, not every EU member state has national legislation in place for all materials, see Simoneau et al. (2016, especially Table 4). The question is: "How to ensure that chemical compounds in food do not cause adverse effects on health, as required by law, when their structure, concentration, and hazards are completely unknown?" In other words: when dealing with a non-harmonised material that can contains possibly harmful chemicals, e.g., NIAS, printing inks, adhesives, how are producers supposed to demonstrate safety for use in the EU? The capacity to demonstrate either safety or non-safety is greatly limited by the fact that NIAS are not known or regulated and a number of compounds are exempted from regulation.

2.5.2. Risk assessment as tool for safety

Risk Assessment (RA) of chemical compounds plays a very important role in food safety, and is the tool of choice to control and reduce the adverse effects of chemical compounds on humans. The risk a chemical may pose is estimated by combining the exposure to the chemical, i.e., *exposure assessment,* and the adverse effects of the chemical, i.e., *hazard characterization* (WHO/IPCS 2009). Here, exposure is the intake of compounds over a certain amount of time through consumption of food. Exposure levels are related to the consumption pattern of individuals and the concentrations of chemical compounds in the product, and should be made available from a diet study or from modelling (Fryer et al. 2006). Hazard is the potentially adverse effect a chemical may cause on human health, which can range from minor to fatal effects often determined from assays or animal studies (O'Brien et al. 2006).

In combination, exposure and hazard dose-response relationship provide a quantitative basis to assess possible risk which is the result of a RA. The availability of this data is critical in to ensure adequate risk characterization to manage the risk. The need to reach a consensus decision between experts, based on appropriate data, is one of the reasons RA is a time-consuming and resource-intensive process. Despite the historical use of RA, a recent study initiated by the Food Packaging Forum (FPF) concluded that current procedures of RA are ineffective at protecting public health by relying strongly on self-regulation by industry and availability of standards (Muncke et al. 2017), similar to previous conclusions on the topic (Grob 2009). Consequently, RA is often reactive to existing problematic chemical compounds in food rather than proactive in identifying risk of new hazards, and it is questionable that a reactive approach by RA is suitable to ensure continued food safety.

There is a larger awareness nowadays of the risk imposed by NIAS, which have directed the development of tools capable of understanding those compounds. In this light, the use of the Threshold of Toxicological Concern (TTC) concept is getting increased attention due to its capacity to deal with unknowns much more easily than traditional risk assessment (EFSA & WHO 2016; Kroes et al. 2004). However, the TTC approach still requires an assessment of exposure and also some chemical structure information, which may not be easily attained for unknown chemical compounds. For gathering exposure data, screening

strategies are suggested (de Fátima Poças & Hogg 2007), but these still rely strongly on available quantification methods. Koster et al. (2014) have shown that a TTC-like approach (named CoMSAS) including preliminary risk assessment is possible using a combination of analytical methods. This approach is promising as it assumes a worst-case TTC scenario; thereby it doesn't directly need initial identification. However, the approach is unsuitable for testing chemical compounds incompatible with gas chromatography (GC), cannot ensure there are no genotoxic effects, and assumes that the internal standard is worst-case. In addition, the approach contributes little to overall knowledge in NIAS, as it excludes many NIAS without identification or suitable quantification and is limited to predominantly GC-compatible chemical compounds and general purpose LC-compatible compounds.

However, risk assessment and even the TTC strongly rely on available data, but do not provide strategies for obtaining data. The fundamental idea of Koster et al. (2014) is to obtain more knowledge on NIAS, and this is badly needed judging from the current status quo. Approaches similar to CoMSAS can be valuable tools in the discovery of new chemical compounds and new risks. More knowledge on available NIAS supports making better decisions, so acquisition of knowledge should be the priority of analytical methodology, rather than continuously screening for the same well-known compounds (Grob 2009). With improved knowledge, compounds can be prioritized according to expected risk, which results in a better allocation of (limited) resources. Here, a stepwise and knowledge-based approach allows the prioritization of known compounds, but also allows the acquisition of new data needed for prioritization of less-known compounds. Essentially, the goal of knowledge-building analytical strategies should not be to replace existing RA methodology, but to improve the quality of choices that occur before RA.

2.5.3. Tentative data: the value of more knowledge

In order to start a systematic investigation of NIAS, there is need for much more knowledge on this group of compounds. Knowledge can include occurrence of chemical compounds in FCM (identity), at what concentration level they are expected in either FCM or in food (quantity), and what possible harmful effects the chemical compounds can potentially have (hazard). As discussed, currently all three of these information pillars are

equally important in risk assessment (also see Figure 2 on page 5) and rely strongly on targeted methods (e.g., quantification by standards or assay testing), or are depending on each other (like identification is needed prior to quantification). Recent studies appear to better capable to decouple these dependencies and thereby study risk more efficiently, for example by Bengtström et al. (2016) and Rosenmai et al. (2017). Here, an approach that evaluates the entire FCM for activity on gene assays, and subsequently identifies some of the chemical compounds responsible for this activity. This hazard-based approach provides valuable information on the hazard-related activity of the entire FCM or fractions thereof, and can be used for prioritization. However, the testing and identification of chemical compounds is still somewhat limited due to the need for reference standards.

Because even novel investigative methods with high potential, e.g., Bengtström et al. (2016) and Rosenmai et al. (2017), are limited by targeted analysis it may be viable to investigate chemical compounds in FCM directly by tentative data obtained via untargeted analysis. Tentative data is often the result of estimates, contains uncertainty, and thereby differs from confirmed data. In previous sections some examples of tentative data have been discussed, e.g., semi-quantitative concentration estimates (2.3.3), chemical structural elucidation (2.4.2), and computer-based hazard prediction (2.4.4). There are similar characteristics to these methods: the acquisition is less resource-intensive, methods are non-specific and are largely designed to minimize labour, and ideally should be a fair representation of the actual data. Generally, the quality of tentative data is a compromise: the less uncertainty tentative data has, usually more resources are required, more specific methods are needed, and/or there is more manual labour.

The inherent uncertainty of tentative data makes its use in RA a complicated matter, because RA relies on accurate and representative data that is not easily disputed. As a result, the use of tentative data in risk assessment is uncommon, and there is a strong dependency on specific data. However, tentative data may be able to aid RA in various ways. Firstly, the tentative data can be used in chemical risk prioritization (Figure 18), which is especially valuable for unassessed and undiscovered NIAS. Secondly, tentative data can be used for establishing databases of discovered chemical compounds which can aid future investigations, and also in generating a “chemical fingerprint” of typical samples. Finally, the acquisition of fundamental knowledge on a chemical compound is

valuable as future developments may discover new hazards associated to known chemical compounds, in which case the already established presence of chemical compounds in FCM will aid RA. In this thesis, the use of tentative data in risk prioritization is primarily discussed, but the use for knowledge-building may be even more important.

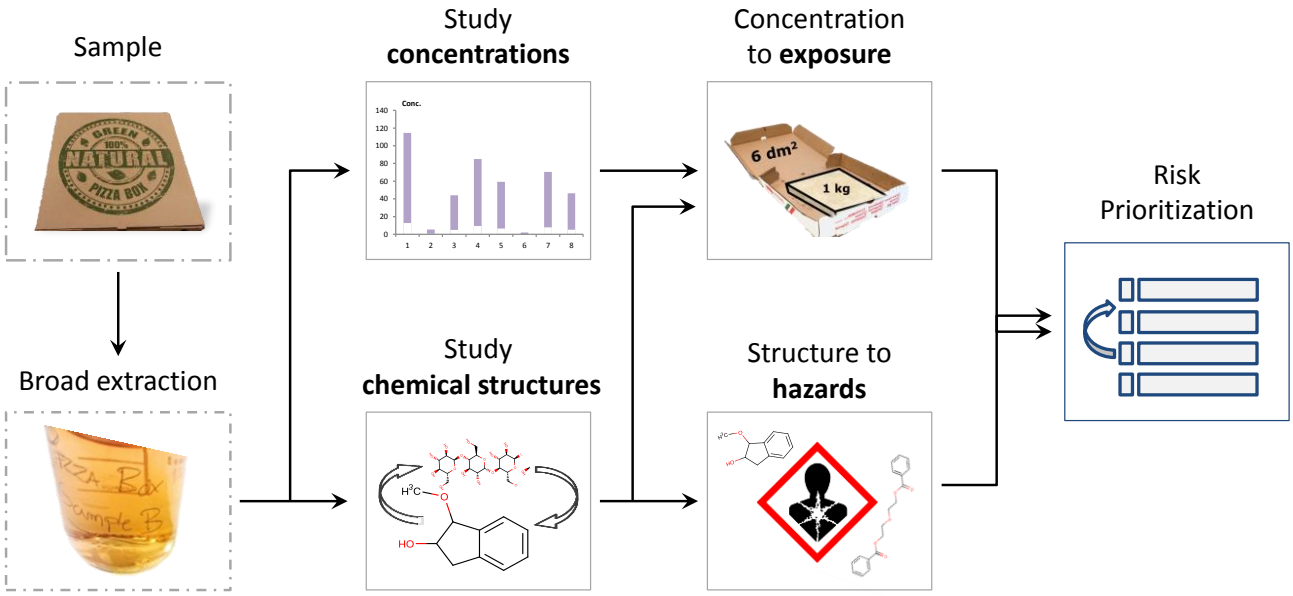


Figure 18: Simplified work scheme for exploration of FCM compounds. First, data on compounds is acquired via untargeted methods like semi-quantification and tentative identification. This data is converted to figures relevant to risk assessment, after which risk prioritization can be performed.

The collection of tentative data, whether for prioritization or knowledge-building, needs to occur in a comprehensive way that excludes as few substances as possible. The migration from paper and board is not due to a single migration mechanism and can therefore vary per application and food type. Hence, it is better to broadly screen the FCM itself for all possible migration chemical compounds and then analyse for all possible types of compounds, e.g., via GC-MS and LC-MS. This provides a broad concept of chemical compounds that can possibly migrate from the FCM to food, but is not necessarily true migration. Consequently, the emphasis of tentative data is to identify chemical compounds in FCM with potential for migration and that may be of risk (Figure 18), rather than to provide an accurate and final figure of risk as a result of predefined migration routes.

Chapter 3

Scientific Research

This chapter briefly summarizes the outcome of the research performed in the course of this project. The main results from the scientific work are shown as supplemented in two published manuscripts and one manuscript in preparation for submission.

3.1. An analytical strategy for risk prioritization

Closing the knowledge gap between occurrence of unknown chemical compounds in FCM and safe usage of packaging materials is a complex undertaking with no simple solution. Due to the size and complexity of the problem: many and diverse materials, rapid changes, complex chemistry, many stakeholders, there is not one method that can resolve the complete picture of risk assessment of unknown compounds known to occur. Consequently, because of the scale of this problem very few public inventories exist of what chemical compounds may be found in the packaging since constructing these would require considerable effort, so it is questionable that it can match the value, but the absence of this knowledge can make regulation and research difficult.

In general both research and regulations are retrospective, first we find a problem, then study it, and finally legislate as a preventive measure. This retrospective nature is problematic because the chemical puzzle represented by FCM is not static: the chemical composition of food contact materials will shift in time as a result of innovation, globalisation, and other changes in the manufacturing process. Since there is not a good overview of what we *currently* find in FCM, changes in the chemical composition will go unnoticed until they show up as a specific chemical threat. This strong focuses on specific compounds that may be harmful challenges the general understanding of the chemical risk of food packaging including combination effects.

As an alternative to the targeted approach investigating specific compounds in food packaging materials, we propose an untargeted screening strategy. This approach is not designed to monitor compliance with legislation, but to build knowledge on chemical risks from FCMs first-hand. It is much more viable to have a greater understanding of FCM prior to designing a comprehensive legislation and efficient to monitor compliance. In the experimental work two essential parts of knowledge-building are considered: first to explore the sample for chemical compounds (Exploration – qualitative and semi-

quantitative), and second to rank chemical compounds based on tentative risk derived from exploration and predictive hazard tools, like toxicological modelling (Risk Prioritization).

The first step of exploration is collecting fundamental chemical knowledge on chemical compounds that are present, e.g., in a sample. FCM exploration consists of a series of methodologies that allow collection of basic but fundamental information on concentration and chemical structure of the chemical compounds found in FCM. Yet, the principle of exploration is that the resulting data does not have a predefined use, so it can be used in a number of further applications. By decoupling exploration from an outcome use, it becomes a valuable tool for screening for unknown chemical compounds that is not necessarily related to a certain field or a certain purpose.

A parallel approach rather than a serial approach is essential for general usable exploration data. The parallel approach requires that the dependency between data is sufficiently reduced so that each experiment can be considered independent. For example, quantification requires knowledge about chemical identity, as quantification is usually performed by comparing the measured signal to a known amount of reference standard. This implies that if structural information is unavailable, there are few to no viable methods for quantification: quantification is dependent on identification. This is discussed in greater detail in Manuscript A and B, see section 3.4 and 3.5. This dependency means part of exploration is simplifying the acquisition of data of undiscovered chemical compounds by eliminating dependencies between different types of data.

The second step is to do a risk prioritization using the information gained from exploration, a setting which is ultimately useful for risk assessment (RA). Here, a number of assumptions is required that may not be valid in all practical applications, but are needed to facilitate a conversion of limited chemical analytical results to provide usable risk predictions. Hence, the risk prioritization that may result from such data is not a definitive assessment of risk, but rather a ranking of possible and perceived chemical risks based on the limited but broad-scope data. In addition, risk prioritization helps to translate the somewhat abstract analytical results of concentration and structure into a tangible format of expected exposure and expected hazards.

Because risk prioritization as discussed here relies on the data of exploration, it inherits the uncertainties from the data exploration. Hence, the challenge here is to interpret the data: what is perception of risk, given all assumptions and uncertainties? Moreover, these uncertainties imply even further that a risk study based on explorative result is not a definitive answer to a question of RA. Instead, risk prioritization in essence making sense of explorative data in a RA setting, where the resulting answers will provide a preliminary estimate of risk.

3.2. Exploration of unknown chemical compounds

3.2.1. Semi-quantification of unknown compounds

In order to quantify unknown compounds, we need to address the core nature of quantification: the relative nature that requires comparing to authentic standards. The prime challenge is that quantification requires the chemical compound to be structurally identified we need authentic standards. Yet, the process of identifying an unknown chemical structure is extremely labour-intensive, subject to availability of chemically pure standards, and requires comprehensive chemical data from several domains. Dealing with a large number of chemical compounds as those extracted from FCM is nearly impossible. This is discussed more in Manuscript B, section 3.5.

The alternative to classical quantification that may work for unknown chemical compounds is a semi-quantifications strategy that exploits a different chemical as surrogate for an authentic standard. Naturally, this requires the measured response factor (RF) of the surrogate chemical to be similar to the quantified chemical in order to obtain reasonable estimations, which would require some compound-dependent data. These surrogate standards, quantification markers (QM), should not be tailored to match the selected analytes of interest (AOI) because this would require some sort of identification of the AOI. Hence, a method was required that could generate AOI and QM matches that were similar in RF without having detailed knowledge about the AOI.

LC-ESI-MS is well known to show significant variation in the RF of different analytes (Espinosa et al. 2015), therefore it is not straightforward to quantify a selected AOI using a

specific QM. A part of RF variation is observed in the typical optimisation studies that are used in LC-MS; here, one selected AOI is typically optimised for maximum response, while this may result in lowered response of other possible analytes. In untargeted screening all analytes have to be considered as possible AOI. Therefore, the responses cannot be maximised for each, but rather using normalised optimisation across the chromatogram. The objective of this optimization is to minimize the impact of different responses between QM and AOI. To achieve a suitable normalisation, careful study of the effects of each instrumental parameter is required. Considering extractable compounds from paper and cardboard materials it was possible to optimise the source conditions to achieve a wide maximisation (compound-wise), and where low-response analytes were favoured over high-response analytes.

Optimising the selection of AOI and QM pairs needs to be based on readily available information from the experiment to ensure it is not limited to AOI that have been identified, but can also be applied to AOI that are not identified. Studies have shown that retention time may be an adequate predictor for response since it is correlated to chemical properties like polarity and hydrophobicity (Cech et al. 2001), while it is also readily available from chromatographic experiments. In addition, the molecular volume may also be an important parameter for ESI (Chalcraft et al. 2009), which is related to the molecular mass for most relevant compounds. We demonstrated that the retention time was a better parameter than accurate mass in AOI and QM pairing, resulting in lower prediction errors and better overall predictions.

By exploiting these concepts, semi-quantification was made possible on almost any detected analyte in FCM extracts as seen the chromatogram from LC-QTOF-MS datamining (Pieke, Granby, et al. 2017). The errors of the predicted concentration were at maximum up to 3-fold error with average around up to 2-fold error. While this is much larger than the errors in traditional quantification used to check compliance, it does not require authentic standards, and all analytes could be quantified in bulk (see Figure 19). Hence, the method provides an estimated concentration in untargeted screening even for a possibly unknown chemical compound. The trade-off for working both untargeted and beyond the traditional methodology is a lower precision and lower accuracy in the concentration estimate.

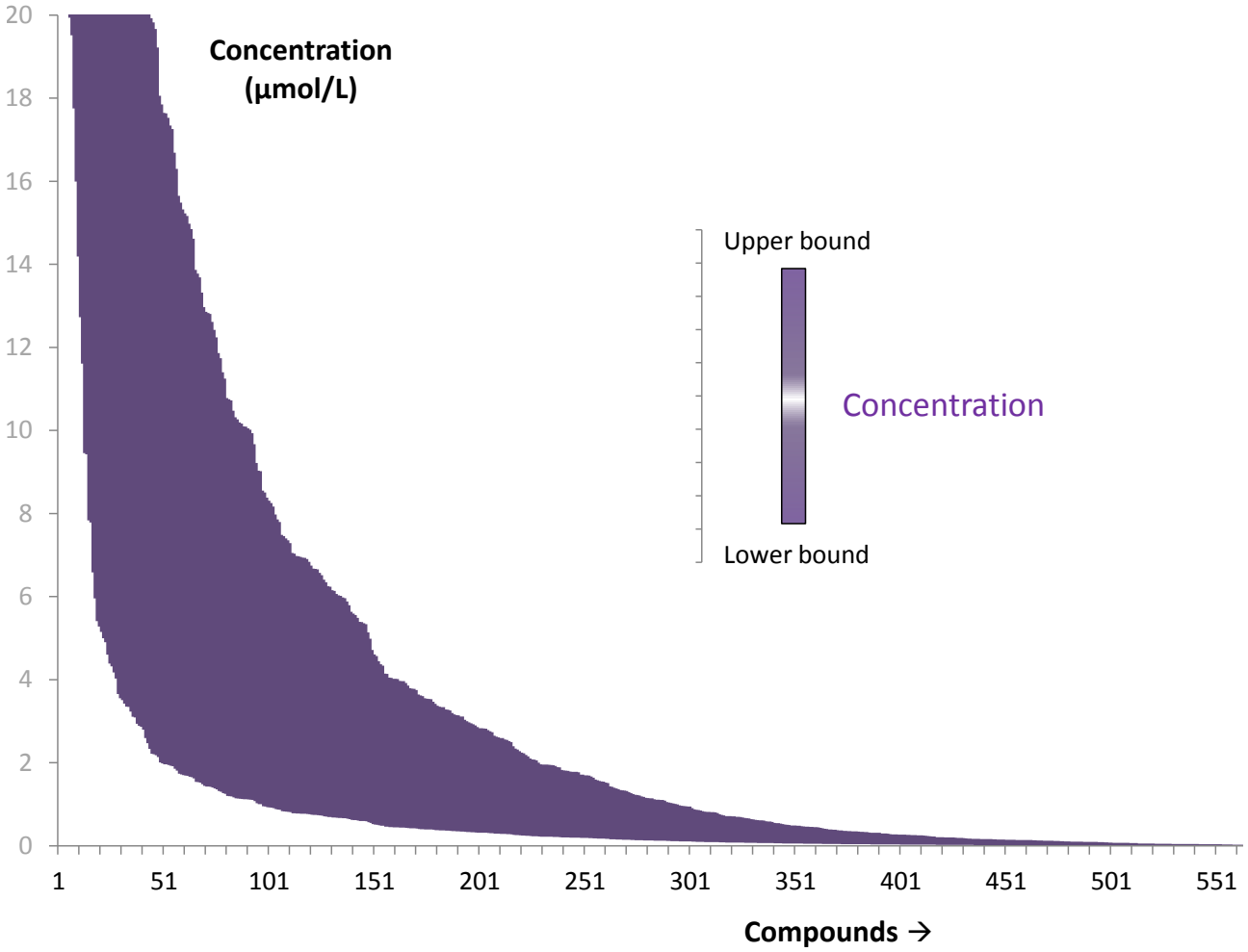


Figure 19: Semi-quantitative results for over 500 compounds originating from a single sample. The concentration of each compound is shown as a bar with lower and upper bound estimation errors set at the maximum observed error of 3-fold from relevant research (Pieke, Granby, et al. 2017).

3.2.2. Chemical identification of unknown compounds

Chemical identification for an unknown compound is often a laborious process that requires well-trained analysts and multiple sources of data, e.g., NMR spectroscopy or mass spectrometry. As the complexity of the molecular structure increases, so does the amount of effort needed to elucidate these. An elucidation of the chemical structure often involves the acquisition of fragmentation spectra from MS/MS experiments, which are relatively easy to obtain when coupled with chromatography. However, MS/MS fragmentation spectra for a chemical compound do not directly provide the chemical structure of the compound. In fact, it is seldom possible to generate a chemical structure

from purely MS/MS because the possible number of structures is much larger than what can be easily derived from an MS/MS spectrum. Hence, reference standards or recorded spectra are often required, which can be problematic to obtain for novel compounds or for large number of compounds. Consequently, there is need for fast and practical analytical strategies that can identify potentially unknown compounds without the need for reference standards.

To translate MS/MS fragmentation data into a structural assessment without comparison to standards is often a human task that requires the expertise of a skilled analyst. Because the number of spectra can quickly exceed what is feasible to assess manually, the expert-based evaluation of each spectrum is impractical, and instead automation of chemical structure elucidation is highly desirable. For this, spectrum interpretation by correlation between measured MS/MS spectra and known or generated fragmentation spectra e.g., by *in silico* procedures, may be useful. By correlating MS/MS spectra on bond breaking, isotopic patterns and spacing, fragmentation, and accurate mass to existing or generated spectra of existing chemical compounds, it may be possible to find the best-matching chemical compound structure to the MS/MS data from a pool of spectra. Here, software that automatically correlates fragmentation patterns to *in silico* bond breaking data for a database of compounds can be used (Hill & Mortishire-Smith 2005).

In order to ensure that chemical identification is applicable to non-target compounds, e.g., those that are not initially expected to be present, the commonly targeted nature of MS/MS strategies needs to be altered. In preference to targeted MS/MS methodology, where the goal of the analysis is defined before the analysis, the use of modern quadrupole time-of-flight (Q-TOF) mass spectrometry equipment to automatically select precursor ions is promising. Here, continuous switching between MS mode to determine the most relevant ion and subsequent MS/MS for obtaining fragmentation spectra of the ion is used in a process called Data-Dependant Acquisition (DDA). The real-time TOF signal of analytes eluting from LC are passed through a DDA algorithm that evaluates each detected *m/z* value for its relevance, i.e., a newly emerging *m/z* or a high-intensity *m/z* are considered relevant. If a certain *m/z* is considered relevant, MS/MS fragmentation spectra are collected. By configuring the DDA, it effectively turns Q-TOF into a semi-targeted analyser that collects MS/MS spectra from emerging chemical compounds separated by the LC.

Correlations and subsequent prediction of molecular structure are better when the mass spectra contain high-quality information. The requirement for high spectral quality is somewhat problematic when combined with DDA, as the most abundant ions commonly selected by DDA do not always produce the highest-quality spectrum. Generally, more and larger fragments improve the quality of the spectrum, whereas fewer and smaller fragments contribute less to quality. The most abundant ions are primarily selected by DDA, but there may be lower-intensity ions, e.g., different salt adducts of the same compounds that produce better-quality spectra. To negate some of these limitations imposed by untargeted screening, it is possible to incorporate a basic level of targeting in algorithm, in the form of a preference list, to aim for better-quality spectra. However, this approach requires prior knowledge about the presence of possible ions and their spectral quality, which can partly be obtained by performing an analysis in MS-only mode and investigating the ion traces.

We studied the concept of using DDA with automatic structure correlation as an untargeted screening method to discover the chemical identify of compounds in FCM (*Manuscript C*). The method provides an untargeted approach to chemical identification without requiring standards or deep pre-existing knowledge, so it is ultimately suited for chemical discovery. However, the approach is in most cases unable to provide an accurate structure prediction; instead, it provides a likely structure which could differ from the actual structure. As a source of chemicals structures for structure correlation, a combination of seven different data sources (DS) were combined into five databases (DB). The requirements for DS are relatively simple: only a chemical structure is needed. This is a novelty compared to common MS/MS databases, which need existing mass spectrometric information that is expensive to obtain for a large number of chemical compounds. By performing the same correlation on different DB, correlative matching can be based on a large pool of expected and unexpected chemical compounds. In summary, the untargeted identification strategy does not produce a comprehensive identification of chemical compounds, but is instead very well-suited for screening new chemical compounds, obtaining predicted chemical structures, exploration analysis, and generating new data on the chemical profile of a sample.

3.3. Risk Prioritization

Exploration provides invaluable data estimates, which makes it a promising tool for discovery of new compounds. However, it is not obvious how explorative data can be utilized for risk assessment (RA). For RA, an exposure assessment and a hazard characterization are needed; yet, the predicted concentration by semi-quantification does not directly relate to human exposure, and the prediction of the chemical structure is not equivalent for a hazard assessment. Consequently, explorative data requires conversion before it can be utilized in risk assessment.

3.3.1. Exposure assessment for risk prioritization

Conversion of semi-quantitative data into exposure requires a number of steps. First, the usage of FCM for food needs to be estimated. These need to be corresponding to what is realistically anticipated contact with food, but in actuality this can vary widely depending on food and FCM used. EFSA proposes the standard to be for an average adult consuming 1 kg of food in contact with 6 dm^2 of plastic food packaging each day, but this has been suggested as being an underestimation and is only used for plastic materials (Bouma et al. 2003). However, the estimates by Bouma et al. for contact of a broad range of packaging, 10–30 dm^2 day^{-1}, and the actual fraction of paper and board in packaging, 10 to 20% (FDA 2007; Duffy et al. 2007), can be used to estimate a paper and board usage. As a result, the calculated exposure based on 10–30 dm^2 day^{-1} paper and board with 10–20% usage is 1–6 dm^2 day^{-1}, which is not substantially different from the 6 dm^2 day^{-1} standard used by EFSA. Second, the measured amount of compound needs to correspond to migration. Essentially, the extraction procedure prior to semi-quantification must be related with anticipated migration. The applied semi-quantitative extraction procedure (full immersion; 24 hr at 40°C) seem similar to the methods used by the U.S. Food and Drug Administration (FDA) for testing uncoated paper (FDA 2007), and also the European Food Safety Authority (EFSA) standard for testing plastic materials in contact with water/oil emulsions (European Parliament and Council of the European Union 2011a). Consequently, the semi-quantified amount in the extract is considered overestimated for compounds with the potential to migrate, but is not as severe as a full chemical extraction of the material.

The exposure estimate can be found by converting the semi-quantified concentration to an intake by including the assumptions presented before: the FCM contact is 6 dm^2 $person^{-1}$ day^{-1} and observed concentrations resulting from the extraction procedure are conceivable, but overestimated. To obtain the intake I in µg $person^{-1}$ day^{-1}, equations (3.1) to (3.3) can be used. In these Equations, C_i is the semi-quantified concentration of analyte i in the extract in µmol $liter^{-1}$, M_i is the molecular mass or accurate mass of molecule i in µg $µmol^{-1}$, V_e is the total volume used for extraction in liter, A_{PB} the surface area of the tested sample in dm^2, and E_S the average contact exposure in dm^2 $person^{-1}$ day^{-1}. (3.1) is the single formula for calculating estimated daily intake from a semi-quantitative experiment. (3.2) is the unit-based formula representation of (3.1). Finally, (3.3) is yielded from simplification of (3.2), showing that estimated intake is the product of the amount of compound per dm^2 contact multiplied by the estimated exposure of contact.

$$\mathrm{I} = \frac{\mathrm{C_i} * \mathrm{M_i} * \mathrm{V_e}}{\mathrm{A_{PB}}} * \mathrm{E_S} \qquad \textbf{(3.1)}$$

$$\mathrm{I} = \frac{\frac{\mu\mathrm{mol}}{\mathrm{liter}} * \frac{\mu\mathrm{g}}{\mu\mathrm{mol}} * \mathrm{liter}}{\mathrm{dm^2}} * \frac{6\ \mathrm{dm^2}}{\mathrm{person} * \mathrm{day}} \qquad \textbf{(3.2)}$$

$$\mathrm{I} = \frac{\mu\mathrm{g}}{\mathrm{dm^2}} * \frac{6\ \mathrm{dm^2}}{\mathrm{person} * \mathrm{day}} \qquad \textbf{(3.3)}$$

3.3.1. Hazard character for risk prioritization

Obtaining a hazard characterization requires conversion or interpretation of the predicted structures. However, manual lookup of hazards, e.g., in literature, for every predicted structure is extremely onerous. Promising tools for generating toxicity information in novel hazard assessments are Quantitative Structure-Activity Relationship (QSAR) models that use statistical tools to relate observed hazards to chemical structural features. Because QSAR models require only an input molecule, they are highly compatible with tentative data available from exploration. QSAR has recently been applied to FCM chemicals in order to prioritize based on hazard, so the application is promising (van Bossuyt et al. 2017). However, QSAR results can vary greatly on the quality, scope, and application of the model. In addition, QSAR models are not available for every possible endpoint and may not provide a dose-dependent relationship. Consequently, the use of QSAR in explorative analysis is well-suited towards highlighting active compounds, but cannot be directly used as a tool to confirm safety.

Here, all predictions obtained for a single compound are entered into a number of QSAR models in order to obtain as much information as possible prior to deciding the possible activity of the molecule. The process is currently limited to three well-investigated endpoints: Carcinogenicity, Mutagenicity, and Reproductive toxicity (known as CMR). In addition, a number of models are used per endpoint to minimize the effect of model outliers and to use consensus evaluation on the average prediction result. The result is a complex data matrix of *n* endpoint predictions by *m* models on *p* molecules, which may ultimately be too difficult to use in light of a hazard characterization. However, the matrix can be simplified by reducing dimensions applying a consensus model on the *m* variant of models. Here, the largest vector of prediction outcome plus prediction strength can be used as an average predictor across the models, which simplifies the hazard characterization. Consequently, the final result is a table of activity probabilities per molecule per endpoint.

3.3.2. Applying risk prioritization based on explorative data

The conversion of semi-quantification to exposure and identification to hazards provides an early setup for a possible risk prioritization, but since the data is explorative (i.e., contains uncertainties) it should be applied with caution. The propagation of uncertainties, possibly low-quality predictions of structures, and strong disagreements between structure predictions or QSAR predictions invalidates the data for direct use in automated processes. Risk prioritization could be considered similar to risk assessment: an expert decision reached by consensus after evaluating all available data. Therefore, while all steps leading up to risk prioritization have been focused strongly on using automated processes, the risk prioritization itself should be mainly a human-based decision. There is insufficient scientific support to perform risk prioritization or otherwise risk assessment by an inflexible systematic approach rather than expert-based.

Despite the inability to automate risk prioritization, decision thresholds for exposure and hazard are still required. The intake and hazard of a chemical compound need to be compared to what constitutes as acceptable or non-acceptable. A concept of considerable merit and possibly applicable to explorative data is the Threshold of Toxicological Concern (TTC) approach (Kroes et al. 2004; EFSA & WHO 2016). This approach is based on the notion that for a newly discovered chemical it is possible to define an exposure limit based on chemical structure, below which there is no foreseeable harm given that the compound is not carcinogenic, mutagenic, bio accumulative, or otherwise a potent toxicant. Hence, the TTC includes both an exposure assessment and a hazard assessment, without providing a definitive conclusion on risk, which makes it suitable to be applied on explorative data.

To facilitate risk prioritization on explorative compound data, a TTC-like decision approach was combined with manual expert assessment (*Manuscript C*). The classification approach follows a stepwise evaluation of all available information from exploration, like semi-quantification and tentative identification, but the final risk profile is decided by an expert assessor. First, the predictions of chemical structures are evaluated for the number and quality of relevant predictions, which predictions make sense chemically, and how each prediction compares to other predictions in order to establish a chemical picture of the most likely chemical structure. When the structure predictions are considered

sufficiently informative, the corresponding QSAR predictions are evaluated for any possible CMR alerts that indicate hazard and which may exempt the compound from exposure assessment, similar to the TTC approach. Following no alerts, the calculated intake is compared to thresholds assigned for each respective Cramer class, gained from a structure assessment.

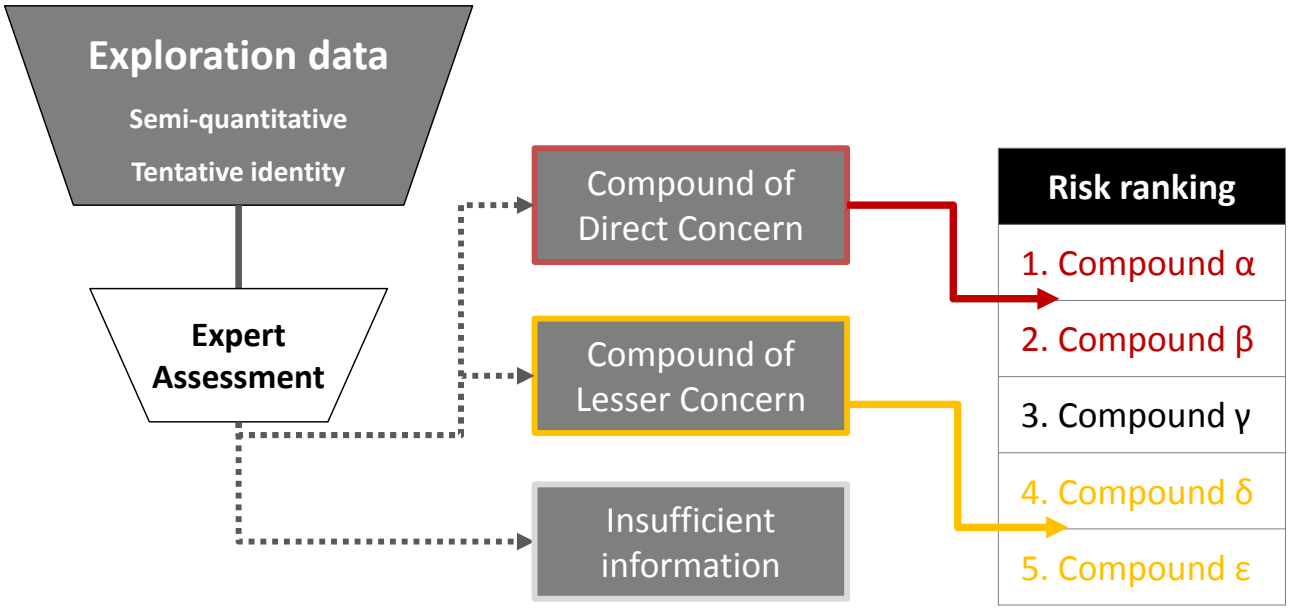


Figure 20: Schematic displaying the workflow from explorative data towards risk ranking, which is used in risk prioritization. Figure is adapted from Manuscript C.

The risk prioritization strategy presented groups the discovered compounds according to the estimated risk based on tentative data. The grouping process classifies chemical compounds, based on both data-driven decisions and expertise-based decisions, into three risk profiles: compounds of direct concern, compounds of lesser concern, and compounds that lack sufficient data (Figure 20). The classification is broad by using only three classes; yet, from a RA point of view this less desirable, most notably by omitting an "in-between" option. However, a broad classification is unavoidable because the uncertainties in the data do not allow for sharp boundaries between the categories. Consequently, the presented risk prioritization approach identifies high-priority and lower-priority chemical compounds without assigning a quantitative risk.

3.4. Manuscript A: Exploration - Semi-quantification

Title:	A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionisation mass spectrometry
Journal:	*Analytica Chimica Acta*, Volume 975, 2017, Pages 30-41, ISSN 0003-2670, https://doi.org/10.1016/j.aca.2017.03.054
Author(s):	Eelco N. Pieke, Kit Granby, Xenia Trier, Jørn Smedsgaard
Keywords:	semi-quantification; untargeted analysis; liquid chromatography-mass spectrometry; electrospray ionisation; method optimisation; screening
Contact:	Eelco N. Pieke: enpi@food.dtu.dk, eelco.pieke@gmail.com
Contributions:	Design of the experimental design, experimental work, data analysis, and writing of the manuscript was done by EP. The overall concept of the study was developed collaboratively by XT and EP. Revision and correction of manuscript proofs was performed by EP, KG, JS, and XT.
Supplement:	Supplementary text and visual material to this manuscript is available in the Appendix, starting at page 147. Supplementary non-text files can be downloaded from: https://doi.org/10.1016/j.aca.2017.03.054

Analytica Chimica Acta 975 (2017) 30–41

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry

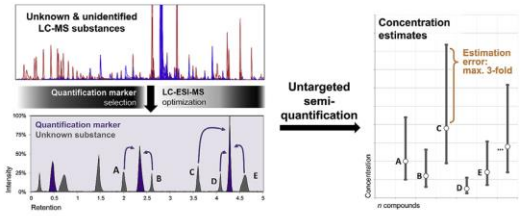
Eelco N. Pieke*, Kit Granby, Xenia Trier [1], Jørn Smedsgaard

Research Group for Analytical Food Chemistry, National Food Institute, Technical University of Denmark, Kemitorvet Building 202, Kgs. Lyngby, DK-2800, Denmark

HIGHLIGHTS

- A semi-quantitative procedure for LC-MS is proposed and evaluated.
- The procedure is optimized for predictions on potentially unknown analytes.
- ESI parameters are optimized for semi-quantitative applications.
- The need of identification prior to quantification is eliminated.
- Untargeted results indicate a maximum of factor 3 prediction error range.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 December 2016
Received in revised form
27 February 2017
Accepted 31 March 2017
Available online 10 April 2017

Keywords:
Semi-quantification
Untargeted analysis
Liquid chromatography-mass spectrometry
Electrospray ionization
Method optimization
Screening

ABSTRACT

Risk assessment of exposure to chemicals from food and other sources rely on quantitative information of the occurrence of these chemicals. As screening analysis is increasingly used, a strategy to semi-quantify unknown or untargeted analytes is required. A proof of concept strategy to semi-quantifying unknown substances in LC-MS was investigated by studying the responses of a chemically diverse marker set of 17 analytes using an experimental design study. Optimal conditions were established using two optimization parameters related to weak-responding compounds and to the overall response. All the 17 selected analytes were semi-quantified using a *different* analyte to assess the quantification performance under various conditions. It was found that source conditions had strong effects on the responses, with the range of low-response signals varying from −80% to over +300% compared to centerpoints. Positive electrospray (ESI+) was found to have more complex source interactions than negative electrospray (ESI-). Choice of quantification marker resulted in better quantification if the retention time difference was minimized (12 out of 12 cases error factor < 4.0) rather than if the accurate mass difference was minimized (7 out of 12 cases error factor < 4.0). Using optimal conditions and retention time selection, semi-quantification in ESI+ (70% quantified, average prediction error factor 2.08) and ESI- (100% quantified, average prediction error factor 1.74) yielded acceptable results for untargeted screening. The method was successfully applied to an extract of food contact material containing over 300 unknown substances. Without identification and authentic standards, the method was able to

* Corresponding author.
E-mail addresses: enpi@food.dtu.dk (E.N. Pieke), kgra@food.dtu.dk (K. Granby), Xenia.Trier@eea.europa.eu (X. Trier), smeds@dtu.dk (J. Smedsgaard).
[1] Present address: The European Environment Agency, Kgs. Nytorv 6, Copenhagen K, DK-1050, Denmark.

http://dx.doi.org/10.1016/j.aca.2017.03.054

estimate the concentration of a virtually unlimited number of compounds thereby providing valuable data to prioritize compounds in risk assessment studies.

1. Introduction

Quantifying unknown substances without available suitable standards seems to be a predicament in most modern analytical protocols. For a large section of analytical methods, dealing with unknown substances is notoriously difficult. Therefore, a large majority of modern analytical methods are performed targeted, selected analytes of interest (AOI) a priori by optimizing methods specifically for these AOI. The targeted analytical methodology is the basis for nearly all quantitative measurements using authentic references or standards, leaving very little room for unknown substances. Consequently, it is not uncommon that unknown substances are avoided or ignored due to the difficulty associated with resolving them. However there is an increasing demand to quantify new emerging compounds, sometime also at low concentration levels even though reference compounds are not available. While performance of modern analytical instruments enables such analyses, the quantification still rely on availability of target compounds as standards. Furthermore, the number of detectable analytes quickly exceeds the practical limit of pure authentic standards that can be included. The consequence is a knowledge gap for dealing with unknown compounds in routine analyses, which is especially true for liquid chromatography – mass spectrometry (LC-MS). Therefore the challenge to deal with un-targeted and unknown compounds, e.g., as needed in food safety [1] or metabolomics [2], demands ever more data from the analytical methods like LC-MS so that it requires re-thinking the analytical approach.

Firstly, the usage of standards for analysis needs revision as the number of detectable compounds in untargeted analyses exceeds the number of quantifiable compounds in targeted analyses. Targeted quantification rely on the relationship between a detector response and the analyte concentration in a standard to provide a response factor (RF), which needs to be established for every AOI. In the situation where not one, but tens or hundreds additional untargeted AOIs are included, the number of standards needed increases equally. It can easily become too costly or simply impossible to obtain a standard for all possible AOI. Furthermore, the variation in chemical structure in untargeted analyses may be much larger than in targeted analyses.

Secondly, in complex samples, hundreds or perhaps thousands of analytes can be present, and not all of these analytes will have been identified. A targeted analysis usually focuses method development on a limited number of substances, therefore not requiring an identification procedure. However, to include newly identified (or unknown) compounds in the method, their RF needs to be known for quantification – this is currently impossible if the structure is unknown. In case of several unidentified compounds, the chemist developing the method has to choose: to allocate resources to identify the compounds, or to allocate resources to obtain the standards for those already identified ignoring all others?

Clearly, given such a choice, the targeted analyses are unsuitable as the method-of-choice, especially in fields that require detailed quantifications. Ideally, a strategy that estimates the concentration of an AOI is preferred. An alternative method applied in gas chromatography (GC) is semi-quantification by comparison to a different analyte [3,4]. Semi-quantification (SQ) relies on the idea that any unknown analyte can be quantified by a different known analyte, albeit with larger uncertainty. To perform SQ, a quantification marker (QM) is used as surrogate standard for an authentic standard AOI. In the simplest application, SQ estimates the concentration of an AOI by using the relative response ratio between QM and AOI. Based on this approach, the magnitude in the prediction error will be directly proportionate to the difference in RF between QM and AOI. Using SQ instead of targeted analyses adds a larger prediction error; however, a larger prediction error is preferable to no data at all, where latter is usually the case when dealing with unknowns in targeted analyses.

Despite the potential of SQ, a number of problems challenge a widespread use in LC-MS. First, the quantification marker (QM) and the target (AOI) need to have the same RF in order to obtain acceptable predictions. However, in LC-MS, the RF of any given analyte varies strongly with the chemical structure, concentration and the various analytical conditions [5–10], resulting in a much wider range of RFs. The large range in response factors require several QM at different RF levels to compare with. To date, some studies have attempted to predict or correlate the RF of analytes [7,11–14]. The results are interesting, but require identification, are often specific to compound classes, or cannot be implemented with ease. Secondly, variations in RF depend on instrument, leading back to why RF models are often not usable. In LC-MS, the electrospray ionization (ESI) interface greatly affects RF via an interplay of electrochemical, physical and physical-chemical processes particularly via suppression and matrix effects [15,16], so that the observed RF may be a function of source design and analyte properties [13,17–19]. Thirdly, optimization studies in ESI aim to maximize the signal of specific AOI or observe a global average maximum for a group of analytes [20–22]. However, favoring the response of one analyte over others increases the difference in RF, and this will increase the prediction error obtained in SQ. Finally, targeted methods generally have narrow chemical scope, mostly on one or a few analytes simultaneously. Since SQ aims to include untargeted compounds, the chemical scope needs to be widened, which means that chromatographic separation needs to be more efficient, and that difficult to detect analytes should have priority in detection. This also requires a different approach than maximization.

This study present a basic proof of concept strategy to develop untargeted semi-quantitative methods based on LC-MS, investigating some of the problems mentioned: ESI source optimization, QM selection, and concentration prediction. An experimental design study was used to examine the effect of ESI parameters on normalization of responses and maximizing the response of low-responsive compounds. Rather than using a single QM as typical in GC, several QM were chosen for semi-quantification to improve predictions. Then, QM selection was investigated via two methods requiring minimal compound information. Finally, semi-quantification was performed under optimized conditions with the best possible QM selection. Overall, the goal of this paper is not to report optimal conditions applicable in any study, but to understand critical system parameters that affect semi-quantification.

2. Materials and methods

2.1. Chemical standards

In order to evaluate the performance of SQ in a wide chemical scope, a range of standards exhibiting diverse chemical properties and functionalities was chosen. The mixtures were designed to mimic difficult samples in this case an extract from crude paper and cardboard, which is chemically diverse and structurally complex, also containing sulfur- or chlorine-atoms [23]. The ESCO WG list [24] was used for inspiration in the selection of compounds. The following chemicals were purchased from Sigma Aldrich in the highest available purity: 2-isopropylthioxanthen-9-one (ITX), 4-chloro-3-methylphenol (p-Chlorocresol), syringaldehyde, abietic acid, benzisothiazolinone, benzophenone, dicyclohexyl phthalate, diethylenetriamine, diphenyl phthalate, distearyl thiodipropionate, ethylenediamine, methyl stearate, octadecyl 3-(3,5-di-*tert*- butyl-4-hydroxyphenyl)propanoate (IRGANOX 1076), octanol, oleic acid, sodium 2-dodecylbenzenesulfonate (SDS), L-threonine, and α'α-trehalose dihydrate (trehalose). Other chemicals obtained were: chlorothalonil (Dr Ehrenstorfer, 98.8%), 2,6- diisopropylnaphthalene (DIPN) (Acros Organics, 99.0%), melamine (EURL standard), Methamidophos (Dr Ehrenstorfer, 98.5%), Flumethrin (Dr Ehrenstorfer), and pentachlorophenol (Dr Ehrenstorfer, 98.5%). All chemicals were used without further purification.

2.2. Solutions

MilliQ grade water (>18.2 MΩ cm resistivity, <5 ppb Total Organic Carbon) was used for LC analysis. Acetonitrile (ACN) and methanol (MeOH) were of high-grade purity suitable for LC-MS (LiChrosolv®, Merck). For MS calibration, an ESI-L Low Concentration Tuning Mix (PN: G1969-85000, Agilent, Santa Clara, CA, USA) was diluted 50 times with water/ACN (90/10). Eluent buffer was prepared by dissolving 9.67 g formic acid (analytical grade, Sigma-Aldrich) and 2.36 g ammonium formate (analytical grade, Sigma-Aldrich) in 500 mL water. Calculated pH of the buffer was approximately 3.1, which was confirmed by pH paper. Eluent A was prepared by dissolving 2.5 mL buffer and 7.5 mL ACN in 490 mL water (approx. 1.5 vol.% ACN in water), and was refreshed bi-daily. Eluent B was pure unmodified ACN.

Analyte stock solutions were prepared in concentrations of approximately 5 or 10 mM in MeOH using an analytical balance, which were further diluted to the required concentration by MeOH. Two test mixture solutions were prepared: one for negative mode analysis and one for positive mode analysis. The ESI+ and ESI− mixes contained 18 and 16 compounds, respectively, in a concentration range of 0.25–1.0 μmol L^{-1}. The exact concentrations and compounds for each solution are available in the Supporting Information (Table A.1).

2.3. Sample preparation

To show analytical concept, an illustrative demonstration case from food contact material was selected. A printed carton box containing dry muesli for breakfast was chosen as demonstration sample obtained from a local retail store (Copenhagen, Denmark). To evaluate substances migrating from the board, crude extracts of the carton box were prepared by soaking 1 dm^2 cutouts in 50v% ethanol/water at 40 °C for 24 h. The extract was filtered over a 0.20 μm Phenex PTFE syringe filter (Phenomenex, Vaerloese, Denmark). 0.5 mL of filtered extract and 1.0 mL of Quantification Marker mixture were mixed in an analytical vial (LC-MS vials, Agilent, Glostrup, Denmark). The vials were capped, shaken, and analyzed in positive mode ESI. Blank extracts were prepared by performing the entire extraction procedure without cardboard cutouts.

2.4. Liquid chromatography

Liquid chromatography was performed on an Agilent Infinity 1260 series LC equipped with a 1290 binary pump, 1260 autosampler with a 900 μL metering device, and 1260 Thermostatted Column Compartment (Agilent, Glostrup, Denmark). Separation was achieved on a Waters ACQUITY UPLC® CSH™ C_{18} column (2.1 × 150 mm, 130 Å, d_p = 1.7 μm), protected by a Waters ACQUITY UPLC® CSH™ C_{18} Van-Guard pre-column (2.1 × 5 mm, 130 Å, d_p = 1.7 μm). 12.5 μL sample was injected for each analysis by the metering device followed by a back-flush with 100% eluent A to the column for 1.5 min at 0.15 mL min^{-1} (18 sample volumes). The metering device was bypassed after 1.5 min and gradient elution started. The column was held isothermal at 60 °C during analysis.

Both the solvent composition and flow were programmed as gradient. Solvent gradient increased eluent B fraction from 0% after an initial isocratic period of 1.5 min to reach 99% B at 28.0 min. Following, fraction B was increased from 99% at 28.0 min to 100% at 36.0 min, and held isocratic 100% from 36.0 min to 37.0 min. To reduce analysis time the eluent flow rate was increased over the course of the gradient: from 1.5 min to 28.0 min the flow linearly increased from 0.15 mL min^{-1} to 0.20 mL min^{-1}, and was further increased to be 0.30 mL min^{-1} at 36.0 min. In the Supporting Information an example of a negative ion chromatogram (Figure A.1) and the gradient elution (Figure A.2) are displayed.

2.5. Mass spectrometry

Mass analysis was performed using an Agilent 6550 Q-TOF mass spectrometer (Agilent, Santa Clara, CA, USA) equipped with Agilent JetStream (AJS) electrospray ionization (ESI) interface. Dry nitrogen was used as sheath- and drying gas. Fragmentor voltage was set to 350 V for all experiments. The settings for capillary voltage, nebulizer pressure, drying gas temperature, sheath gas temperature, drying gas flow, sheath gas flow, and Nozzle Voltage were varied according to experimental design, see Table 1. The Q-TOF MS was auto-tuned for mass accuracy and resolution between *m/z* 100 to *m/z* 1700 prior to each sequence, and mass axes were calibrated using the calibration mixture.

The Q-TOF MS was operated in MS mode between *m/z* 45–800 to achieve higher sensitivity, but broader mass ranges can be applied for different applications. Residual mass error was after post-calibration less than 0.5 ppm on average. For each of the detected compound the sum of most common ionization products: $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, and $[M+K]^+$ as response in ESI+, and $[M-H]^-$ and $[M+COO]^-$ as responds in ESI−, were used. Wherever possible, neutral loss of H_2O and/or CO_2 with possible adducts were included. To facilitate normalization, signals from all

Table 1
Parameters used in the optimization study.

Parameter	Term	Lvl −1	Lvl 0	Lvl +1
Nozzle Voltage (kV)	Noz-V	0	0.5	1
Sheath Gas Temp. (°C)[a]	SHG-T	150	240	350
Sheath Gas Flow (L min^{-1})[a]	SHG-F	4	8	12
Capillary Voltage (kV)	Cap-V	2	3	4
Nebulizer Pressure (psi)	Neb-P	20	35	50
Drying Gas Temp. (°C)[b]	DrG-T	150	210	290
Drying Gas Flow (L min^{-1})[b]	DrG-F	11	14.9	20

[a] Linearly constrained via: { [SHG-T − 50 · SHG-F] ≤ −50 }.
[b] Linearly constrained via: { [DrG-T + 15 · DrG-F] ≤ 505 }.

experiments were divided by the analyte concentration (in μmol L^{-1}) to obtain response factors (RF) in signal per concentration.

2.6. Software

The Marvin software (Marvin 16.10.10, 2016, ChemAxon, http://www.chemaxon.com) was used for drawing, displaying and characterizing chemical structures, substructures and reactions; and the chemical properties of the analytes were predicted by Calculator Plugin(s). MassHunter Acquisition B.06.00 and MassHunter Qualitative B.07.00 (Agilent Technologies) were used for controlling the Q-TOF, data acquisition and data extraction. JMP®, Version 12 (SAS Institute Inc., Cary, NC, 1989–2015) was used for experimental design, data analysis and calculation of Lenth t-ratio. Microsoft Office Excel 2010 was used for analyzing, graphing and tabulating data.

2.7. Experimental design and data analyses

A custom screening design, based on a fractional factorial design of approximately resolution IV, was generated using JMP®. Seven factors were entered as continuous 2-level parameters, shown in Table 1. The design could not be generated independently on the factors since not all combinations of factors are allowed, e.g., a combination of high SHG-T and low SHG-F resulted in an instrumental error. The implications of these constraints are further shown in Supporting Information Figures A.3.1 and A.3.2. To prevent gaps in the experimental design, two linear constraints were added prior to generating the design, shown as footnote in Table 1.

Eight centerpoints (CP) and ten replicates were added to the design, resulting in a total of 56 experiments. The design was then generated from a random seed (885474) with 15,000 starts. The full design matrix is available as Table A.2 in Supporting Information. Each experiment in the design was performed once for each ESI polarity and once for each sample, where a full sequence was performed consequently in one ESI polarity. In total, two times 56 (112) experimental runs were performed. Each experimental condition is noted by EX and a two digit number defined from the run order, e.g., EX15 regards the 15th experimental design run. Within 56 experimental conditions, 39 experimental conditions were considered unique combinations of factors. In between each run, the re-equilibration step (4 min) at the end of the gradient run was also used to equilibrate the source parameters.

3. Results and discussion

3.1. Chromatography

The chromatographic conditions were kept in all experiments to achieve an acceptable separation, as the main focus of this study was not to achieve optimal separation but rather optimal MS detection settings. The markers used in this study were chosen to have a large variety in chemical properties to simulate complex samples, while in addition they were considered useful for their future application as QM. Additionally, with prospect of future applications, achieving a wide elution window with large peak capacity to minimize co-elution was taken into consideration.

First, a simple linear elution program from 0% to 99% organic phase over 20 min was evaluated for both positive and negative ionization. The separation was found workable but did not perform well in the early eluting region and did not elute all compounds within the time allocated. The gradient was improved by increasing in gradient slope length to 26 min and allowing a relatively long time of 10 min at 99–100% organic eluent. This hold time was necessary for the most nonpolar structure, IRGANOX 1076, to elute within gradient time. An example of a chromatogram of the negative mixture is available in Supporting Information Figure A.1. While the separation could be further improved and the analysis time further reduced, this was not considered a priority in this study.

3.2. Response maximization

To demonstrate the effects of maximization in SQ, a maximization strategy was applied to ESI+ data to investigate analyte RF under different ESI conditions. By comparing the results of each experiment to centerpoints (reference analysis) the optimal conditions were investigated. A gain/loss factor was calculated for each analyte in each experiment by subtracting the average RF in centerpoint runs ($n = 8$) from the RF in each experiment. The resulting table contained a gain/loss factor for each analyte, and the overall gains/losses were calculated per experiment by averaging the gain/loss for all analytes, see Table A.3 in Supporting Information.

Experiments where RF changes exceeded five times the centerpoint variability were investigated. In total, eight experiments exceeded the gain threshold (>+50%). The largest increases were found in the conditions used in Experiment 20 (EX20), +131%, followed by EX40, +106%. EX20 altered the response range of analytes from −44% to +355% compared to the centerpoint conditions, and two out of ten RF were lowered. EX40 showed response changes from −71% to +322%, and similarly two RFs were diminished. EX20 performed better than EX40 in two ways: the gains were generally higher, and the loss for two analytes was smaller, so EX20 was initially considered as the most suitable condition for maximization. However, a deeper look on the data in EX20 reveals that RF decreases coincide with the compounds already having the lowest response, hence that under EX20 conditions these compounds will be harder to detect (already difficult under centerpoint conditions). Furthermore, EX20 only enhanced other low-responding compounds slightly. In contrast, EX40 expresses responses that are much more desirable for the weakest-responding analytes, +80% on average. As a result, EX40 is actually more suitable than EX20 for detecting low-responding compounds, but the maximization strategy could not reveal this. The implications of choice of optimal conditions to semi-quantification are quite pronounced: worst-case semi-quantification by using a high-response QM (dicyclohexyl phthalate) to quantify a low-response AOI (α,α-trehalose), the result in EX20 is at least a factor 100 underestimation, whereas using EX40 will give “only” a factor 36 underestimation.

3.3. Response normalization

Response normalization was investigated as possible alternative to response maximization. Normalization was considered successful if the response differences between low- and high-response analytes were minimized, but also if the response of weak-responding analytes could be increased. To evaluate the specific response of low- and high-response analytes, two new variables were defined for optimization: S_{Low} (S_L) as the response of the weak response analytes and S_{High} (S_H) for high response analytes. The variables were calculated by taking the product of the n lowest or highest responding compounds, ($S_1 * S_2 * \dots * S_n$), where n was chosen to represent 25% of the compound total ($n = 3$), so that approx. 50% of the compounds were congregated in either S_L or S_H. This aggregation resulted in a lowest/highest response variable that was considered a dynamic selection, whereas pre-selecting the compounds in either S_L or S_H is a static approach that may exclude relevant analytes if the response substantially changes. Furthermore, S_L and S_H were calculated as the relative overall gain ($S_1 * S_2 *$

S_3) in preference to the absolute gain ($S_1 + S_2 + S_3$), because the absolute gain is biased towards the largest values, while masking small but significant effects.

In order to measure the degree of normalization obtained, a new parameter was defined: the normalization ratio (N_i). As previously discussed, a normalized response is obtained when the differences in RF across different chemical compounds are minimized. So, the normalization ratio could be considered the ratio between the dataset extremes, i.e., highest and lowest RF. We can therefore write the normalization ratio as $N_i = S_L/S_H$. Also, at perfectly normalized conditions $S_L = S_H$, therefore the normalization ratio assumes unity. Consequently, the optimization endpoint can be defined as $\mathbf{N_i} = \mathbf{1}$ and $\mathbf{S_L} = \mathbf{max}$. Maximized S_L corresponds with larger RF for low responding analytes, and closer to unity N_i corresponds with smaller difference between RF. Since the definition of N_i limits its maximum value to unity, maximizing N_i and S_L simultaneously results in optimal normalized conditions. Under these conditions, the RF for low-responding analytes is increased while the RF difference is minimized.

S_L, S_H, and N_i were calculated for each experiment in both negative and positive mode. To identify the experimental conditions that gave the most improved results, S_L was plotted against N_i to illustrate the effect on the normalization and the maximization of low-response compounds. Fig. 1 shows values of N_i (x-axis, scale 10^{-4}) versus of S_L (y-axis, scale 10^{18}) for positive mode ESI (ESI+, Fig. 1a) and negative mode ESI (ESI−, Fig. 1b). Note that the scale of the data is enlarged due to the product-based (S_1 *S_2 * … * S_n) calculations. It is clearly observed that most experimental conditions performed inadequately seen as a “cloud” near the axial intersection of experiments with low S_L and low N_i. For each polarity, the most extreme effects and desirable conditions were identified by their position on the plot in Fig. 1. A large S_L is preferred to improve the signal of the weakest compounds, and a large N_i is preferred to achieve more equal RF between the low- and high-response compounds. In addition, ESI+ and ESI− had different plots, so the data was treated per polarity.

In ESI+, there was a gap in the data distribution: the experimental conditions appeared to favor either N_i or S_L. For example, EX40/EX47 show increased S_L indicating that low-responding compounds have higher RF, yet provides little gain in the normalization N_i. Under these conditions, the gain in RF from centerpoint (CP) conditions is +46% for L-threonine and +80% for trehalose. In contrast, EX37/EX46 result in increased N_i, but do not improve S_L. Under EX37 conditions, the response deviation from centerpoint conditions was −47% for L-threonine and −15% for trehalose. The improvement in N_i provides a potentially better quantification, but at the cost of total signal intensity (sensitivity). However, the signal losses in EX37/EX46 were worse than the centerpoint, but still improved over most conditions, e.g., the losses for EX22 were −71% for L-threonine and −80% for threhalose. There were no experimental conditions that favored both S_L and N_i. Beside those mentioned here, few experiments in ESI+ stood out as being potentially useful.

In ESI−, some experimental conditions provided more desirable results, i.e., large values for N_i and S_L. EX40/EX47 provides large values for S_L similarly to ESI+. In EX40, the gain in RF compared to centerpoint for low-response analytes is a +147% for L-threonine and +335% for syringaldehyde, indicating a large gain. However, particularly interesting responses are also observed under the conditions in EX10/EX48, EX54, but also EX56. These experimental conditions seem to provide large increase in values for *both* N_i and S_L. Furthermore, very few experiments reach similar normalization as EX10/EX48 does. In EX10, the response was +111% for L-threonine and +180% for syringaldehyde compared to centerpoint conditions. Under EX10 conditions, low-response analytes are easier to detect therefore a lower semi-quantification error was expected due to the improved N_i.

3.4. Statistical analysis on effect of ESI source parameters

In order to better understand the source effects on the responses N_i and S_L, a statistical factor analysis was performed to obtain a quantitative comparison, as source parameters are critical to successful method optimization, perhaps even more so in SQ than in targeted analyses. A Response Screening Analysis (RSA, 2^n-model) was used to compare seven instrumental variables as factors X (see Table 1) to experimental responses using four different response types: positive mode S_L ($+S_L$), positive mode N_i ($+N_i$), negative mode S_L ($-S_L$), and negative mode N_i ($-N_i$). The experimental design space allowed full explorative analysis of 1st order factors and exploration on most of the 2nd order interactions in-between factors. Hence, in the RSA the factors interactions investigated were: first order (A); most second order (A*B); some quadratic interactions (A*A); and few third order interactions (A*B*C; A*A*B; or A^3). For this study, the investigation of first and second order alone was considered sufficient, as third order parameters generally express only few significant effects [13,25].

The output of the RSA was a significance level per factor, expressed as the Lenth t-ratio [26,27]. The Lenth t-ratio (LtR) is comparable to the p-value, which is similarly based on probability figures. Here, an LtR of 2.0 approximated a p-value of 0.05, whereas larger LtR indicated lower p-values. Yet, in contrast to p-values the LtR value has both a magnitude and sign. The LtR magnitude correlates to the significance: larger is more significant. The LtR sign (+ or −) indicates whether the effect was positive, i.e., the response variable increased with higher parameter level, or a negative effect, i.e., response variable decreased with higher parameter level.

Fig. 2 shows the parameter-wise comparison of LtR values for all 1st order parameters for all four responses. Out of the seven 1st order parameters, six were significant in $+S_L$, five in $+N_i$, five in $-S_L$, and two in $-N_i$. Beside the 1st order interactions, most of the interactions between 1st order parameters, i.e. 2nd order effects were also investigated. The full dataset on 2nd order and quadratic interactions is available per response type in Supporting Information Table A.5.1 to A.5.4. Overall, the 2nd order effects exhibited less significant effects than 1st order effects, yet a few 2nd order factors were strongly significant. For each source parameter, relevant effects and suggested explanations are given below.

3.4.1. Nozzle voltage (Noz-V)

Initial single factor analysis revealed that the Noz-V was a critical parameter in optimizing both S_L and N_i. Notably, exclusively positive and large LtR values (>+6) for all responses were observed. Furthermore, Noz-V showed several second-order interactions. Both in ESI+ and ESI− for S_L there was an interaction between Nozzle Voltage and Sheath Gas Temperature (Noz-V * ShG-T), indicating that S_L is optimal when both ShG-T and Noz-V are at high settings. However, the (Noz-V * ShG-T) interaction in ESI+ normalization ($+N_i$ LtR = −5) resulted that the normalization is significantly penalized when S_L is maximized. Consequently, the strong increase of N_i at higher Nozzle Voltage suggests that S_L increases stronger than S_H, i.e., the low-response analytes are increased more than high-response analytes.

The exact mechanism behind the effect of Nozzle Voltage on normalization is not yet well understood. As both S_L and N_i increased, S_L increased more than S_H. It was proposed that the compound types that are typically among the low-RF: often small, (semi-)polar molecules, are expected to ionize poorly in ESI because of their tendency to favor the less-charged core of droplets, or because of the solvent composition, where mixtures of

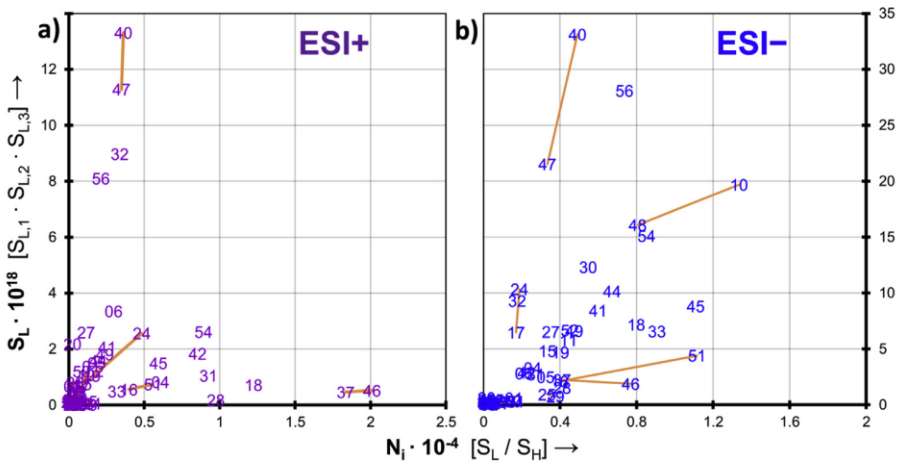


Fig. 1. Experimental values for normalization (N_i) plotted against the lowest signals (S_L) for a) ESI+ and b) ESI−. Values are plotted as their respective experiment number. Ten duplicates are marked with a connection line (orange) illustrating the precision of the experiment. Note that axes for ESI+ and ESI− are not identically scaled. The ideal situation for normalization is where both N_i and S_L are large positive values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

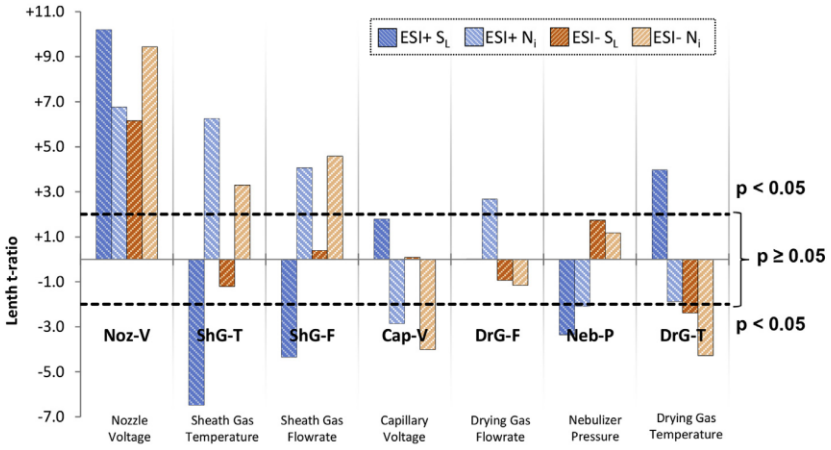


Fig. 2. Lenth t-ratio (LtR) values demonstrating the significance of 1st order parameters obtained from the experimental screening design. For each combination of ESI polarity and optimization criterion one LtR value is shown. LtR values of |LtR| > 2 are considered to have significant (p-value < 0.05) influence on the response factors in S_L and N_i.

chromatographic eluent (acetonitrile or methanol with water) with larger water content lower the signal intensity [28]. Necessity of the Nozzle Voltage follows from high-flow ESI designs, which require intensive gas drying and can generate an enveloping heated curtain of gas, which can isolate the electrostatic field by a kinetic barrier [29]. The Nozzle Voltage may overcome this barrier by increasing the electric field at the spray tip inside of the curtain, and may assist overall ionization by providing more highly-charged droplets [30].

3.4.2. Sheath Gas Temperature (ShG-T)

Sheath Gas Temperature was a significant parameter and forced a compromise between N_i and S_L. As shown in Fig. 2, a high temperature was favorable for improving low-level responses (S_L), while a low temperature favorable for improving normalization (N_i). This means that the ShG-T setting will improve *either* N_i or S_L, but not both, which assists explaining the dichotomy between N_i and S_L in Fig. 1. However, the LtR values for ShG-T were smaller in ESI- than in ESI+. As consequence, Sheath Gas Temperature governed the choice between maximization and normalization in ESI+, but not in ESI- due to the lower significance.

The most probable mechanisms underlying these data are that increasing the temperature will increase response of all analytes, but that of high-RF more strongly, so effectively this directs towards maximization. The second-order effects displayed a similar pattern. Firstly, there is a strong interaction between sheath gas flow and temperature (ShG-F and ShG-T), which originates from their interdependency. However, the interactions ShG-F*ShG-T in $+N_i$ (+2.04) and S_L (+6.35) reveal that lowering both temperature and flow simultaneously greatly reduces N_i and S_L. This may be obvious as when both flow and temperature are lowered the solvent drying capacity drops sharply, so only few ions are generated.

3.4.3. Sheath Gas Flow (ShG-F)

The sheath gas flow had a moderate impact on results. Its effect was significant in 3 out of 4 responses, and considered to be suitable for optimization studies. However, the profile of response

alterations appears identical to the Sheath Gas Temperature and, due to the strong dependency of flow and temperature, separating these two parameters as truly unique may not be desirable. Similarly to ShG-T, low-level responses (S_L) are optimized at high settings, but this subsequently lowers the normalization ratio N_i. Apart from the interaction with Sheath Gas Temperature and Nozzle Voltage, there appear to be very few consistent interactions with other parameters.

3.4.4. Capillary Voltage (Cap-V)

The capillary voltage had moderate effect on responses. Data suggested that increasing the capillary voltage negatively affected the ion generation, therefore reducing the responses of all analytes. This was deducted from the fact that S_L is lowered as function of increasing Cap-V, while the normalization (N_i) is increased slightly, so that indicates S_H is penalized stronger than S_L. However, the effect of capillary voltage was significant but not decisive, and in these experiments increasing the Cap-V to 4000 V resulted in only minor losses.

Beside first order effects, only few secondary interactions were significant for the capillary voltage. A second order interaction was significant between the Drying Gas Temperature (DrG-T) and the Capillary Voltage, but it was not significant for every type of response. This interaction might be due to parameters having opposite effects: the drying gas works as a repulsive force for the MS, i.e., removing neutrals, solvents, but also ions, whereas the capillary voltage draws ions towards the MS. If the drying gas prevents ions from entering the MS, a higher capillary voltage may overcome this.

3.4.5. Nebulizer Pressure (Neb-P)

The nebulizer pressure had minor effect. Uniquely to this parameter, the responses were related the ESI polarity. As a result, positive ESI favored a low pressure and negative ESI a higher pressure. A surprising significant second order interaction of Nebulizer Pressure was with the Nozzle Voltage in positive ESI mode: it produced a significant negative LtR between Noz-V and Neb-P, so that these two settings should not be maximized simultaneously. From a practical point of view, however, maximizing Neb-P reduces performance in ESI+, so the interaction only reinforces the need for a low Neb-P. In addition, this interaction was absent in ESI-, so no compromise was required.

3.4.6. Drying Gas Temperature (DrG-T)

The drying gas temperature had a minor effect. In general, a low temperature was optimal, since higher temperatures reduced responses. Only $+N_i$ seemed to truly be linked DrG-T but was attributed to thermal stability effect of some of the compounds, mostly the phthalates. The chemical degradation of high-response analytes, in particular the phthalates, caused S_H to drop sharply, causing N_i to increase. The effect is unfavorable due to its unpredictability, and the DrG-T should be minimized in all cases. Furthermore, even quadratic interactions were found significant for Drying Gas Temperature, further stressing the need to reduce the temperature to retain acceptable responses.

3.4.7. Drying Gas Flow (DrG-F)

The drying gas flow was found to be the least important parameter, exerting very little effect on the responses. DrG-F showed a preference for high values in order to improve $+S_L$, probably by improving the desolvation of the polar analytes, but this effect was not observed in ESI-. Surprisingly, no interaction between the drying gas flow and drying gas temperature was discovered. For future optimization studies, it can be proposed to omit this parameter entirely.

3.5. Selection of suitable quantification markers

The selection of proper QMs for screening analysis may be as critical as operating under optimal source conditions. Optimal results are obtained when the QM mimic the properties of the AOI as closely as possible, so that the response variations are likely to be minimized. In practice, the desired properties of a QM are similar to those of an Internal Standard (IS). Ideally, QM should not already be present in the sample (overlap), should have suitable stability, spread out over the retention time range, and should accurately represent the sample or compounds of interest. In the case of unknown analyses, overlap is especially important to avoid, as it voids the use of the QM. In addition, if a compound is present in both sample and as QM, it cannot be directly quantified besides being invalid as QM.

However, the major analytical challenges are targeted screening, i.e., looking for groups of compounds that may have significant structural differences, or where analytes of interest are data-dependent, as it is impossible to a priori choose a perfect QM for every AOI. Regardless of the a priori sample knowledge, finding compounds for the QM representing the AOI becomes increasingly difficult as the variation of compounds to investigate increases. More QM can be used in case of a complex spectrum of AOI, but care should be taken not to overload the chromatogram with QM peaks and to avoid overlapping a QM with an AOI present. Here, the best choice of QMs should be made based on the knowledge of the sample combined with a qualitative screening analysis. With sufficient sample knowledge, the QMs may be selected to be similar to the AOI, e.g., isomers of existing substances, phased-out substances, or isotope labelled substances. These are all viable options if the natures of the AOI are known beforehand, e.g., in phthalate or pesticide analysis.

Yet, when the target AOIs are entirely unknown or undecided beforehand, which is often the case in exploratory analyses; the selection of QM requires some extra considerations. Here, the QM profile should be tailored to the sample and not necessarily towards the AOI, since the latter are unknown. For example, in the case used here, the QMs were selected based on their known occurrence in paper and board food contact materials. Available databases were consulted to obtain an overview of possible AOI present [24]. In general suitable QM can be selected for various AOI groups (e.g., biocides, natural products, additives), by selecting substances that are analogous, phased out, or that have very low migration limits/occurrence, so that the probability of overlap is minimized. This type of selection may not achieve good quantification for all substances however this is the challenge of working in an unknown space. A similar approach could also be applied to other samples, e.g., plant extracts, or environmental samples.

Choosing QM in cases where the target AOIs are not known beforehand requires considerable more work, but the QM will then have a wider scope. Once a set of QM is established, it can be used for any similar samples, and it allows for a quick screening of similar samples to the original samples. In addition, using an established set of QM provide several advantages. First, the QM set may serve as system performance check: shifts in retention times and ion intensity are quickly observed. In addition, it may highlight problems like ion suppression in ESI. The usage of a QM set for each sample subtype is not unrealistic, and allows for reproducibility across both labs and instruments. In light of this method's applicability, extra research regarding QM selection in various applications would be valuable.

3.6. Semi-quantification performance

Semi-quantification is mostly used in cases where limited a

priori knowledge on compound response is available, standards not available or there are too many compounds. Using the optimized source settings semi-quantification of selected AOI was performed to demonstrate the feasibility. As the structural variety and the range of RF in many real-world analyses are wide, it is unlikely that a single quantification marker, QM, can provide good quantification for all types of compounds. Hence, several QMs with varying physical-chemical properties need to be selected instead of a single QM as it more likely to find suitable QM for relevant AOI. In the worst-case scenario, the only information available for an AOI is the retention time (RT) and molecular mass or accurate molecular mass (AM). Either of these two parameters can be used to perform adequate quantifications alone, however semi-quantitative analysis by LC-MS can be achieved by the method proposed here. In these cases the selection of QM can be made significantly less strict, based only on RT or AM similarity, rather than complex physical-chemical properties. Note that for ease *maximization* in this paragraph refers to *maximizing* S_L, not to traditional optimization studies.

For each compound, AOI, a quantification marker QM was selected using either minimum difference in RT, ΔRT or in mass, ΔAM. Semi-quantification was achieved by assuming that the selected QM has the same RF as the AOI. To demonstrate the performance of semi-quantification, four case studies were performed using the data from the optimization experiments. Based on the response normalization (3.3), EX40 and EX47 were chosen as maximization candidates in ESI− and ESI+ respectively, and the choices for normalization were EX37 in ESI+ and EX10 in ESI−. Table 2 shows case-study results for three indicatively quantified compounds: one early eluting, one late eluting, and one center eluting compound. In each ionization polarity, all compounds were quantified using the nearest neighbor of either RT or AM. For every AOI, a suitable QM was selected by minimizing retention time (ΔRT) or mass difference (ΔAM). Additionally, a scoring factor was used which was derived from z-scoring the errors in predictions; the derivation of the score is described in Section A.3 in the Supporting Information. The scoring factor was designed to penalize concentration predictions exceeding factor 5 ($AOI_{predicted}$ <0.20 or >5.00) and strongly penalize concentration predictions that exceeded factor 10 ($AOI_{predicted}$ <0.10 or >10.0). In summary, scores above 80 indicate very good predictions, between 70 and 25 adequate, while scores below 10 indicate inadequate performance. Scores of 0 indicated very bad quantification performances.

Minimizing ΔRT (Table 2, right column set) gave adequate results in all selected cases with an average score of 49.5, and no score below 10. In ΔRT total, 12 out of 12 cases were quantified with scores >20 (factor < 4.00 errors). On average, the over- or under-estimation of AOI concentration is approximately factor 2 via ΔRT selection. For individual cases, the best-case prediction provided an error of only 6% (score 93.8), and the worst-case prediction had factor 3 error (score 21.4). In contrast, the selection of QM based on minimum mass difference ΔAM (Table 2, left column set) appeared to be hit or miss: the average score is acceptable (44.7), but many quantification attempts score zero. In ΔAM total, 7 out of 12 cases were quantified with scores >20. Best-case prediction was only 4% error (score 96.1), but the worst cases ranged from factor 12.5 (score 2.8) to a factor 25 error (score 0.0). In AM selection, the results are combinations of good and unacceptable quantification errors, which is undesirable in a setting where consistent predictions are required. Concluding, the selection of AOI and QM should be made on the minimal retention time difference, which is also indirectly supported by studies on similar topics [6,31,32].

As the RT based selection is a viable semi-quantification strategy, the effects on quantification of operating under (S_L) maximized and N_i normalized conditions were investigated and compared to the center points. Table 3 displays the predicted AOI concentration factor and the quantification score for ESI+ under the three different conditions. Notably, ESI+ quantifications benefit only slightly from moving from maximized RF (EX47) to normalized RF (EX40). Further predictions do not become significantly better as seen by no change in quantification scores. In addition, the scores under either normalized or maximized conditions are not very dissimilar from the center point scores, so the optimization seems to have little effect on the quantification results. This may be related to a large number of compromises in ESI+ due to conflicting effects of source parameter. As a result, normalization (EX37) seems to provide the widest range of compounds to be quantified with acceptable errors, but it suffers from a partly worse sensitivity for S_L (Fig. 1A). In perspective, the quantification performance in S_L maximization (EX47) is acceptable in most but a few cases, and would give good sensitivity for low-responses analytes. Based on these data, it makes more sense to choose EX47 as optimum condition to benefit maximum from the extra sensitivity. When using EX47 as benchmark, 7 out of 10 compounds are quantified within factor 10 prediction error and the average prediction error factor is 2.08. Extrapolating these findings it is expected that at least 70% of compounds can be estimated within factor 2 prediction errors. Better results may be possible with a different marker set, as melamine and IRGANOX 1076 were not quantified properly by any of the experiments. In case of melamine, it was found to be due to a co-elution at the peak maximum, and IRGANOX was not detected

Table 2
Case-based results for marker selection in semi-quantification. The score parameter was calculated using the z-score of the prediction errors (see Section A.3 in the Supporting Information), and represents the overall acceptability of the semi-quantification, of which >80 is excellent, 70 to 30 is good, and <10 is inadequate.

Case	Mode	AOI selected	ΔAM AOI_{pred}[a]	ΔAM Score	ΔRT AOI_{pred}[a]	ΔRT Score
Max. *(EX47)*	ESI+	Benzisothiazolinone	1.04	96.1	2.56	35.8
		2-isopropylthioxanthen-9-one	26.17	0.0	1.97	48.3
		L-threonine	0.04	0.0	0.47	44.5
Norm. *(EX37)*	ESI+	Benzisothiazolinone	0.87	86.7	2.07	45.7
		2-isopropylthioxanthen-9-one	1.12	88.8	0.51	48.7
		L-threonine	0.04	0.0	0.40	37.2
Max. *(EX40)*	ESI−	Benzisothiazolinone	2.04	46.4	1.06	93.8
		Oleic acid	0.82	81.3	0.38	34.6
		α,α-Trehalose	0.08	2.8	3.95	21.4
Norm. *(EX10)*	ESI−	Benzisothiazolinone	1.34	73.4	1.17	84.9
		Oleic acid	1.66	58.3	0.69	66.9
		α,α-Trehalose	0.08	2.8	2.76	32.9

[a] Predicted concentration of AOI under condition concentration[AOI]true = 1.00.

Table 3
Semi-quantification results in ESI+ for all analytes. Results are shown as AOI concentration predicted when the true AOI concentration is 1.00, and the prediction score is given in brackets. Factors and scores are displayed for centerpoint (CP) conditions (EX52), normalized conditions (EX37), and S_L maximized conditions (EX47).

AOI in ESI+	QM selected	CP (EX52)	Norm (EX37)	Max (EX47)
L-Threonine	α,α-Trehalose	0.51 (48)	0.40 (37)	0.47 (45)
Benzisothiazolone	Octanol	1.97 (48)	2.07 (46)	2.56 (36)
Melamine	L-Threonine	89.3 (0)	27.8 (0)	26.7 (0)
Octanol	Methamidophos	0.25 (21)	0.42 (39)	0.41 (37)
Methamidophos	Octanol	4.06 (21)	2.37 (39)	2.47 (37)
Benzophenone	Diphenyl phthalate	0.10 (6)	0.12 (7)	0.89 (88)
ITX[a]	Dicyclohexyl phthalate	0.87 (86)	0.51 (49)	1.97 (48)
Diphenyl phthalate	ITX[a]	0.57 (54)	0.89 (89)	0.04 (0)
Dicyclohexyl phthalate	ITX[a]	1.15 (86)	1.95 (49)	0.51 (48)
α,α-Trehalose	L-Threonine	1.97 (48)	2.48 (37)	2.11 (45)
IRGANOX 1076	Dicyclohexyl phthalate	0.00 (0)	0.01 (0)	0.00 (0)

[a] 2-isopropylthioxanthen-9-one.

Table 4
Semi-quantification results in ESI- for all analytes. Results are shown as AOI concentration predicted when the true AOI concentration is 1.00, and the prediction score is given in brackets. Factors and scores are displayed for centerpoint (CP) conditions (EX52), normalized conditions (EX10), and S_L maximized conditions (EX40).

AOI in ESI-	QM selected	CP (EX52)	Norm (EX10)	Max (EX40)
L-Threonine	α,α-Trehalose	0.29 (25)	0.36 (33)	0.25 (21)
Benzisothiazolone	Syringaldehyde	2.05 (46)	1.17 (85)	1.06 (94)
Abietic acid	Oleic acid	2.62 (35)	1.46 (67)	2.64 (35)
SDS[a]	Pentachlorophenol	0.20 (16)	0.22 (18)	0.34 (31)
Chlorothalonil	p-Chlorocresol	2.13 (44)	1.38 (71)	4.19 (20)
Oleic acid	Abietic acid	0.38 (35)	0.69 (67)	0.38 (35)
p-Chlorocresol	Chlorothalonil	0.47 (44)	0.73 (71)	0.24 (20)
Pentachlorophenol	SDS[a]	4.93 (16)	4.50 (18)	2.93 (31)
Syringaldehyde	Benzisothiazolone	0.49 (46)	0.86 (85)	0.94 (94)
α,α-Trehalose	L-Threonine	3.44 (25)	2.76 (33)	3.96 (21)
IRGANOX 1076	Oleic acid	1.58 (61)	1.17 (85)	0.94 (94)

[a] Sodium 2-dodecylbenzenesulfonate.

stably and is more suitable as ESI- compound.

Table 4 displays the predicted AOI concentration factor and the quantification score for ESI- under the three different conditions. The small differences observed in ESI+ were not found in ESI-. Instead, the normalized condition in ESI- (EX10) is a significant improvement over both the centerpoint (EX52) and the maximized conditions (EX40) as seen in Table 4. In almost all cases, EX10 performs superior on the quantification part. Unlike ESI+, moving to normalized conditions in ESI- will improve the quantification scores without affecting the sensitivity too much (Fig. 1B). The strong improvement in scores in EX10 makes it the best possible choice for conditions. Using EX10, 10 out of 10 compounds are quantified within factor 5 prediction errors, while the average prediction error is below a factor 1.75 hence most compounds can be estimated with less than factor 2 error in prediction. This is significantly better than ESI+, probably due to a much more complex mechanisms in ESI+, as well the chemical stability that may be an issue with phthalates in ESI+.

Maximizing the normalization factor N_i is not always optimal for semi-quantification, as it may result in lowering S_L hence lowering the overall sensitivity of the method. As an example, the RF of L-threonine (a generally low-response compound) was reduced in ESI+ by 60% going from EX47 to EX37, however the quantification score was increased from 37.2 to 45.3. The similar change in ESI− resulted only in a 15% loss in L-threonine RF from EX40 to EX10, but an increase in score from 21.4 to 32.9. Indeed, when N_i is improved at the cost of S_L, a choice between sensitivity and quantification accuracy is evident. Defining a limit at which the loss or gain in sensitivity exceeds the importance of quantification is beyond the scope of this work. In general, if the detection limit is sufficient, it might be preferable to improve the quantification by optimizing N_i. However, reflecting back on the fields that could potentially use such methodology, extra sensitivity for weak-responding compounds is a welcome feature. Hence, in this case, the optimal conditions selected were EX10 in ESI- and EX47 in ESI+.

Finally, whether to use semi-quantification or traditional targeted quantification with authentic standards is outside the scope of this paper: the choice of suitable method depends strongly on the application. When using standard maximization for untargeted analyses the semi-quantification results are most likely unreliable, because the system was not optimized to work on a chemical diverse range of compounds. There is a clear need to optimize systems to work with unknowns. It appears that when attempting to perform semi-quantification, using tools like S_L maximization and N_i normalization seem to provide an edge since they allow an analyst to focus on what is relevant: improving the low response analytes and allowing a fair range of concentration estimates for unknowns.

3.7. Applying semi-quantification to samples

The method is demonstrated on practical food contact material (FCM) application where an extract from a breakfast cereals box was extracted and analyzed according to the Methods section. The QTOF-MS data was analyzed using MassHunter Qualitative Analysis (Qual). The chromatograms, shown in Supporting Information Figure A.4, were found acceptable. However, in TOF-MS analysis, many potential compounds of interest could be masked by a strong background signal in the Total Ion Chromatogram (TIC). Hence, extracted compounds were discovered using the data mining algorithms Molecular Feature Extractor (MFE) in Qual. The ion chromatogram from MFE is shown in Supporting Information Figure A.5. The blank, containing common contaminats from extraction process, was here used as a negative reference for

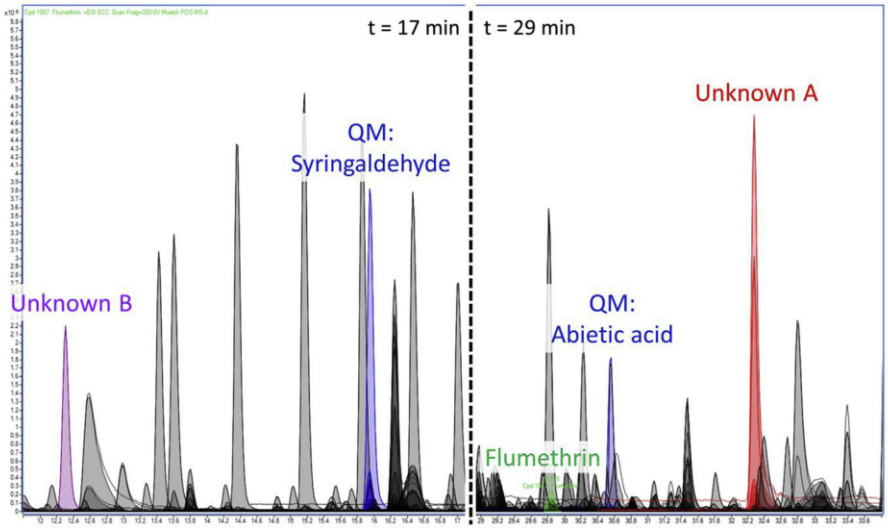


Fig. 3. Compound chromatogram of the two RT windows, RT = 12–17 min and RT = 29–34 min, that contained the three selected AOI and QM (for the full compound chromatogram, see Supporting Information Figure A.5). The AOI are highlighted in color: unknown A (red), unknown B (magenta), and Flumethrin (green). Of the three AOI, only concentration of Flumethrin was known. The QMs, shown in blue, were selected on minimal retention time differences for each AOI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

compounds of interest. The MFE results for blank and sample were overlaid and colored differently to contrast differences. The sample chromatograms from the FCM sample had a large number of new substances compared to the blank, and the number of compound groups that could be identified as unique in the extract surpassed 300.

For demonstration purpose, three compounds were selected for quantification although all compound extracted by MFE could be quantified in a similar way. Fig. 3 details parts of the compound chromatogram that contained the three compounds of interest (marked by red, green and magenta). The three substances chosen were Flumethrin (RT = 29.83 min), an insecticide added to the sample as means of verification, and two completely unknown substances in the sample matrix, Unknown A at RT = 32.27 min and Unknown B at RT = 12.30 min. Only for the added substance Flumethrin could the estimate be verified directly, as the concentration was known. No prior information on Unknown A and B were available, but fragmentation spectra of Unknown A (not shown) revealed intense fragments around *m/z* 163.04 and *m/z* 149.025 with its parent ion at *m/z* 391.2845, typical for phthalates [33]; the accurate mass appears to suggest dioctyl phthalate or an isomer of this. The fragmentation spectra of Unknown B (not shown) show a parent ion at *m/z* 239.1495, and fragmented as repeating mass unit 44 as $[M+H]^+$ adducts (45, 89, 133, …) which are indicative of polyethylene glycol (PEG) [34].

The concentrations of all substances were estimated using the methodology of section 3.6, where the QM with minimal retention difference shown as blue in Fig. 3 was used for AOI quantification. The volume[2] ratio between AOI and QM was multiplied with the QM concentration in the vial to obtain the analyte concentration in the vial. Note that to calculate the concentration in the extract, the concentration in the vial needed to be corrected for the dilution caused by the QM mixture. However, since this factor is the same for all compounds, it was not needed for the comparison. The results of the semi-quantification of the three model compounds are summarized in Table 5.

As it appears, the semi-quantitative estimate for Flumethrin is good, as the concentration estimate is well within a factor 1.5, which is in line with the expectations obtained by the QM comparison. The predictions of the unknown substances cannot be verified as their identity is not elucidated and concentration known. However, with the concentration and structural information obtained, it can be concluded that Unknown B is of significantly less interest to investigate than Unknown A, because the concentration of A is ten times higher than of B – which is significant even in light of a possible factor 3 prediction estimate error. In a similar fashion, it is possible to rank all chemicals in the sample on their semi-quantitative concentration, and investigate those with high concentrations first. Furthermore, using mass data and elution time can be used to evaluate bioactivity by parameters like log D. This basic compound information can be used to support decisions to focus on compounds of significant importance. As an example, PEG and PEG-like substances (as suspected Unknown B) are not known for their human toxicity. However, some phthalates (as suspected Unknown A) are regarded as potentially having adverse health effects, so their presence in food packaging can be seen as a flag for follow-up research [35], especially in (semi-quantitative) concentrations that exceed legal limits. This is highly valuable information as structure elucidation/identification often is very time consuming.

3.8. Limitations

Using general screening method requires several considerations to determine the application scope. Firstly, not all compounds will be ionized therefore will not be detected: the absence of analyte signal does not correspond with the absence of an analyte. This is a general problem in all analytical approaches: one tool will never see all possible compounds. In small molecule analysis multiple analytical approaches needs to be combined if extensive compound coverage is needed. Beside LC-ESI-MS, other techniques like GC-MS,

[2] Volume is the MFE equivalent of area, but includes all possible adducts responsible for a compound.

Table 5
Semi-quantification in a paperboard sample extract for a known substance added to the QM and two unknown compounds in the extract. Concentrations are reported as the respective concentration in the sample vial.

Substance	RT	QM	Predicted Conc. (μM)	True Conc. (μM)
1) Flumethrin	29.84	Abietic acid	0.104	0.146
2) Unknown A	32.27	Abietic acid	1.986	not known
3) Unknown B	12.30	Syringaldehyde	0.199	not known

or different ionization methods like chemical ionization or photoionization are needed to improve coverage. Despite the selective nature of ESI, it still is a widely used technique. However, atmospheric pressure chemical ionization (APCI) might be a good supplementary technique to ESI for use in semi-quantification due to a very different ionization mechanism.

Secondly, the need to use several QMs in LC-MS techniques can be seen as a disadvantage compared to GC-MS that can do functional semi-quantification with a single QM. However, LC-MS is capable of analyzing a much wider range of compounds, both volatile and non-volatile, small molecules to peptides, and so on. This, in combination with the difficulty to predict behavior of analytes in ESI forces the need for extra QM. Still, in this study a very wide selection of QM were chosen to quantify a wide range of compounds, and most of these QM were dissimilar from the AOI but still gave acceptable results. A relatively small set of QMs spread out over the retention range will suffice for most applications. Increasing the number of QM is likely to improve the predictions since the differences in retention time are minimized, but care should be taken not to overload a sample with extra peaks that may contribute to ion suppression or overlap.

Finally, the added complexity of optimizing source parameters and finding QM markers in order to enable good semi-quantification conditions may prevent widespread use. However, if dealing with a large number of samples that use the same LC program; this optimization needs to be performed only once. The selection of markers need to span the retention time window, as it was demonstrated that even with a wide array of chemical substructures, for most compounds semi-quantification was viable. Furthermore, an extensive experimental design as demonstrated in this paper likely is not needed; the first order effects were most significant, so a greatly simplified experimental design would reach an optimum similarly.

4. Conclusion

The continuous need for authentic standards is preventing the ability to deal with unknown substances. LC-MS quantification is complicated due to highly compound- and condition-depending response factors. It is these limitations that have kept generic screenings methods by LC-MS for being a truly powerful method for dealing with quantification of initially unknown substances.

A strategy that allows unknowns to be (semi-)quantified while bypassing the identification step was presented. The true power of this method lies in that identification and quantification was decoupled, hence providing some quantitative information without authentic standards. This makes the method useful in areas where identification is not always possible, or at best tentative. However this is at cost in form of a certain error in the concentration prediction. Yet, the ability to provide estimates of concentration in situations where no data is present is invaluable in untargeted analyses.

Using this method, it was possible to estimate the concentration of well over 300 potential AOI without requiring prior identification or standard-matching. The potential of this application is most useful in fields where a large number of analytes can be present in a sample at any time. Notably, in screening analyses it is useful to prioritize compounds based on their concentration, and identification can be performed primarily on those with higher concentrations. With regards to the standards chosen, chemical food safety is one of the areas that can benefit from semi-quantification.

Further work focus on reduction of the estimation errors, to improve the suitability of QM, and to generate suitable sets of QM for sample types as well as different analytical challenges. There are many considerations regarding QM selection, and extra knowledge in this field is likely to assist in improving the predictions from semi-quantifications. Future studies plan to use the methodology discussed here in an exploration tool for discovering and identifying unknown substances that may pose adverse health effects to humans.

Acknowledgements

Thomas Glauner and Simon Krabbe of Agilent Technologies are thanked for their insight in ESI source parameters. An anonymous reviewer is thanked for extensive feedback on earlier versions of this work. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.aca.2017.03.054.

References

[1] U. Honkalampi-Hämäläinen, E.L. Bradley, L. Castle, I. Severin, L. Dahbi, O. Dahlman, J.-C. Lhuguenot, M.A. Andersson, P. Hakulinen, D. Hoornstra, J. Mäki-Paakkanen, M. Salkinoja-Salonen, L. Turco, A. Stammati, F. Zucco, A. Weber, A. von Wright, Safety evaluation of food contact paper and board using chemical tests and in vitro bioassays: role of known and unknown substances, Food Addit. Contam. Part A 27 (2010) 406–415, http://dx.doi.org/10.1080/19440040903401358.

[2] N. Garg, C.A. Kapono, Y.W. Lim, N. Koyama, M.J.A. Vermeij, D. Conrad, F. Rohwer, P.C. Dorrestein, Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures, Int. J. Mass Spectrom. (2014), http://dx.doi.org/10.1016/j.ijms.2014.06.005.

[3] Q. Bu, D. Wang, X. Liu, Z. Wang, A high throughout semi-quantification method for screening organic contaminants in river sediments, J. Environ. Manage 143 (2014) 135–139, http://dx.doi.org/10.1016/j.jenvman.2014.05.009.

[4] K. Banerjee, S. Utture, S. Dasgupta, C. Kandaswamy, S. Pradhan, S. Kulkarni, P. Adsule, Multiresidue determination of 375 organic contaminants including pesticides, polychlorinated biphenyls and polyaromatic hydrocarbons in fruits and vegetables by gas chromatography–triple quadrupole mass spectrometry with introduction of semi-quantificatio, J. Chromatogr. A 1270 (2012) 283–295, http://dx.doi.org/10.1016/j.chroma.2012.10.066.

[5] R. Bonfiglio, R. King, T. Olah, K. Merkle, The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds., Rapid Commun, Mass Spectrom. 13 (1999) 1175–1185, http://dx.doi.org/10.1002/(SICI)1097-0231(19990630)13, 12<1175::AID-RCM639>3.0.CO;2–0.

[6] N.B. Cech, J.R. Krone, C.G. Enke, Predicting electrospray response from chromatographic retention time, Anal. Chem. 73 (2001) 208–213, http://dx.doi.org/10.1021/ac0006019.

[7] N.B. Cech, C.G. Enke, Relating electrospray ionization response to nonpolar character of small peptides, Anal. Chem. 72 (2000) 2717–2723, http://dx.doi.org/10.1021/ac9914869.

[8] P.J.R. Sjöberg, C.F. Bökman, D. Bylund, K.E. Markides, Factors influencing the determination of analyte ion surface partitioning coefficients in electrosprayed droplets, J. Am. Soc. Mass Spectrom. 12 (2001) 1002–1010, http://dx.doi.org/10.1016/S1044-0305(01)00281-1.
[9] I. Leito, K. Herodes, M. Huopolainen, K. Virro, A. Künnapas, A. Kruve, R. Tanner, Towards the electrospray ionization mass spectrometry ionization efficiency scale of organic compounds, Rapid Commun, Mass Spectrom. 22 (2008) 379–384, http://dx.doi.org/10.1002/rcm.3371.
[10] A. Kruve, K. Kaupmees, J. Liigand, I. Leito, Negative electrospray ionization via deprotonation: predicting the ionization efficiency, Anal. Chem. 86 (2014) 4822–4830, http://dx.doi.org/10.1021/ac404066v.
[11] M.A. Raji, P. Frycák, C. Temiyasathit, S.B. Kim, G. Mavromaras, J.-M. Ahn, K.A. Schug, Using multivariate statistical methods to model the electrospray ionization response of GXG tripeptides based on multiple physicochemical parameters, Rapid Commun. Mass Spectrom. 23 (2009) 2221–2232, http://dx.doi.org/10.1002/rcm.4141.
[12] S. Caetano, T. Decaestecker, R. Put, M. Daszykowski, J. Van Bocxlaer, Y. Vander Heyden, Exploring and modelling the responses of electrospray and atmospheric pressure chemical ionization techniques based on molecular descriptors, Anal. Chim. Acta 550 (2005) 92–106, http://dx.doi.org/10.1016/j.aca.2005.06.069.
[13] M. Soledad, L. Folguera, J. Federico, P. Alejandra, Exploring analyte response in an ESI-MS system with different chemometric tools, Chemom. Intell. Lab. Syst. 146 (2015) 120–127, http://dx.doi.org/10.1016/j.chemolab.2015.05.004.
[14] K.R. Chalcraft, R. Lee, C. Mills, P. Britz-McKibbin, Virtual quantification of metabolites by capillary electrophoresis-electrospray ionization-mass spectrometry: predicting ionization efficiency without chemical standards, Anal. Chem. 81 (2009) 2506–2515, http://dx.doi.org/10.1021/ac802272u.
[15] P.J. Taylor, Matrix effects: the achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry, Clin. Biochem. 38 (2005) 328–334, http://dx.doi.org/10.1016/j.clinbiochem.2004.11.007.
[16] C. Müller, P. Schäfer, M. Störtzel, S. Vogt, W. Weinmann, Ion suppression effects in liquid chromatography-electrospray-ionisation transport-region collision induced dissociation mass spectrometry with different serum extraction methods for systematic toxicological analysis with mass spectra libraries, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 773 (2002) 47–52, http://dx.doi.org/10.1016/S1570-0232(02)00142-3.
[17] M. Oss, A. Kruve, K. Herodes, I. Leito, Electrospray ionization efficiency scale of organic compound, Anal. Chem. 82 (2010) 2865–2872, http://dx.doi.org/10.1021/ac902856t.
[18] S. Zhou, M. Hamburger, Effects of solvent composition on molecular ion response in electrospray mass spectrometry: investigation of the ionization processes, Rapid Commun. Mass Spectrom. 9 (1995) 1516–1521, http://dx.doi.org/10.1002/rcm.1290091511.
[19] S. Zhou, K.D. Cook, A mechanistic study of electrospray mass spectrometry: charge gradients within electrospray droplets and their influence on ion response, J. Am. Soc. Mass Spectrom. 12 (2001) 206–214, http://dx.doi.org/10.1016/S1044-0305(00)00213-0.
[20] N.C. Maragou, N.S. Thomaidis, M.A. Koupparis, Optimization and comparison of ESI and APCI LC-MS/MS methods: a case study of irgarol 1051, diuron, and their degradation products in environmental samples, J. Am. Soc. Mass Spectrom. 22 (2011) 1826–1838, http://dx.doi.org/10.1007/s13361-011-0191-z.
[21] C. Zwiener, T. Glauner, F.H. Frimmel, Method optimization for the determination of carbonyl compounds in disinfected water by DNPH derivatization and LC-ESI-MS-MS, Anal. Bioanal. Chem. 372 (2002) 615–621, http://dx.doi.org/10.1007/s00216-002-1233-y.
[22] G. Greco, A. Boltner, T. Letzel, Optimization of jet stream ESI parameters when coupling agilent 1260 infinity analytical SFC system with agilent 6230 TOF LC/MS, Tech. Overv. (2014). https://www.agilent.com/cs/library/applications/5991-4510EN.pdf (Accessed 7 February 2017).
[23] L. Bengtström, Chemical Identification of Contaminants in Paper and Board Food Contact Materials, Technical University of Denmark, 2014.
[24] European Food Safety Authority, Report of ESCO WG on non-plastic food contact materials, 2012. Parma, Italy, http://www.efsa.europa.eu/.
[25] M.A. Raji, K.A. Schug, Chemometric study of the influence of instrumental parameters on ESI-MS analyte response using full factorial design, Int. J. Mass Spectrom. 279 (2009) 100–106, http://dx.doi.org/10.1016/j.ijms.2008.10.013.
[26] SAS Institute Inc., Estimates, JMP 12 Online Doc, 2015. http://www.jmp.com/support/help/Estimates.shtml#742843 (accessed October 12, 2016).
[27] R.V. Lenth, Quick and easy analysis of unreplicated factorials, Technometrics 31 (1989) 469–473, http://dx.doi.org/10.1080/00401706.1989.10488595.
[28] T. Henriksen, R.K. Juhler, B. Svensmark, N.B. Cech, The relative influences of acidity and polarity on responsiveness of small organic molecules to analysis with negative ion electrospray ionization mass spectrometry (ESI-MS), J. Am. Soc. Mass Spectrom. 16 (2005) 446–455, http://dx.doi.org/10.1016/j.jasms.2004.11.021.
[29] T. Glauner, A.P. Zavitsanos, Electrospray operational parameters in TOF-MS, in: A.R. Fernandez-Alba (Ed.), TOF-MS within Food Environ. Anal, Elsevier B.V., Edinburgh, 2012, pp. 273–305, http://dx.doi.org/10.1016/B978-0-444-53810-9.00003-1.
[30] A. Mordehai, ESI technology with thermal gradient focusing - theoretical and practical aspects, in: ASMS Annu. Conf. Mass Spectrom. Allied Top, 2009. Philadelphia, Pennsylvania, https://www.agilent.com/cs/library/technicaloverviews/Public/5990-3494en_lo CMS.pdf.
[31] B.J.A. Berendsen, L.A.M. Stolker, M.W.F. Nielen, The (un)certainty of selectivity in liquid chromatography tandem mass spectrometry, J. Am. Soc. Mass Spectrom. 24 (2013) 154–163, http://dx.doi.org/10.1007/s13361-012-0501-0.
[32] C.F. Bökman, D. Bylund, K.E. Markides, P.J.R. Sjöberg, Relating chromatographic retention and electrophoretic mobility to the ion distribution within electrosprayed droplets, J. Am. Soc. Mass Spectrom. 17 (2006) 318–324, http://dx.doi.org/10.1016/j.jasms.2005.11.006.
[33] P. Yin, H. Chen, X. Liu, Q. Wang, Y. Jiang, R. Pan, Mass spectral fragmentation pathways of phthalate esters by gas chromatography–tandem mass spectrometry, Anal. Lett. 47 (2014) 1579–1588, http://dx.doi.org/10.1080/00032719.2013.879658.
[34] B.O. Keller, J. Sui, A.B. Young, R.M. Whittal, Interferences and contaminants encountered in modern mass spectrometry, Anal. Chim. Acta 627 (2008) 71–81, http://dx.doi.org/10.1016/j.aca.2008.04.043.
[35] J. Muncke, Phthalates, Food Packag. Forum, 2012. http://www.foodpackagingforum.org/food-packaging-health/phthalates (accessed February 7, 2017).

3.5. Manuscript B: Exploration - Tentative Identification

Title:	Exploring the chemistry of complex samples by tentative identification and semi-quantification: a food contact material case
Journal:	*Journal of Mass Spectrometry*, Volume xxx, 2018, Pages xx-xx, ISSN 1096-9888, http://dx.doi.org/10.1002/jms.4052
Author(s):	Eelco N. Pieke, Jørn Smedsgaard, Kit Granby
Keywords:	structure assessment; mass spectrometry; semi-quantification; exploration; food contact materials
Contact:	Eelco N. Pieke: enpi@food.dtu.dk, eelco.pieke@gmail.com
Contributions:	Design of the experimental design, experimental work, data analysis, and writing of the manuscript was done by EP. The overall concept of the study was developed by EP with input from JS and KG. Revision and correction of manuscript proofs was performed by EP, KG, and JS.
Supplement:	Supplementary text and visual material to this manuscript are available in the Appendix, starting at page 166.

Received: 15 August 2017 | Revised: 21 November 2017 | Accepted: 26 November 2017

DOI: 10.1002/jms.4052

RESEARCH ARTICLE

WILEY Journal of MASS SPECTROMETRY

Exploring the chemistry of complex samples by tentative identification and semiquantification: A food contact material case

Eelco N. Pieke | Jørn Smedsgaard | Kit Granby

National Food Institute, Research Group for Analytical Food Chemistry, Technical University of Denmark, Kemitorvet Building 202, Kgs. Lyngby DK-2800, Denmark

Correspondence
Eelco N. Pieke, National Food Institute, Research Group for Analytical Food Chemistry, Technical University of Denmark, Kemitorvet Building 202, Kgs. Lyngby DK-2800, Denmark.
Email: enpi@food.dtu.dk

Funding information
Danish Veterinary and Food Authority (DVFA), Grant/Award Number: Fødevareforlig 3; Technical University of Denmark (DTU), Grant/Award Number: DTU Scholarship

Abstract

In fields such as food safety and environmental chemistry, ensuring safety is greatly challenged by large numbers of unknown substances occurring. Even with current state-of-the-art mass spectrometers, dealing with nonidentified substances is a very laborious process as it includes structure elucidation of a vast number of unknowns, of which only a fraction may be relevant. Here, we present an exploration and prioritization approach based on high-resolution mass spectrometry. The method uses algorithm-based precursor/product-ion correlations on quadrupole time-of-flight tandem mass spectrometry data to retrieve the most likely chemical match from a structure database. In addition, time-of-flight-only data are used to estimate analyte concentration via semiquantification. The method is demonstrated in recycled paper food contact material. Here, 585 chromatographic peaks were discovered, of which 117 were unique to the sample and could be tentatively elucidated via accurate mass, isotopic pattern, and precursor/product-ion correlations. Nearly 85% of these 117 peaks were matched with database entries, which provided varying certainty of information about the analyte structure. Semiquantitative concentration ranges of investigated compounds were between 0.7 and 1600 μg dm^{-2}. With these data, a subgroup of chemicals was risk-categorized and prioritized by using the most likely candidate structure(s) obtained. Prioritization based on expected health impact was possible by using the tentatively assigned data. Overall, the described method not only is a valuable chemical exploration tool for nonidentified substances but may also be used as a preliminary prioritization tool for substances expected to have the highest health impact, for example, in food contact materials.

KEYWORDS

exploration, food contact materials, mass spectrometry, semiquantification, structure assessment

1 | INTRODUCTION

The increasing awareness of adverse effects of chemicals in our environment, especially in food, has led to ever more demanding questions about occurrence of unwanted chemicals. Traditionally, analytical methods are designed to detect specific groups of chemicals by using targeted methods. However, these methods are limited to the original method scope so fail to detect chemicals outside this scope. Fortunately, newer different analytical methods may address the challenge of unknown compounds that are not known to be present or where the structure is not initially expected to be present. This challenge is prevalent in many areas, eg, metabolomics,[1] food safety,[2,3] and environmental research.[4,5] Unknown chemicals shift the research question from "can *this* substance be found in this sample?" toward "what *else* can be found in this sample?" While targeted analyses are known for high sensitivity and useful for routine analysis,[6] they can similarly provide a falsified argument of safety: Absence of evidence is not equal to evidence of absence. Food scandals, notably the melamine case,[7,8] have proved that targeted methodology cannot always ensure safety due to the large number of substances not included.

When one takes into account the nontargeted substances, orders of 100,000s of chemical substances may be found in food contact materials (FCMs), which may have the potential to migrate to food in concentrations of concern.[9] Research is often directed according to the "popularity" of a substance: much for well-studied compounds but few or none for unknown compounds. The lack of knowledge on

unexpected compounds in food and environment has driven a need to look past the targeted methodology. Yet, even untargeted methods still have a too strong dependence on preexisting knowledge,[10] and available sources of data require extensive cleaning and pruning prior to use in untargeted methodology.[11] Dealing with complex samples, especially to ensure human safety, requires a level of discovery that targeted or even some untargeted analyses do not currently provide.[12]

Modern accurate MS provides the capability to predict molecular structures from isotopic patterns, MS-MS fragmentation, retention behavior, and elemental composition.[13,14] Even so, MS is rarely sufficient for a full identification, and authentic standards or other elucidation methods are still required, eg, purification followed by analysis by nuclear magnetic spectroscopy.[15-17] Recently, Kind and Fiehn[18] reviewed unknown elucidation and summarized promptly that there is currently no software or tool that can truly predict an unknown structure from MS spectra. Hence, the unknown chemicals tend to direct to the void between targeted analysis (not included in method) and untargeted analysis (unable to be identified or quantified).

Food science (flavor, safety, etc.), metabolomics, and environmental chemistry often use MS combined with reference libraries of known chemicals.[19-22] Food safety, risk assessment relies on the potential hazard and the exposure to these compounds[23]; hence, the challenge is in 2 questions: *what* is present in the sample and *how much*? We recently showed that quantifications could be decoupled from identification,[24] and this allows for rough concentration estimates of virtually an unlimited amount of unidentified substances. Furthermore, studies by Bengtström et al[25] and Rosenmai et al[26] showed that untargeted screening can be used in a bio-guided strategy to find hazardously compounds present in extracts from cardboard food packaging materials. This approach aims to avoid wasting effort identifying and quantifying irrelevant compounds by employing prioritization.[27,28] Yet, ideally elucidation, quantification, and prioritization should be performed on *any* present substance to assess which substances are risk relevant or risk irrelevant. It may be viable to adapt a strategy similar to recent literature[25,26] and combine it with the latest semiquantification (SQ) approach[24] to obtain a chemical exploration tool that can prioritize chemical substances based on adjustable endpoints.

In this paper, we present a fast workflow based on high-resolution mass spectrometry to combine structure prediction methodology with semiquantitative methodology to generate a profile of potential hazardous compounds in complex samples. To demonstrate the feasibility of such a method, we apply it in a FCM setting. The investigated workflow provides a much-needed tool for complex sample exploration: the abundances and chemical profiles that can be used for further investigations.

2 | MATERIALS AND METHODS

2.1 | Software

Marvin (16.6.13.0, 2016, ChemAxon, http://www.chemaxon.com) and Calculator Plugins were used for drawing, displaying, and characterizing chemical structures; calculating properties; and converting names. JChem Standardizer (JChem 16.5.2.0, 2016, ChemAxon) was used for structure cleanup and pruning. JChem for Office (16.1.1100.489, 2016, ChemAxon) was used for chemical reporting in Excel. Agilent MassHunter Molecular Structure Correlator (MSC, 8.0.36.0, 2016, Agilent Technologies, Santa Clara, CA, USA) was used to correlate accurate mass MS/MS fragment ions to databases to search for a most likely match. Agilent MassHunter Acquisition (B.06.01, 2016, Agilent Technologies) was used for data acquisition. Agilent MassHunter Qualitative Analysis (B.07.0.7024.29, 2016, Agilent Technologies) was used for data analysis via built-in Molecular Feature Extraction (MFE) and further data treatment of time-of-flight (TOF) MS and quadrupole (Q)-TOF MS/MS data.

2.2 | Instrumental setup

2.2.1 | Electrospray ionization quadrupole time-of-flight mass spectrometry

Mass analysis was performed by using an Agilent 6550 Q-TOF MS (Agilent Technologies, Santa Clara, CA, USA) equipped with Agilent JetStream electrospray ionization (ESI) interface. Equipment setup and optimization follow recent work.[24] The optimized parameters for the ESI source are shown in Table 1. The Q-TOF MS was auto-tuned for mass accuracy and resolution between m/z 100 to m/z 3200 prior to each sequence, and mass axes were 2-point calibrated real time by using supplied ESI calibration mixture.

Mass spectral experimental were run 2 modes: TOF only (MS) and Q-TOF (MS/MS). In TOF mode, acquisition was performed in the mass range 50 to 2500 m/z at 1.3 Hz with no collision gas or collision energy. During product-ion acquisition, the mass range was reduced to 50 to 2400 m/z at 2.5 Hz with the quadrupole isolation width set to "narrow" (1.3 m/z width) and at varying collision energy voltages: 10, 20, and 40 V. The auto-MS/MS feature was used to isolate precursor ions dynamically in real time. Common background ions were added to the auto-MS/MS "exclusion" list to improve spectral quality. Other settings of the auto-MS/MS were tuned for this specific instrument and the samples.

2.2.2 | Liquid chromatography

Ultrahigh-performance liquid chromatography mass spectrometry was performed on an Agilent 1290 system (Agilent, Waldron, Germany). For separation, a separation strategy was adopted as described in Pieke et al[24] by using 2 ultrahigh-performance liquid chromatography columns connected serially to improve separation power and

TABLE 1 Optimized settings employed in the electrospray ionization module. Settings are given for both positive mode electrospray (ESI+) and negative mode electrospray (ESI−)

Setting	Unit	ESI+	ESI−
Nozzle voltage	Volt (V)	1000	1000
Sheath gas temperature	°C	350	350
Sheath gas flow	L min^{-1}	12	10
Capillary voltage	Volt (V)	2000	2000
Nebulizer pressure	psi	25	40
Drying gas temperature	°C	150	150
Drying gas flow	L min^{-1}	14	11

selectivity. Both columns were held isothermally at 50°C. Injection volume was 5 µL. Eluent flow was 0.2 mL min^{-1}.The first column was a Phenomenex Luna Omega Polar C18 (100 Å, 1.6 µm, 100 × 2.1 mm) (Phenomenex, Denmark). The second column was a Waters ACQUITY UPLC CSH C18 (130 Å, 1.7 µm, 100 × 2.1 mm) (Waters, Denmark). Columns were protected by using the Phenomenex SecurityGuard™ ULTRA (Phenomenex, Denmark) fitted with Omega Polar C18 phase. Eluent A was MilliQ-grade (Millipore Corp., Bedford, MA, USA) water, and eluent B was TOF-grade methanol (Fluka, Denmark). Both eluents A and B contained 2-mM ammonium formate buffer (pH 3.5).

The columns were selected for improved selectivity toward polar compounds (Luna Omega) and free acids (ACQUITY) while retaining high capacity C18-like separation. Methanol was chosen for compatibility with buffer salts, a favorable interaction with sugars, and being a generally weaker eluent allowing a smoother gradient. Eluent gradient was as follows: after injection gradient was held for 3 minutes at 0% B, then increased to 90% B at 25 minutes, then ramped to 99% B at 33 minutes, and finally set to 100% B at 33.01 minutes, performing isocratic elution for 5 minutes until 38 minutes. Posttime cleaning (4 min) involved flushing the injection loop and column with 100% A.

2.3 | Sample preparation

2.3.1 | Sampling

Samples were taken from an unused recycled food cardboard item (pizza box) used in retail applications. The sample was chosen as recent work has shown that substances of concern can migrate from these types of samples.[29] From the center of the bottom of the box, a 1-dm^2 board piece was cut out by using a cleaned surgical knife. Sample cutout was resized to 4 identically sized strips of 2.5 cm × 10 cm. Before extraction, all nondisposable glassware, eg, glass extractor vessel, glass vials (PYREX), and measuring glasses, were cleaned and heated to 400°C for a minimum of 12 hours. All plastic equipment was soaked in 95% ethanol for 24 hours before use. Liquid chromatography mass spectrometry vials (Agilent Technologies) and Pasteur pipettes were used as received.

Four sample strips were placed in an extraction vessel, and 100 mL of warm (40-50°C) extraction solvent was added. The extraction vessels were capped and sealed with a metal clamp, transferred into a thermally sealed airtight bag, and placed in an oven for 24 hours at 40°C. After 24 hours, the vessels were removed from the oven, and the liquid content was poured into 100-mL glass bottles where it was cooled down to room temperature. A blank extract was prepared by following the extraction procedure without sample.

The resulting extract was filtered in triplicate and transferred to LC-MS compatible vials for analysis. Each sample was filtered by using a plastic 2-mL BD Discardit™ syringe (Becton Dickinson, USA) and plastic-PTFE 0.2 µm Phenex™ filters (Phenomenex, Denmark), thoroughly rinsed with 50% ethanol. The syringe was used in a way that ensured there was an air cavity between plunger and solvent to prevent contamination. First, 0.9-mL extract was filtered to waste followed by 0.5 mL of extract filtered into the LC-MS vial. Approximately 1.0 mL of quantification marker (QM) mixture was added to each LC-MS vial. The contents of the QM mixture are described in Table A.2.1.

2.4 | Molecular structure correlation

The Agilent MassHunter MSC program proposes a structure suggestion of an analyte of interest (AOI) by comparing MS/MS data containing accurate mass, fragment ions, and isotopic patterns with local or online databases of available chemical structures. Each potential molecular formula of an AOI is compared against a database of theoretically fragmented chemical structures as described by Hill and Mortishire-Smith.[30] As experimental spectra are compared with theoretical spectra, a matching score is calculated based on fragment matching, mass accuracy matching of precursor and fragment ions, and the overall percentage of fragment ions that can be plausibly explained.[31] As a result, the MSC reports the most plausible chemical structure inside the consulted database that fits best to the experimental MS/MS data.

2.5 | Database preparation

Seven data sources (DS) were used to generate a set of 5 databases. Three DSs were chosen for high specificity to sample types. Data sources #1 and #2: compiled list from research regarding contaminants in paper and board material.[29,32] Data source #3: the US FDA database "Indirect Additives used in Food Contact Substances."[33] Two DSs were selected for moderate relevance to the field of usage. Data source #4: REACH Annex III database containing chemicals used in industry that have some associated safety concerns.[34] Data source #5: Agilent MassHunter Extractables and Leachables library, containing extractable and leachable compounds found in components of (mostly plastic) food and drug packaging and medical devices.[35] Finally, 2 large DSs were consulted online. Data source #6: ChemSpider (http://www.chemspider.com/). Data source #7: PubChem.[36]

Before use as database, DS #1 to #4 were converted into a structure format and pruned as described below. In addition, DS #1 to #3 were merged into a single database (PBDB) due to large overlap. Data sources #5 to #7 needed no further modification. The conversion and pruning for each DS first involved converting entries (CAS or name) to *simplified molecular-input line-entry system* by using a combination of ChemAxon software and in-house Excel scripts. After conversion, the simplified molecular-input line-entry system database was pruned for duplicate, erroneous, or empty entries. The conversion process is described in greater detail in Supplementary Material A.1 and Figure A.1.1.

3 | RESULTS AND DISCUSSIONS

3.1 | Databases and pruning performance

The mechanism of MSC is designed to propose the best-matching structure inside a database of structures, thereby critically dependent on the appropriateness, quality, and size of the databases. Hence, the *database scope* and *database specificity* are decisive quality factors in structural elucidation. In Table 2, a quality assessment of different databases is presented. Here, the *scope* of database is proportional to the database size and uniqueness: a larger scope increases the likelihood of obtaining a positive or false positive match from a mass spectrum. The *specificity* of the database is based on the relevance of

TABLE 2 Qualitative assessment of databases: scope, specificity, and authenticity

Database	Number	Locality	# Entries	Spec.	Scope	Auth.
Paper and Board Database (PBDB)	#1 - #3	Local	4,353	++	+	++
REACH	#4	Local	55,648	+	+	+
Agilent Extractables and Leechables (AgEL)	#5	Local	1,840	±	−	±
ChemSpider (CS)	#6	Online	> 58 million†	− −	++	− −
PubChem (PC)	#7	Online	> 91 million†	− −	++	− −

Specificity (spec.) is the relevance of the database to the field used; scope is based on the uniqueness and size of the database. Authenticity (auth.) is the overall reliability of predictions using this database. Database properties were qualitative marked; ++, very good; +, good; ±, moderate; −, low; − −, very low.
†Online as of June 2017.

database to the sample type, ie, whether the substances in the database are representative to the analytical problem and sample types. Database *authenticity* is a qualitative parameter based on the *size* and *specificity*. Hence, authenticity is a measure of the trustworthiness of an MSC prediction from that database.

In Table 2, the PBDB and the REACH database scored well on authenticity due to a mid-sized scope along with a high specificity. The Agilent PCDL had limited use due to a relatively small scope and also a low specificity because it is designed primarily for plastic substances. The ChemSpider and PubChem databases could not be considered authentic despite the largest scope, as these databases included structures from a highly diverse pool of chemicals from many possible applications. Here, structural correlations may shift to overfitting data: small measurement errors in the data can be fitted on the near-unlimited number of theoretical spectra present in these databases. This does not imply that correlation results from ChemSpider and PubChem are not usable, but it does imply that the user needs to be critical on the matching results.

3.2 | The role of databases in structure correlation

The usefulness of structure predictions is related to the database used to predict the structure. Databases PBDB, REACH, and AgiEL are strongly relevant to the investigated samples, and the structure predictions can be used without reassessment. However, for CS and PC, the predictions require manual assessment of the proposed structure, as these predictions were found to contain anomalies, eg, excess triple bonds, prevalent radicals, or carbon rings of n < 5. Entries containing said anomalies required reassessment because the presence in the REACH Annex III database, representative of a large proportion of industrial chemicals in use, was uncommon: only 0.64% of the entries contained a single triple bond and much less multiple, while only 0.037% contained a ring with n = 3, and 0.019% a ring with n = 4. Thus, only if the best-scoring hit in CS or PC contained no structural anomalies it was accepted, but if anomalies were present the next-best hit with no anomalies was accepted.

The issue with anomaly entries highlights a somewhat larger issue in using untargeted predictions for structural elucidation. First, available databases are used as-is: the search range is limited by the content of the database. Therefore, chemical structures that are not originally found in the database cannot be proposed. Consequently, if the actual AOI structure is not in any database, then no exact structure is obtained and, at most, a similar structure may be found.

Second, fitting experimental data to nonspecific large databases involves a high risk of false-positive results. The nonspecific public online databases (such as ChemSpider and PubChem) are useful for elucidating unknown chemicals, but there is also a need for specific databases. Specific databases require extensive knowledge on the sample type but are relatively easy to construct, inexpensive, and unlikely to comprise proprietary materials as the only input required are molecular structures. Hence, specific databases might be constructed in collaboration with industrial companies that often have deep knowledge on the samples.

3.3 | Using structural elucidation in untargeted work

3.3.1 | Data-dependent acquisition

Multistage mass spectrometry (eg, QqQ, Q-TOF, or ion traps) are often associated with targeted analyses: the mass-to-charge (m/z) of AOI is preselected, and fragmentation spectra are collected only for preselected fragments (Selected- or Multiple Reaction Monitoring [SRM or MRM]). Because of preselection, SRM/MRM results in product ion spectra of sufficient abundance and high quality. However, SRM/MRM is not well suited for any sort of discovery that involves unknowns (Figure 1a). In fact, SRM/MRM is blind to any mass features not included in preselection as illustrated in Figure 1b.

Due to the limitation of SRM as required for untargeted analyses alternative data acquisition methods are in development such as data-dependent acquisition (DDA).[37,38] In DDA, real-time data are evaluated for precursor ions based on TOF MS data, and an algorithm selects relevant precursors for fragmentation. This permits the investigation of unknown analytes, as no prior definition is needed. Another methodology is data-independent acquisition (DIA),[39,40] in which all ions are fragmented simultaneously without selection. DIA results in complex product-ion spectra, and data analysis is challenging, while spectral quality is not optimal.[41,42] This study exclusively uses DDA due to the added complexity of performing DIA.

A limitation associated with DDA is the compromise needed for complex samples containing many unknown substances.[43] The number of precursor ions to be analyzed by fragmentation is limited by time required per scan cycle, so only a limited number of precursors can be included within a given time range.[41] If a large number of viable precursors are discovered, those with highest abundance are preferentially selected while the remainder is discarded (Figure 1c), which is problematic for chromatographic areas with many eluting substances, or where background peaks are dominant.

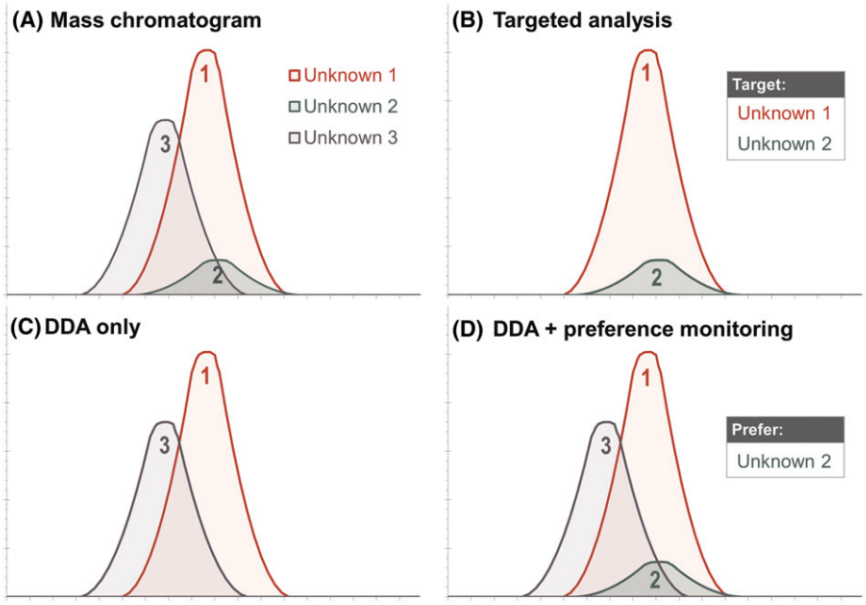


FIGURE 1 Mass chromatogram ion traces demonstrating preselection and active selection acquisition. a) Full mass chromatogram illustrating three different overlapping peaks of varying abundance. b) Analysis via targeted methods exclusively fragments the preselected precursor ions (1 and 2), but excludes all ions not listed (3). c) Analysis via DDA results in product-ion spectra via dynamic precursor selection, but excludes low-abundance ions (2) by preferentially selecting high-abundance ions (1 and 3) in case there are large numbers of ions. d) Analysis via DDA augmented with preference monitoring permits dynamic precursor selection without sacrificing low-abundance ions, provided these are included in the preference list

To control the selection of ions a combination of DDA and preference/exclusion monitoring based on a preset list of priority ions (preference list) was used (Figure 1d). The design of a preference list requires TOF-only (MS) analysis to identify the time and *m/z* of interesting peaks, and these data were already available from the quantification experiments. Preference monitoring ensures that certain ions and/or fragments have fragmentation spectra collected, which is beneficial for investigating low-abundance analytes. Unfortunately, designing the preference list is a manual task and, for large chromatograms, can be time intensive. Consequently, the investigative power of the method is enhanced by combining real-time targeting with a preselected list of ions of interest.

The experimental results indicated that DDA deals poorly with different adducts generated by the ESI process. The DDA algorithm cannot unambiguously identify the adduct type of the precursor, and the most abundant adduct ion did not consistently produce high-quality fragmentation spectra. In Figure 2, the product-ion spectrum for 3 different adducts for a single structure in positive ESI is shown. Fragmentation spectra based on protonated ($[M + H]^+$), deprotonated ($[M - H]^-$), or ammoniated ($[M + NH_4]^+$) adducts were generally of high quality, while adducts based on alkaline salts (Na^+ or K^+) produced little to no fragmentation. This signifies a flaw in DDA: some compounds ionize dominantly as salt adducts that are consecutively selected as precursor ion due to the high intensity. Yet, in order for fragmentation to be effective, the precursor ion must not be sodium ($[M + Na]^+$), potassium ($[M + K]^+$), or formate ($[M\text{-}COO]^-$) adducts. Ensuring proper adduct selection can partly be achieved by using DDA preference monitoring to select optimal adducts for fragmentation spectra despite having a lower initial abundance.

3.4 | Food contact material analysis: Identification and quantification

3.4.1 | Analysis workflow

To evaluate the method performance on a real sample, a recycled cardboard pizza box was investigated. Pizza boxes contain both virgin and recycled board fibers, so they are expected to contain a diverse set of chemicals with a potentially large probability of containing unknowns.[44,45] Figure 3 shows the simplified workflow from sample workup to quantification and elucidation. The extracted sample was analyzed in triplicate (A1 to A3) in both positive and negative mode ESI. Due to the size of the dataset produced, only a single sample in positive ESI (A1) is discussed here to demonstrate the method.

After sample preparation, sample and blanks were run in TOF mode. Time-of-flight MS data were analyzed by the MFE algorithm in MassHunter Qualitative Analysis software, and the results were overlaid with the blank extraction. A DDA preference monitoring list was generated by recording the retention time and precursor mass of analytes that were distinct in the sample compared with the blank. When several adducts were observed for a preferred compound, a maximum of 2 precursor ions were selected with a priority order of $[M + H]^+ > [M + NH_4]^+ > [M + Na]^+ > [M + K]^+$. Adduct prioritization was not necessary in negative ESI mode due to the lower number of possible adducts.

Semiquantification was limited to the largest 1200 MFE groups (ie, possible compounds) per sample to reduce computation times for the MFE algorithm. First, mass spectrometric data were investigated for the MFE groups of the QMs. Molecular Feature Extraction groups consist of multiple ion traces originating from the same analyte. For each

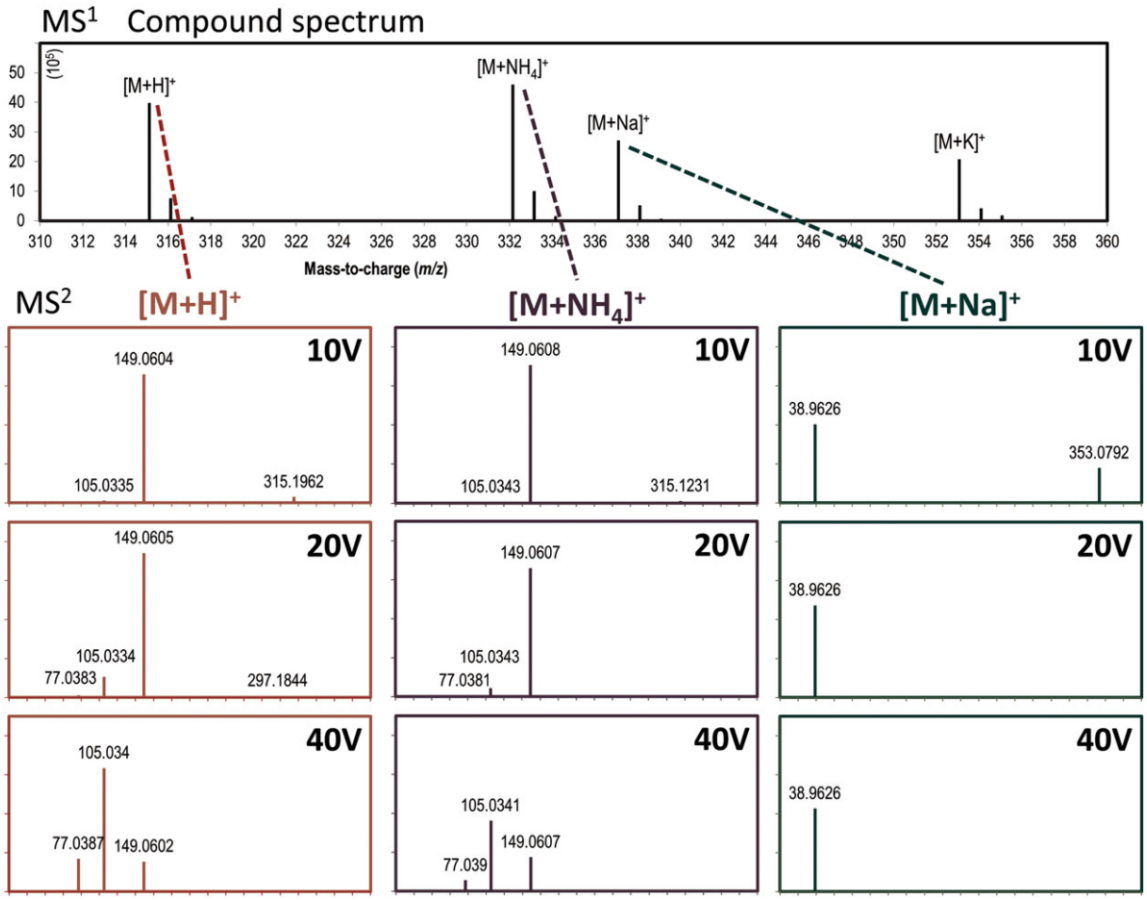


FIGURE 2 The fragmentation spectrum of an unknown compound demonstrates the adverse effect of alkaline ion adducts on spectral quality. The original mass spectrum is shown at the top for phthalate-like compound with four different precursor adducts. In MS/MS, fragmentation of the molecular adducts $[M+H]^+$ and $[M+NH_4]^+$ resulted in high quality spectra, containing numerous fragments of varying intensity depending on the energy applied. Fragmentation of molecular adducts $[M+Na]^+$ and $[M+K]^+$ (data not shown) resulted in spectra with little fragmentation unchanged by varying collision energies, thereby containing little structural information on the compound [Colour figure can be viewed at wileyonlinelibrary.com]

compound, the summed group ion traces were recognized as being the true abundance of the analyte in the sample. Actual SQ was achieved by comparing each analyte to its nearest QM.[24]

Finally, compounds detected in both Q-TOF (identification) and TOF (quantification) were coupled by using Excel VBA scripts. The coupling occurred only when both methods resulted in similar in retention time and similar *m/z* of principal ions.

3.4.2 | Identification

Triplicate analysis of the pizza box sample (A1, A2, and A3) resulted in 133, 131, and 131 compounds, respectively. To facilitate structure assignment, MSC predictions were categorized as either authentic predictions, based on PBDB, AgiEL, or REACH, or nonauthentic predictions, based on PubChem or ChemSpider. Two simplified rules were used to decide whether a structure assignment by MSC could be passed as relevant:

1. A threshold average prediction score was above 70: at least 70% of the observed Q-TOF spectrum was explained by the database structure.
2. A compound required a match from at least 2 databases, and at least 1 match originated from an "authentic" source (PBDB, REACH, or AgiEL).

The performance of the assignments was assessed by the total prediction statistics as summarized in Figure 4 for sample A1.

Assessing the quality of the predictions revealed that nearly 85% of the predictions (Figure 4A) were of sufficient quality usable for interpretation (score > 70). The average prediction score was 78; thus, on average, more than three-quarters of the fragmentation spectrum could be explained. A substantial part of the predictions scored between 70 and 80 (37.6%), but most scored 80 or above (47.4%). Some predictions were considered of insufficient quality (score < 70) for interpretation (15%), and most of these prediction scores were between 50 and 70 (11.3%).

When looking at the prediction quantity n (Figure 4B), more than 83% of possible AOI had n ≥ 3 predictions that could support structure assignment. For n = 2, these predictions were exclusively from CS and PC. As discussed, CS and PC are probable to contain false-positive predictions, so the MSC results were evaluated manually. However, a

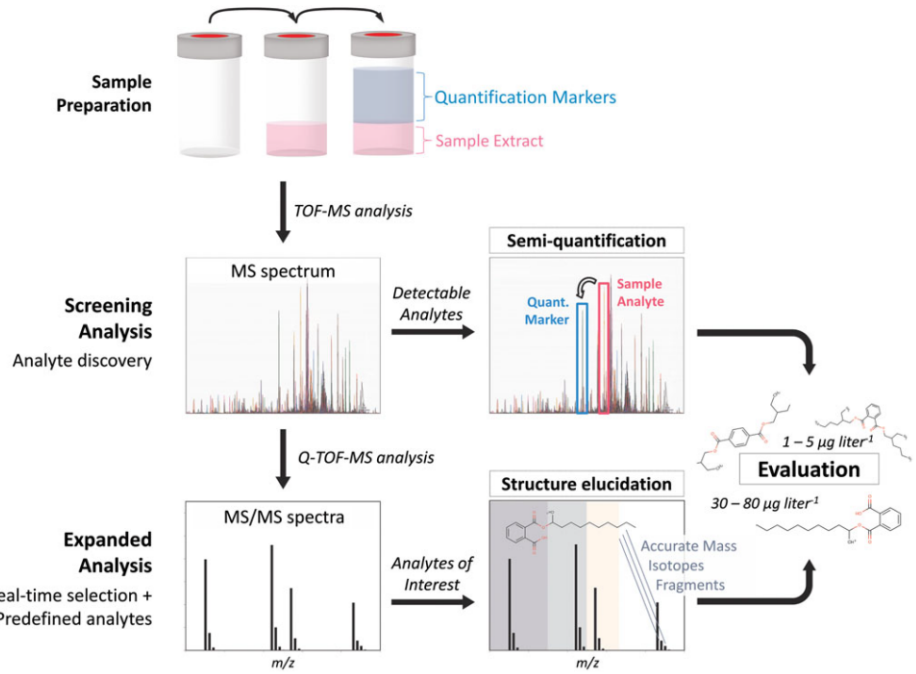
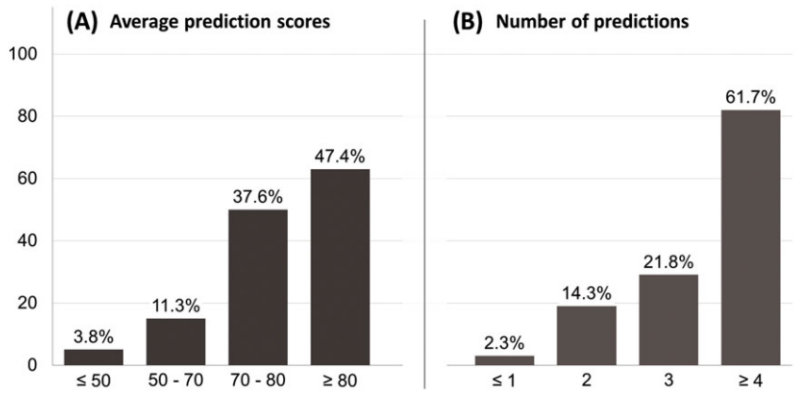


FIGURE 3 Workflow for structure elucidation and for semi-quantification. First, sample cleanup is required before most LC-MS analysis. During this step, the quantification markers (QM) are added to all samples. Semi-quantification data and preference lists are retrieved via screening analysis. Structural prediction is done by structure correlations on product-ion spectra of listed and detected analyte in the sample. The semi-quantification and structural prediction are combined into an evaluation of the substance

FIGURE 4 Graph indicating the distribution of prediction scores and number of effective predictions. Data is shown for sample A1, a recycled pizza box with total number of compounds n = 133. A) Distribution of prediction matching scores. The scores were divided into four bins centered on the acceptance threshold of score ≥ 70. B) Distribution of the number of viable predictions from various databases (max. 5). At minimum two predictions were needed to consider the elucidation as adequate for further work, where one prediction must have originated from a high-specificity database



lack of authentic predictions may indicate that these were truly unknown compounds not yet available in any database. Cases where insufficient prediction reliability (n = 0 or n = 1) occurred were rare: only in 3 of 133 cases.

3.4.3 | Quantification

To simplify the description and tables in these results, only data from A1 are used here. In sample A1, 585 potential peaks were semiquantified. Only compounds for which the concentration could be estimated in at least 2 of 3 samples were included. A coefficient of variation (CV) was calculated for 109 compounds by using replicate experiments A1, A2, and A3. The median CV was 10%, while the average CV was 17%. However, compounds eluting before 3.5-minute retention (gradient start) contributed more to the observed variation; if these were excluded, the median CV was similarly 10%, but the average CV reduced to 13%.

Figure 5 summarizes the semiquantitative results. The range of concentrations calculated by SQ was about 5 orders of magnitude between 2.7 and 159 µmol L^{-1}, which approximately corresponded to a range from 470 ppb to 17 ppm. The median of concentration estimates was 68 nmol L^{-1}. The range of concentrations of entries that were successfully linked between quantification and identification (marked in Figure 5 by filled points) was between 19 and 159 µmol L^{-1}. As the masses of the linked compounds are known from the identification, the concentration range could also be expressed as mass unit per dm^2: 0.7 to 1653 µg dm^{-2}.

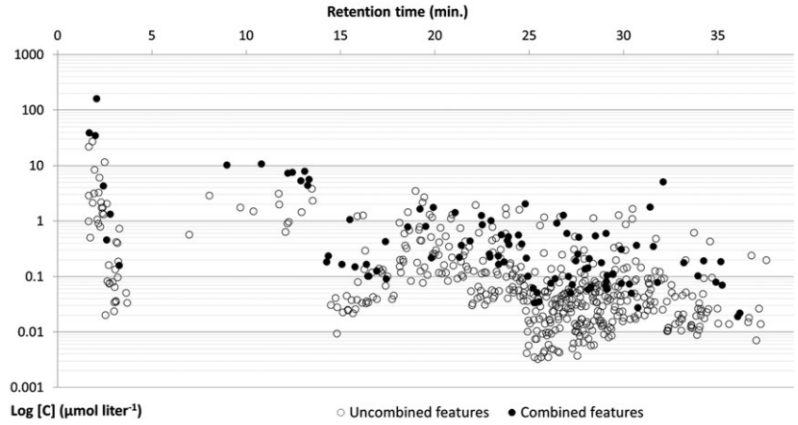


FIGURE 5 Retention time and semi-quantified concentration of 585 chromatographic peaks in a recycled board sample. Peaks include those also found in the blank. Of these 585 peaks, 117 peaks were linked to elucidation data in subsequent MS/MS experiments, which are marked as filled points. The remaining unlinked peaks (blank points) either were considered irrelevant in the MS/MS experiments or did not have suitable fragmentation spectra recorded

3.4.4 | Combining semiquantification and identification data

The 585 SQs were compared with the 133 structure elucidations, which resulted in 117 compounds with both structure and concentration assignments. The unlinked 468 SQs, marked in Figure 5 by transparent points, could not be coupled to existing MS/MS data. The 16 identified but not quantified compounds were probably due to a targeted ion being of such low abundance that the data processing did not detect the principal ion.

There appeared to be a trend in the coupling of SQ with identification where the successful links are often at the higher range of concentrations. This effect was expected: DDA can only simultaneously evaluate a limited quantity of ions, and during saturation the higher abundance signals take priority. This indicates the balance needed between operating in MS mode and MS/MS mode: fragmentation spectra (MS/MS) may be used to evaluate new candidate ions, but shortening the acquisition time in MS/MS reduces the quality of the spectra due to lower abundance. Hence, to collect spectra from sub-ppm level compounds or those with elution overlap, the precursor ions need to be part of the DDA preference list.

Alternatively, the elution time window could be broader so that DDA can utilize the longer time to evaluate more ions. A broader elution window could be beneficial as the DDA algorithm excludes peaks after a set number of MS/MS spectra, so a longer time would free up the queue for lower abundance precursor ions. However, broader peaks are also inversely related to abundance and generally unfavorable as the peak overlap increases significantly, thus reinforcing the initial problem.

3.5 | Assessment of quantification and identification of a paper and board food contact material

In risk studies, the most prominent decision is whether the presence of a certain substance is acceptable (ie, safe and/or legal) or unacceptable (ie, present a risk and/or is illegal) and needs to be considered for further study. Targeted methods are more sensitive than untargeted screening,[20] so highly potent toxic substances are not properly assessed via the latter methodology because they require sufficient sensitivity.[46] Assessing risk and/or legal status for the remaining chemical substances requires a combination of knowledge about hazard characterization and structure: to assess health effects or determine if it is a regulated substance, and about the concentration: to determine if the human exposure poses a risk or is above the legal threshold.

Initially, the number of discovered compounds in this study is not as high as literature might suggest.[9] The real number of compounds is most likely higher if DDA was optimized to include more low-abundance peaks. However, a change in analytical methodology will also increase the number and diversity of chemicals detected, eg, different HPLC columns or different ionization principles. In this case study, analytes that are strongly nonpolar (eg, polycyclic aromatic hydrocarbons and mineral oils) are excluded by the use of ESI, while those that are very small or very polar (eg, ionic substances, monomers, and solvents) could be excluded by using C_{18} liquid chromatography.

Not all substances are discussed or reported in this paper due to the size of the total dataset (133 compounds in triplicate); however, illustrative examples are presented to document the method and output. A subset (25) of compounds was investigated sorted by concentration, of which 20 have the highest estimated concentration and 5 have the lowest estimated concentration. The results (Table 3) show the estimated concentration (without the 3-fold error from Pieke et al[24]) in both µmol L^{-1} and µg dm^{-2}, together with the structures obtained from MSC. Empty structures imply that no structure was predicted by using this database.

3.5.1 | Case: Insufficient information

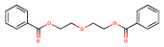
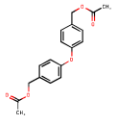
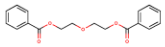
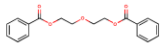
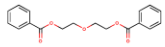
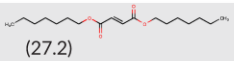
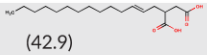
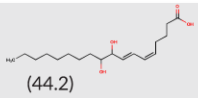
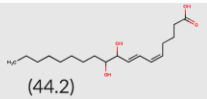
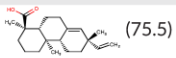
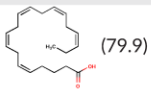
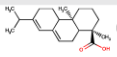
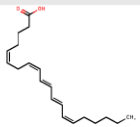
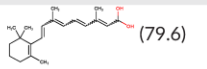
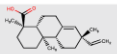
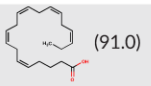
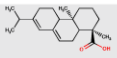
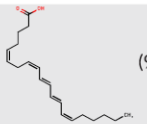
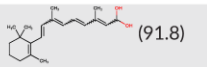
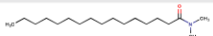
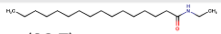
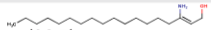
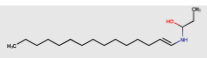
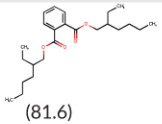
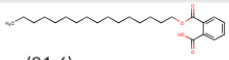
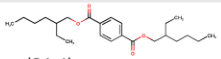
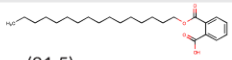
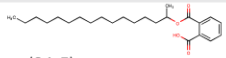
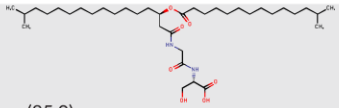
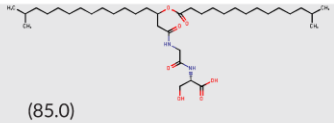
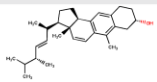
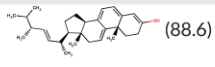
The evaluation of data was somewhat complicated when structure predictions contained little agreement and/or low scores. There is not sufficient evidence to prioritize further studies if databases predicted inconsistent or low-score structures, or if the estimation of concentration is unavailable. This is the case for compounds C.109 and C.110, C.65, C.39, C.34, C.36, and C.14, which are all at relatively low concentration with ambiguous structure assignment. The most straightforward option will be to treat the results as tentative and conclude that no consensus can be reached. This solution, however, is controversial as there is no prioritization, and it implies postponing the assessment. A second approach is to reanalyze the sample under more specific targeting conditions. However, this may not actually lead to a better result, especially if the actual chemical simply is not available in the databases and/or as reference standard. In light of a safety assessment, a more practical approach may be to

TABLE 3 Structural identifications via Molecular Structure Correlator (MSC) combined with semiquantitative data for a selection of compounds in a board sample

C.	RT (min)	Mass (amu)	Level (μg dm^{-2})	Level (μM)	Paper and Board Database (PBDB) (score)	REACH (score)	Agilent Extractables and Leechables (score)	ChemSpider (score)	PubChem (score)
1	1.696	344.875	1336	38.7	nd	nd	nd	nd	nd
2*	2.003	134.117	460	34.3	(93.2)	(93.7)	nd	(93.8)	(93.8)
3*	2.091	104.107	1653	159	(93.7)	(93.7)	nd	(93.9)	(93.9)
4	2.247	620.277	265	4.28	nd	nd	nd	(77.7)	(77.7)
6*	2.799	360.150	48	1.33	(89.1)	(89.5)	nd	(91.5)	(91.7)
9*	8.989	116.016	118	10.1	(81.5)	(81.5)	nd	(90.2)	(87.8)
10*	10.83	195.123	207	10.6	(94.3)	(91.4)	(94.3)	(94.0)	(94.0)
11*	12.21	239.149	173	7.24	nd	(93.8)	(93.8)	(94.2)	(94.2)
13*	12.52	208.133	156	7.48	nd	(79.7)	nd	(74.1)	(81.9)
14	12.92	114.091	60	5.29	(66.2)	(80.3)	(66.2)	(80.8)	(81.6)
15	13.10	183.065	143	7.82	(57.3)	(75.6)	nd	(80.5)	(80.5)
16*	13.27	169.0497	74	4.37	(77.2)	(77.6)	(77.2)	(79.5)	(79.4)
18*	13.34	300.201	169	5.61	(88.1)	(88.1)	nd	(88.3)	(88.3)
34	18.98	211.097	35	1.63	(86.7)	(86.9)	(80.3)	(87.6)	(88.2)
36	19.37	401.159	70	1.74	nd	(82.5)	nd	(87.1)	(88.1)
39	21.08	165.091	23	1.41	(70.4)	(74.6)	(72.2)	(75.2)	(75.4)

(Continues)

TABLE 3 (Continued)

C.	RT (min)	Mass (amu)	Level (μg dm^{-2})	Level (μM)	Paper and Board Database (PBDB) (score)	REACH (score)	Agilent Extractables and Leechables (score)	ChemSpider (score)	PubChem (score)
61	24.82	332.150	67	2.03	 (75.3)	 (66.9)	 (75.0)	 (75.0)	 (75.0)
65	25.28	330.264	1.1	0.03	 (27.2)	 (42.9)	 (39.0)	 (44.2)	 (44.2)
109	30.38	303.232	0.8	0.03	 (75.5)	 (79.9)	 (75.5)	 (79.4)	 (79.6)
110	30.52	303.233	0.8	0.03	 (85.8)	 (91.0)	 (85.7)	 (91.5)	 (91.8)
115*	31.47	284.295	51	1.78	 (93.6)	 (93.7)	 (93.6)	 (93.7)	 (93.7)
116*	30.66	284.295	51	1.78	 (80.8)	 (81.0)	 (80.8)	 (80.4)	 (80.8)
118*	32.17	391.285	198	5.07	 (81.6)	 (81.6)	 (81.6)	 (81.5)	 (81.5)
133	36.2	655.524	1.4	0.02	nd	nd	nd	 (85.0)	 (85.0)
132	36.13	393.315	0.7	0.02	nd	nd	nd	 (88.1)	 (88.6)

The top 20 highest concentrations and 5 lowest concentrations are included in the table. For each compound, its serial number, retention time (RT), accurate mass, and estimated concentration are reported. In addition, structure predictions for each database are given along with the prediction score (scale 0 to 100). Entries with no suitable prediction are marked with nd. Entries discussed in the text are marked with an asterisk (*).

use the available structures, in context of the worst-possible application. As human safety is the top concern, one could consider the worst-case chemical proposed as the leading case when in doubt, and then proceed to evaluate the worst-case scenario. This would often provide an overestimation of risk, but most risk-based evaluations consider that the preferred method. Unfortunately, there is not yet a solution if there is lacking or conflicting data, and it remains the decision of the user in light of the application to decide the best course of action.

3.5.2 | Case: Low expected risk

A number of found chemicals were classified as low risk. For example, the compounds with the highest estimated concentration C.3 and C.2 appear to be an amino-ethanol structure, with high concentrations (including uncertainty) in the interval from 0.15 to 4.5 mg dm^{-2}. The actual structure appears fairly certain across database predictions. These chemicals are often used in the production processes and have very little known toxicological effects attributed to them.[47] The LD50 for rats is remarkably high at 2.46 mL kg^{-1}.[48] Similarly, C.115 appears to be an aliphatic carboxylic amide, like stearamide, between 20 and 150 μg dm^{-2}. These types of compounds are also known to have a relative low toxicity.[49] Other examples of low-risk substances are compounds C.10, C.11, and C.18, which indicate polyethylene glycol at relatively high concentrations between 185 and 1630 μg dm^{-2}. Polyethylene glycol is historically known to have few issues associated with its intake.[50,51] Finally, compound C.6 seems to be a sugar-like compound that is unlike to be a risk.

3.5.3 | Case: Possible or uncertain expected risk

Some chemicals cannot be directly classified as a low or high risk. An example is compound C.16 where structure prediction suggests dehydroacetic acid, which is a fungicide and common industrial chemical, or maltol acetate, which also is an industry chemical, at concentrations from 25 to 225 μg dm^{-2}. The assigned structures to C.16 did not have high scores, so the suggested and actual structure is uncertain. Dehydroacetic acid is permitted in squash at maximum 65 ppm[52] and as adhesive in packaging[53] by the US FDA, but dehydroacetic acid was recently discovered in coatings of cheese in the EU, which caused a withdrawal.[54] There are no known regulations for maltol acetate. Because the identification did not have a high score and the use of the possible chemicals is debatable, C.16 is an example of a possible but no certain risk. Another case in this category is compound C.13. The structure proposed by 1 authentic database is ethyl *N*-benzyl-*β*-alaninate at 50 to 450 μg dm^{-2}, and PC predicted a similar compound, while CS proposed a lower-score alternative. Unfortunately, neither have toxicological data or known uses for the proposed structures while the concentration is high (2.5 to 22 μmol L^{-1}), which merit further studies. As discussed before, in cases of uncertainty, the safe approach is more appropriate, which marks classifying C.13 as possible risk.

3.5.4 | Case: Certain expected risk

Only a few of the predicted structures seem to call for attention. A notably entry is C.118, that all predictions indicate as a phthalate-like substance with good prediction scores (>80), concentration levels between 70 and 600 μg dm^{-2}, and good agreement between structure predictions. Due to the large number of rather similar phthalates, determining the exact structure is not possible. Phthalates are found frequently in FCMs due to world-wide use and migration of phthalates into food from plastics, but to a lesser extent also from a paper-based material, eg, recycled paper and board.[55] In addition, it has been argued that the continuous exposure to phthalates is often related to adverse health effects in humans,[56,57] so some risk may be expected. A final case discussed here is compound C.9, where the structure assignment points to methylisothiazolinone based on authentic database predictions with scores >80, supported by nonauthentic predictions showing highly similar structures. This substance is frequently used in cosmetics, water tanks, and paper mills as biocide or preservative and is considered to cause skin sensitization[58,59] and other effects.[60] It is among the substances with higher concentration (40 to 360 μg dm^{-2}), yet very little testing on chronic dietary exposure has been performed, so the presence in FCMs could be of concern.

4 | CONCLUSION

In this study, we investigated combining SQ with structural elucidation as a tool for exploration of chemical risk. We have shown that an exploration tool based on elucidative and SQ principles can be used on complex samples to give an essential overview of present chemical substances and concentration levels occurring, here demonstrated by using a paper and board FCM extract. The method is designed to be applicable on most samples analyzed by LC-MS because it does not rely on availability of a large number of standards. Instead, the method is capable to operate on structure libraries without extensive MS/MS data, and these can be created inexpensively as long as information on the sample type is available, for example, via collaboration with product manufacturers or authorities. In essence, the method permits chemical exploration of the world of chemically diverse and complex samples with simplicity, speed, and only few requirements for prior experiments or standardized methods.

Evidently, no "one size fits all" methodology for comprehensive unknown screening can be provided yet. Instead, exploration should be performed with the aim to obtain as much and as broad information as possible, followed by selecting the relevant information to evaluate further. Indeed, an indication of concentration and identity is already useful data to any field dealing with unknown substances. With both quantitative data and indications of chemical structures available, decisions can be taken at a much greater knowledge level than possible with the lack of knowledge that analysis of unknowns currently imposes.

The dependence on standards or existing knowledge on the sample is not completely eliminated because relevant structure databases are essential for results, and this marks 1 of the method limitations. Substances that are not (yet) in a database are hard to retrace, and predicting true unknowns with the current data tools is not possible. However, this is a first step toward a simplification of identification databases, having already reduced the need for fully recorded MS/MS spectra and instead relying on structure-only databases, but MS/MS spectra are still the preferred source of data. Ideally, future software tools will be able to understand MS data better and synthesize new

structures based purely on mass spectra. This means that truly unknown substances could be eligible for elucidation. In addition, it would be valuable to see DIA applied in this methodology for its improved precursor throughput, although the issues with low-quality fragment spectra would need to be resolved. Finally, the use of explorative data in different fields or in risk assessments needs to be evaluated.

ACKNOWLEDGEMENTS

The authors would like to thank Signe Helle Ptak for her contributions to data acquisition. The authors report no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors would like to thank the Technical University of Denmark (DTU) and the Danish Veterinary and Food Authority (DVFA) for financial support to enable this study.

ORCID

Eelco N. Pieke http://orcid.org/0000-0001-8939-6264

REFERENCES

1. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev*. 2007;26:51-78. https://doi.org/10.1002/mas.20108
2. Malik AK, Blasco C, Picó Y. Liquid chromatography-mass spectrometry in food safety. *J Chromatogr A*. 2010;1217:4018-4040. https://doi.org/10.1016/j.chroma.2010.03.015
3. Biedermann M, Grob K. Assurance of safety of recycled paperboard for food packaging through comprehensive analysis of potential migrants is unrealistic. *J Chromatogr A*. 2013;1293:107-119. https://doi.org/10.1016/j.chroma.2013.04.009
4. Viant MR, Sommer U. Mass spectrometry based environmental metabolomics: a primer and review. *Metabolomics*. 2013;9:144-158. https://doi.org/10.1007/s11306-012-0412-x
5. Richardson SD. Environmental mass spectrometry: emerging contaminants and current issues. *Anal Chem*. 2012;84:747-778. https://doi.org/10.1021/ac202903d
6. Thoren KL, Colby JM, Shugarts SB, Wu AHB, Lynch KL. Comparison of information-dependent acquisition on a tandem quadrupole TOF vs a triple quadrupole linear ion trap mass spectrometer for broad-spectrum drug screening. *Clin Chem*. 2016;62:170-178. https://doi.org/10.1373/clinchem.2015.241315
7. Chen J. A worldwide food safety concern in 2008—melamine-contaminated infant formula in China caused urinary tract stone in 290,000 children in China. *Chin Med J (Engl)*. 2009;122:243-244. https://doi.org/10.3760/cma.j.issn.0366-6999.2009.03.001
8. Handford CE, Campbell K, Elliott CT. Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. *Compr Rev Food Sci Food Saf*. 2016;15:130-142. https://doi.org/10.1111/1541-4337.12181
9. Grob K. Work plans to get out of the deadlock for the safety assurance of migration from food contact materials? A proposal. *Food Control*. 2014;46:312-318. https://doi.org/10.1016/j.foodcont.2014.05.044
10. Schymanski EL, Singer HP, Slobodnik J, et al. Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. *Anal Bioanal Chem*. 2015;407:6237-6255. https://doi.org/10.1007/s00216-015-8681-7
11. Bologa CG, Olah MM, Oprea TI. Chemical database preparation for compound acquisition or virtual screening. *Methods Mol Biol*. 2006;316:375-388. https://doi.org/10.1385/1-59259-964-8:375
12. Muncke J, Backhaus T, Geueke B, et al. Scientific challenges in the risk assessment of food contact materials. *Environ Health Perspect*. 2017;125:1-9. https://doi.org/10.1289/EHP644
13. Ibáñez M, Sancho JV, Pozo ÓJ, Niessen W, Hernández F. Use of quadrupole time-of-flight mass spectrometry in the elucidation of unknown compounds present in environmental water. *Rapid Commun Mass Spectrom*. 2005;19:169-178. https://doi.org/10.1002/rcm.1764
14. Kind T, Fiehn O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics*. 2006;7:234. https://doi.org/10.1186/1471-2105-7-234
15. Steinbeck C. Recent developments in automated structure elucidation of natural products. *Nat Prod Rep*. 2004;21:512-518. https://doi.org/10.1039/b400678j
16. Foti C, Alsante K, Cheng G, Zelesky T, Zell M. Tools and workflow for structure elucidation of drug degradation products. *TrAC - Trends Anal Chem*. 2013;49:89-99. https://doi.org/10.1016/j.trac.2013.06.005
17. Elyashberg M, Blinov K, Molodtsov S, Smurnyy Y, Williams AJ, Churanova T. Computer-assisted methods for molecular structure elucidation: realizing a spectroscopist's dream. *J Chem*. 2009;1:1-26. https://doi.org/10.1186/1758-2946-1-3
18. Kind T, Fiehn O. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal Rev*. 2010;2:23-60. https://doi.org/10.1007/s12566-010-0015-9
19. Bletsou AA, Jeon J, Hollender J, Archontaki E, Thomaidis NS. Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment. *TrAC - Trends Anal Chem*. 2015;66:32-44. https://doi.org/10.1016/j.trac.2014.11.009
20. Krauss M, Singer H, Hollender J. LC-high resolution MS in environmental analysis: from target screening to the identification of unknowns. *Anal Bioanal Chem*. 2010;397:943-951. https://doi.org/10.1007/s00216-010-3608-9
21. Hoffmann T, Krug D, Hüttel S, Müller R. Improving natural products identification through targeted LC-MS/MS in an untargeted secondary metabolomics workflow. *Anal Chem*. 2014;86:10,780-10,788. https://doi.org/10.1021/ac502805w
22. García-Reyes JF, Hernando MD, Molina-Díaz A, Fernández-Alba AR. Comprehensive screening of target, non-target and unknown pesticides in food by LC-TOF-MS. *TrAC Trends Anal Chem*. 2007;26:828-841. https://doi.org/10.1016/j.trac.2007.06.006
23. Muncke J. Endocrine disrupting chemicals and other substances of concern in food contact materials: an updated review of exposure, effect and risk assessment. *J Steroid Biochem Mol Biol*. 2011;127:118-127. https://doi.org/10.1016/j.jsbmb.2010.10.004
24. Pieke EN, Granby K, Trier X, Smedsgaard J. A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry. *Anal Chim Acta*. 2017;975:30-41. https://doi.org/10.1016/j.aca.2017.03.054
25. Bengtström L, Rosenmai AK, Trier X, et al. Non-targeted screening for contaminants in paper and board food-contact materials using effect-directed analysis and accurate mass spectrometry. *Food Addit Contam Part A*. 2016;33:1080-1093. https://doi.org/10.1080/19440049.2016.1184941
26. Rosenmai AK, Bengtström L, Taxvig C, Trier X, Petersen JH, Svingen T, Binderup M-L, Vugt-Lussenburg van BMA, Dybdahl M, Granby K, Vinggaard AM, An effect-directed strategy for characterizing emerging chemicals in food contact materials made from paper and board, Submitted. 2017. https://doi.org/10.1016/j.fct.2017.05.061
27. Ibáñez M, Sancho JV, Hernández F, McMillan D, Rao R. Rapid non-target screening of organic pollutants in water by ultraperformance liquid chromatography coupled to time-of-light mass spectrometry. *TrAC - Trends Anal Chem*. 2008;27:481-489. https://doi.org/10.1016/j.trac.2008.03.007
28. Schymanski EL, Singer HP, Longrée P, et al. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. *Environ Sci Technol*. 2014;48:1811-1818. https://doi.org/10.1021/es4044374

29. Bengtström L. Chemical identification of contaminants in paper and board food contact materials, Technical University of Denmark, 2014.
30. Hill AW, Mortishire-Smith RJ. Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Commun Mass Spectrom*. 2005;19:3111-3118. https://doi.org/10.1002/rcm.2177
31. Agilent Technologies Inc., Agilent MassHunter Molecular Structure Correlator (MSC) software; 2011.
32. Pivnenko K, Eriksson E, Astrup TF. Waste paper for recycling: overview and identification of potentially critical substances. *Waste Manag*. 2015;45:134-142. https://doi.org/10.1016/j.wasman.2015.02.028
33. U.S. Food and Drug Administration, List of indirect additives used in food contact substances; 2012. https://www.accessdata.fda.gov/scripts/fdcc/?set=IndirectAdditives
34. European Chemicals Agency, Preparation of an inventory of substances suspected to meet REACH Annex III criteria; 2016. https://echa.europa.eu/information-on-chemicals/annex-iii-inventory (accessed April 25, 2017)
35. Agilent Technologies Inc., MassHunter Extractables and Leachables PCDL; 2015.
36. Kim S, Thiessen PA, Bolton EE, et al. PubChem substance and compound databases. *Nucleic Acids Res*. 2016;44:D1202-D1213. https://doi.org/10.1093/nar/gkv951
37. Thomas MC, Dunn SR, Altvater J, Dove SG, Nette GW. Rapid identification of long-chain polyunsaturated fatty acids in a marine extract by HPLC-MS using data-dependent acquisition. *Anal Chem*. 2012;84:5976-5983. https://doi.org/10.1021/ac3006523
38. Schwudke D, Liebisch G, Herzog R, Schmitz G, Shevchenko A. Shotgun Lipidomics by tandem mass spectrometry under data-dependent acquisition control. *Methods Enzymol*. 2007;433:175-191. https://doi.org/10.1016/S0076-6879(07)33010-3
39. Gillet LC, Navarro P, Tate S, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics*. 2012;11:O111.016717. https://doi.org/10.1074/mcp.O111.016717
40. Egertson JD, MacLean B, Johnson R, Xuan Y, MacCoss MJ. Multiplexed peptide analysis using data-independent acquisition and skyline. *Nat Protoc*. 2015;10:887-903. https://doi.org/10.1038/nprot.2015.055
41. Zhu X, Chen Y, Subramanian R. Comparison of information-dependent acquisition, SWATH, and MS(All) techniques in metabolite identification study employing ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Anal Chem*. 2014;86:1202-1209. https://doi.org/10.1021/ac403385y
42. Xiao JF, Zhou B, Ressom HW. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC - Trends Anal Chem*. 2012;32:1-14. https://doi.org/10.1016/j.trac.2011.08.009
43. Hopfgartner G, Tonoli D, Varesio E. High-resolution mass spectrometry for integrated qualitative and quantitative analysis of pharmaceuticals in biological matrices. *Anal Bioanal Chem*. 2012;402:2587-2596. https://doi.org/10.1007/s00216-011-5641-8
44. Suciu NA, Tiberto F, Vasileiadis S, Lamastra L, Trevisan M. Recycled paper-paperboard for food contact materials: contaminants suspected and migration into foods and food simulant. *Food Chem*. 2013;141:4146-4151. https://doi.org/10.1016/j.foodchem.2013.07.014
45. Ozaki A, Yamaguchi Y, Fujita T, Kuroda K, Endo G. Chemical analysis and genotoxicological safety assessment of paper and paperboard used for food packaging. *Food Chem Toxicol*. 2004;42:1323-1337. https://doi.org/10.1016/j.fct.2004.03.010
46. Grob K, Biedermann M, Scherbaum E, Roth M, Rieger K. Food contamination with organic materials in perspective: packaging materials as the largest and least controlled source? A view focusing on the European situation. *Crit Rev Food Sci Nutr*. 2006;46:529-535. https://doi.org/10.1080/10408390500295490
47. U.S. Environmental Protection Agency, 2-[2-(Dimethylamino)ethoxy]ethanol, Chem. Dashboard. (n.d.). https://comptox.epa.gov/dashboard/DTXSID1027427 (accessed July 22, 2017).
48. Carpenter CP, Weil CS, Smyth HF. Range-finding toxicity data: list 8. *Toxicol Appl Pharmacol*. 1974;28:313-9. http://www.ncbi.nlm.nih.gov/pubmed/4854023. accessed June 22, 2017
49. Patty FA. Industrial hygiene and toxicology. Vol. II. 2nd rev. ed. Frank A. Patty, Editor. John Wiley & Sons, Inc., 605 Third Ave., New York 16, N. Y., 1963. xxxi + 1546 pp. 15 1/2 × 24 cm. Price $40.00, 1963. https://doi.org/10.1002/jps.2600520932
50. Working PK, Newman MS, Johnson J, Cornacoff JB. Safety of poly(ethylene glycol) and poly(ethylene glycol) derivatives. *ACS Symp Ser*. 1997;680:45-57. https://doi.org/10.1021/bk-1997-0680.ch004
51. Smyth HF, Carpenter CP, Weil CS. The toxicology of the polyethylene glycols*. *J Am Pharm Assoc (Scientific Ed)*. 1950;39:349-354. https://doi.org/10.1002/jps.3030390615
52. Federal Register, Food additives permitted for direct addition to food for human consumption, 21 C.F.R. § 172.130, United States of America, 1988.
53. Federal Register, Indirect food additives: adhesives and components of coatings. Subpart B—substances for use only as components of adhesives, 21 C.F.R. § 175.105, United States of America, 2017.
54. European Commission, 2017.0798: unauthorised food additive E 265—dehydroacetic acid (presence/kg) in external coating for cheeses from Spain, 2017. https://webgate.ec.europa.eu/rasff-window/portal/index.cfm?event=notificationDetail&NOTIF_REFERENCE=2017.0798
55. Binderup ML, Pedersen GA, Vinggaard AM, Rasmussen ES, Rosenquist H, Cederberg T. Toxicity testing and chemical analyses of recycled fibre-based paper for food contact. *Food Addit Contam*. 2002;19 Suppl:13-28. https://doi.org/10.1080/02652030110089878
56. Heudorf U, Mersch-Sundermann V, Angerer J. Phthalates: toxicology and exposure. *Int J Hyg Environ Health*. 2007;210:623-634. https://doi.org/10.1016/j.ijheh.2007.07.011
57. Muncke J. Phthalates, food packag. Forum; 2012. http://www.foodpackagingforum.org/food-packaging-health/phthalates (accessed February 7, 2017)
58. Yazar K, Lundov MD, Faurschou A, et al. Methylisothiazolinone in rinse-off products causes allergic contact dermatitis: a repeated open-application study. *Br J Dermatol*. 2015;173:115-122. https://doi.org/10.1111/bjd.13751
59. Isaksson M, Hauksson I, Hindsén M, Pontén A, Svedman C, Bruze M. Methylisothiazolinone contact allergy is rising to alarming heights also in southern Sweden. *Acta Derm Venereol*. 2015;95:31-34. https://doi.org/10.2340/00015555-1844
60. Connor TH, Tee PG, Afshar M, Connor KM. Mutagenicity of cosmetic products containing Kathon®. *Environ Mol Mutagen*. 1996;28:127-132. https://doi.org/10.1002/(SICI)1098-2280(1996)28:2<127::AID-EM9>3.0.CO;2-C

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Pieke EN, Smedsgaard J, Granby K. Exploring the chemistry of complex samples by tentative identification and semiquantification: A food contact material case. *J Mass Spectrom*. 2018;1–13. https://doi.org/10.1002/jms.4052

3.6. Manuscript C: Risk prioritization strategies

Title:	Prioritization before Risk Assessment: the viability of uncertain data on food contact materials
Journal:	Manuscript submitted to *Regulatory Toxicology and Pharmacology*.
Author(s):	Eelco N. Pieke, Kit Granby, Bruno Teste, Jørn Smedsgaard, Gilles Rivière
Keywords:	risk prioritization; FCM; structure assessment; semi-quantification; exposure assessment; hazard assessment
Contact:	Eelco N. Pieke: enpi@food.dtu.dk, eelco.pieke@gmail.com
Contributions:	Experiments, data analysis, and writing of the manuscript were done by EP. The decision process and logic behind it were developed by EP, GR, and BT collaboratively. The concept idea of the research, proofreading, and revision was the responsibility of all authors (EP, BT, GR, JS, KG).
Supplement:	Supplementary material to this manuscript in preparation is in Excel format, hence excluded from the Appendix. The material can be downloaded electronically from a dedicated Google Drive link.
	Shortened link: https://goo.gl/Q21bJF
	Full length link: https://drive.google.com/uc?export=download&id=1ICBEBqREMVVNVpDCXc8VNW6Mhcawxq3E

Prioritization before Risk Assessment: the viability of uncertain data on food contact materials

Eelco N. Pieke [*1], Kit Granby [1], Bruno Teste [2], Jørn Smedsgaard [1], Gilles Rivière [2]

[1] Research Group for Analytical Food Chemistry, National Food Institute, Technical University of Denmark (DTU), Kemitorvet, 2800 Kgs. Lyngby, Denmark

[2] French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Maisons-Alfort Cedex, France

The views and opinions expressed in this article are those of the authors.

[*] Corresponding Author: Eelco Nicolaas Pieke; eelco.pieke@gmail.com; +45 51 86 42 05

Abbreviations: ADI: applicability domain index; CMR: carcinogenicity, mutagenicity, reproductive toxicity; FCM: food contact material; (N)IAS: (non-)intentionally added substances; QSAR: quantitative structure-activity relationship; RA: risk assessment; TMC: total migratable content; TTC: threshold of toxicological concern.

Abstract

The shortage of data on non-intentionally added substances (NIAS) present in food contact materials (FCM) limits the ability to ensure food safety. Recent strategies in analytical method development allow investigating NIAS by using chemical exploration; but this has not been sufficiently investigated in risk assessment context. Here, exploration is applied on two paperboard FCM samples followed by risk prioritization for chemicals that can potentially migrate to food. Concentration estimates from exploration are converted into a tentative exposure assessment, while predicted chemical structures are assessed using quantitative structure-activity relationships (QSAR) models for carcinogenicity, mutagenicity, and reproductive toxicity. A selection of 60 chemical compounds from two FCM is assessed by four risk assessors to classify chemical compounds based on probable risk. For 60% of cases, the assessors classified compounds as either high priority or low priority. Unclassified compounds are due to disagreements between experts or due to a lack of data. Among the high priority substances were high concentration compounds, benzophenone derivatives, and dyes. The low priority compounds contained e.g. oligomers from plasticizers and linear alkane amides. The classification scheme was demonstrated to provide valuable information based on tentative data, able to prioritize discovered chemical compounds for pending risk assessment.

***Keywords**: risk prioritization; FCM; structure assessment; semi-quantification; exposure assessment; hazard assessment*

1. Introduction

An all-time debated source of human health risk is the chronic long-term exposure to chemical compounds due to presence in food. One important source of chemicals in food is due their migration from food packaging materials (Arvanitoyannis and Bosnea, 2004; Castle, 2006; Grob, 2014; Jickells, 2007). Investigations into the safety of food contact materials, especially those non-harmonized in legislations affirm that thousands of possible chemicals may be present in paper and board packaging alone (Bengtström et al., 2016; Biedermann et al., 2011; Biedermann and Grob, 2013; Binderup et al., 2002; Ozaki et al., 2005; Triantafyllou et al., 2007), while only a minor fraction of these chemicals have been successfully identified and risk assessed (Geueke et al., 2014). In addition, some chemical compounds originating from paper and board have been shown to have biological activity and therefore are of concern to human health (Bengtström et al., 2016; Honkalampi-Hämäläinen et al., 2010; Rosenmai et al., 2017). As a result, packaging contaminates food with uncharacterized chemicals that may exert significant adverse effects (Gallart-Ayala et al., 2013), yet the extent or nature of the chemical migration is not well-defined because it depends on many parameters, e.g., the packaging material, contact type, temperature, and food type (Barnes et al., 2007; Hauder et al., 2013; Poças et al., 2011).

The regular approach to chemicals in food is to perform a specific risk assessment (RA) for each individual chemical, see Figure 1. However, determining the risk character is convoluted when there is a shortage of available data on migrating compounds (Skjevrak et al., 2005). For the commonly investigated Intentionally Added Substances (IAS), there is often data available from prior research or via accredited methods, but for the more elusive Non-Intentionally Added Substances (NIAS), there is rarely relevant data (Driffield et al., 2016; Grob, 2014; Koster et al., 2015; Pivnenko et al., 2015). In fact, most NIAS do not have assigned chemical structures, concentration data, or characterization of hazards, and few methods are capable to obtain these data for such a large group of chemicals. The sheer amount of possible compounds prohibits performing a dedicated safety evaluation on each compound, and it significantly challenges analytical methods to provide adequate data to perform RA. Consequently, some researchers recently concluded that the existing frameworks RA are inadequate to ensure food safety (Muncke et al., 2017).

The knowledge gap for NIAS and other chemical compounds needs to be reduced in order to incorporate them in legislation. We recently investigated the use of explorative methods to discover chemical compounds in FCM and concluded that untargeted analytical strategies are useful and efficient to estimate the concentration and chemical structure of unknowns (Pieke et al., 2018,

2017). However, it is unrealistic to perform comprehensive analysis on all compounds discovered via exploration (Biedermann and Grob, 2013), so some sort of risk prioritization is required to ensure resources are dedicated to compounds most likely to introduce adverse health effects (Barlow, 2009). One of the core requirements of risk prioritization is to determine a risk character of a chemical compound that is in line with common risk assessment (Guillén et al., 2012; Schymanski et al., 2014).

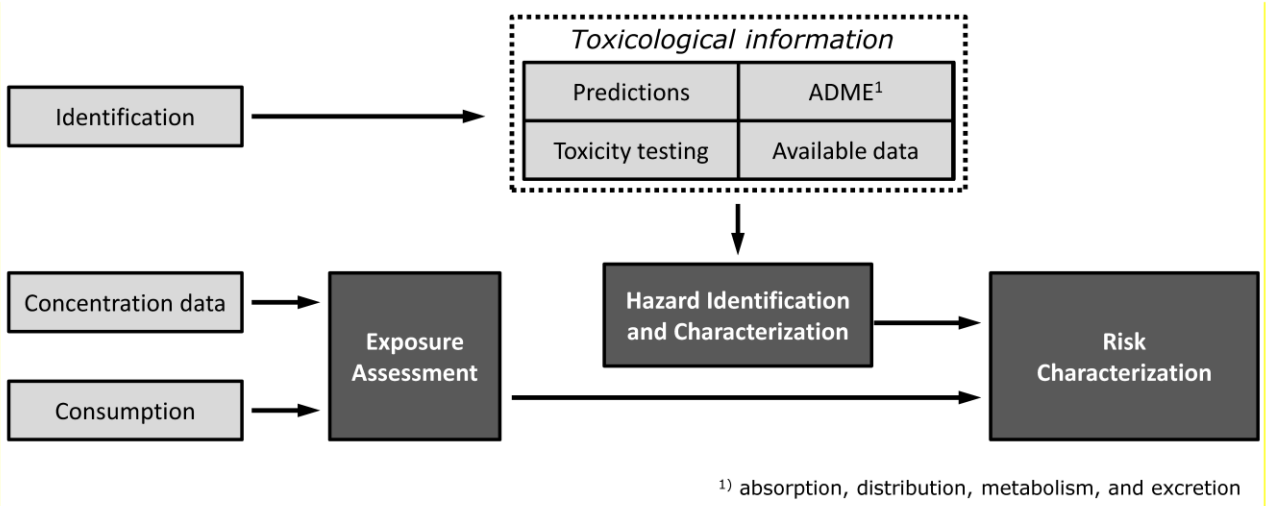


Figure 1: The characterization of risk is a result from highly specific data, which are combined into exposure assessment and hazard identification and characterization. Obtaining the data needed for these assessments is resource-intensive, especially for larger number of compounds with existing data gaps.

The Threshold of Toxicological Concern (TTC) concept has been adopted within European Union (EU) legislation as a tool to better deal with NIAS and other unknown chemical compounds (EFSA and WHO, 2016; Kroes et al., 2004). The TTC concept uses tentative exposure data to assess if intake is below an accepted threshold of no concern, defined by assigning a Cramer class based on the chemical structure. Hence, TTC is an exposure-based assessment has been applied in a strategy for NIAS discovery by Koster et al. (2014), and may be viable for the exploration approaches shown recently by Pieke et al. (2018, 2017). However, TTC requires compounds to show no genotoxicity (e.g., mutagenicity) or do not exceed an exposure of 0.15 µg person^{-1} day^{-1}. Hence, if the exposure exceeds 0.15 µg person^{-1} day^{-1} genotoxicity testing is required, which is problematic for the large number of compounds that may exceed this threshold. Quantitative Structure Activity Relationship (QSAR) modeling of chemical hazard may provide substitute toxicity data if testing is prohibitive, which has successfully been applied to FCM for hazard-based

assessment and prioritization by van Bossuyt et al. (2017). However, a limitation in hazard-based approaches is that these generally do not always consider occurrence, migration, and exposure.

In present article, we aim to develop a strategy for risk prioritization of chemical compounds in FCM following their prior discovery by exploration strategies. For this, we aim to establish the link between tentative data, e.g., semi-quantification and tentative identification, and existing hazard-based and exposure-based assessment tools, e.g., TTC and QSAR, to perform qualitative risk prioritization. The risk prioritization tool is designed to mimic conventional risk assessment, identically obtaining exposure assessment and hazard assessment, followed by an expertise decision on risk. The tool proposed here is not suggested as a definite method for performing qualitative risk prioritizations, but emphasizes the need and possibility for using tentative data in a risk assessment perspective.

2. Methods

2.1. Analysis

Analysis is performed as reported in two previous studies (Pieke et al., 2018, 2017). In brief: UHPLC-MS was performed on an Agilent 1290 system (Agilent Technologies, Santa Clara, CA, USA). Two UHPLC columns were used serially (Phenomenex Luna Omega Polar C18 100 Å, 1.6 µm, 100 x 2.1 mm (Phenomenex, Denmark) and Waters ACQUITY UPLC CSH C18 130 Å, 1.7 µm, 100 x 2.1 mm (Waters, Denmark)). Mass analysis post-UHPLC was performed using an Agilent 6550 Quadrupole-Time of Flight (Q-TOF) mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) equipped with Agilent JetStream electrospray ionization (ESI) interface. The optimization, operating conditions, data collection, and data interpretation are discussed in previous studies (Pieke et al., 2018, 2017).

Semi-quantification was used to determine estimated concentration of chromatographically eluting chemical substances within a threefold error (Pieke et al., 2017). The semi-quantification was limited to the 1200 largest eluting peaks and to detectable analytes in the sample. The chemical structures of compounds in the extract of the sample were tentatively identified by recording fragmentation spectra and using structure correlations to propose a best matching chemical compound (Pieke et al., 2018). The tentative identification results (five predicted chemical structures) were later combined with the semi-quantification results by comparing exact mass and retention times.

2.2. Construction and evaluation of a decision unit

The decision unit for risk prioritization and risk profile classification boundaries was designed by discussing with various interdisciplinary experts at the “Risk assessment for substances and processes submitted to human food regulation” panel at the French Agency for Food, Environmental and Occupational Health & Safety (ANSES). Based on the feedback of the expert panel, the decision unit was designed to involve automation (data-based decisions) and manual assessing (expertise-based decisions).

To test the classification scheme the semi-quantification and tentative identification results of two paper and board FCM samples were used. 30 identified compounds were selected per sample, of which 20 from ESI+ and 10 from ESI−, resulting in a total test set of 60 chemical compounds. The selected entries were evaluated to avoid including chemicals which would not produce a meaningful classification, e.g., no predicted structures. The chemical compounds were gathered in a single Excel-based program available as Supplementary Information. The file contained the predicted structure(s), QSAR consensus and individual prediction by the QSAR models, estimated intake compared to a defined threshold (TTC Excess), absolute estimated intake, and finally the predicted Cramer class from the TTC methodology (Cramer et al., 1976).

The 60 entries were assessed by four individual assessors using the decision unit. Each assessor was tasked with classifying 60 compounds via the decision unit into one of the three risk profile classes: high expected risk ([A]), low expected risk ([B]), or insufficient data ([C]). Prior to classification, each assessor obtained documented instructions on how to work with the Excel program and decision unit. Following, each assessor individually classified the chemical subset. The assessors were specifically instructed to use the decision unit as much as possible, but also to deviate from the decision unit in case their opinion would conflict with the decision unit result.

2.3. Quantitative Structure-Activity Relationship (QSAR)

Possible adverse health effects of tentatively identified chemicals were predicted using Quantitative Structure-Activity Relationship (QSAR) models and software. Three endpoints were defined: Carcinogenicity, Mutagenicity, and Reproductive Toxicity; abbreviated as CMR. To predict CMR activity, the VEGA-QSAR platform (Benfenati et al., 2013) was employed using the included four models for carcinogenicity (CAESAR, ISS, IRFMN/Antares, IRFMN/ISSCAN-CGX), four Ames models for mutagenicity (CAESAR, SarPy/IRFMN, ISS, KNN/Read-Across), and two models for reproductive toxicity (PG Toxicity Library, CAESAR). The VEGA-QSAR platform predicted only the likely activity of the chemical compound, not a dose-response relationship. The Applicability

Domain Index (ADI) was used as performance criterion to define the quality of prediction (Istituto di Ricerche Farmacologiche Mario Negri Milano, 2017).

An in-house solution was applied to integrate the QSAR results from VEGA-QSAR into the decision unit. For each prediction the result, active (+) or inactive (−), and the prediction applicability domain index (ADI) were extracted. A QSAR consensus score was calculated from each endpoint results and accompanying ADI score. Each QSAR model applied contributed a fraction to the total consensus score, e.g., for carcinogenicity four models were used, so each model contributed a maximum ±0.25 score. The score was corrected for lower ADI (i.e. prediction certainty) so that less certain prediction had lower weight in the consensus score. The consensus was calculated by the biggest sum for either positive values (active effect) or negative values (no active effect). Hence, the result of the consensus calculation was a value between −1 and +1, in which a negative number indicated a predicted non-active effect and a positive number indicated an active effect. Values closer to the extremes −1 or +1 were results of good agreement between model predictions on the same endpoint; values close to zero indicated a poor agreement.

2.4. Sample selection and preparation

Two paper and board samples were analyzed. The first sample was a recycled unused carton pizza box, similar to the sample used in (Pieke et al., 2018), because these are known to contain many extractable compounds (Bengtström et al., 2016). The second sample was a carton sheet part of the packaging of luxury chocolates. The sheet was folded in a way to compartment separate chocolates, thereby being in contact with the chocolates on multiple sides. The sheet was unfolded before preparing the sample. From each sample, a 10 cm x 10 cm (1 dm^2) sample was prepared using a clean knife.

Each of the 10 cm x 10 cm samples were cut into four identically sized pieces of 2.5 cm x 10 cm and inserted into a glass vessel. The Total Migratable Content (TMC, see 3.1) was recovered by adding 100 mL of warm (40–50 °C) Food Simulant D1 (50 v/v ethanol water) to the vessel. The vessel was closed and sealed into a calibrated thermostat compartment at 40°C. The setup was left to soak for 24 hours, after which the food simulant was removed from the vessel and allowed to cool to room temperature. Proceeding, the food simulant was filtered and prepared for semi-quantification and identification as described in recent work (Pieke et al., 2017).

3. Results and discussion

3.1. The total migratable content (TMC)

TENAX is frequently used for paper and board migration testing, but shows different behaviour for polar and non-polar compounds depending on vapour pressure (Poças et al., 2011). In addition, the use of TENAX implies limited direct contact transfer, but it has been shown that migration from direct contact is not negligible for paper and board and migration can occur even for non-volatile compounds (Biedermann-Brem et al., 2012; Triantafyllou et al., 2007). In addition, there are examples of food contact by paperboard that question the assumption of exclusively dry indirect migration like TENAX simulates, e.g., pizza boxes, snacks, fast food, or fruits (Binderup et al., 2002; Bradley, 2006). Some of the test methods presented in Commission Regulation no. 10/2011 regarding plastic FCM might be used for paper FCM (European Parliament and Council of the European Union, 2011). However, the usage pattern of paper and plastic is different: plastic is often used in longer term storage of a wide array of products, whereas paper is used for shorter contact times or for freezing boxes, e.g., fast food or prepared foods. To use the plastic migration test conditions (10 days at 40°C) on paper materials may not be representative, and this perception is supported by U.S. FDA recommendations proposing migration studies on uncoated paper at 40°C for 24 hours (FDA, 2007). Therefore, here a smaller testing window of 24 hours was used for paper and board FCM.

No migration tests exist for paper and board FCM, so the intended food simulant should ideally have a broad extractable range and compatibility for further analysis by LC-MS. From the analytical and investigative perception in this study these two criteria are met using water/ethanol mixtures. However, using water/ethanol mixtures in contact with paper or board for 24 hours is not a migration test. Noticeably compared to plastics, paper is porous, inhomogeneous, and poorly resistant to liquids, which lead to large numbers of extractives (FDA, 2007). Hence, we consider these testing conditions to be somewhat more severe than a migration test, yet less severe than a complete extraction, as the material integrity is preserved. Instead, we defined the tests performed here as Total Migratable Content (TMC). The TMC contains chemical compounds from the FCM that can reasonably be expected to migrate into food, but is an overestimation of actual-use migration levels. Consequently, TMC implies a thorough screening of extractable chemical compounds, which when observed in the simulant can – but not necessarily will – migrate.

3.2. Risk characterization of tentative data

3.2.1. Tentative hazard identification

The identification technique recently published allows high-throughput tentative elucidation of the chemical structure of a potentially unknown compound, but does not provide an unambiguous chemical structure, instead presenting several chemical candidate structures (Pieke et al., 2018). Finding existing toxicity data on multiple structures is convoluted. Here, we applied predictive hazard modeling by Quantitative Structure-Activity Relationship (QSAR). Because QSAR assumes that similar molecules likely have the same effect (Raies and Bajic, 2016), it is compatible with the concept of tentative identification: if the structure prediction closely resembles the actual molecule, the QSAR prediction results will likely be similar. A precaution in using QSAR is that the application of different models can produce different and sometimes conflicting results. To minimize the leverage of a single model in cases where the model performed inadequate, several models are used in parallel for the same endpoint on the same molecule. This presented a battery of results for each prediction, of which the average prediction can cancel the effect of single outliers or false predictions (Benfenati et al., 2013). Consequently, the average prediction of these models (the QSAR consensus) is more likely to contain accurate information than any model alone.

The in-house consensus model closely mimics those presented by the VEGA software (Benfenati et al., 2013). To evaluate thresholds for consensus relevance, the consensus approach was applied on chemical compounds of IARC's Group 1, 2A, and 2B of known, probable, and possible carcinogens list (International Agency for Research on Cancer, 2017). To ensure a strict consensus, the VEGA QSAR results were compared to the assumption chemicals on the extracted carcinogen list (n = 204) are active carcinogens. The threshold for false negative prediction results was set to 2.5%. The results indicated that a consensus score of at least +0.40 was required to minimize the chance of a false negative prediction. This value can be logically evaluated to make sense: +0.40 only be obtained by two or more models predicting the same results, considering the best-case predictions can only contribute +0.25 per model.

Characterizing the hazard as demonstrated here is limited to interpretation of the QSAR evaluation on Carcinogenicity, Mutagenicity, and Reproductive Toxicity (CMR) prediction models. However, there are other toxicity endpoints that influence the probable risk of a substance, e.g., hepatotoxicity, neurotoxicity, or endocrine disruption, but these are not well-studied and few broad range QSAR models exist for these. In addition, CMR is already incorporated in the TTC approach, and a CMR substance has the most strict exposure limit (0.15 µg person^{-1} day^{-1}). Consequently, a

reliable CMR alert from QSAR is sufficient to assign the hazard characterization of the substance as a high priority substance.

3.2.2. Tentative exposure assessment

Semi-quantification reports a concentration per volume or per surface with a maximum uncertainty of threefold (Pieke et al., 2017). The content per surface area cannot be directly used for assessing exposure, because the contact factor of the FCM is usually unknown. According to European regulations, it is usually considered that an average person has a body weight of 60 kg and consumes 1 kg of food containing the substance daily in contact with a plastic FCM with 6 dm^2 packaging (European Parliament and Council of the European Union, 2011). However, other studies have shown that actual food contact is likely in the range of 10–14 dm^2 (Bouma et al., 2003; Duffy et al., 2007; ILSI Europe Packaging Material Task Force, 1996), and in some cases even higher at 30–40 dm^2 (Bouma et al., 2003). However, paper and board FCM constitute only a limited fraction of 10–20% of the total used packaging materials (Duffy et al., 2007; FDA, 2007). Hence, by applying a usage reduction factor of 10–20% on the worst-case estimate of packaging results in an estimated contact range of 3–8 dm^2 $person^{-1}$ day^{-1}, which is close to EFSA assumptions of 6 dm^2 $person^{-1}$ day^{-1} and likely to be sufficiently conservative. Hence, by adopting the standard used by EFSA, the semi-quantitative concentration data in µg dm^{-2} can be converted to µg $person^{-1}$ day^{-1} by multiplying with 6 dm^2 $person^{-1}$ day^{-1}.

Due to the similarities of the data in this study with that needed in the Threshold of Toxicological Concern (TTC) approach (EFSA and WHO, 2016; Kroes et al., 2004), parts of the TTC strategy are applicable here. Notably, the division of chemical compounds into Cramer classes is useful, because it provides an exposure limit below which likelihood of adverse effect is considered to be very low: for Class I compounds max. 1800 µg $person^{-1}$ day^{-1}; Class II compounds max. 540 µg $person^{-1}$ day^{-1}; and for Class III compounds max. 90 µg $person^{-1}$ day^{-1} (Cramer et al., 1976; Kroes et al., 2004).

Here, the exposure (in µg $person^{-1}$ day^{-1}) from semi-quantification is compared to the limit imposed by the Cramer class assignment calculated from the tentative identification. The result is the TTC Excess factor, which is the fraction of exposure compared to the threshold, i.e., TTC Excess of 100% means the predicted intake is equal the threshold from the TTC approach. However, not every structural prediction was successful where the structure of the chemical compound was unresolved or largely uncertain. For those cases, we considered the worst-case scenario excluding carcinogenicity by assigning Class III. Considering uncertainty of the concentration estimate, worst case ± 3-fold, TTC Excess above 300% would most probably

indicate that the TTC would be exceeded, whereas below 33% indicates that most probably the TTC would not be exceeded. Values within this range are to be decided on a case by case basis.

3.3. Risk Prioritization based on tentative data

3.3.1. Risk profile classification

Prioritization based on semi-quantification and structure predictions is convoluted: even if a complete “picture” of exposure and hazard is available, these still contain considerable uncertainty. Consequently, it is not recommended to perform a quantitative risk assessment (RA) on these data, and it should be more feasible to use qualitative risk prioritization, where all variables are evaluated stepwise in order to determine a likely risk profile of the chemical compound. While it would be convenient to classify chemical compounds into subgroups with well-described risk profiles and priority, it is practically less achievable. Here, prioritization is likely to produce only broad risk profiles of chemical compounds, because uncertainties in the estimated hazard and estimated exposure assessment do not support clear boundaries for risk profile classes. The concept for broad classification with uncertain data is not new, as the TTC approach effectively only uses two Cramer classes: Low (Class I) and High (Class III), supplemented by the highly specific Intermediate (Class II).

Only three classes are used in this prioritization approach, shown in Figure 2: [A] — Compound of Direct Concern; [B] — Compound of Lesser Concern; and [C] — insufficient information available. It could be argued that a class between [A] and [B] is needed that defines moderate concern. However, more than two risk profile classes require the capability to define a clear distinction between classes. This is not straightforward due to uncertainty in the data, and a large number of substances might not be classified properly when too many classes are present, thereby making the decision process much more complicated for the assessor. The ultimate goal of risk prioritization is to categorize chemical compounds for probable risk, so a limited decision choice of two risk profile classes was thought to be sufficient for this purpose, whereas for actual risk assessment more classes would be desirable.

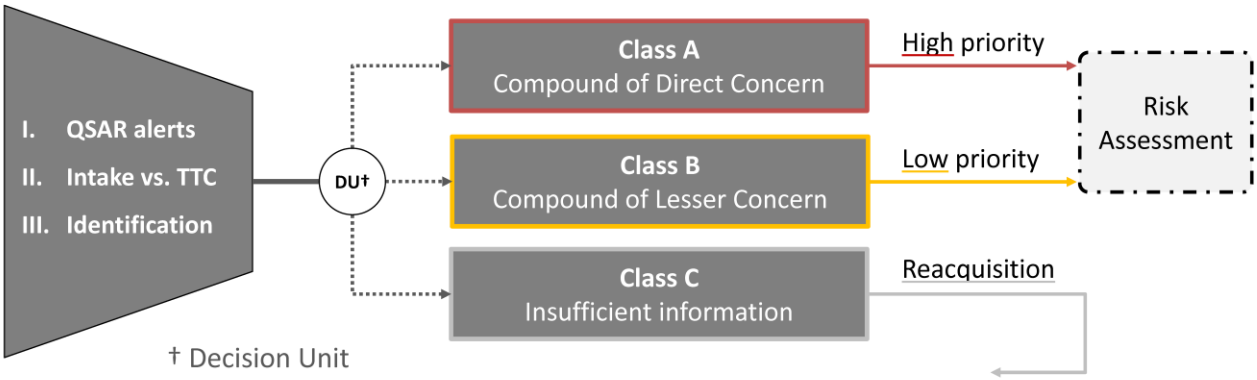


Figure 2: Framework representing one of the possible approaches to incorporate tentative data from exploration in risk assessment principles. The chemical compounds are subdivided into three priority classes following a decision unit (DU), which is an expertise-driven decision tool. The resulting risk profile classes can be used to prioritize further risk assessments.

3.3.2. Design of a decision unit

To facilitate the assignment of a risk profile class to a chemical compound, a decision unit was designed to incorporate all available data from the tentative exposure assessment and the tentative hazard identification, as shown in Figure 2. The goal of the decision unit is to provide a simple, unified, and reproducible workflow for risk assessors to evaluate input data from exploration experiments into a risk profile classification. Input data for the decision unit consisted of tentative exposure, i.e., estimated intake, Cramer Class exposure limit, and resulting TTC Excess; and tentative hazard identification, i.e., predicted structures, structure correlation scores, QSAR CMR predictions, and QSAR consensus. Due to the tentative nature of the data, the input data can contain variations especially in structure predictions, which affect hazard predictions and intake limit by Cramer class. Hence, it is important to note that small changes in chemical structure may affect different Cramer classifications and exposure limits, so these values should always be seen in context, e.g., evaluation of the actual intake in addition to the TTC Excess.

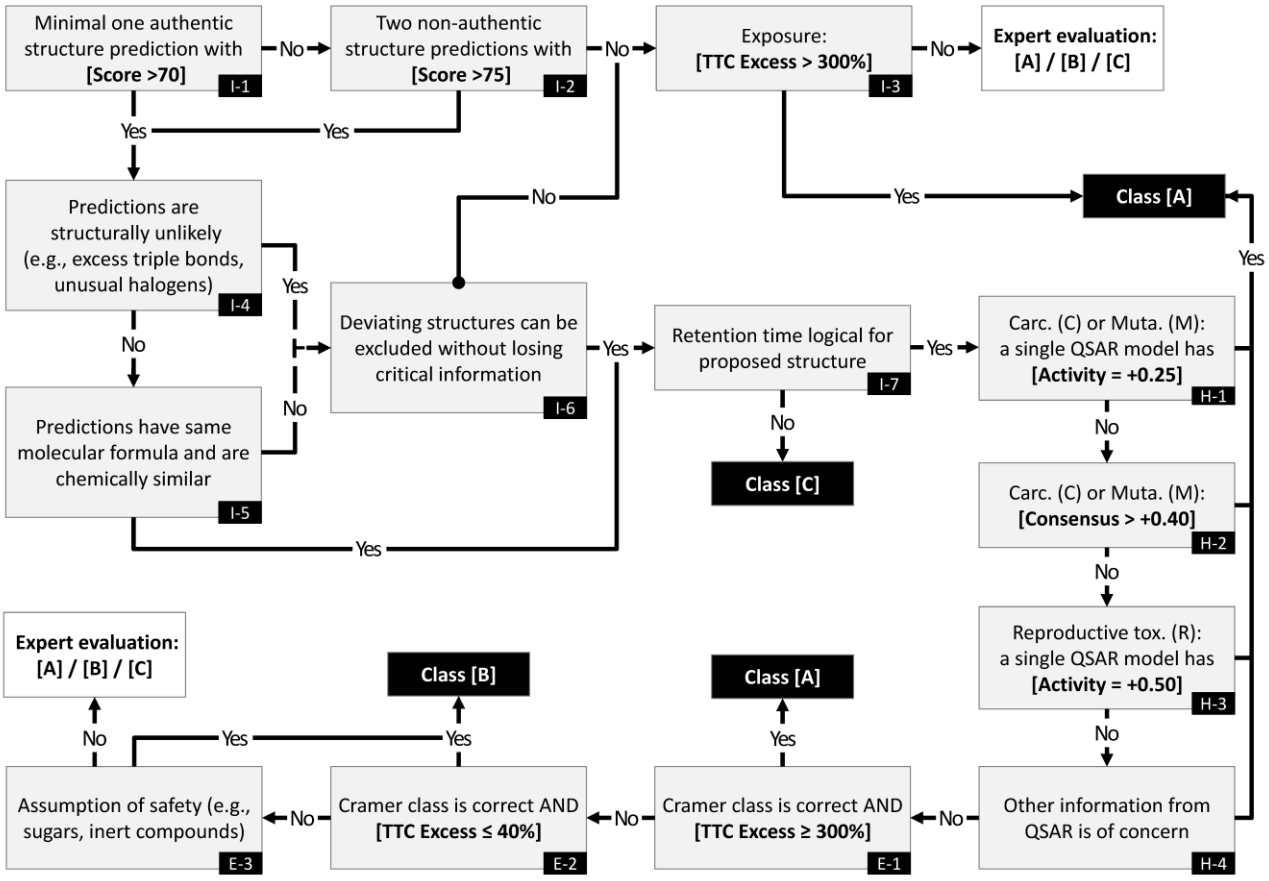


Figure 3: Implementation of a decision unit for risk prioritization. The decision unit is designed as a decision tree that is evaluated by an expert for each node. The result from the decision unit is risk profile class: [A] high priority, [B] low priority, or [C] insufficient data. The risk profile can be determined either data-driven or via expert decision, in which an experienced assessor decides the class based on all available data.

The decision unit was constructed like a decision tree, as shown in Figure 3, built up from the structure prediction by tentative identification (I), the hazard prediction by QSAR prediction (H), and the TTC Excess exposure prediction (E). The decision unit consists of 14 decision-nodes and 6 risk profile classification end-nodes. Decision nodes systematically evaluate all input data and provide a path to the most appropriate end-node. End-nodes within the decision unit result in the assignment of a risk profile to a compound (discussed in 3.3.1), but are always expert judgements. In some cases the end-node provides a non-binding advice for the most likely risk profile class considering the data. In case no such advice is attainable, i.e., where data interpretation cannot unambiguously result in classification, the final risk profile is a decision that needs to be made by the assessor. Consequently, the decision unit is not designed as an automated data evaluation tool

although it contains decisions based purely on data, but more as a guide for assessors to stepwise evaluate all available data.

3.3.3. Design of data-driven decision modules

The first nodes in the decision unit are to assess the quality and appropriateness of the predicted structure(s) by tentative identification. The nodes in the tree (see Figure 3) evaluate structure and quality parameters, but require assessor feedback and insight. As discussed in by Pieke et al. (2018) at least one authentic prediction (I–1) or two non-authentic predictions with sufficient prediction score are required (I–2). When there is insufficient chemical structure information, the exposure (TTC Excess) should be evaluated for exceeding the TTC threshold (I–3). When there are sufficient structural predictions, the predicted structures should be evaluated for chemically unlikely features (I–4), molecular mass and similar chemical structure (I–5), and sufficient chemical information (I–6). Finally, the polarity and molecular weight of the predictions should be proportional to the chromatographic retention time (I–7).

Following, the predicted chemical structures are evaluated for exerting possible CMR activity. If there is sufficient QSAR data that suggests CMR activity, the compound is immediately classified as [A]. The QSAR results are checked for experimental data on possible carcinogenicity (H–1), mutagenicity (H–1), or reproductive toxicity (H–3) evident from a maximum reliability score. In addition, the prediction consensus for C and M — but not for R, only limited to two models — is evaluated for exceeding the threshold >0.40 (H–2). The final node is an expert assessment on concerns with the chemical structure regarding hazard, or below-threshold QSAR alerts that promotes concern for safety and should therefore be classified as [A] (H–4).

Finally, the Exposure Module is evaluated by means of the Cramer class and TTC Excess. By comparing the estimated intake with the intake threshold the acceptability of exposure can be decided. First, the intake is assessed compared to the threshold beyond the uncertainty of the exposure measurement, i.e., more than 300% TTC Excess, in which case it will be risk profile [A] (E–1), or below the uncertainty, i.e., less than 40% TTC Excess, in which case it is risk profile [B] (E–2). Next, there is a final expert evaluation node that confirms that the given substance is not known for likely safety, like sugars or inert materials, because the derived TTC limit may be too strict for these, especially since the Cramer classification is often Class III (High) if the structure deviates slightly from a well-known Class I (Low) structure (E–3).

3.3.4. Incorporation of expert decisions

Several nodes in the decision unit are based on human evaluation by requiring expert input. Expert-based decisions are included in the decision unit for two reasons: First, they are a result of discussions with risk assessment expert panels, which summarized that the need for an expert to control the final decision is critical. Second, a simple decision tree is not able to assess the multiplicative effects of several parameters, or capable to assess the data as a whole instead of individually. Hence, expert judgment is required for cases where data obtained by QSAR and/or quantitative methods are inconclusive (Lester et al., 2018). The use of a human assessor within the decision unit fulfills the need for control, but also mitigates the limitations of simplistically-designed decision units, and can thereby help improve decisions. However, it also requires the full attention of a trained risk assessor throughout the entire decision unit, which is problematic with a very large number of substances. Advances in computer sciences, such as advanced machine learning neural networks, may provide an outcome for this in the future (Ru et al., 2017). Consequently, the outcome from the decision unit is codependent on assessor expertise, which in fact closely resembles the methods for traditional RA.

Within the decision unit, the expert assessments are generally called upon in situations where simple data evaluation did not result in a classification. In other words, most expert decisions are needed when no immediate hazard or exposure of concern is detected. In those cases, a comprehensive picture of all available data followed by an expertise decision is required. For example, there may be stacked evidence for classification without exceeding any of the defined thresholds in prior nodes, e.g., a QSAR consensus of 0.39 for both carcinogenicity and mutagenicity. None of the nodes H–1 to H–3 will have marked this compound as a possible risk, but the expertise decision node H–4 likely will via human evaluation. The expert decision nodes at the end of the tree are needed in case iterating through single descriptors such as exposure or hazard identity did not lead to a proper classification. It is impossible to model every likely scenario into the decision unit and retain its accessibility. In addition, an automated decision unit cannot effectively decide whether the available data is sufficient for classification. Expert decisions are consequently the only decisions that can result into a [C] classification for a lack of information.

3.4. Applying risk prioritization explorative data

3.4.1. Application and results of the decision unit

The decision unit (Figure 3) was applied to a set of data obtained from exploration experiments on two different paperboard FCM samples described in the Experimental section. Assessment results of the 60 discovered chemical compounds are summarized in Table 1. The full dataset, which

includes all predicted chemical structures, QSAR predictions, and estimated exposure of these 60 compounds are given in the Supplementary Information. Note that the total number of chemical compounds per sample targeted for structural elucidation was 249 for the pizza box sample and 161 for the chocolate box sample, 410 in total, so the 60 compounds represented here are only a fraction of the total number of discovered compounds.

To convert the assessment into risk ranking, a score was calculated based on the assessors' answers. The score is on a scale from −100, low priority, to +100, high priority. Scores near zero were those either that showed no consensus between assessors or where data was inadequate for classification. Calculation of the score is performed according to Equation 1, where n_X represents the occurrence count of each classification x = [A], [B], or [C] per compound. The formula has deliberately not been simplified for clarification: the first part penalizes differences between [A] and [B], while the second part penalizes a lack of consensus. Hence, more contrast in the classification results in a ranking score closer to zero.

$$\text{Rank} = \frac{n_A - n_B}{n_A + n_B + n_C} * \frac{\max[n_A, n_B, n_C]}{n_A + n_B + n_C} \qquad \text{Equation 1}$$

The threshold of priority and no consensus was set at a score of ±30. This marked the point where above which at least three assessors assigned the same risk profile, but one assessor assigned a conflicting profile or indicated insufficient data, e.g., AAAB or AAAC. If an assessment contained two or more entries of [C] these were marked as uncertain, since at least 50% of assessors indicated that available data was not sufficient to take an appropriate decision.

The overall results from the assessment in Table C.1 reveal that approximately 60% of the chemical compounds were eligible for prioritization as a result of the evaluation, while 40% of the substances either have insufficient data for prioritization, or displayed conflicts in assignments by different assessors. The results show an almost even distribution of cases between high priority (29%), insufficient data (23%), no consensus (18%), and low priority (30%). A number of compounds were unanimously ranked by all assessors as high risk or low risk for 13% and 13% of the cases, respectively.

Some illustrative examples of each consensus result are discussed in order to understand some of the choices behind the classification. For a visualization of the chemical structures discussed, the reader is referred to the Supplementary Information.

***Table 1**: Assessments results of four assessors on 60 different chemical compounds from two different samples. Assessors were tasked to assign one of three risk profiles to the chemical substance. The ranking score is calculated from the ratio of the risk profiles reaching four different consensus results, where a score of at least ±30 was considered consensus. When two or more assessors assigned [C], the entry was considered to be deficit in information.*

ID.	Sample	ESI Polarity	Ret. time (min)	1	2	3	4	Ranking score	**Consensus**
10	Pizza	ESI−	18.582	A	A	A	A	100	High priority
13	Choc	ESI−	18.572	A	A	A	A	100	High priority
15	Pizza	ESI−	20.385	A	A	A	A	100	High priority
16	Pizza	ESI+	35.247	A	A	A	A	100	High priority
19	Pizza	ESI+	34.313	A	A	A	A	100	High priority
24	Pizza	ESI+	13.270	A	A	A	A	100	High priority
36	Pizza	ESI+	23.521	A	A	A	A	100	High priority
49	Choc	ESI+	2.088	A	A	A	A	100	High priority
1	Choc	ESI+	3.652	C	A	A	A	56.25	High priority
30	Choc	ESI+	24.146	A	A	C	A	56.25	High priority
32	Pizza	ESI+	8.989	A	A	A	C	56.25	High priority
34	Pizza	ESI−	22.801	A	A	C	A	56.25	High priority
45	Choc	ESI+	26.474	A	A	C	A	56.25	High priority
54	Choc	ESI+	15.071	A	A	C	A	56.25	High priority
5	Pizza	ESI+	34.689	A	B	A	A	37.5	High priority
53	Choc	ESI+	8.988	A	A	A	B	37.5	High priority
57	Choc	ESI+	19.585	A	A	B	A	37.5	High priority
39	Pizza	ESI−	27.602	C	A	C	A	25	Insufficient data
20	Pizza	ESI−	11.742	B	A	C	A	12.5	No consensus
40	Choc	ESI−	17.711	B	A	C	A	12.5	No consensus
6	Pizza	ESI+	27.600	B	C	C	A	0	Insufficient data
11	Choc	ESI+	27.018	C	C	B	A	0	Insufficient data
21	Pizza	ESI+	19.644	C	A	C	B	0	Insufficient data

ID.	Sample	ESI Polarity	Ret. time (min)	1	2	3	4	Ranking score	**Consensus**
23	Choc	ESI+	10.820	B	A	B	A	0	No consensus
27	Choc	ESI+	24.373	B	A	B	A	0	No consensus
35	Pizza	ESI−	17.712	B	A	B	A	0	No consensus
38	Choc	ESI−	3.473	C	C	C	C	0	Insufficient data
42	Choc	ESI+	30.173	C	C	C	C	0	Insufficient data
43	Choc	ESI+	33.504	C	C	C	C	0	Insufficient data
46	Choc	ESI−	30.398	C	C	B	A	0	Insufficient data
56	Pizza	ESI+	28.307	B	A	B	A	0	No consensus
58	Pizza	ESI+	10.831	B	A	B	A	0	No consensus
25	Choc	ESI+	13.095	C	A	B	B	−12.5	No consensus
26	Pizza	ESI+	3.283	B	B	C	A	−12.5	No consensus
50	Pizza	ESI−	24.446	B	B	C	A	−12.5	No consensus
52	Pizza	ESI+	36.131	B	B	C	A	−12.5	No consensus
2	Choc	ESI+	31.825	C	C	C	B	−18.75	Insufficient data
51	Choc	ESI−	29.098	C	C	C	B	−18.75	Insufficient data
3	Choc	ESI+	16.438	B	C	C	B	−25	Insufficient data
7	Choc	ESI+	14.272	B	C	C	B	−25	Insufficient data
12	Choc	ESI+	27.434	C	C	B	B	−25	Insufficient data
17	Pizza	ESI−	28.550	C	C	B	B	−25	Insufficient data
14	Choc	ESI+	17.416	B	A	B	B	−37.5	Low priority
22	Choc	ESI−	20.588	B	A	B	B	−37.5	Low priority
31	Choc	ESI+	13.335	B	B	B	A	−37.5	Low priority
44	Pizza	ESI−	2.790	B	A	B	B	−37.5	Low priority
60	Choc	ESI−	18.167	B	A	B	B	−37.5	Low priority
4	Choc	ESI−	19.438	B	B	C	B	−56.25	Low priority
28	Choc	ESI−	14.359	B	C	B	B	−56.25	Low priority
33	Pizza	ESI+	15.083	B	B	B	C	−56.25	Low priority

ID.	Sample	ESI Polarity	Ret. time (min)	1	2	3	4	Ranking score	**Consensus**
37	Pizza	ESI+	31.396	B	B	B	C	−56.25	Low priority
59	Pizza	ESI−	14.359	B	C	B	B	−56.25	Low priority
8	Pizza	ESI+	27.289	B	B	B	B	−100	Low priority
9	Choc	ESI−	20.881	B	B	B	B	−100	Low priority
18	Pizza	ESI+	26.490	B	B	B	B	−100	Low priority
29	Choc	ESI+	24.943	B	B	B	B	−100	Low priority
41	Pizza	ESI+	28.868	B	B	B	B	−100	Low priority
47	Pizza	ESI+	33.264	B	B	B	B	−100	Low priority
48	Pizza	ESI+	15.083	B	B	B	B	−100	Low priority
55	Pizza	ESI+	34.327	B	B	B	B	−100	Low priority

3.4.2. Compounds with high priority

One notable entry unanimously marked as risk profile [A] is ID 10 and ID 13, which is in fact the same chemical compound observed in different samples. The chemical structure suggests a benzophenone-like compound at relatively high exposure levels of 360−445 µg person^{-1} day^{-1}, excluding the three-fold semi-quantitative uncertainty, compared to the 90 µg person^{-1} day^{-1} TTC limit of Cramer Class III compounds, with QSAR alerts that indicate carcinogenicity and mutagenicity. There is another instance of a benzophenone-like compound among the high priority compounds: ID 15, which has a less unambiguous predicted structure which occurs at nearby retention time. The presence of benzophenone compounds in paper and board is known mostly due to recycling of printed board (Anderson and Castle, 2003), so detection of benzophenone-like substances is not unexpected; however, the concentration estimates indicate a relatively high exposure potential. This potential can also be limited by the overestimation in the TMC, but it is nevertheless a compound of concern.

Another entry clearly marked as risk profile [A] is ID 19, which strongly represents an azo dye Pigment Red 2. The chemical compound could exceed TTC limits with an estimated intake of 50 µg person^{-1} day^{-1}, excluding uncertainty, compared to 90 µg person^{-1} day^{-1} defined by Cramer Class III. However, QSAR results clearly indicate a possible carcinogenicity and mutagenicity, which would exempt the compound from Class III limits and instead impose the stricter limit of 0.15

µg person^{-1} day^{-1}. The presence of pigments, especially pigment red, has been observed before (Bengtström, 2014). Azo dyes are capable of breaking down into carcinogenic substances like amines and aromatic amines, which can be cause for concern, e.g., in cosmetic products (SCCNFP, 2002).

Compound ID 32 and ID 53, both marked as risk profile [A], represent an isothiazolinone fungicide compound present in both samples at similar retention times. For the pizza box, it exceeds the maximum exposure significantly: 700 µg person^{-1} day^{-1} excluding uncertainty, but for the chocolate box semi-quantification was unsuccessful. When strictly following the decision unit, the presence in the chocolate box would likely be marked as risk profile [B] or [C] because none of the thresholds are explicitly exceeded, but since it was classified as risk profile [A] by most experts this demonstrates the added value of an expert decision. Here, the expert decision rightly classified the same chemical compound with the same priority, despite differences in available data. This substance has been discovered and more extensively discussed in previous work (Pieke et al., 2018).

3.4.3. Compounds with low priority

ID 47 and 55 represent two compounds marked as risk profile [B], and are chemically similar long-chain amides originating from the pizza box. Estimated intake of these substances is significantly below a TTC Class III compound at 33–38 µg person^{-1} day^{-1}, excluding uncertainty, which is unlikely to exceed 90 µg person^{-1} day^{-1} when including uncertainties. In addition, there are no QSAR alerts for these compounds. These substances have previously been identified (Pieke et al., 2018) in a similar sample, where these were also considered unlikely to pose a risk. The consensus of the risk prioritization here emphasises that the previous assessment is probably correct, and this type of compound is not anticipated to be at risk by different evaluators.

ID 31, ID 33, and ID 48 represent polyethylene glycol (PEG) oligomers, while ID 18 represents dipropylene glycol dibenzoate. These are all commonly used plasticizers. The intake for these compounds is relatively high compared to other compounds listed here, but these compounds are not commonly associated with any hazardous effects. It was shown here that the expert decisions play a critical role in ensuring the proper class assignment, e.g., ID 31 has large TTC Excess values because the compound is marked as Cramer Class III. Despite that, three out of four assessors marked the compound as risk profile [B] because the chemical structure was known to them as a PEG oligomer, for which a Cramer Class I is more likely appropriate. All of these compounds are relatively inert plasticizers with no QSAR alerts, and especially PEG oligomers are unlike to pose a risk at these concentrations.

ID 4, ID 8, ID 37, ID 41, and ID 60 are a number of diverse, yet chemically similar and simple structures that are each marked as risk profile [B] by most assessors, indicating a low priority. These compounds are characterized by a generally low exposure estimate, simple chemical structures composed predominantly of C, H, and O, and few carbon-rings, and rarely contain any QSAR alerts for CMR. A number of the predicted chemical structures are classified as Cramer Class I, which increases the exposure limit significantly to 1800 µg person^{-1} day^{-1}, but for most of these compounds the 90 µg person^{-1} day^{-1} is not exceeded.

3.4.4. Compounds with no assigned priority

The compounds that did not have a prioritization can be separated in two main groups: compounds with insufficient data, or compounds with mixed information containing both elements of high priority and low priority, which prevented consensus. Compounds with insufficient data are marked if at least half of the assessors indicated that the available data is insufficient to assign a risk profile [A] or [B], e.g., ID 51, ID 2, ID 38, and ID 43. These cases are not discussed extensively, but reassessment should only occur upon obtaining additional or improved data.

Interesting cases of non-consensus compounds are ID 23 and ID 58. Some structure predictions seem to indicate a polyethylene glycol (PEG) oligomer, similar to ID 31, ID 33, ID 48 previously discussed. However, the exposure is significantly higher: 1240 µg person^{-1} day^{-1} for ID 58 and 520 µg person^{-1} day^{-1} for ID 23. In addition, some of the predicted structures seem to be PEG derivatives or unrelated structures, which have more severe QSAR alerts and TTC Excess due to being Cramer Class III. Different assessors interpreted this information differently: two considered this a high priority substance and two considered this a low priority substance. Based on the exposure, it is sensible to consider these substances as high priority, but on the other hand the knowledge of PEG oligomers can render these compounds as low priority. Consequently, the lack of consensus can warrant the need for a discussion on the substances to clarify where differences in opinion are originating from.

Another case where no consensus could be achieved is for ID 56. The predicted structures greatly varied between the authentic and non-authentic databases, where structure predictions of the latter appeared unlikely, but the predictions from the first were of relatively low confidence. The estimated exposure was 83 µg person^{-1} day^{-1}, which is close to the limit of 90 µg person^{-1} day^{-1} of a Cramer Class III compound. There are no obvious QSAR alerts that indicated direct concern. Here, the assessment of the compound was primarily based on expert decision, and assessors are unable to agree. ID 56 is an example of a group of compounds that have very little information, or where the information shows conflict between different predictions, so a decision for low- or high-

priority is not straightforward. Some other examples are ID 20, ID 26, ID 27, and ID 35. The proper classification of these compounds may require additional information, a stricter decision unit, or more assessors.

In a number of cases some assessors considered the information to be not sufficiently informative, but others tried to give a classification. These cases are characterized by an equal distribution in assessors indicating [C] and [A] or [B]. An illustrative case is ID 3, which initially does not appear to lack information. However, the structure predictions are varying greatly and are accompanied by low confidence, so no good structural image can be obtained. Because there is no structural image, QSAR alerts cannot be considered reliable. Yet, the exposure to this compound is low: 12 μg person^{-1} day^{-1}, which is including uncertainty well below the TTC limit for a Class III compound. As a result, half the assessors indicated [B] for no likely risk due to the low exposure, but the other half indicated [C] likely due to the poor quality of structure predictions. There are some other examples where this occurred, e.g., ID 59, ID 6.

3.5. The implementation of risk prioritization tools

Based on the results in 3.4, a different course of action is required for each assigned priority and rank score. Compounds that show the maximum rank score do not require much discussion, as these are classified similarly by all assessors, so their priority for risk assessment is fairly unanimous. For compounds that do not score maximally, but are still classified as low- or high-priority (e.g., AAAC), it is suggested to discuss these entries briefly to understand the reason for reaching a less than maximum consensus. Unless there is a good reason to deviate from the advice of the consensus, it should be maintained. Results with rank scores close to zero need to be investigated: either the available data is insufficient or has too many uncertainties, or the assessors disagree on the risk profile. In the first case, insufficient data, more data will need to be gathered or, as discussed in the next paragraph, the quality of results needs to be improved. The latter case, no consensus, requires discussion and is cause for concern. In some cases, the differences occur as a result of data weight: some assessors weigh the exposure heavier than hazards. Disagreements between assessors will need to be better understood in order to improve the decision unit.

Presently designed decision unit was found to be suitable for assessing a small to moderate number of chemical compounds with tentative data. The decision unit is currently an expert-based model in which the decision tree is a helpful tool for the experts to reach classification. The value of the expert decision was shown throughout the data, e.g., in classifying ID 32 and ID 53. However, for a larger number of compounds the workload on the assessors increases similarly, so the

current design may not be suitable for a very large number of chemical compounds. For this, automation may be a solution, but discussions with risk assessment experts indicated caution to changing the decision of risk to a fully automated process. In addition, automated decisions for tentative data are complex since they require a multivariate approach that can incorporate multiple uncertainties, whereas it also must be able to derive decisions from experience as humans do. Hence, while automated decisions are desirable for many compounds, these should be developed with caution to the human expertise needed to classify compounds.

Further needs are to improve the input data on which decisions are based, which will reduce the number of non-consensus and insufficient data prioritizations. For example, the inclusion of more and improved *in silico* models (e.g., Expert Model, QSAR, or hybrids) may allow a better decision process, as more hazards can be included in the assessment possibly with higher prediction certainties. In addition, a reduction in uncertainty originating from tentative data is beneficial, e.g., lower error in concentration estimation or improved structure predictions. Suggestions for improving the strategies for semi-quantification and tentative identification have been provided in the respective research (Pieke et al., 2018, 2017). Both, however, highlight that these explorative methods are relatively novel in applications, and will need substantial further developments. Finally, the inclusion of more assessors can improve the classification results. More assessors permit more combinations of risk profile assessments, which will improve the amount of ranks that are available, as well as allowing a better investigation and discussion of compounds that did not reach a risk prioritization consensus.

4. Conclusion

A strategy for risk prioritization based on tentative data is demonstrated for ranking the tentative risk of discovered compounds. This tool is based on a simple and low cost approach. The classification/prioritization of 60 substances was performed in a short time (less than 1h). The strategy is demonstrated to be capable to discriminate sufficiently (>60%) within a test set of 60 compounds between low priority compounds expected not to be of concern, and high priority substances expected to be of concern or demonstrating indications of concern. The tool is validated on compounds previously reported in literature as being of concern, so the strategy is able to sort relevant results. Consequently, the tool can easily be transposed on the total set of 400 compounds discovered by exploration to greatly improve the chemical knowledge on complex samples from a risk assessment perspective.

Automation of part of the decision process may be needed to ensure more rapid decisions for larger sets of data. However, implementation of automated processes is complicated because the current presentation of data is reliant on interpretation and experience, for which dedicated *in silico* models would be required. However, improving the quality of the tentative data, e.g., by reducing uncertainties, will be helpful in reducing the number of compounds that remain unclassified after prioritization, and will also assist in improving the quality of the decisions.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors confirm that there are no known conflicts of interest associated with this publication. The authors are grateful to the members of the ANSES panel “Risk assessment for substances and processes submitted to human food regulation” for scientific discussions, namely: Fabien Bolle, Jalloul Bouajila, Nicolas Cabaton, Marie-Christine Chagnon, Dany Chevalier, Véronique Coma, Luc Fillaudeau, Angel Gil-Izquierdo, Florence Lacoste, Claude Lambre, Michel Laurentie, Jean-Michel Maixent, Anne Platel, Philippe Saillard, Patrick Sauvegrain, and François Zuber.

Supplementary Material

The quantification, identification, expert assessment, and risk prioritization data results are available as supplementary Microsoft Office Excel file (*.xlsx*).

References

Anderson, W.A.C., Castle, L., 2003. Benzophenone in cartonboard packaging materials and the factors that influence its migration into food. Food Addit. Contam. 20, 607–618. doi:10.1080/0265203031000109486

Arvanitoyannis, I.S., Bosnea, L., 2004. Migration of Substances from Food Packaging Materials to Foods. Crit. Rev. Food Sci. Nutr. 44, 63–76. doi:10.1080/10408690490424621

Barlow, S.M., 2009. Risk assessment of food-contact materials: past experience and future challenges. Food Addit. Contam. Part A 26, 1526–1533. doi:10.1080/02652030903233231

Barnes, K., Sinclair, R., Watson, D. (Eds.), 2007. Chemical Migration and Food Contact Materials.

Benfenati, E., Manganaro, A., Gini, G., 2013. VEGA-QSAR: AI inside a platform for predictive toxicology, in: CEUR Workshop Proceedings. pp. 21–28.

Bengtström, L., 2014. Chemical identification of contaminants in paper and board food contact

materials. Technical University of Denmark, Kongens Lyngby, Denmark.

Bengtström, L., Rosenmai, A.K., Trier, X., Jensen, L.K., Granby, K., Vinggaard, A.M., Driffield, M., Højslev Petersen, J., 2016. Non-targeted screening for contaminants in paper and board food-contact materials using effect-directed analysis and accurate mass spectrometry. Food Addit. Contam. Part A 33, 1080–1093. doi:10.1080/19440049.2016.1184941

Biedermann-Brem, S., Kasprick, N., Simat, T., Grob, K., 2012. Migration of polyolefin oligomeric saturated hydrocarbons (POSH) into food. Food Addit. Contam. - Part A Chem. Anal. Control. Expo. Risk Assess. 29, 449–460. doi:10.1080/19440049.2011.641164

Biedermann, M., Grob, K., 2013. Assurance of safety of recycled paperboard for food packaging through comprehensive analysis of potential migrants is unrealistic. J. Chromatogr. A 1293, 107–119. doi:10.1016/j.chroma.2013.04.009

Biedermann, M., Uematsu, Y., Grob, K., 2011. Mineral oil contents in paper and board recycled to paperboard for food packaging. Packag. Technol. Sci. 24, 61–73. doi:10.1002/pts.914

Binderup, M.L., Pedersen, G.A., Vinggaard, A.M., Rasmussen, E.S., Rosenquist, H., Cederberg, T., 2002. Toxicity testing and chemical analyses of recycled fibre-based paper for food contact. Food Addit. Contam. 19, 13–28. doi:10.1080/02652030110089878

Bouma, K., Stavenga, K., Draaijer, A., 2003. Domestic Use of Food Packaging Materials in the Netherlands.

Bradley, E., 2006. Case study: Chemical migration from snack and take-away food packaging, Chemical Migration and Food Contact Materials. Woodhead Publishing Limited. doi:10.1533/9781845692094.3.416

Castle, L., 2006. Chemical migration into food: An overview, Chemical Migration and Food Contact Materials. Woodhead Publishing Limited. doi:10.1533/9781845692094.1

Cramer, G.M., Ford, R.A., Hall, R.L., 1976. Estimation of toxic hazard-A decision tree approach. Food Cosmet. Toxicol. 16, 255–276. doi:10.1016/S0015-6264(76)80522-6

Driffield, M., Bradley, E.L., Castle, L., 2016. Safety Assessment of Food Contact Materials : The Role of High-resolution Mass Spectrometry in the Comprehensive Analysis of the Total Migrate.

Duffy, E., Hearty, A.P., McCarthy, S., Gibney, M.J., 2007. Estimation of exposure to food packaging materials. 3: Development of consumption factors and food-type distribution factors from data collected on Irish children. Food Addit. Contam. 24, 63–74. doi:10.1080/02652030600865475

EFSA, WHO, 2016. Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree. EFSA Support. Publ. 2016EN-1006 1–50. doi:10.2903/SP.EFSA.2016.EN-1006

European Parliament and Council of the European Union, 2011. Commission Regulation (EU) No

10/2011 of 14 January 2011 on plastic materials and articles intended to come into contact with food. Off. J. Eur. Union 50, 12–88.

FDA, 2007. Guidance for Industry : Preparation of Premarket Submissions for Food Contact Substances : Chemistry Recommendations 1–18.

Gallart-Ayala, H., Núñez, O., Lucci, P., 2013. Recent advances in LC-MS analysis of food-packaging contaminants. TrAC Trends Anal. Chem. 42, 99–124. doi:10.1016/j.trac.2012.09.017

Geueke, B., Wagner, C.C., Muncke, J., 2014. Food contact substances and chemicals of concern: A comparison of inventories. Food Addit. Contam. - Part A Chem. Anal. Control. Expo. Risk Assess. 31, 1438–1450. doi:10.1080/19440049.2014.931600

Grob, K., 2014. Work plans to get out of the deadlock for the safety assurance of migration from food contact materials? A proposal. Food Control 46, 312–318. doi:10.1016/j.foodcont.2014.05.044

Guillén, D., Ginebreda, A., Farré, M., Darbra, R.M., Petrovic, M., Gros, M., Barceló, D., 2012. Prioritization of chemicals in the aquatic environment based on risk assessment: Analytical, modeling and regulatory perspective. Sci. Total Environ. 440, 236–252. doi:10.1016/j.scitotenv.2012.06.064

Hauder, J., Benz, H., Rüter, M., Piringer, O.-G., 2013. The specific diffusion behaviour in paper and migration modelling from recycled board into dry foodstuffs. Food Addit. Contam. Part A 30, 599–611. doi:10.1080/19440049.2012.762605

Honkalampi-Hämäläinen, U., Bradley, E.L., Castle, L., Severin, I., Dahbi, L., Dahlman, O., Lhuguenot, J.-C., Andersson, M.A., Hakulinen, P., Hoornstra, D., Mäki-Paakkanen, J., Salkinoja-Salonen, M., Turco, L., Stammati, A., Zucco, F., Weber, A., von Wright, A., 2010. Safety evaluation of food contact paper and board using chemical tests and in vitro bioassays: role of known and unknown substances. Food Addit. Contam. Part A 27, 406–415. doi:10.1080/19440040903401358

ILSI Europe Packaging Material Task Force, 1996. SUMMARY OF A WORKSHOP HELD 15 JULY 1996, BRUSSELS, in: Food Consumption and Packaging Usage Factors.

International Agency for Research on Cancer, 2017. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, years 1972 - 2017.

Istituto di Ricerche Farmacologiche Mario Negri Milano, 2017. VEGA interpretation [WWW Document]. URL https://www.vegahub.eu/download/vega-interpretation/

Jickells, S., 2007. Chemical migration from secondary packaging into foods, in: Chemical Migration and Food Contact Materials. Elsevier, pp. 395–415. doi:10.1533/9781845692094.3.395

Koster, S., Bani-Estivals, M.-H., Bonuomo, M., Bradley, E., Chagnon, M.-C., Garcia, L.M., Godts, F., Gude, T., Helling, R., Paseiro-Losada, P., Pieper, G., Rennen, M., Simat, T., Spack, L., 2015. Guidance of Best Practices on the Risk Assessment of NIAS in Food Contact Materials

and Articles, International Life Science Institute. doi:10.1146/annurev.phyto.40.120301.093728

Koster, S., Rennen, M., Leeman, W., Houben, G., Muilwijk, B., van Acker, F., Krul, L., 2014. A novel safety assessment strategy for non-intentionally added substances (NIAS) in carton food contact materials. Food Addit. Contam. Part A 31, 422–443. doi:10.1080/19440049.2013.866718

Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., Van Schothorst, F., Vos, J.G., Würtzen, G., 2004. Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. Food Chem. Toxicol. 42, 65–83. doi:10.1016/j.fct.2003.08.006

Lester, C., Reis, A., Laufersweiler, M., Wu, S., Blackburn, K., 2018. Structure activity relationship (SAR) toxicological assessments: The role of expert judgment. Regul. Toxicol. Pharmacol. 92, 390–406. doi:10.1016/j.yrtph.2017.12.026

Muncke, J., Backhaus, T., Geueke, B., Maffini, M. V, Martin, O.V., Myers, J.P., Soto, A.M., Trasande, L., Trier, X., Scheringer, M., 2017. Scientific Challenges in the Risk Assessment of Food Contact Materials. Environ. Health Perspect. 125, 1–9. doi:10.1289/EHP644

Ozaki, A., Yamaguchi, Y., Fujita, T., Kuroda, K., Endo, G., 2005. Safety assessment of paper and board food packaging: Chemical analysis and genotoxicity of possible contaminants in packaging. Food Addit. Contam. 22, 1053–1060. doi:10.1080/02652030500090885

Pieke, E.N., Granby, K., Trier, X., Smedsgaard, J., 2017. A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry. Anal. Chim. Acta 975, 30–41. doi:10.1016/j.aca.2017.03.054

Pieke, E.N., Smedsgaard, J., Granby, K., 2018. Exploring the chemistry of complex samples by tentative identification and semiquantification: A food contact material case. J. Mass Spectrom. 53, 323–335. doi:10.1002/jms.4052

Pivnenko, K., Eriksson, E., Astrup, T.F., 2015. Waste paper for recycling: Overview and identification of potentially critical substances. Waste Manag. 45, 134–142. doi:10.1016/j.wasman.2015.02.028

Poças, M. de F., Oliveira, J.C., Pereira, J.R., Brandsch, R., Hogg, T., 2011. Modelling migration from paper into a food simulant. Food Control 22, 303–312. doi:10.1016/j.foodcont.2010.07.028

Raies, A.B., Bajic, V.B., 2016. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip. Rev. Comput. Mol. Sci. 6, 147–172. doi:10.1002/wcms.1240

Rosenmai, A.K., Bengtström, L., Taxvig, C., Trier, X., Petersen, J.H., Svingen, T., Binderup, M.-L.L., van Vugt-Lussenburg, B.M.A., Dybdahl, M., Granby, K., Vinggaard, A.M., Barbara

Medea Alice, van V.L., Dybdahl, M., Granby, K., Vinggaard, A.M., 2017. An effect-directed strategy for characterizing emerging chemicals in food contact materials made from paper and board. Food Chem. Toxicol. 106, 250–259. doi:10.1016/j.fct.2017.05.061

Ru, G., Crescio, M.I., Ingravalle, F., Maurella, C., Gregori, D., Lanera, C., Azzolina, D., Lorenzoni, G., Soriani, N., Zec, S., Berchialla, P., Mercadante, S., Zobec, F., Ghidina, M., Baldas, S., Bonifacio, B., Kinkopf, A., Kozina, D., Nicolandi, L., Rosat, L., 2017. Machine Learning Techniques applied in risk assessment related to food safety. EFSA Support. Publ. 14. doi:10.2903/sp.efsa.2017.EN-1254

SCCNFP, 2002. Safety review of the use of certain azo-dyes in cosmetic products.

Schymanski, E.L., Singer, H.P., Longrée, P., Loos, M., Ruff, M., Stravs, M.A., Ripollés Vidal, C., Hollender, J., 2014. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. Environ. Sci. Technol. 48, 1811–8. doi:10.1021/es4044374

Skjevrak, I., Brede, C., Steffensen, I.L., Mikalsen, A., Alexander, J., Fjeldal, P., Herikstad, H., 2005. Non-targeted multi-component analytical surveillance of plastic food contact materials: Identification of substances not included in EU positive lists and their risk assessment. Food Addit. Contam. 22, 1012–1022. doi:10.1080/02652030500090877

Triantafyllou, V.I., Akrida-Demertzi, K., Demertzis, P.G., 2007. A study on the migration of organic pollutants from recycled paperboard packaging materials to solid food matrices. Food Chem. 101, 1759–1768. doi:10.1016/j.foodchem.2006.02.023

van Bossuyt, M., van Hoeck, E., Raitano, G., Manganelli, S., Braeken, E., Ates, G., Vanhaecke, T., Van Miert, S., Benfenati, E., Mertens, B., Rogiers, V., 2017. (Q)SAR tools for priority setting: A case study with printed paper and board food contact material substances. Food Chem. Toxicol. 102, 109–119. doi:10.1016/j.fct.2017.02.002

Chapter 4

Discussion, Conclusion, and Outlook

4.1. Non-targeted analytical strategies for exploration

4.1.1. Semi-quantification

Semi-quantification allows quantification of any detected substance without needing reference standards, instead using a predefined mixture of markers as quantification endpoints. This eliminates the need for quantification to obtain a reference standard for every chemical compound. Hence, semi-quantitative methods are interesting to estimate concentration of compounds in a number of fields, in particular when dealing with unknown compounds as the procedure does not require identification of chemical compounds. This is for example not the case using response factor (RF) prediction in ESI (Caetano et al. 2005; Kruve et al. 2014), where the structure must be known before the response can be predicted. Ensuring sufficiently reliable semi-quantification requires the use of representative quantification markers (QMs) and a thorough optimization of ESI source parameters to minimize the differences in response factors. However, semi-quantification requires some technical expertise in mass spectrometry and some pre-existing knowledge about the sample. Consequently, semi-quantification can be optimized for the sample of interest where it provides an untargeted and simplified concept for quantitative experiments on unknown substances that require minimal prior knowledge on the chemical compounds.

Before semi-quantification can be considered as a suitable alternative to traditional quantification, there are a number of aspects that need to be investigated. The fundamental limitation of semi-quantification is that the prediction error in concentration is still relatively large. To reduce this error, the role of QMs needs to be better studied, especially in regards to what makes a QM suitable or unsuitable, and on which properties the QM should be preferred. Secondly, the prediction error may be reduced by improving chromatography by the use of hydrophobic Interaction chromatography (HILIC), or perhaps by further ESI source optimization, like the using nano-ESI to improve responses and reduce the effects of suppression (Moreno-González et al. 2017). Finally, a correction for suppression effects in the LC-MS for the AOI and QM may also be used, e.g., like Kaufmann & Butcher (2005), as this can improve the quantification results.

4.1.2. Tentative identification

Identification of chemical compounds via mostly untargeted screening approach is a viable strategy to obtain more information on the chemical composition of a sample. For most of the chemical compounds selected by dynamic targeting, the interpretation of mass spectra provided an appreciable level of information on the chemical structure. For a number of compounds, it was possible to obtain a chemical structure with very high certainties. Most of the predicted structures contained some uncertainty, which is expected using a limited amount of information; however, the predicted structures provided an overall impression of the chemical “picture”, which is significantly better than having no structural information. Due to the use of different databases with chemical compounds, the tentative identification was capable of identifying both well-known and unexpected compounds. As the requirements used in tentative identification are fairly low, e.g., not using a reference standard or isolating the compound for structure confirmation, it provides a substantial gain in speed and capacity while information on an unknown compound is not comprehensive.

While tentative identification provides valuable data, there are still a number of improvements needed. The automated interpretation of mass spectra by fragment correlation can especially limit the range and quality of the results. First, the interpretation relies on a number of existing databases; therefore, the quality of prediction is inherently restricted by the quality and appropriateness of the databases. However, databases for tentative identification are considerably easier to develop than those needed for targeted analysis. Secondly, correlation is a simplified approach for a complex concept like fragmentation, and this requires manual confirmation to confirm the absence of overfitting mass spectra on a chemically unlikely substance. The use of advanced data interpretation based on artificial intelligence, like neural networks, could greatly improve automatic interpretation of mass spectra, since these systems can be designed to interpret a spectrum like a human would. This could eliminate the need for existing databases since the system can be taught to create chemical structures based on likeliness, thereby also reducing the possibility of overfitting.

4.2. The role of exploration: now and in the future

The development of explorative strategies (exploration) in risk assessment is characterized by minimizing prerequisite information and avoiding time and resource constraints imposed by compound-specific or targeted methods. Exploration as a high-throughput untargeted method fills a critical analytical gap in research fields that call for rapid identification and characterization of chemical compounds in complex or poorly-studied samples without the availability of reference standards. The primary goal of exploration is to discover chemical compounds that may be of high relevant but where little prior information is available for, or we do not have specific analysis methods for, rather than continue investigating already well-known chemical compounds. One application for exploration included in this study is the discovery of chemical compounds that have potential to migrate to food and pose a certain food safety risk. Other examples of research fields that would benefit from exploration are environmental analysis or metabolomics used to unravel biochemical pathways. The exact implementation of the exploration results depend strongly on the field of research and the applicable goal of the study. Here, the results are applied to risk prioritization, but in metabolomics they could be used for generating a biochemical fingerprint.

An important aspect of exploration studies is the capacity to build a continuous and profound knowledge of the samples investigated even when no direct application of the data is defined. The chemical untargeted nature of exploration makes the strategy well-suited to discover and record new chemical compounds and provide new insights on concentration and identification. Even though the data from exploration has inherent uncertainty, the increase in knowledge from having no information to a predicted chemical structure and/or concentration estimate is invaluable. Moreover, the knowledge-building function of exploration can be used in development of databases. These databases of discovered chemical compounds can be used in a non-target screening approach on new or similar samples for these discovered compounds, or for newly validated QSAR models to test for any new hazards. Consequently, generation of chemical databases can also be used to make future screening faster, and even to identify new hazards.

It is likely that explorative studies will play a much greater role in FCM analysis in the future, because they are complimentary to tools like QSAR and migration modelling. One

of the major challenges in FCM analysis is the need for automated and high-throughput methods to cope with the huge diversity and rapid changes in the field of new materials. An example is the availability of a migration model that works for multiple FCMs that might make specific migration tests redundant, which makes the most critically needed information is what chemical compounds are present and in which concentration in the food contact material. In addition, the development of QSAR models for an increased number of available toxicity endpoints with more reliable results combined with improved tentative identifications enable better identification of hazards from a wide range of chemical compounds without the need for individual testing.

4.3. The application of risk prioritization based on tentative data

The need for a risk prioritization framework that permits early hazard identification for possible chemical exposure is because comprehensive risk assessment of chemical substances is slow and costly, therefore unrealistic for a large number of chemical substances. By performing an early-stage prioritization of chemical risk based on tentative data from exploration experiments, it is possible to rank chemical compounds on their tentative risk profile. In this study a risk prioritization tool has been developed in the form of a decision tree that is demonstrated in combination with expert decisions. The results show that the decision tree approach including expert decisions by an assessor is suitable for classifying compounds discovered by exploration, but that differences in risk profiles occur due to different data interpretations by assessors, or due to differences in confidence on uncertainty in data.

Risk prioritization as demonstrated requires expert assessment in most of the steps, and is thus difficult to implement on a very large number of compounds. As each decision in the approach must be evaluated by an expert, evaluating of a large number of compounds can be limited by time, or by disagreement in the assigned risk profile between experts. Nevertheless, expert disagreements are not necessarily unique to risk prioritization, as may similarly occur in RA. Hence, disagreements are not necessarily as a disadvantage or limitation of the approach, as they permit dialogue on risk-related decisions. Yet, in order to improve capacity of the approach for more compounds, the possibility of automated decisions within the tool has been discussed extensively with risk assessors. One of the

key learnings was that most risk assessors are uncomfortable with having too little control in the decision process. While it may be possible to automate a substantial part of the risk profile decision, this would also limit the expert influence and thereby reduce the experts' trust in the final decision.

Aside from a reduction in expert trust upon automation of the decision tree, it is also complicated to implement automation when uncertainties have to be included. For example, the recorded exposure to a chemical compound with no significant Carcinogenicity, Mutagenicity, or Reproductive Toxicity (CMR) alerts can exceed the given TTC limit, e.g., at 150% of the limit. However, this excess still falls within the error margin of semi-quantification. In those cases, it would be difficult to automatically assign a risk profile since the prediction results fall within the uncertainty of the methods. Consequently, all other sources of data will need to be collectively studied for a risk profile decision. These other sources include the non-significant predictions by QSAR, quality of the chemical structure prediction, evaluation of assigned TTC class, and experience with similar substances. Including these parameters in a decision is not readily achieved by a simple decision model, and some of the parameters are for human evaluation. Hence, while simplification of the decision tree by automating improves the speed of decisions, it is difficult to implement. The use of more advanced machine learning — like neural networks — may provide an outcome in this respect similarly as for the identification.

The tool as demonstrated here provides a valuable framework for the development of newer generations of risk prioritization approaches. While the implementation of risk prioritization based on tentative data can be challenged, there is need for a strategy like this. Improving the data quality obtained by exploration will assist in reducing the disagreements in especially human evaluated risk profile decisions, therefore the prioritization should be available more rapidly (see also discussion in Manuscript A (3.4), Manuscript B (0), and discussion of results (4.1)). However, all chemical compounds will still need to be assessed comprehensively as required by the current principles of RA, but the order in which to do so can be based on available risk profiles obtained from prioritization studies. Consequently, risk prioritization based on tentative data does not eliminate the need for comprehensive chemical RA, but instead ensures that RA of high-risk substances is done with a higher priority.

4.4. Final remarks

The outcome of this thesis is for a large part the result of many scientific discussions and presentations on the current status of risk assessment, and how to move forward. There has been significant discussion on the technical aspects of what moving forward will constitute of and what it will require, but perhaps less on a visualization of the problem at hand. An image that leads to considerable discussion on this topic is the crude representation of choices as illustrated in Figure 21: “the choice to do nothing should be deliberate.” This reflects the current situation of NIAS where most of the unknown NIAS are forced down the “do nothing” path as there is simply insufficient information to do something. This is — of course — undesirable for numerous reasons which have been the driving force behind this research, and the reasons are repeatedly highlighted throughout this thesis. Yet, even though the current situation with unknown chemicals is unacceptable in a number of ways, it is actually common. In order to truly ensure safety, each decision, “do something” or “do nothing”, should be motivated by adequate information rather than a lack of information.

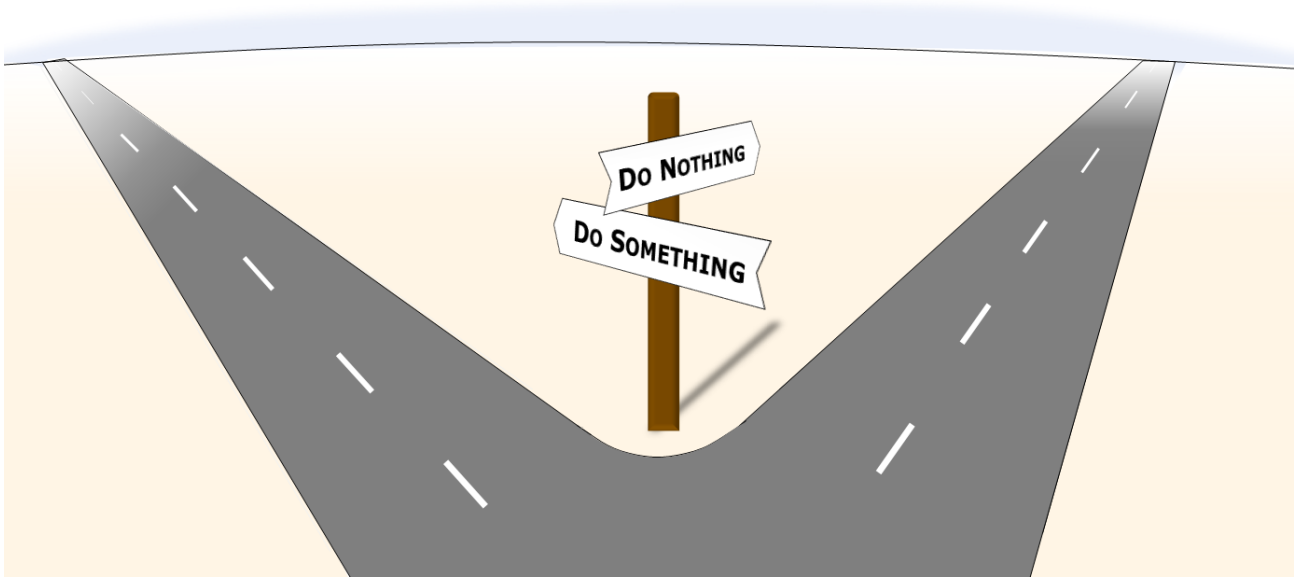


Figure 21: For chemical contamination in food, the choice can be greatly simplified to taking no action or taking action, regardless of whether this decision is deliberate or not.

Analytical strategies and perhaps research in general tend to favour the “easy” approach by focusing primarily on a limited number of well-studied compounds. When a chemical compound or group of compounds are identified as a risk, this leads to a significant increase in research on that particular risk. Some examples of this system are bisphenol A, benzophenone, acrylamide, and currently mineral oils. Because resources are devoted to these “high value” topics, substantially less research is being performed on the huge number of unknown chemical compounds, e.g., in FCM. Essentially, this behaviour observed in food safety follows a hype profile, and in fact it is harmful to establishing a broader concept of food safety that includes all possible chemical interactions. Clearly, specific and in-depth research on hazardous chemicals for a proper risk assessment is needed, but the possible very high risk of a large amount of non-identified chemical compounds should not be neglected in the process.

Trusting food is paramount to our society, and safety is a critical component of this trust. To ensure the current and future food safety, we need to change our mind-set on analytical strategies, but also in the way of thinking about risk assessment and mitigating risk. The core idea of the work presented here is to present how a change in mind-set may look like: a paradigm shift in the conceptualization of data and how data-driven approaches can be utilized to fill the information gaps in risk assessment for mitigation and control. How these shifts are to be practically realised is of course a matter of debate even at the completion of this study, but the need for this change should be obvious: tentative and explorative data is needed because the current alternative is no data.

References and cited materials

Ackerman, L.K., Noonan, G.O. & Begley, T.H., 2009. Assessing direct analysis in real-time-mass spectrometry (DART-MS) for the rapid identification of additives in food packaging. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, 26(12), pp.1611–1618.

ALL4PACK, 2016. *Packaging: Market And Challenges In 2016*, Available at: https://www.all4pack.com/ [Accessed January 26, 2018].

Amad, M.H. et al., 2000. Importance of gas-phase proton affinities in determining the electrospray ionization response for analytes and solvents. *Journal of Mass Spectrometry*, 35(7), pp.784–789. Available at: http://doi.wiley.com/10.1002/1096-9888%28200007%2935%3A7%3C784%3A%3AAID-JMS17%3E3.0.CO%3B2-Q.

Anderson, W.A.C. & Castle, L., 2003. Benzophenone in cartonboard packaging materials and the factors that influence its migration into food. *Food Additives and Contaminants*, 20(6), pp.607–618.

Arvanitoyannis, I.S. & Bosnea, L., 2004. Migration of Substances from Food Packaging Materials to Foods. *Critical Reviews in Food Science and Nutrition*, 44(2), pp.63–76.

Aurela, B., Kulmala, H. & Soderhjelm, L., 1999. Phthalates in paper and board packaging and their migration into Tenax and sugar. *Food Additives and Contaminants*, 16(12), pp.571–577. Available at: http://www.tandfonline.com/doi/abs/10.1080/026520399283713.

Barnkob, L.L. & Petersen, J.H., 2013. Effect of relative humidity on the migration of benzophenone from paperboard into the food simulant Tenax® and modelling hereof. *Food Additives & Contaminants: Part A*, 30(2), pp.395–402. Available at: http://www.tandfonline.com/doi/abs/10.1080/19440049.2012.741717.

Bengtström, L., 2014. *Chemical identification of contaminants in paper and board food contact materials*. Kongens Lyngby, Denmark: Technical University of Denmark.

Bengtström, L. et al., 2016. Non-targeted screening for contaminants in paper and board food-contact materials using effect-directed analysis and accurate mass spectrometry. *Food Additives & Contaminants: Part A*, 33(6), pp.1080–1093. Available at: http://www.tandfonline.com/doi/full/10.1080/19440049.2016.1184941.

Bentayeb, K. et al., 2013. Non-visible print set-off of photoinitiators in food packaging: Detection by ambient ionisation mass spectrometry. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, 30(4), pp.750–759.

Berendsen, B.J.A., Stolker, L.A.M. & Nielen, M.W.F., 2013. The (Un)Certainty of Selectivity in Liquid Chromatography Tandem Mass Spectrometry. *Journal of The American Society for Mass Spectrometry*, 24(1), pp.154–163. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23345060.

Biedermann-Brem, S. et al., 2012. Migration of polyolefin oligomeric saturated hydrocarbons (POSH) into food. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, 29(3), pp.449–460.

Biedermann, M. et al., 2013. Migration of mineral oil from printed paperboard into dry foods: Survey of the German market. Part II: Advancement of migration during storage. *European Food Research and Technology*, 236(3), pp.459–472.

Biedermann, M. & Grob, K., 2013. Assurance of safety of recycled paperboard for food packaging through comprehensive analysis of potential migrants is unrealistic. *Journal of Chromatography A*, 1293, pp.107–119. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23639127.

Biedermann, M., Uematsu, Y. & Grob, K., 2011. Mineral oil contents in paper and board recycled to paperboard for food packaging. *Packaging Technology and Science*, 24(2), pp.61–73. Available at: http://onlinelibrary.wiley.com/doi/10.1002/pts.893/abstract.

Biedermann, S. et al., 2013. Migration of cyclo-diBA from coatings into canned food: Method of analysis, concentration determined in a survey and in silico hazard profiling. *Food and Chemical Toxicology*, 58, pp.107–115. Available at: http://dx.doi.org/10.1016/j.fct.2013.04.004.

Biermann, C.J., 1996. *Handbook of Pulping and Papermaking (Second Edition)* Second Edi. C. J. Biermann, ed., San Diego: Academic Press. Available at: http://www.sciencedirect.com/science/article/pii/B9780120973620500005.

Bilbao, A. et al., 2015. Processing strategies and software solutions for data-independent acquisition in mass spectrometry. *Proteomics*, 15(5–6), pp.964–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25430050.

Binderup, M.L. et al., 2002. Toxicity testing and chemical analyses of recycled fibre-based paper for food contact. *Food Additives & Contaminants*, 19(sup1), pp.13–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11962701.

Blades, A.T., Ikonomou, M.G. & Kebarle, P., 1991. Mechanism of electrospray mass spectrometry. Electrospray as an electrolysis cell. *Analytical Chemistry*, 63(19), pp.2109–2114. Available at: http://pubs.acs.org/doi/abs/10.1021/ac00019a009.

Blok-Tip, L. et al., 2004. Structure elucidation of sildenafil analogues in herbal products. *Food Additives and Contaminants*, 21(8), pp.737–748.

Boccacci Mariani, M., Chiacchierini, E. & Gesumundo, C., 1999. Potential migration of Diisopropyl naphthalenes from recycled paperboard packaging into dry foods. *Food Additives and Contaminants*, 16(5), pp.207–213.

Bökman, C.F. et al., 2006. Relating chromatographic retention and electrophoretic mobility to the ion distribution within electrosprayed droplets. *Journal of the American Society for Mass Spectrometry*, 17(3), pp.318–324.

Bonfiglio, R. et al., 1999. The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds. *Rapid Communications in Mass Spectrometry*, 13(12), pp.1175–1185. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10407294.

van Bossuyt, M. et al., 2017. (Q)SAR tools for priority setting: A case study with printed paper and board food contact material substances. *Food and Chemical Toxicology*, 102, pp.109–119. Available at: http://dx.doi.org/10.1016/j.fct.2017.02.002.

Bouma, K., Stavenga, K. & Draaijer, A., 2003. *Domestic Use of Food Packaging Materials in the Netherlands*,

Bradley, E., 2006. *Case study: Chemical migration from snack and take-away food packaging*, Woodhead Publishing Limited. Available at: http://dx.doi.org/10.1533/9781845692094.3.416.

Brandsch, J. et al., 2002. Migration modelling as a tool for quality assurance of food packaging. *Food additives and contaminants*, 19(January), pp.29–41.

Broecker, S. et al., 2011. Semi-quantitative Determination of Concentrations in Systematic Toxicological Analysis by LC-QTOF-MS. *Toxichem Krimtech*, 78(Special Issue), pp.302–305.

Bruins, A.P., 1998. Mechanistic aspects of electrospray ionization. *Journal of Chromatography A*, 794(1–2),

pp.345–357. Available at: http://www.sciencedirect.com/science/article/pii/S0021967397011102.

Bu, Q. et al., 2014. A high throughout semi-quantification method for screening organic contaminants in river sediments. *Journal of Environmental Management*, 143, pp.135–139. Available at: http://dx.doi.org/10.1016/j.jenvman.2014.05.009.

Caetano, S. et al., 2005. Exploring and modelling the responses of electrospray and atmospheric pressure chemical ionization techniques based on molecular descriptors. *Analytica Chimica Acta*, 550(1–2), pp.92–106. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0003267005011578.

Castle, L., 2006. *Chemical migration into food: An overview*, Woodhead Publishing Limited. Available at: http://dx.doi.org/10.1533/9781845692094.1.

Castle, L., Offen, C.P., et al., 1997. Migration studies from paper and board food packaging materials. 1. Compositional analysis. *Food Additives and Contaminants*, 14(1), pp.35–44. Available at: http://www.tandfonline.com/doi/abs/10.1080/02652039709374495.

Castle, L., Damant, A.P., et al., 1997. Migration studies from paper and board food packaging materials. Part 2. Survey for residues of dialkylamino benzophenone UV-cure ink photoinitiators. *Food Additives and Contaminants*, 14(1), pp.45–52. Available at: http://www.tandfonline.com/doi/abs/10.1080/02652039709374496.

Cech, N.B. & Enke, C.G., 2000. Relating electrospray ionization response to nonpolar character of small peptides. *Analytical Chemistry*, 72(13), pp.2717–2723. Available at: http://dx.doi.org/10.1021/ac9914869.

Cech, N.B., Krone, J.R. & Enke, C.G., 2001. Predicting electrospray response from chromatographic retention time. *Analytical Chemistry*, 73(2), pp.208–213. Available at: http://dx.doi.org/10.1021/ac0006019.

Chalcraft, K.R. et al., 2009. Virtual quantification of metabolites by capillary electrophoresis- electrospray ionization-mass spectrometry: Predicting ionization efficiency without chemical standards. *Analytical Chemistry*, 81(7), pp.2506–2515. Available at: http://pubs.acs.org/doi/abs/10.1021/ac802272u.

Chen, J., 2009. A worldwide food safety concern in 2008--melamine-contaminated infant formula in China caused urinary tract stone in 290,000 children in China. *Chinese medical journal*, 122(3), pp.243–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19236797.

Cherkasov, A. et al., 2014. QSAR Modeling: Where Have You Been? Where Are You Going To? *Journal of Medicinal Chemistry*, 57(12), pp.4977–5010. Available at: http://pubs.acs.org/doi/10.1021/jm4004285.

Cincinelli, A. et al., 2012. Compound Specific Isotope Analysis (CSIA) for chlorine and bromine: A review of techniques and applications to elucidate environmental sources and processes. *Environmental Pollution*, 169, pp.112–127. Available at: http://dx.doi.org/10.1016/j.envpol.2012.05.006.

Coma, V., 2008. Bioactive packaging technologies for extended shelf life of meat-based products. *Meat science*, 78(1–2), pp.90–103. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22062099.

Dallyn, H. & Shorten, D., 1988. Hygiene aspects of packaging in the food industry. *International Biodeterioration*, 24(4–5), pp.387–392.

Derraik, J.G.B., 2002. The pollution of the marine environment by plastic debris: A review. *Marine Pollution Bulletin*, 44(9), pp.842–852.

DG Sanco, 2014. *Union Guidelines on Regulation (EU) No 10/2011 on plastic materials and articles intended to come into contact with food This*, Available at: http://ec.europa.eu/food/food/chemicalsafety/foodcontact/docs/10-2011_plastic_guidance_en.pdf.

Dole, M. et al., 1968. Molecular Beams of Macroions. *The Journal of Chemical Physics*, 49(5), pp.2240–2249. Available at: http://aip.scitation.org/doi/10.1063/1.1670391.

Driscoll, J.N. et al., 1978. Gas chromatographic detection and identification of aromatic and aliphatic hydrocarbons in complex mixtures by coupling photoionization and flame-ionization detectors. *Journal of Chromatography A*, 158, pp.171–180. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0021967300899653.

Duffy, E. et al., 2007. Estimation of exposure to food packaging materials. 3: Development of consumption factors and food-type distribution factors from data collected on Irish children. *Food Additives and Contaminants*, 24(1), pp.63–74.

Eberlin, M.N., 2006. Structurally diagnostic ion/molecule reactions: Class and functional-group identification by mass spectrometry. *Journal of Mass Spectrometry*, 41(2), pp.141–156.

EFSA & WHO, 2016. Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree. *EFSA Supporting Publication 2016:EN-1006*, (February), pp.1–50. Available at: http://www.efsa.europa.eu/sites/default/files/corporate_publications/files/1006e.pdf.

Eicher, A. et al., 2015. Migration by “direct” or “indirect” food contact? “Dry” and “wetting” foods? Experimental data for “touching” contact of dry foods with paper and board. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, 32(1), pp.110–119. Available at: http://dx.doi.org/10.1080/19440049.2014.975753.

Ellen MacArthur Foundation, 2016. *The New Plastics Economy: Rethinking the future of plastics*, Available at: https://newplasticseconomy.org/report-2016.

Enke, C.G., 1997. A Predictive Model for Matrix and Analyte Effects in Electrospray Ionization of Singly-Charged Ionic Analytes. *Analytical Chemistry*, 69(23), pp.4885–4893. Available at: http://pubs.acs.org/doi/abs/10.1021/ac970095w.

Espinosa, M.S. et al., 2015. Exploring analyte response in an ESI-MS system with different chemometric tools. *Chemometrics and Intelligent Laboratory Systems*, 146, pp.120–127. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0169743915001173.

European Chemicals Agency, 2008. Guidance on information requirements and chemical safety assessment: QSARs and grouping of chemicals. In *Guidance for the implementation of REACH*. p. 134. Available at: https://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf.

European Food Safety Authority, 2012. *Report of ESCO WG on non-plastic Food Contact Materials*, Parma, Italy. Available at: http://www.efsa.europa.eu/.

European Paper Recycling Council, 2017. New 74% paper recycling target set for 2020. *Press Release*. Available at: http://www.paperforrecycling.eu/.

European Parliament and Council of the European Union, 2011a. Commission Regulation (EU) No 10/2011 of 14 January 2011 on plastic materials and articles intended to come into contact with food. *Offical Journal of the European Union*, 50(1), pp.12–88. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:012:0001:0089:EN:PDF.

European Parliament and Council of the European Union, 2002. Regulation (EC) N^{o} 178/2002 of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. *Official Journal of the European Communities*, L31, pp.1–24. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:031:0001:0024:EN:PDF.

European Parliament and Council of the European Union, 2006. REGULATION (EC) No 1907/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency. In *Official Journal of the European Communities*. pp. 1–15.

European Parliament and Council of the European Union, 2004. Regulation (EC) No 1935/2004 “on

materials and articles intended to come into contact with food and repealing Directives 80/590/EEC and 89/109/EEC." *Official Journal of the European Union*, L338/4(1935), pp.4–17.

European Parliament and Council of the European Union, 2011b. Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, an. *Official Journal of the European Union*, (1169), pp.18–63.

Eurostat, 2017. *Packaging waste statistics*, Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Packaging_waste_statistics.

de Fátima Poças, M. & Hogg, T., 2007. Exposure assessment of chemicals from packaging materials in foods: a review. *Trends in Food Science and Technology*, 18(4), pp.219–230. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0924224407000118.

FDA, 2007. Guidance for Industry : Preparation of Premarket Submissions for Food Contact Substances : Chemistry Recommendations. , pp.1–18.

Fiselier, K. & Grob, K., 2012. Barriers against the Migration of Mineral Oil from Paperboard Food Packaging: Experimental Determination of Breakthrough Periods. *Packaging Technology and Science*, 25(5), pp.285–301. Available at: http://onlinelibrary.wiley.com/doi/10.1002/pts.893/abstract.

Focant, J.-F. et al., 2004. High-throughput analysis of human serum for selected polychlorinated biphenyls (PCBs) by gas chromatography-isotope dilution time-of-flight mass spectrometry (GC-IDTOFMS). *The Analyst*, 129(4), pp.331–336. Available at: http://xlink.rsc.org/?DOI=B313675B.

Food Safety and Veterinary Affairs (FSVO), 2017. *Annex 10 of the Ordinance of the FDHA on materials and articles intended to come into contact with food-stuffs*, Switzerland.

Fryer, M. et al., 2006. Human exposure modelling for chemical risk assessment: a review of current approaches and research and policy implications. *Environmental Science & Policy*, 9(3), pp.261–274. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1462901106000141.

Furey, A. et al., 2013. Ion suppression; A critical review on causes, evaluation, prevention and applications. *Talanta*, 115, pp.104–122. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0039914013001975.

Gaines, L.L., Mintz, M.M. & Shepherd, P.B., 1994. Energy implications of glass-container recycling. , p.68.

Gallart-Ayala, H. et al., 2011. Preventing false negatives with high-resolution mass spectrometry: The benzophenone case. *Rapid Communications in Mass Spectrometry*, 25(20), pp.3161–3166.

García-Reyes, J.F., Hernando, M.D., Molina-Díaz, A., et al., 2007. Comprehensive screening of target, non-target and unknown pesticides in food by LC-TOF-MS. *TrAC Trends in Analytical Chemistry*, 26(8), pp.828–841. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0165993607001434.

García-Reyes, J.F., Hernando, M.D., Ferrer, C., et al., 2007. Large Scale Pesticide Multiresidue Methods in Food Combining Liquid Chromatography– Time-of-Flight Mass Spectrometry and Tandem Mass Spectrometry. *Analytical Chemistry*, 79(19), pp.7308–7323. Available at: http://pubs.acs.org/doi/abs/10.1021/ac070855v.

Gärtner, S. et al., 2009. Analysis and migration of phthalates in infant food packed in recycled paperboard. *Journal of Agricultural and Food Chemistry*, 57(22), pp.10675–10681.

Gomez, A. & Tang, K., 1994. Charge and fission of droplets in electrostatic sprays. *Physics of Fluids*, 6(1), pp.404–414. Available at: http://aip.scitation.org/doi/10.1063/1.868037.

Goodson, W.H. et al., 2015. Assessing the carcinogenic potential of low-dose exposures to chemical mixtures in the environment: The challenge ahead. *Carcinogenesis*, 36, pp.S254–S296.

Grob, K., 2009. Analytical food control does not live up to expectations - What to do? *Journal of Separation*

Science, 32(21), pp.3575–3578. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20029905.

Grob, K. et al., 2006. Food Contamination with Organic Materials in Perspective: Packaging Materials as the Largest and Least Controlled Source? A View Focusing on the European Situation. *Critical Reviews in Food Science and Nutrition*, 46(7), pp.529–535. Available at: http://www.tandfonline.com/doi/abs/10.1080/10408390500295490.

Grob, K., 2014. Work plans to get out of the deadlock for the safety assurance of migration from food contact materials? A proposal. *Food Control*, 46, pp.312–318. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0956713514003077.

Gros, M., Petrović, M. & Barceló, D., 2006. Development of a multi-residue analytical methodology based on liquid chromatography-tandem mass spectrometry (LC-MS/MS) for screening and trace level determination of pharmaceuticals in surface and wastewaters. *Talanta*, 70(4), pp.678–690.

Halloran, A. et al., 2014. Addressing food waste reduction in Denmark. *Food Policy*, 49(P1), pp.294–301.

Han, B. et al., 2016. Migration of photoinitiators from paper to fatty food simulants: experimental studies and model application. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, 33(5), pp.876–884. Available at: http://dx.doi.org/10.1080/19440049.2016.1166524.

Handford, C.E., Campbell, K. & Elliott, C.T., 2016. Impacts of Milk Fraud on Food Safety and Nutrition with Special Emphasis on Developing Countries. *Comprehensive Reviews in Food Science and Food Safety*, 15(1), pp.130–142. Available at: http://doi.wiley.com/10.1111/1541-4337.12181.

Hauder, J. et al., 2013. The specific diffusion behaviour in paper and migration modelling from recycled board into dry foodstuffs. *Food Additives & Contaminants: Part A*, 30(3), pp.599–611. Available at: http://www.tandfonline.com/doi/abs/10.1080/19440049.2012.762605.

Hawkridge, A.M., 2014. Practical Considerations and Current Limitations in Quantitative Mass Spectrometry-based Proteomics. In C. E. Eyers & S. Gaskell, eds. *Quantitative Proteomics*. The Royal Society of Chemistry, pp. 1–25. Available at: http://dx.doi.org/10.1039/9781782626985-00001.

Heudorf, U., Mersch-Sundermann, V. & Angerer, J., 2007. Phthalates: toxicology and exposure. *International journal of hygiene and environmental health*, 210(5), pp.623–34. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1438463907001125 [Accessed June 28, 2017].

Hill, A.W. & Mortishire-Smith, R.J., 2005. Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Communications in Mass Spectrometry*, 19(21), pp.3111–3118. Available at: http://doi.wiley.com/10.1002/rcm.2177.

van der Hooft, J.J.J. et al., 2013. Structural elucidation of low abundant metabolites in complex sample matrices. *Metabolomics*, 9(5), pp.1009–1018.

Hopfgartner, G., Tonoli, D. & Varesio, E., 2012. High-resolution mass spectrometry for integrated qualitative and quantitative analysis of pharmaceuticals in biological matrices. *Analytical and bioanalytical chemistry*, 402(8), pp.2587–96. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22203371.

van Den Houwe, K. et al., 2017. The Use of Tenax® as a Simulant for the Migration of Contaminants in Dry Foodstuffs: A Review. *Packaging Technology and Science*. Available at: http://doi.wiley.com/10.1002/pts.2320.

Hubbe, M.A., Venditti, R.A. & Rojas, O.J., 2007. What happens to cellulosic fibers during papermaking and recycling? A review. *BioResources*, 2(4), pp.739–788.

Ibáñez, M. et al., 2005. Use of quadrupole time-of-flight mass spectrometry in the elucidation of unknown compounds present in environmental water. *Rapid Communications in Mass Spectrometry*, 19(2), pp.169–178.

Iribarne, J. V., 1976. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64(6), p.2287. Available at: http://scitation.aip.org/content/aip/journal/jcp/64/6/10.1063/1.432536.

Iribarne, J. V., Dziedzic, P.J. & Thomson, B.A., 1983. Atmospheric pressure ion evaporation-mass spectrometry. *International Journal of Mass Spectrometry and Ion Physics*, 50(3), pp.331–347.

Jaeger, C. et al., 2016. Automated Annotation and Evaluation of In-Source Mass Spectra in GC/Atmospheric Pressure Chemical Ionization-MS-Based Metabolomics. *Analytical Chemistry*, 88(19), pp.9386–9390.

Jaeger, C. et al., 2017. Compound annotation in liquid chromatography/high-resolution mass spectrometry based metabolomics: robust adduct ion determination as a prerequisite to structure prediction in electrospray ionization mass spectra. *Rapid Communications in Mass Spectrometry*, 31(15), pp.1261–1266.

Jickells, S., 2007. Chemical migration from secondary packaging into foods. In *Chemical Migration and Food Contact Materials*. Elsevier, pp. 395–415. Available at: http://dx.doi.org/10.1533/9781845692094.3.395.

Jickells, S.M. et al., 2005. Migration of contaminants by gas phase transfer from carton board and corrugated board box secondary packaging into foods. *Food Additives and Contaminants*, 22(8), pp.768–782. Available at: http://www.tandfonline.com/doi/abs/10.1080/02652030500151992.

Johns, S.M. et al., 2000. Studies on functional barriers to migration. 3. Migration of benzophenone and model ink components from cartonboard to food during frozen storage and microwave heating. *Packaging Technology and Science*, 13(3), pp.99–104. Available at: http://doi.wiley.com/10.1002/1099-1522%28200005%2913%3A3%3C99%3A%3AAID-PTS499%3E3.0.CO%3B2-K.

Kaufmann, A. & Butcher, P., 2005. Segmented post-column analyte addition; a concept for continuous response control of liquid chromatography/mass spectrometry peaks affected by signal suppression/enhancement. *Rapid Communications in Mass Spectrometry*, 19(5), pp.611–617. Available at: http://doi.wiley.com/10.1002/rcm.1832.

Kebarle, P., 2000. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *Journal of Mass Spectrometry*, 35(7), pp.804–817. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10934434.

Kebarle, P. & Tang, L., 1993. From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry. *Analytical Chemistry*, 65(22), p.972A–986A. Available at: http://pubs.acs.org/doi/abs/10.1021/ac00070a001.

Kind, T. & Fiehn, O., 2010. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanalytical Reviews*, 2(1), pp.23–60.

Kind, T. & Fiehn, O., 2006. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics*, 7(1), p.234. Available at: http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-234.

KnowThis.com, 2017. Factors in Packaging Decisions. *Product Decisions Tutorial*. Available at: https://www.knowthis.com/product-decisions/factors-in-packaging-decisions [Accessed August 28, 2017].

Koster, S. et al., 2014. A novel safety assessment strategy for non-intentionally added substances (NIAS) in carton food contact materials. *Food Additives & Contaminants: Part A*, 31(3), pp.422–443. Available at: http://www.tandfonline.com/doi/abs/10.1080/19440049.2013.866718.

Koster, S. et al., 2010. *Guidance On Best Practices On the Risk Assessment of Non Intentionally Added Substances (NIAS) in Food Contact Materials And Articles*,

Kovalchik, K.A. et al., 2017. Standard method design considerations for semi-quantification of total naphthenic acids in oil sands process affected water by mass spectrometry: A review. *Frontiers of*

Chemical Science and Engineering, 11(3), pp.497–507.

Krauss, M., Singer, H. & Hollender, J., 2010. LC-high resolution MS in environmental analysis: From target screening to the identification of unknowns. *Analytical and Bioanalytical Chemistry*, 397(3), pp.943–951.

Kroes, R. et al., 2004. Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology*, 42(1), pp.65–83.

Kruve, A. et al., 2014. Negative electrospray ionization via deprotonation: Predicting the ionization efficiency. *Analytical Chemistry*, 86(10), pp.4822–4830. Available at: http://dx.doi.org/10.1021/ac404066v.

De La Mora, J.F., 1992. The effect of charge emission from electrified liquid cones. *Journal of Fluid Mechanics*, 243(1), p.561. Available at: http://www.journals.cambridge.org/abstract_S0022112092002829.

Lago, M.A. & Ackerman, L.K., 2016. Identification of print-related contaminants in food packaging. *Food Additives & Contaminants: Part A*, 33(3), pp.518–529. Available at: http://dx.doi.org/10.1080/19440049.2015.1136435.

Lee, D.H. et al., 2016. Association between background exposure to organochlorine pesticides and the risk of cognitive impairment: A prospective study that accounts for weight change. *Environment International*, 89–90, pp.179–184. Available at: http://dx.doi.org/10.1016/j.envint.2016.02.001.

Leito, I. et al., 2008. Towards the electrospray ionization mass spectrometry ionization efficiency scale of organic compounds. *Rapid Communications in Mass Spectrometry*, 22(3), pp.379–384. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18183635.

Lin, Q.-B. et al., 2011. Kinetic migration of isothiazolinone biocides from paper packaging to Tenax and Porapak. *Food Additives & Contaminants: Part A*, 28(9), pp.1294–1301. Available at: http://www.tandfonline.com/doi/abs/10.1080/19440049.2011.584071.

Loscertales, I.G. & Fernández de la Mora, J., 1995. Experiments on the kinetics of field evaporation of small ions from droplets. *The Journal of Chemical Physics*, 103(12), pp.5041–5060. Available at: http://aip.scitation.org/doi/10.1063/1.470591.

Malik, A.K., Blasco, C. & Picó, Y., 2010. Liquid chromatography-mass spectrometry in food safety. *Journal of Chromatography A*, 1217(25), pp.4018–4040.

Mays, C., Benfenati, E. & Pardoe, S., 2012. Use and perceived benefits and barriers of QSAR models for REACH: Findings from a questionnaire to stakeholders. *Chemistry Central Journal*, 6(1), p.1. Available at: Chemistry Central Journal.

Moreno-González, D. et al., 2017. Matrix-effect free quantitative liquid chromatography mass spectrometry analysis in complex matrices using nanoflow liquid chromatography with integrated emitter tip and high dilution factors. *Journal of Chromatography A*, 1519, pp.110–120. Available at: https://doi.org/10.1016/j.chroma.2017.09.006.

Muncke, J. et al., 2017. Scientific Challenges in the Risk Assessment of Food Contact Materials. *Environmental Health Perspectives*, 125(9), pp.1–9. Available at: http://ehp.niehs.nih.gov/EHP644.

Nerín, C. & Asensio, E., 2007. Migration of organic compounds from a multilayer plastic-paper material intended for food packaging. *Analytical and Bioanalytical Chemistry*, 389(2), pp.589–596. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17680237.

NIST Mass Spectrometry Data Center, 2017. NIST Standard Reference Database.

O'Brien, J. et al., 2006. Approaches to the risk assessment of genotoxic carcinogens in food: A critical appraisal. *Food and Chemical Toxicology*, 44(10), pp.1613–1635.

O'Leary, L. & Manavalan, R., 2015. One victim of falling oil prices? Recycling. *Marketplace*. Available at: https://www.marketplace.org/2015/09/04/sustainability/one-victim-falling-oil-prices-recycling.

Osaka, I. & Takayama, M., 2014. Influence of hydrophobicity on positive- and negative-ion yields of peptides in electrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry*, 28(20), pp.2222–2226.

Pieke, E.N., Granby, K., et al., 2017. A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry. *Analytica Chimica Acta*, 975, pp.30–41. Available at: http://linkinghub.elsevier.com/retrieve/pii/S000326701730452X.

Pieke, E.N., Smedsgaard, J. & Granby, K., 2017. Exploring the chemistry of complex samples by tentative identification and semi-quantification: a food contact material case. *Journal of Mass Spectrometry*. Available at: http://doi.wiley.com/10.1002/jms.4052.

Piringer, O., 2007. Mathematical modelling of chemical migration from food contact materials. In *Chemical migration and food contact materials*. Woodhead Publishing Limited, pp. 180–202.

Pivnenko, K., Eriksson, E. & Astrup, T.F., 2015. Waste paper for recycling: Overview and identification of potentially critical substances. *Waste management (New York, N.Y.)*, 45, pp.134–142. Available at: http://dx.doi.org/10.1016/j.wasman.2015.02.028.

Poças, M. de F. et al., 2011. Modelling migration from paper into a food simulant. *Food Control*, 22(2), pp.303–312. Available at: http://dx.doi.org/10.1016/j.foodcont.2010.07.028.

Raji, M.A. & Schug, K.A., 2009. Chemometric study of the influence of instrumental parameters on ESI-MS analyte response using full factorial design. *International Journal of Mass Spectrometry*, 279(2–3), pp.100–106.

Rayleigh, Lord, 1882. XX. *On the equilibrium of liquid conducting masses charged with electricity*. *Philosophical Magazine Series 5*, 14(87), pp.184–186. Available at: http://www.tandfonline.com/doi/abs/10.1080/14786448208628425.

Reepmeyer, J.C. & Woodruff, J.T., 2006. Use of liquid chromatography-mass spectrometry and a hydrolytic technique for the detection and structure elucidation of a novel synthetic vardenafil designer drug added illegally to a "natural" herbal dietary supplement. *Journal of Chromatography A*, 1125(1), pp.67–75.

Renou, S. et al., 2008. Landfill leachate treatment: Review and opportunity. *Journal of Hazardous Materials*, 150(3), pp.468–493.

Roberts, J.C., 1996. *The Chemistry of Paper*, Cambridge: Royal Society of Chemistry. Available at: http://ebook.rsc.org/?DOI=10.1039/9781847552068.

Robertson, G.L., 2010. *Food Packaging and Shelf Life : A practical guide* G. L. Robertson, ed.,

Rosenberg, A., 2017. *Development and application of QSAR models for mechanisms related to endocrine disruption*. Technical University of Denmark.

Rosenmai, A.K. et al., 2017. An effect-directed strategy for characterizing emerging chemicals in food contact materials made from paper and board. *Food and Chemical Toxicology*, 106, pp.250–259. Available at: http://dx.doi.org/10.1016/j.fct.2017.05.061.

Rosenmai, A.K. et al., 2014. Are structural analogues to bisphenol a safe alternatives? *Toxicological Sciences*, 139(1), pp.35–47.

Sajilata, M.G. et al., 2007. Scalping of flavors in packaged foods. *Comprehensive Reviews in Food Science and Food Safety*, 6(1), pp.17–35.

Scholz, S. et al., 2013. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regulatory Toxicology and Pharmacology*, 67(3), pp.506–530. Available at: http://dx.doi.org/10.1016/j.yrtph.2013.10.003.

Schwudke, D. et al., 2007. Shotgun lipidomics by tandem mass spectrometry under data-dependent acquisition control. *Methods in enzymology*, 433(7), pp.175–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17954235.

Schymanski, E.L. et al., 2015. Non-target screening with high-resolution mass spectrometry: Critical review using a collaborative trial on water analysis. *Analytical and Bioanalytical Chemistry*, 407(21), pp.6237–6255.

Schymanski, E.L. et al., 2014. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. *Environmental science & technology*, 48(3), pp.1811–8. Available at: http://pubs.acs.org/doi/abs/10.1021/es4044374.

Seltenrich, N., 2015. A hard nut to crack: Reducing chemical migration in food-contact materials. *Environmental Health Perspectives*, 123(7), pp.A175–A179. Available at: http://ehp.niehs.nih.gov/123-A174.

Silano, V. et al., 2008. Note for Guidance For the Preparation of an Application for the Safety Assessment of a Substance to be used in Plastic Food Contact Materials. *EFSA Journal*, 6(7). Available at: http://doi.wiley.com/10.2903/j.efsa.2008.21r.

Silvenius, F. et al., 2014. The Role of Household Food Waste in Comparing Environmental Impacts of Packaging Alternatives. *Packaging Technology and Science*, 27(4), pp.277–292. Available at: http://onlinelibrary.wiley.com/doi/10.1002/pts.893/abstract.

Simoneau, C. et al., 2016. *Non-harmonised food contact materials in the EU: regulatory and market situation*,

Sjöberg, P.J.R. et al., 2001. A method for determination of ion distribution within electrosprayed droplets. *Analytical Chemistry*, 73(1), pp.23–28.

Sjöberg, P.J.R. et al., 2001. Factors influencing the determination of analyte ion surface partitioning coefficients in electrosprayed droplets. *Journal of the American Society for Mass Spectrometry*, 12(9), pp.1002–1010.

Smithers Pira, 2014. *The Future of Luxury Packaging to 2019*, Available at: http://www.smitherspira.com/news/2015/march/luxury-packaging-market-forecast.

Snedeker, S.M., 2014. Benzophenone UV-Photoinitiators Used in Food Packaging: Potential for Human Exposure and Health Risk Considerations. In S. M. Snedeker, ed. *Toxicants in Food Packaging and Household Plastics: Exposure and Health Risks to Consumers*. London: Springer London, pp. 151–176. Available at: https://doi.org/10.1007/978-1-4471-6500-2_6.

Stapleton, H.M. et al., 2011. Identification of flame retardants in polyurethane foam collected from baby products. *Environmental Science and Technology*, 45(12), pp.5323–5331.

Straub, R.F. & Voyksner, R.D., 1993. Negative ion formation in electrospray mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 4(7), pp.578–587. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24227644.

Suciu, N.A. et al., 2013. Recycled paper-paperboard for food contact materials: Contaminants suspected and migration into foods and food simulant. *Food Chemistry*, 141(4), pp.4146–4151. Available at: http://www.sciencedirect.com/science/article/pii/S0308814613009357.

Summerfield, W. & Cooper, I., 2001. Investigation of migration from paper and board into food-development of methods for rapid testing. *Food Additives and Contaminants*, 18(1), pp.77–88. Available at: http://www.tandfonline.com/doi/abs/10.1080/02652030010004674.

Taflin, D.C., Ward, T.L. & Davis, E.J., 1989. Electrified droplet fission and the Rayleigh limit. *Langmuir*, 5(2), pp.376–384. Available at: http://pubs.acs.org/doi/abs/10.1021/la00086a016.

Tang, K., Page, J.S. & Smith, R.D., 2004. Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 15(10), pp.1416–1423.

Tang, L. & Kebarle, P., 1991. Effect of the conductivity of the electrosprayed solution on the electrospray current. Factors determining analyte sensitivity in electrospray mass spectrometry. *Analytical Chemistry*, 63(23), pp.2709–2715. Available at: http://pubs.acs.org/doi/abs/10.1021/ac00023a009.

Taylor, P.J., 2005. Matrix effects: The Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry. *Clinical Biochemistry*, 38(4), pp.328–334.

Tehrany, E.A. & Desobry, S., 2004. Partition coefficients in food/packaging systems: A review. *Food Additives and Contaminants*, 21(12), pp.1186–1202.

Thomas, M.C. et al., 2012. Rapid identification of long-chain polyunsaturated fatty acids in a marine extract by HPLC-MS using data-dependent acquisition. *Analytical chemistry*, 84(14), pp.5976–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22816781.

Triantafyllou, V.I., Akrida-Demertzi, K. & Demertzis, P.G., 2007. A study on the migration of organic pollutants from recycled paperboard packaging materials to solid food matrices. *Food Chemistry*, 101(4), pp.1759–1768. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0308814606001294.

U.S. Environmental Protection Agency, 2016. *Paper Recycling Facts and Figures*. Available at: https://archive.epa.gov/wastes/conserve/materials/paper/web/html/faqs.html.

USDA, 1997. Food Industry Costs, Profits, and Productivity. *Food Cost Review, 1950-97*, pp.12–19.

Vellini, M. & Savioli, M., 2009. Energy and environmental analysis of glass container production and recycling. *Energy*, 34(12), pp.2137–2143. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360544208002375.

Venelampi, O. et al., 2003. The biodegradation and disintegration of paper products in the composting environment. *Compost Science and Utilization*, 11(3), pp.200–209.

Vera, P., Canellas, E. & Nerin, C., 2013. Identification of non-volatile compounds and their migration from hot melt adhesives used in food packaging materials characterized by ultra-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. *Analytical and Bioanalytical Chemistry*, 405(14), pp.4747–4754.

Viant, M.R. & Sommer, U., 2013. Mass spectrometry based environmental metabolomics: A primer and review. *Metabolomics*, 9(SUPPL.1), pp.144–158.

Wagner, C.C., 2014. Non-intentionally added substances (NIAS). *Food Packaging Forum*. Available at: http://www.foodpackagingforum.org/food-packaging-health/non-intentionally-added-substances-nias [Accessed January 5, 2018].

WHO/IPCS, 2009. *Principles and Methods for the Risk Assessment of Chemicals in Food*, Stuttgart, Germany: EHC 240.

Williams, H. et al., 2012. Reasons for household food waste with special attention to packaging. *Journal of Cleaner Production*, 24, pp.141–148. Available at: http://dx.doi.org/10.1016/j.jclepro.2011.11.044.

Wilm, M., 2011. Principles of Electrospray Ionization. *Molecular & Cellular Proteomics*, 10(7), p.M111.009407. Available at: http://www.mcponline.org/lookup/doi/10.1074/mcp.M111.009407.

World Health Organization, 2015. IARC Monographs evaluate consumption of red meat and processed meat and cancer risk. *International Agency of Research on Cancer*, (October), pp.1–2.

Yabannavar, A. & Bartha, R., 1993. Biodegradability of some food packaging materials in soil. *Soil Biology and Biochemistry*, 25(11), pp.1469–1475.

Zhu, X., Chen, Y. & Subramanian, R., 2014. Comparison of information-dependent acquisition, SWATH, and MS(All) techniques in metabolite identification study employing ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Analytical chemistry*, 86(2), pp.1202–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24383719.

Zou, S. et al., 2017. Semi-quantitative analysis of multiple chemical mixtures in solution at trace level by surface-enhanced Raman Scattering. *Scientific Reports*, 7(1), pp.1–9. Available at: http://dx.doi.org/10.1038/s41598-017-06543-y.

Zülch, A. & Piringer, O., 2010. Measurement and modelling of migration from paper and board into foodstuffs and dry food simulants. *Food Additives & Contaminants: Part A*, 27(9), pp.1306–1324. Available at: http://www.tandfonline.com/doi/abs/10.1080/19440049.2010.483693.

Appendix

Appendix contents:

Manuscript A — Supporting Information

A framework to estimate concentrations of potentially unknown substances by semi-quantification in Liquid Chromatography Electrospray Ionization Mass Spectrometry

Eelco N. Pieke[1], Kit Granby, Xenia T. Trier[a], Jørn Smedsgaard

Research Group for Analytical Food Chemistry, National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, Søborg, DK-2860, Denmark

Email contact:

Kit Granby: kgra@food.dtu.dk
Xenia Trier: Xenia.Trier@eea.europa.eu
Jørn Smedsgaard: smeds@dtu.dk

[1] Corresponding author; mail: enpi@food.dtu.dk, phone: +45 518 642 05
[a] Present address: The European Environment Agency, Kgs. Nytorv 6, Copenhagen K, DK-1050, Denmark

Manuscript A — Supporting Information

A.1. Supplementary Figures

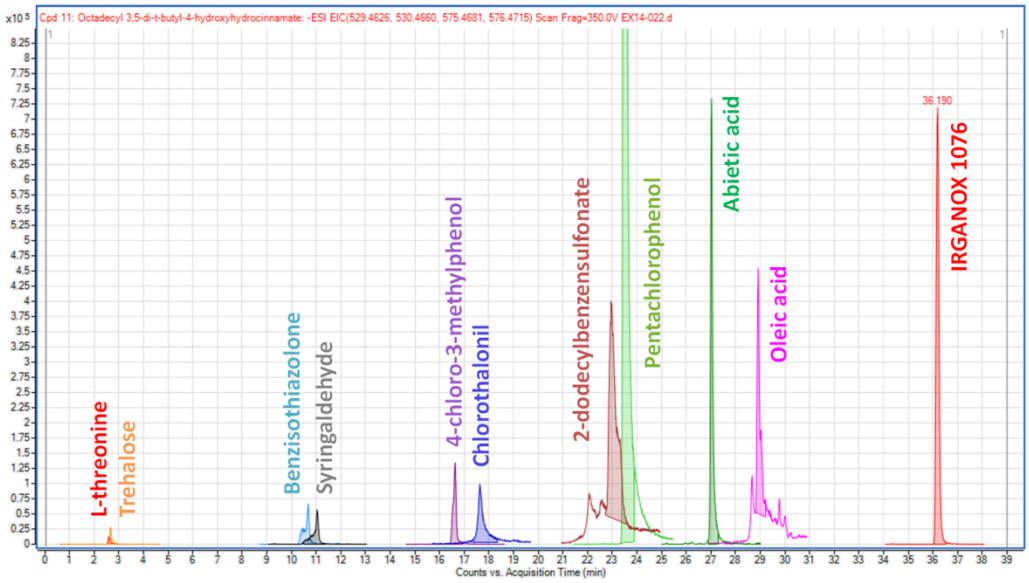


Figure A.1. Chromatogram obtained from analysis the ESI(-)mix in ESI⁻ SIM. Extracted ion chromatograms are the $[M-H]^-$ and $[M+COO]^-$ adducts predominantly. From left to right: L-threonine, α, α –trehalose (trehalose), benzisothiazolone, syringaldehyde, 4-chloro-3-methylphenol, chlorothalonil (as $[M-Cl+OH]^-$), 2-dodecylbenzensulfonate, pentachlorophenol, abietic acid, oleic acid, IRGANOX 1076. Co-eluting compounds at the start of the chromatogram are L-threonine and trehalose.

Manuscript A — Supporting Information

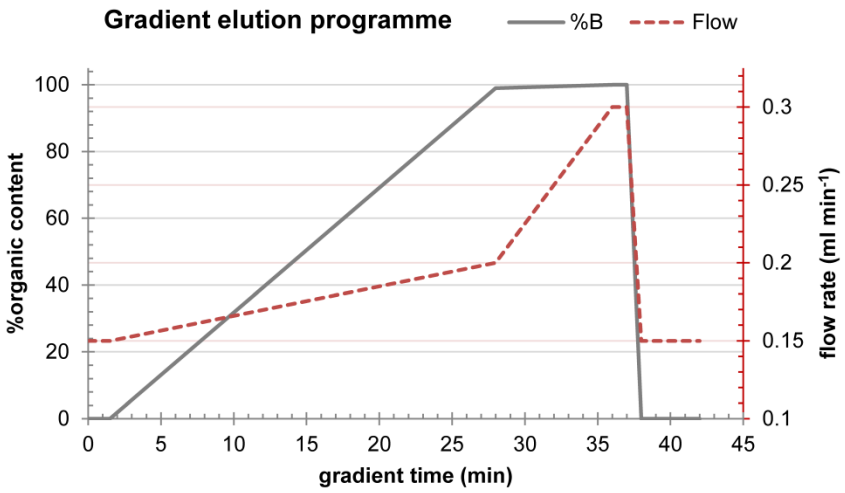


Figure A.2. Elution program used for separating the test mixtures. Both the flow and the percentage of organic content were varied in the run in order to more rapidly elute the most nonpolar compounds.

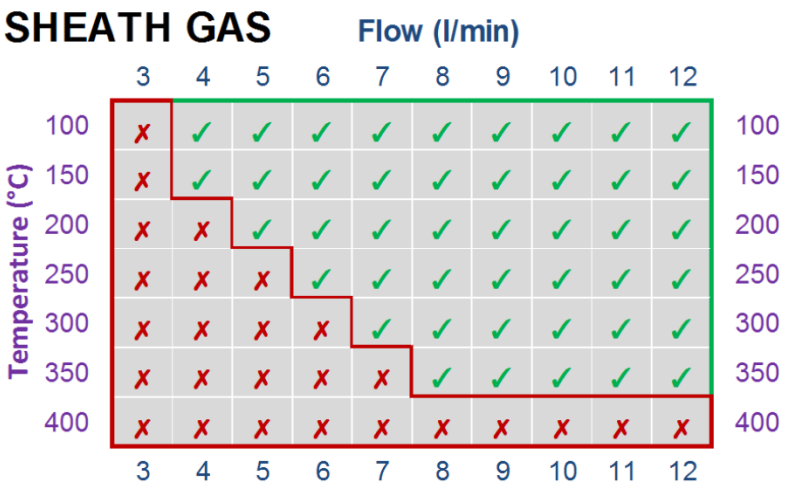


Figure A.3.1. Interaction table between Sheath Gas Flow (ShG-F) and Sheath Gas Temperature (ShG-T). Red crosses indicate a setting mismatch that would not result in a proper experiment. Adding a linear constraint to the possible combinations of ShG-T and ShG-F in the experiment design was required to obtain a functional design. The constraint was defined as: $[\text{ShG-T} - 50 \cdot \text{ShG-F}] \leq -50$

Manuscript A — Supporting Information

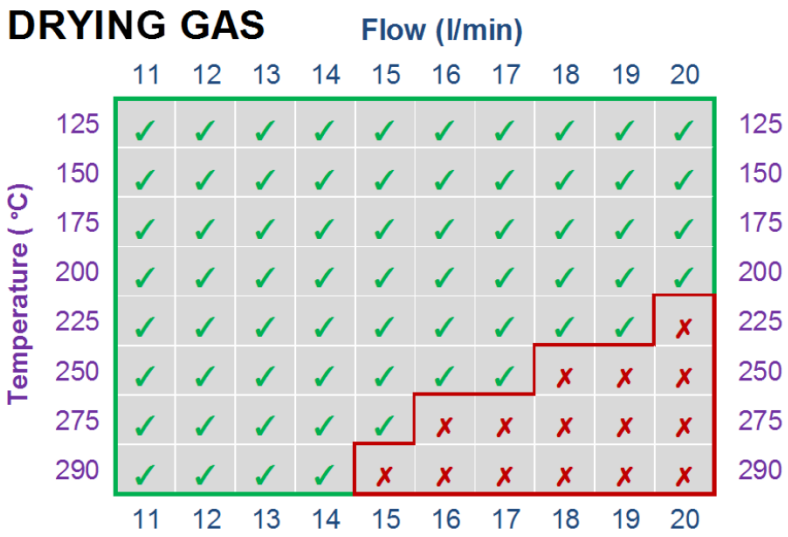


Figure A.3.2. Interaction table between Drying Gas Flow (DrG-F) and Drying Gas Temperature (DrG-T). Red crosses indicate a setting mismatch that would not result in a proper experiment. Adding a linear constraint to the possible combinations of DrG-T and DrG-F in the experiment design was required to obtain a functional design. The constraint was defined as: $[\text{DrG-T} + 15 \cdot \text{DrG-F}] \leq 505$

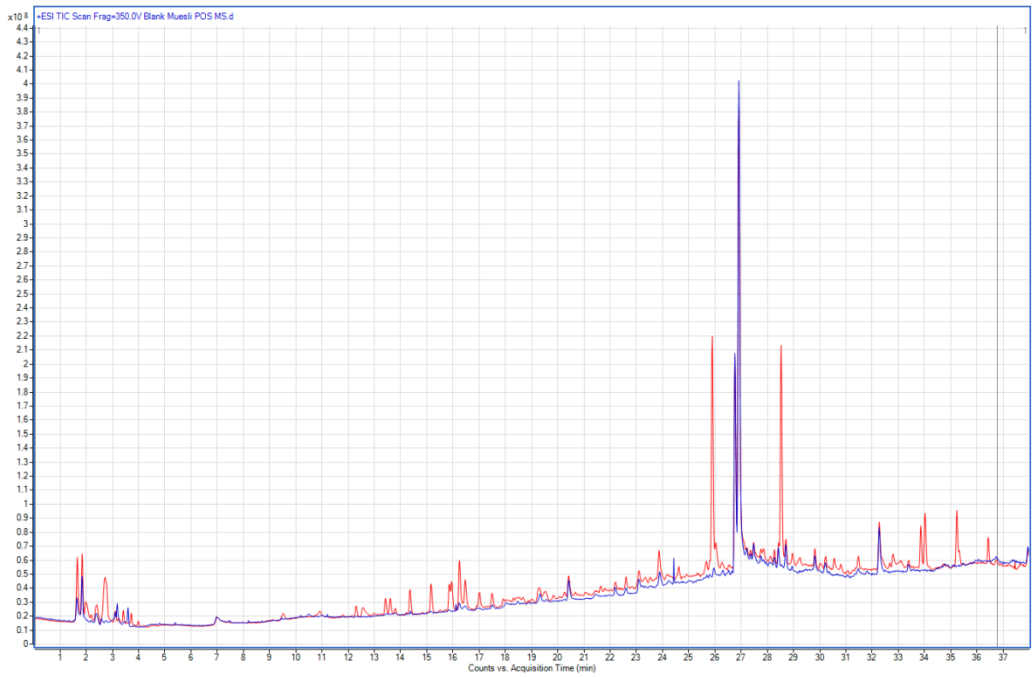


Figure A.4. Total Ion Chromatograph (TIC) of muesli box extract analysis (red) with subsequent method blank (blue). In the chromatogram, a notable increase of ion count in the sample extract compared to the blank was observed. Due to the usage of Time of Flight mass spectrometry, large background is persistent and more compounds may be present below the baseline.

Manuscript A — Supporting Information

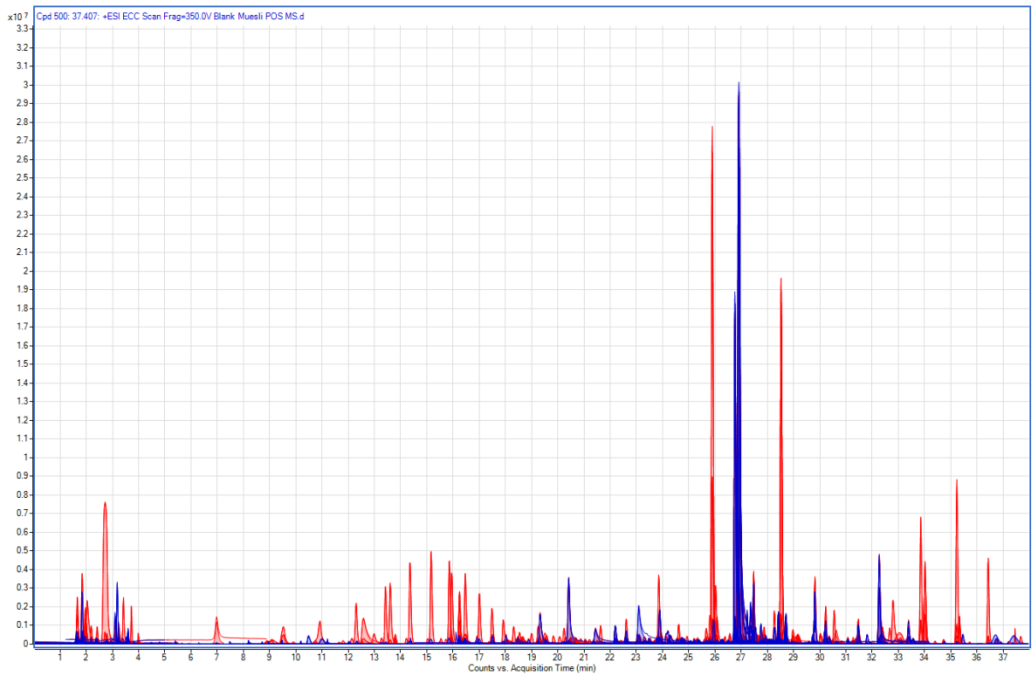


Figure A.5. Extracted compounds in the muesli box sample (red) and the method blank (blue) via the Molecular Feature Extractor. Extracted ion chromatograms of ion groups possibly representing compounds are shown for the extract (red) and for the blank (blue). Each peak represents a compound, so that peaks shown in red but missing are blue are unique identities to the sample. The number of unique compounds surpassed 300, but many peaks are not visible without zooming in on the axes.

A.2. Supplementary Tables

Table A.1. Test mixture standards used for the experimental design. The presence of a compound in a mixture is signified by its concentration. For detected compounds, a retention time is given. However, not all compounds in the mixtures were detected. The ESI^+-mix contained 18 compounds of which 10 were consistently detected, while the ESI^--mix contained 16 compounds of which 11 were detected. The identifier was used as short 5 symbol acronym for tabulating later properties of the compounds.

			ESI^+ mix		**ESI^- mix**	
Name	Identifier	Accurate mass	µmol L^{-1}	RT	µmolL^{-1}	RT
2,6-diisopropylnaphthalene	2,6-d.	212.1565	0.50	-	-	-
2-isopropylthioxanthen-9-one	2-iso.	254.077	0.28	24.558	-	-
4-chloro-3-methylphenol	4-chl.	142.0185	-	-	0.60	16.625
Abietic acid	Abiet.	302.2246	-	-	0.59	27.003
Benzisothiazolione	Benzi.	151.0092	0.51	10.644	0.60	10.578
Benzophenone	Benzo.	182.0732	0.50	19.468	-	-
Chlorothalonil	Chlor.	263.8816	-	-	0.59	17.681
Dicyclohexyl phthalate	Dicyc.	330.1831	0.51	25.947	0.60	-
Diethylenetriamine	Dieth.	103.1109	0.54	-	-	-
Diphenyl phthalate	Diphe.	318.0892	0.51	22.396	0.60	-
Distearyl thiodipropionate	Diste.	682.5934	0.28	-	0.33	-
Ethylenediamine	Ethyl.	60.0687	1.01	-	-	-
Hexatriacontane	Hexat.	506.5791	-	-	0.51	-
L-Threonine	L-Thr.	119.0582	0.56	2.642	0.66	2.619
Melamine	Melam.	126.0654	0.50	2.035	-	-
Methamidophos	Metha.	141.0013	0.53	6.660	-	-
Methyl stearate	Methy.	298.2872	0.50	-	-	-
Octadecane	Octad.	254.2974	0.51	-	-	-
Irganox 1076	Irgan.	530.4699	0.29	35.22†	0.35	36.112
Octanol	Octan.	130.1358	0.50	7.100	0.59	-
Oleic acid	Oleic.	282.2559	-	-	0.61	28.884
Pentachlorophenol	Penta.	263.8470	-	-	0.62	23.470
sodium 2-dodecylbenzenesulfonate	sodiu.	348.1735 325.1842‡	-	-	0.27	22.147
Syringaldehyde	Syrin.	182.0579	-	-	0.60	10.961
Tetracosane	Tetra.	338.3913	0.28	-	-	-
α,α-Trehalose	α,α-T.	342.1162	0.26	2.751	0.31	2.661

† Irganox 1076 was not consistently detected in Positive ESI mode and therefore excluded for optimization
‡ sodium 2-dodecylbenzenesulfonate dissociates in water, so the sodium ion should not be accounted for

Manuscript A — Supporting Information

Table A.2. Experimental design generation table. A custom screening design generated using JMP®, Version *12*. (SAS Institute Inc., Cary, NC, 1989-2015). This fractional factorial-type design has a resolution of approximately IV. Two linear constraints were added to prevent setting mismatching prior to generating the design: $\mathrm{ShG(T)} - [50 \cdot \mathrm{ShG(F)}] \leq -50$ and $\mathrm{DrG(T)} + [15 \cdot \mathrm{DrG(F)}] \leq 505$. The addition of constraints required several points at the maxima of the restrictions, which causes the design to be less efficient. Eight centerpoints and ten replicates were included. Repetitions are marked by a respective number or a dash if the experiment was not repeated. In total 56 experiments were generated, of which 39 can be considered unique combinations of factors. Each condition was tested separately per polarity.

EXP	Rep.	ShG-T (°C)	ShG-F (L min^{-1})	Neb-P (psi)	Cap-V (V)	Noz-V (V)	DrG-T (°C)	DrG-F (L min^{-1})
01	CP*	240	8.9	35	3000	500	210	14.9
02	38	350	8	50	4000	0	290	11.0
03	26	350	8	20	2000	0	150	11.0
04	-	150	12	20	2000	1000	290	11.0
05	CP*	240	9	35	3000	500	210	14.9
06	-	350	8	20	4000	1000	205	20.0
07	35	350	12	20	4000	0	290	11.0
08	-	150	4	50	4000	0	150	11.0
09	-	350	8	50	2000	0	248	17.2
10	48	350	8	50	2000	1000	150	11.0
11	-	350	8	20	4000	1000	150	11.0
12	55	350	8	50	4000	0	150	20.0
13	-	350	12	50	4000	0	150	11.0
14	-	150	4	50	4000	0	254	16.7
15	CP*	240	9	35	3000	500	210	14.9
16	51	150	4	50	4000	1000	150	20.0
17	24	350	8	20	2000	1000	290	11.0
18	-	150	4	20	2000	1000	150	12.0
19	CP*	240	9	35	3000	500	210	14.9
20	-	350	8	20	2000	0	150	20.0
21	-	150	12	50	4000	0	150	20.0
22	-	150	12	50	4000	0	290	11.0
23	-	150	4	50	2000	0	150	20.0
24	17	350	8	20	2000	1000	290	11.0
25	29	150	12	20	4000	0	150	11.0
26	03	350	8	20	2000	0	150	11.0
27	-	350	12	50	4000	1000	290	11.0
28	-	150	4	50	2000	1000	290	11.0
29	25	150	12	20	4000	0	150	11.0
30	-	150	12	20	2000	1000	150	11.0
31	-	150	4	20	2000	1000	237	17.9
32	-	350	12	20	2000	1000	239	17.7
33	-	150	12	50	4000	1000	150	11.0

Continued on next page

Manuscript A — Supporting Information

EXP	Rep.	ShG-T (°C)	ShG-F (L min^{-1})	Neb-P (psi)	Cap-V (V)	Noz-V (V)	DrG-T (°C)	DrG-F (L min^{-1})
34	-	150	4	20	4000	0	150	20.0
35	07	350	12	20	4000	0	290	11.0
36	-	350	12	50	2000	0	290	11.0
37	46	150	4	20	4000	1000	290	11.0
38	02	350	8	50	4000	0	290	11.0
39	-	150	12	20	2000	0	231	18.3
40	47	350	12	20	2000	1000	150	11.0
41	CP*	240	9	35	3000	500	210	14.9
42	-	150	12	20	4000	1000	231	18.3
43	-	150	12	50	2000	0	150	11.0
44	CP*	240	9	35	3000	500	210	14.9
45	-	150	12	50	2000	1000	239	17.7
46	37	150	4	20	4000	1000	290	11.0
47	40	350	12	20	2000	1000	150	11.0
48	10	350	8	50	2000	1000	150	11.0
49	CP*	240	9	35	3000	500	210	14.9
50	-	350	12	20	4000	0	150	20.0
51	16	150	4	50	4000	1000	150	20.0
52	CP*	240	9	35	3000	500	210	14.9
53	-	150	4	20	2000	0	290	11.0
54	-	150	12	20	2000	1000	150	20.0
55	12	350	8	50	4000	0	150	20.0
56	-	350	12	50	2000	1000	150	20.0

*CP: Centerpoints: 01, 05, 15, 19, 41, 44, 49, 52

Table A.3. Maximization data for ESI(+)mix analyses. Given values are gain/loss factors of RF related to the centrepoint, where large positive values indicate a larger analyte RF in that experiment compared to the centrepoint experiments. Average RF gains over +50% are marked in bold, as these signify experimental conditions where a significant gain in RF has been obtained.

RF type:	Low L-Thr.	Low α,α-T.	Low Benzo.	Medium Benzi.	Medium Octan.	Medium Diphe.	Medium Metha.	High 2-iso.	High Melam.	High Dicyc.	Average	Note
EX 001	-2%	+15%	-14%	+10%	-13%	+1%	-2%	+6%	+2%	+11%	+1%	Centre
EX 002	-47%	-68%	-76%	-4%	-30%	-51%	-32%	-5%	-79%	-32%	-42%	
EX 003	**+31%**	**-55%**	**-21%**	**+241%**	**+216%**	**-77%**	**+205%**	**+235%**	**-57%**	**+27%**	**+75%**	
EX 004	+2%	+15%	-51%	-31%	-26%	-56%	-41%	-50%	-50%	-32%	-32%	
EX 005	+18%	+1%	-11%	-22%	-6%	-0%	-13%	+1%	-7%	+9%	-3%	Centre
EX 006	-1%	+55%	+76%	+121%	-14%	-13%	-49%	+164%	-93%	+32%	+28%	
EX 007	-49%	-54%	+14%	+147%	+18%	-72%	-20%	+132%	-93%	-2%	+2%	
EX 008	-67%	-73%	-87%	-90%	-60%	-58%	-56%	-87%	-76%	-71%	-73%	
EX 009	+14%	-68%	-64%	+7%	+27%	-43%	+17%	+53%	+36%	-22%	-4%	
EX 010	+59%	+30%	-64%	+44%	-2%	-52%	+16%	+36%	+18%	-27%	+6%	
EX 011	+24%	+70%	+20%	+101%	-10%	-43%	-98%	+111%	-55%	+8%	+13%	
EX 012	-9%	-59%	-2%	+95%	-5%	+74%	-96%	+90%	+61%	+39%	+19%	
EX 013	+44%	-90%	+17%	+141%	-92%	+1%	-95%	+200%	+97%	+62%	+29%	
EX 014	-63%	-69%	-70%	-72%	-63%	-49%	-67%	-81%	-67%	-59%	-66%	
EX 015	-6%	+0%	-7%	+10%	-7%	+4%	-92%	+3%	-5%	+2%	-10%	Centre
EX 016	-24%	-19%	+73%	-55%	-64%	+6%	-24%	-52%	-59%	-28%	-25%	
EX 017	-65%	+29%	+1%	+36%	+12%	-82%	-28%	+160%	-58%	-7%	-0%	
EX 018	-27%	+2%	-33%	-39%	-46%	-36%	-50%	-64%	-59%	-52%	-40%	
EX 019	-16%	-2%	-12%	-25%	-15%	+2%	+0%	-4%	-7%	-3%	-8%	Centre
EX 020	**+25%**	**-44%**	**+119%**	**+328%**	**+215%**	**-13%**	**+205%**	**+355%**	**+58%**	**+64%**	**+131%**	
EX 021	-65%	-70%	-36%	-41%	-36%	+33%	+12%	-58%	-12%	-22%	-30%	
EX 022	-66%	-80%	-89%	-80%	-57%	-48%	-77%	-87%	-55%	-53%	-69%	
EX 023	-72%	-74%	-66%	-67%	-33%	-44%	+12%	-80%	-67%	-67%	-56%	
EX 024	+25%	+37%	+4%	+102%	-3%	-82%	-24%	+161%	-84%	-11%	+12%	
EX 025	-54%	-53%	-57%	-79%	-2%	-4%	+1%	-62%	-77%	-24%	-41%	
EX 026	**+56%**	**-50%**	**-28%**	**+351%**	**+188%**	**-79%**	**+204%**	**+214%**	**-43%**	**-2%**	**+81%**	

Manuscript A — Supporting Information

RF type:	Low L-Thr.	Low α,α-T.	Low Benzo.	Medium Benzi.	Medium Octan.	Medium Diphe.	Medium Metha.	High 2-iso.	High Melam.	High Dicyc.	Average	Note
EX 027	-15%	+48%	+43%	+65%	-17%	-38%	-44%	+173%	-20%	+29%	+22%	
EX 028	-30%	-34%	-73%	-70%	-51%	-68%	-59%	-85%	-57%	-70%	-60%	
EX 029	-54%	-53%	-61%	-42%	+6%	-8%	+10%	-65%	-23%	-30%	-32%	
EX 030	-3%	+26%	-57%	-34%	-76%	-37%	-100%	-56%	-77%	-29%	-44%	
EX 031	-7%	-16%	+20%	-18%	-47%	-24%	-49%	-43%	-90%	-30%	-30%	
EX 032	**+33%**	**+69%**	**+203%**	**+278%**	**+56%**	**-57%**	**-16%**	**+326%**	**-84%**	**+46%**	**+86%**	
EX 033	-16%	+2%	-61%	-56%	-45%	-17%	-23%	-75%	-38%	-36%	-37%	
EX 034	-65%	-65%	+28%	-38%	-36%	+23%	-97%	-47%	-79%	-34%	-41%	
EX 035	-46%	-89%	+34%	+169%	+22%	-74%	-27%	+124%	-60%	-19%	+3%	
EX 036	+114%	-41%	-27%	+246%	+31%	-77%	-95%	+139%	-5%	-13%	+27%	
EX 037	-37%	-14%	-40%	-46%	-48%	-56%	-68%	-72%	-75%	-54%	-51%	
EX 038	-43%	-72%	-70%	+0%	-17%	-52%	-10%	-7%	-17%	-39%	-33%	
EX 039	-53%	-72%	-1%	+47%	+123%	-12%	+104%	-18%	-6%	-20%	+9%	
EX 040	**+73%**	**+95%**	**+175%**	**+322%**	**+94%**	**-71%**	**+19%**	**+319%**	**-49%**	**+83%**	**+106%**	
EX 041	+23%	-4%	+21%	+21%	+29%	-3%	+24%	+2%	+5%	-6%	+11%	Centre
EX 042	-13%	+12%	+73%	-3%	-26%	-11%	-49%	-25%	-80%	-15%	-14%	
EX 043	-57%	-78%	-95%	-80%	+2%	-64%	-10%	-91%	-3%	-72%	-55%	
EX 044	-0%	-0%	+10%	+4%	+10%	+3%	+75%	-2%	+8%	-3%	+10%	Centre
EX 045	+22%	-24%	+12%	-45%	-11%	-14%	-38%	-44%	-22%	-28%	-19%	
EX 046	-36%	-15%	-35%	-46%	-63%	-54%	-89%	-72%	-75%	-54%	-54%	
EX 047	**+55%**	**+81%**	**+181%**	**+140%**	**+86%**	**-73%**	**+18%**	**+306%**	**-42%**	**+71%**	**+82%**	
EX 048	+54%	+28%	-57%	+27%	-36%	-57%	-98%	+26%	-32%	-41%	-19%	
EX 049	+6%	-2%	+21%	+17%	+25%	-7%	+10%	+1%	+7%	-12%	+7%	Centre
EX 050	**-44%**	**-39%**	**+144%**	**+259%**	**+93%**	**-20%**	**+17%**	**+184%**	**-48%**	**+21%**	**+57%**	
EX 051	-29%	-16%	+99%	-39%	-51%	-2%	-53%	-56%	-50%	-41%	-24%	
EX 052	-17%	-10%	+13%	+2%	+3%	-6%	+7%	-5%	+4%	-10%	-2%	Centre
EX 053	-80%	-77%	-92%	-77%	+12%	-75%	-41%	-90%	-78%	-74%	-67%	
EX 054	-14%	+16%	+107%	-1%	+3%	-1%	-17%	-27%	-40%	-17%	+1%	
EX 055	-30%	-60%	+13%	+77%	+52%	+54%	+57%	+69%	-37%	+10%	+21%	
EX 056	**+102%**	**+56%**	**+234%**	**+284%**	**+1%**	**+24%**	**-25%**	**+304%**	**-28%**	**+65%**	**+102%**	

Table A.4. Calculated values for S_L and S_H per experiment. S_L and S_H were calculated by taking the product of the three lowest and the three highest responses, respectively. N_i, the normalization ratio, was calculated by dividing S_L by S_H. In addition, the compounds responsible for calculating S_L and S_H are given. This data served as response input for the screening design of the parameter effects. CP: Centerpoints: 01, 05, 15, 19, 41, 44, 49, 52

EXP	Rep	ESI^- S_L ($*10^{18}$)	S_1, S_2, S_3	ESI^+ S_L ($*10^{18}$)	S_1, S_2, S_3	ESI^- S_H ($*10^{22}$)	S_1, S_2, S_3	ESI^+ S_H ($*10^{22}$)	S_1, S_2, S_3
01	CP	0.60	L-Thr., Benzi., Syrin.	1.39	L-Thr., α,α-T., Benzo.	3.76	Penta., sodiu., Abiet.	9.55	Dicyc., Melam., 2-iso.
02	38	0.04	L-Thr., α,α-T., Syrin.	0.06	L-Thr., α,α-T., Benzo.	1.40	Penta., sodiu., Abiet.	1.18	2-iso., Dicyc., Diphe.
03	26	0.15	L-Thr., α,α-T., p-Chl.	0.67	α,α-T., L-Thr., Benzo.	7.01	Penta., sodiu., Abiet.	28.82	2-iso., Dicyc., Metha.
04	-	3.31	L-Thr., Syrin., Benzi.	0.82	L-Thr., Benzo., α,α-T.	12.92	Penta., sodiu., Abiet.	1.36	Dicyc., Melam., 2-iso.
05	CP	2.49	L-Thr., Syrin., α,α-T.	1.51	L-Thr., α,α-T., Benzo.	7.69	Penta., sodiu., Abiet.	8.26	Dicyc., Melam., 2-iso.
06	-	2.83	L-Thr., p-Chl., Syrin.	3.35	L-Thr., α,α-T., Octan.	13.44	Penta., sodiu., Abiet.	11.30	2-iso., Dicyc., Diphe.
07	35	0.13	L-Thr., p-Chl., α,α-T.	0.39	L-Thr., α,α-T., Benzo.	1.43	Penta., sodiu., Abiet.	6.37	2-iso., Dicyc., Benzi.
08	-	0.04	L-Thr., α,α-T., Syrin.	0.02	L-Thr., Benzo., α,α-T.	1.65	Penta., sodiu., Abiet.	0.13	Dicyc., Melam., Diphe.
09	-	0.09	L-Thr., α,α-T., Syrin.	0.19	α,α-T., Benzo., L-Thr.	6.04	Penta., sodiu., Abiet.	13.04	Melam., 2-iso., Dicyc.
10	48	19.68	L-Thr., p-Chl., Syrin.	1.08	Benzo., L-Thr., α,α-T.	14.74	Penta., sodiu., Abiet.	9.43	Melam., 2-iso., Dicyc.
11	-	5.80	L-Thr., Benzi., p-Chl.	0.49	Metha., L-Thr., α,α-T.	13.07	Penta., sodiu., Abiet.	8.17	2-iso., Dicyc., Melam.
12	55	0.18	L-Thr., α,α-T., Syrin.	0.15	α,α-T., Metha., L-Thr.	4.25	Penta., sodiu., Abiet.	34.17	Melam., 2-iso., Dicyc.
13	-	0.33	L-Thr., α,α-T., p-Chl.	0.02	α,α-T., Octan., Metha.	5.64	Penta., sodiu., Abiet.	77.14	2-iso., Melam., Dicyc.
14	-	0.04	L-Thr., α,α-T., Syrin.	0.05	L-Thr., α,α-T., Benzo.	1.60	Penta., sodiu., Abiet.	0.32	Dicyc., Melam., Diphe.
15	CP	4.82	L-Thr., Syrin., α,α-T.	0.75	L-Thr., Metha., α,α-T.	14.42	Penta., sodiu., Abiet.	8.07	Dicyc., Melam., 2-iso.
16	51	2.09	L-Thr., Syrin., Chlor.	0.56	L-Thr., α,α-T., Octan.	5.28	Penta., sodiu., Abiet.	1.41	Dicyc., Diphe., Melam.
17	24	6.47	L-Thr., Syrin., p-Chl.	0.66	L-Thr., α,α-T., Benzo.	38.09	Penta., sodiu., Abiet.	8.19	2-iso., Dicyc., Melam.
18	-	7.17	L-Thr., Syrin., Benzi.	0.72	L-Thr., α,α-T., Benzo.	9.00	Penta., sodiu., Abiet.	0.59	Dicyc., Melam., Diphe.
19	CP	4.70	L-Thr., Syrin., p-Chl.	1.04	L-Thr., α,α-T., Benzo.	11.74	Penta., sodiu., Abiet.	6.94	Dicyc., Melam., 2-iso.
20	-	0.61	L-Thr., α,α-T., p-Chl.	2.18	α,α-T., L-Thr., Benzo.	38.52	Penta., sodiu., Chlor.	94.39	2-iso., Dicyc., Melam.
21	-	0.21	L-Thr., α,α-T., Syrin.	0.10	L-Thr., α,α-T., Benzo.	6.04	Penta., sodiu., Abiet.	4.16	Melam., Dicyc., Diphe.
22	-	0.07	L-Thr., α,α-T., Syrin.	0.01	Benzo., L-Thr., α,α-T.	3.03	Penta., sodiu., Abiet.	0.51	Dicyc., Melam., Diphe.
23	-	0.11	L-Thr., α,α-T., Syrin.	0.04	L-Thr., α,α-T., Benzo.	2.77	Penta., sodiu., Abiet.	0.33	Dicyc., Melam., Metha.
24	17	10.34	L-Thr., p-Chl., Benzi.	2.56	L-Thr., α,α-T., Benzo.	54.96	Penta., sodiu., Abiet.	5.30	2-iso., Dicyc., Benzi.
25	29	0.93	L-Thr., α,α-T., Syrin.	0.13	L-Thr., α,α-T., Benzo.	2.79	Penta., sodiu., Chlor.	1.03	Dicyc., Diphe., 2-iso.

Manuscript A — Supporting Information

EXP	Rep	ESI^- S_L ($*10^{18}$)	S_1, S_2, S_3	ESI^+ S_L ($*10^{18}$)	S_1, S_2, S_3	ESI^- S_H ($*10^{22}$)	S_1, S_2, S_3	ESI^+ S_H ($*10^{22}$)	S_1, S_2, S_3
26	03	0.40	L-Thr., α,α-T., p-Chl.	0.81	α,α-T., L-Thr., Benzo.	25.46	Penta., sodiu., Chlor.	20.74	2-iso., Dicyc., Metha.
27	-	6.52	L-Thr., Syrin., p-Chl.	2.59	L-Thr., α,α-T., Benzo.	18.54	Penta., sodiu., Abiet.	22.59	2-iso., Dicyc., Melam.
28	-	1.54	L-Thr., Syrin., Benzi.	0.18	L-Thr., Benzo., α,α-T.	3.74	Penta., sodiu., Abiet.	0.18	Melam., Dicyc., Diphe.
29	25	0.76	L-Thr., Syrin., α,α-T.	0.12	L-Thr., α,α-T., Benzo.	2.02	Penta., sodiu., Abiet.	2.28	Melam., Dicyc., Diphe.
30	-	12.32	L-Thr., Syrin., α,α-T.	0.01	Metha., L-Thr., Octan.	22.57	Penta., sodiu., Abiet.	0.74	Dicyc., 2-iso., Diphe.
31	-	2.74	L-Thr., Syrin., Benzi.	1.05	L-Thr., α,α-T., Octan.	10.04	Penta., sodiu., Abiet.	1.13	Dicyc., 2-iso., Diphe.
32	-	9.31	L-Thr., Syrin., p-Chl.	8.98	L-Thr., α,α-T., Octan.	52.97	Penta., sodiu., Abiet.	26.68	2-iso., Dicyc., Benzi.
33	-	6.57	L-Thr., Syrin., Chlor.	0.48	L-Thr., Benzo., α,α-T.	7.25	Penta., sodiu., Abiet.	1.52	Dicyc., Melam., Diphe.
34	-	0.32	L-Thr., α,α-T., Syrin.	0.03	L-Thr., Metha., α,α-T.	1.93	Penta., sodiu., Abiet.	1.61	Dicyc., Diphe., 2-iso.
35	07	0.28	L-Thr., p-Chl., α,α-T.	0.11	α,α-T., L-Thr., Benzo.	2.27	Penta., sodiu., Abiet.	5.85	2-iso., Dicyc., Melam.
36	-	0.19	L-Thr., α,α-T., p-Chl.	0.64	Metha., α,α-T., Benzo.	15.92	Penta., sodiu., Abiet.	16.00	2-iso., Melam., Dicyc.
37	46	2.24	L-Thr., Syrin., Benzi.	0.47	L-Thr., α,α-T., Benzo.	5.45	Penta., sodiu., Abiet.	0.25	Dicyc., Melam., 2-iso.
38	02	0.17	L-Thr., α,α-T., p-Chl.	0.07	α,α-T., L-Thr., Benzo.	5.47	Penta., sodiu., Abiet.	3.74	Melam., 2-iso., Dicyc.
39	-	0.31	L-Thr., α,α-T., Syrin.	0.19	L-Thr., α,α-T., Benzo.	12.75	Penta., sodiu., Chlor.	4.92	Melam., Dicyc., 2-iso.
40	47	33.14	L-Thr., p-Chl., Syrin.	13.31	L-Thr., α,α-T., Benzo.	67.50	Penta., sodiu., Abiet.	36.85	2-iso., Dicyc., Benzi.
41	CP	8.43	L-Thr., Benzi., Syrin.	2.07	L-Thr., α,α-T., Benzo.	14.09	Penta., sodiu., Abiet.	8.14	Melam., Dicyc., 2-iso.
42	-	2.89	L-Thr., Benzi., Syrin.	1.83	L-Thr., α,α-T., Octan.	12.90	Penta., sodiu., Abiet.	2.15	Dicyc., 2-iso., Diphe.
43	-	0.28	L-Thr., α,α-T., Syrin.	0.01	Benzo., α,α-T., L-Thr.	3.68	Penta., sodiu., Abiet.	0.65	Melam., Dicyc., Metha.
44	CP	10.15	L-Thr., Syrin., α,α-T.	1.59	L-Thr., α,α-T., Benzo.	15.11	Penta., sodiu., Abiet.	8.14	Melam., Dicyc., 2-iso.
45	-	8.82	L-Thr., Syrin., α,α-T.	1.50	L-Thr., α,α-T., Benzo.	7.95	Penta., sodiu., Abiet.	2.53	Melam., Dicyc., 2-iso.
46	37	1.87	L-Thr., Syrin., Benzi.	0.52	L-Thr., α,α-T., Octan.	2.44	Penta., sodiu., Abiet.	0.26	Dicyc., Melam., 2-iso.
47	40	21.53	L-Thr., p-Chl., Benzi.	11.29	L-Thr., α,α-T., Benzo.	64.49	Penta., sodiu., Abiet.	32.37	2-iso., Dicyc., Melam.
48	10	16.08	L-Thr., Syrin., p-Chl.	0.26	Metha., Benzo., L-Thr.	19.97	Penta., sodiu., Abiet.	4.06	2-iso., Melam., Dicyc.
49	CP	6.60	L-Thr., Syrin., p-Chl.	1.82	L-Thr., α,α-T., Benzo.	13.93	Penta., sodiu., Abiet.	7.66	Melam., Dicyc., 2-iso.
50	-	0.34	L-Thr., p-Chl., Syrin.	1.21	L-Thr., α,α-T., Benzo.	2.20	Penta., sodiu., Abiet.	14.30	2-iso., Dicyc., Melam.
51	16	4.41	L-Thr., Benzi., Syrin.	0.74	L-Thr., α,α-T., Octan.	3.95	Penta., sodiu., Abiet.	1.34	Dicyc., Melam., Diphe.
52	CP	6.67	L-Thr., Syrin., α,α-T.	1.22	L-Thr., α,α-T., Benzo.	14.84	Penta., sodiu., Abiet.	7.10	Melam., Dicyc., 2-iso.
53	-	0.07	L-Thr., α,α-T., Benzi.	0.01	L-Thr., Benzo., α,α-T.	3.52	Penta., sodiu., Abiet.	0.09	Dicyc., Melam., Metha.
54	-	15.10	L-Thr., Syrin., α,α-T.	2.62	L-Thr., α,α-T., Octan.	17.74	Penta., sodiu., Abiet.	2.94	Dicyc., Melam., 2-iso.
55	12	0.45	L-Thr., α,α-T., Syrin.	0.46	L-Thr., α,α-T., Benzo.	7.50	Penta., sodiu., Abiet.	10.78	2-iso., Dicyc., Diphe.
56	-	28.09	L-Thr., Benzi., p-Chl.	8.10	L-Thr., α,α-T., Octan.	38.14	Penta., sodiu., Abiet.	38.39	2-iso., Dicyc., Melam.

Manuscript A — Supporting Information

A note regarding tables A.5.1 to A.5.4

Tables A.5.1 to A.5.4 display 2nd order factor interactions observed in the experimental design. Intersections in the table represent an interaction parameter. For example, the intercept between Sheath gas temperature (ShG-T) and Nebulizer pressure (Neb-P) represent the interaction parameter ShG-T*Neb-P.

The significance of an interaction is marked by its LtR value. LtR values larger than 1.65 (p-value ≈ 0.10) are considered active and marked colored in blue for 1st order, and green for 2nd order parameters. Values below this threshold are considered insignificant by the model. Lack of data due to factor space exhaustion is marked with *N.D.*.

The value of LtR indicates the probably of an actual effect by that parameter on the listed response, whereas higher positive or negative numbers correlate with stronger significance: LtR = ±2.00 equals a p-value of 0.05, whereas an LtR of ±1.65 approximates p-value = 0.10. The direction (sign) of the LtR predicts whether the effect is positive, increased signal at higher parameter setting, or negative, decreased signal at higher parameter setting.

More information on the computation and derivation of LtR values is available at http://www.jmp.com/support/help/Estimates.shtml#742843.

Manuscript A — Supporting Information

Table A.5.1: Factor table for **ESI+ S_L** (Positive ionization, low-responding compounds).

		+6.25	+4.06	-2.10	-2.86	+6.76	-1.88	+2.67
	ESI+ S_L	ShG-T	ShG-F	Neb-P	Cap-V	Noz-V	DrG-T	DrG-F
+6.25	ShG-T	*N.D.*	+6.35	-1.50	+0.96	+4.37	-2.10	+1.49
+4.06	ShG-F		*N.D.*	-0.54	-1.78	+2.73	-1.95	-0.12
-2.10	Neb-P			*N.D.*	+1.80	-2.32	+2.27	+0.68
-2.86	Cap-V				*N.D.*	-0.20	+1.62	-0.23
+6.76	Noz-V					*0.00*	-0.97	-0.25
-1.88	DrG-T						*-3.77*	+0.67
+2.67	DrG-F							*-0.25*

Table A.5.2: Factor table for **ESI+** N_i (Positive ionization, normalization ratio).

		-6.48	-4.35	-3.35	+1.79	+10.2	+3.97	+0.02
	ESI+ N_i	ShG-T	ShG-F	Neb-P	Cap-V	Noz-V	DrG-T	DrG-F
-6.48	ShG-T	*N.D.*	+2.04	+1.44	-2.03	-5.02	+0.69	+0.27
-4.35	ShG-F		*N.D.*	+1.03	-0.67	+0.51	-2.28	+2.82
-3.35	Neb-P			*N.D.*	-1.64	-2.80	+0.16	+0.50
+1.79	Cap-V				*N.D.*	+0.78	+3.55	-1.35
+10.2	Noz-V					*+2.10*	+0.17	+0.12
+3.97	DrG-T						*-0.01*	-0.19
+0.02	DrG-F							*-1.51*

Manuscript A — Supporting Information

Table A.5.3: Factor table for **−ESI S_L** (negative ionization, low-responding compounds).

		+3.29	+4.58	+1.18	-4.01	+9.44	-4.27	-1.15
	−ESI S_L	ShG-T	ShG-F	Neb-P	Cap-V	Noz-V	DrG-T	DrG-F
+3.29	ShG-T	*N.D.*	N.D.	+0.97	-1.64	+2.91	-0.87	-0.69
+4.58	ShG-F		*+2.99*	+0.12	-1.54	+0.54	+0.68	-0.11
+1.18	Neb-P			*N.D.*	-0.33	+0.36	-0.29	+1.39
-4.01	Cap-V				*N.D.*	+0.02	+1.84	-0.35
+9.44	Noz-V					*-0.74*	+0.24	-0.25
-4.27	DrG-T						*+1.51*	-0.55
-1.15	DrG-F							*-0.31*

Table A.5.4: Factor table for **−ESI** N_i (negative ionization, normalization ratio).

		-1.21	+0.39	+1.74	+0.09	+6.16	-2.38	-0.93
	−ESI N_i	ShG-T	ShG-F	Neb-P	Cap-V	Noz-V	DrG-T	DrG-F
-1.21	ShG-T	*N.D.*	-0.74	+1.12	-0.04	+0.60	-0.11	+0.02
+0.39	ShG-F		*N.D.*	+0.44	-0.81	-0.50	-0.30	+0.72
+1.74	Neb-P			*N.D.*	-1.25	-0.09	-0.16	0.00
+0.09	Cap-V				*N.D.*	-0.91	+0.85	-0.53
+6.16	Noz-V					*-1.11*	-0.28	-0.32
-2.38	DrG-T						*+0.64*	+0.58
-0.93	DrG-F							*-0.43*

A.3. Supplementary Text

In order to make comparison between predictions easier, an arbitrary but more visual "score" was defined for the prediction result. Normally, standardization (z-scores) is adequate to scale the data. Conventional standardization scales the value on a scale defined by the mean and the standard deviation. However, since the prediction errors are not equally scaled in the positive and negative directly, the ideal value is 1 ranging from 0 to infinite, alternations to the scoring values were required. In addition, a stronger penalizing on low scores is preferable to make good scores stand out. The idea of scoring was to create a scale from 0 to 100 that provided performance information that prediction errors initially did not.

Firstly, all predictions were returned to an absolute format that did not differentiate between positive and negative values. Since errors in either direction are equally important, scaling overestimations and underestimations similarly is needed. Hence, in order to achieve the same scale for overestimations and underestimations, the ratio of values was used. Consequently, all predictions were simplified by dividing the smallest value by the largest value, irrespective of overestimation or underestimations, resulting in the Scaled Error, see below. In the scaled error, an under- or overestimation of 2-fold would both result in 0.5 (1 / 2), and the scaled error is always a value between 0 and 1.

$$\text{Scaled Error} = \frac{C_{\text{small}}}{C_{\text{large}}}$$

In order to improve the readability of the scaled error, the scaled errors were standardized on a 0 to 1 scale. In standardization or z-scoring, all respective values are ranked on their distance to a mean compared to a defined standard deviation, so that values that are exactly 1 standard deviation unit away from the mean are scored as ±1. z-scoring therefore represents the standard deviation unit distance from the mean of a value. As a result, the choice of mean and standard deviation define how strict the scoring system was. Due to the rather unique limitation of the data format, being exclusively in the plane of 0 to 1, standardization can actually be employed as penalization by strongly penalizing low values (bad predictions) but unaffected high values (good predictions). If the mean was chosen as 0 and the standard deviation as 1, the scaled error would return itself as the result of the standardization.

Factually, the mean and standard deviation defined at what level the prediction was still desirable. This question is not answered univocally, but it is clear that predictions with extremely high errors are not desirable. In this case, we decided that a 20-fold error in prediction was borderline unacceptable, but larger or smaller values can be augmented for. The choice of these values is based on the assumption that a single order of magnitude error may still result in valuable data for indicative quantification, but that much larger errors will not actively contribute. However, the choice of standardization parameters is depending entirely on the application. **If it is initially unknown what is acceptable and unacceptable, one should use the scaled error as rating criteria.**

A 20-fold prediction error limit is achieved by choosing the mean 0.05 (1/20) and the standard deviation 0.95 (1-(1/20)), which effectively caused more than 10-fold error to have very low scores (below 5) and values below 0.05 will score negative. Finally, since negative scores are possible, these were set to 0. Since the scaled error is on a 0 to 1 scale, which may be confused with the score, all scores were multiplied by 100 to bring them on a 0 to 100 scale. This effect was only aesthetic, but improved the readability of the tables, i.e., zeroes stood out more. The resulting scoring scale can then be used to compare quantification performance. In general, scores that exceed 80 are considered to be very good, scores between 50 and 79 are good, scores between 20 and 49 are acceptable, and scores below 10 should ideally be avoided. Scores of 0 indicate a bad quantification match.

Manuscript A — Supporting Information

Table A.6 provides a conversion table for different type of error reported in the main article and in this Supplementary Information.

Table A.6. Conversion table for order-of-magnitude error (x-fold), prediction error, scaled error, and score.

$C_{AOI,pred}$	$C_{AOI,true}$	x-fold error	Pred. Error	Scaled Error	Score
0.025	1.000	40-fold	-98%	0.025	-2.6
0.050	1.000	20-fold	-95%	0.050	0.0
0.100	1.000	10-fold	-90%	0.100	5.3
0.250	1.000	4-fold	-75%	0.250	21.1
0.500	1.000	2-fold	-50%	0.500	47.4
0.850	1.000	1.2-fold	-15%	0.850	84.2
1.000	1.000	-	0%	1.000	100.0
1.100	1.000	1.1-fold	10%	0.909	90.4
1.5	1.000	1.5-fold	50%	0.667	64.9
2.5	1.000	2.5-fold	150%	0.400	36.8
4	1.000	4-fold	300%	0.250	21.1
15	1.000	15-fold	1400%	0.067	1.8

Manuscript B — Supporting Information

Supplementary material to:

Exploring the chemistry of complex samples by tentative identification and semi-quantification: a food contact material case

Eelco N. Pieke, Jørn Smedsgaard, and Kit Granby

Technical University of Denmark, National Food Institute, Research Group for Analytical Food Chemistry (DTU), Kemitorvet Building 202, Kgs. Lyngby, DK-2800, Denmark

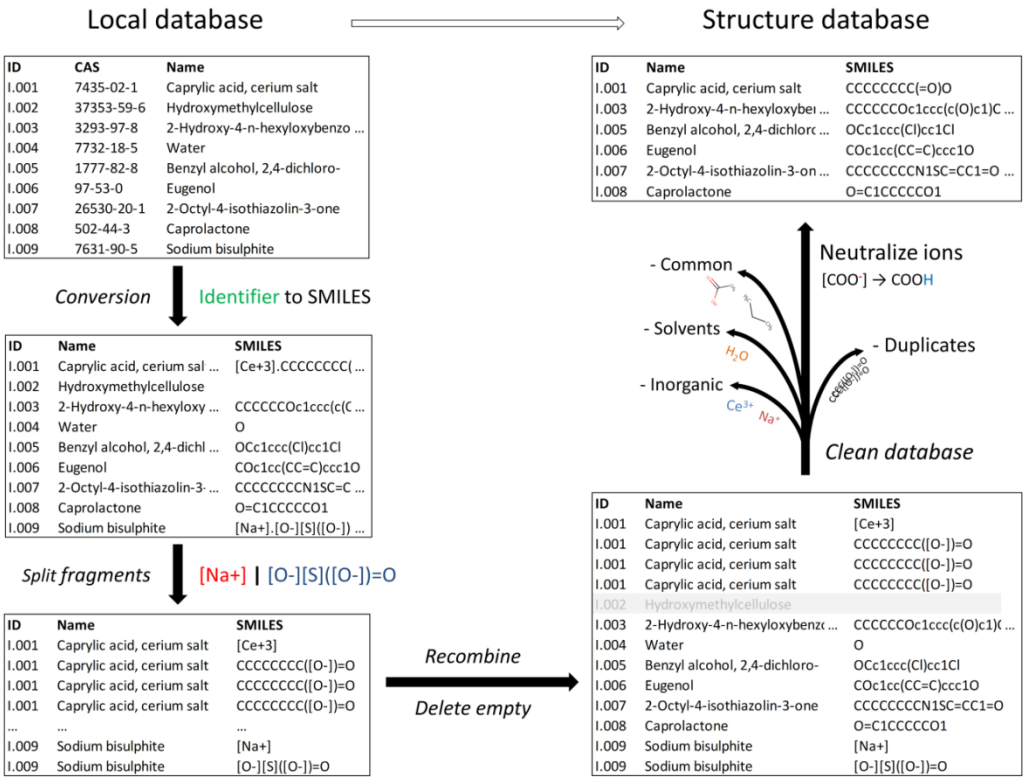
Manuscript B — Supporting Information

A.1. Generation of in-house databases

Each local database was converted to a delimited file where each line represented an entry, and subsequent columns contained structural identifiers or properties. Each row entry was assigned a Unique Identity (UID) containing a database source string, e.g.: DB.001. In total, over 80,000 UID were collected. Each UID was converted from ambiguous identifiers (e.g., molecule name or CAS number) to a *simplified molecular-input line-entry system* (SMILES) format prior to any processing. Conversions were performed by Marvin Calculator Plugins supplemented by in-house developed Excel application scripts written in Visual Basic for Applications (VBA).

First, Marvin was used for identifier conversion in batch mode. Two separate conversions were performed: *CAS to SMILES*, which uses a number of web-service lookups, and *Name to SMILES*, which used internal Marvin repositories. Following Marvin, the in-house Excel VBA script consulted asynchronously web-based lookup services to convert an input chemical identifier to an output SMILES. For this conversion, two web-services used were: Chemical Translation Service (CTS, http://cts.fiehnlab.ucdavis.edu/), and the Chemical Identifier Resolver (CACTUS, https://cactus.nci.nih.gov/chemical/structure). The conversion results were two conversions for Marvin and Excel VBA: CAS to SMILES and NAME to SMILES, totaling four possible conversions per entry.

Following, the best SMILES were selected for each UID, where the Marvin conversions were prioritized over the web service conversions; and, CAS conversions over name conversions. As a result, each UID was stored along with a single SMILES string. However, some of the SMILES strings were composed of multiple fragments, thereby rendering them void. To include these entries, multi-fragment molecules, e.g., salts, or mixtures were split to multiple entries with the same molecular identifier. VBA code was used to split SMILES based on the number of fragments, and each fragment for an UID was stored as a new separate entry with the same UID, e.g., ions would be split to their cationic and anionic parts as separate entry representing the same UID. To improve database quality, the converted databases were “cleaned” prior to use, which is graphically shown in Figure A.1.1. The total list of UID was cleaned using Standardizer software (ChemAxon). The Standardizer removed perceived irrelevant UIDs: atomic entries (e.g., salts), solvents (e.g., water), and common species (e.g., CO_2, acetic acid, propane). Next, charged organic molecules were neutralized by protonation or deprotonation, and retained as neutral form in the database. The final step of pruning involved removing any duplicate or empty SMILES from the database.



ID	CAS	Name
I.001	7435-02-1	Caprylic acid, cerium salt
I.002	37353-59-6	Hydroxymethylcellulose
I.003	3293-97-8	2-Hydroxy-4-n-hexyloxybenzo ...
I.004	7732-18-5	Water
I.005	1777-82-8	Benzyl alcohol, 2,4-dichloro-
I.006	97-53-0	Eugenol
I.007	26530-20-1	2-Octyl-4-isothiazolin-3-one
I.008	502-44-3	Caprolactone
I.009	7631-90-5	Sodium bisulphite

ID	Name	SMILES
I.001	Caprylic acid, cerium sal ...	[Ce+3].CCCCCCCC(...
I.002	Hydroxymethylcellulose	
I.003	2-Hydroxy-4-n-hexyloxy ...	CCCCCCOc1ccc(c(C ...
I.004	Water	O
I.005	Benzyl alcohol, 2,4-dichl ...	OCc1ccc(Cl)cc1Cl
I.006	Eugenol	COc1cc(CC=C)ccc1O
I.007	2-Octyl-4-isothiazolin-3· ...	CCCCCCCCN1SC=C ...
I.008	Caprolactone	O=C1CCCCCO1
I.009	Sodium bisulphite	[Na+].[O-][S]([O-]) ...

ID	Name	SMILES
I.001	Caprylic acid, cerium salt	[Ce+3]
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
...	...	...
I.009	Sodium bisulphite	[Na+]
I.009	Sodium bisulphite	[O-][S]([O-])=O

ID	Name	SMILES
I.001	Caprylic acid, cerium salt	CCCCCCCC(=O)O
I.003	2-Hydroxy-4-n-hexyloxybe ...	CCCCCCOc1ccc(c(O)c1)C ...
I.005	Benzyl alcohol, 2,4-dichlor ...	OCc1ccc(Cl)cc1Cl
I.006	Eugenol	COc1cc(CC=C)ccc1O
I.007	2-Octyl-4-isothiazolin-3-on ...	CCCCCCCCN1SC=CC1=O ...
I.008	Caprolactone	O=C1CCCCCO1

ID	Name	SMILES
I.001	Caprylic acid, cerium salt	[Ce+3]
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
I.001	Caprylic acid, cerium salt	CCCCCCCC([O-])=O
I.002	Hydroxymethylcellulose	
I.003	2-Hydroxy-4-n-hexyloxybenzo ...	CCCCCCOc1ccc(c(O)c1)(...
I.004	Water	O
I.005	Benzyl alcohol, 2,4-dichloro-	OCc1ccc(Cl)cc1Cl
I.006	Eugenol	COc1cc(CC=C)ccc1O
I.007	2-Octyl-4-isothiazolin-3-one	CCCCCCCCN1SC=CC1=O
I.008	Caprolactone	O=C1CCCCCO1
I.009	Sodium bisulphite	[Na+]
I.009	Sodium bisulphite	[O-][S]([O-])=O

Figure A.1.1: Stepwise generation of structure database from most commonly available databases. First, the identifiers (e.g., CAS, molecule name) are converted to structure-based strings. Then, fragments are split to new entries. The new entries are combined with the old and any resulting empty identities are removed. Finally, the database is cleaned of common molecules, solvents, inorganic species, and duplicate entries. Before the database is stored, charged species are neutralized where possible to their common forms.

Manuscript B — Supporting Information

A.2. Contents of the Quantification Marker mixture

Table A.2.1. Quantification marker mixture concentration and measured retention times. POS or NEG: positive or negative mode electrospray ionization. Concentration in sample is the expected value in the extractions, post-dilution by sample extract.

Name	Retention time (min)	Ionization mode	Concentration in stock (µmol liter^{-1})	Concentration in sample (µmol liter^{-1})
L-Threonine	2.42	POS	0.913	0.609
2-diethylaminoethanol	2.70	POS	1.012	0.675
L-Aspartic acid	3.05	NEG	0.626	0.417
α,α-trehalose	3.11	NEG/POS	0.486	0.324
1-Octanol	11.76	POS	0.574	0.383
Adipic acid	11.81	NEG/POS	0.907	0.605
Salicylic acid	13.49	NEG	0.548	0.366
Syringaldehyde	15.85	NEG/POS	0.626	0.417
Benzisothiazolone	16.25	NEG/POS	0.665	0.444
2-Mercaptobenzothiazole	19.86	POS	0.645	0.430
4-Chloro-3-methylphenol	21.93	NEG	0.800	0.533
Benzophenone	23.86	POS	0.664	0.442
Perfluorooctanoic acid	24.74	NEG	0.315	0.210
Dichlorophene	25.04	NEG	0.575	0.383
Diphenyl phthalate	25.90	POS	0.589	0.393
Pentachlorophenol	27.81	NEG	0.553	0.369
Sodium dodecyl benzenesulfonate	28.20	NEG	0.317	0.211
Dicyclohexyl phthalate	28.56	POS	0.577	0.384
Flumethrin	29.87	NEG/POS	0.219	0.146
Hexabromocyclododecane	30.19	NEG	0.311	0.207
Abietic acid	30.50	NEG/POS	0.562	0.375
Oleic acid	31.80	NEG	0.829	0.552
IRGANOX 3114	34.06	NEG/POS	0.302	0.202