Investigation into the use of NMR-based bioinformatics in determining the Composition and Quality of Immune Supplements in Australia.

By Courtney Amber Hall

BSc in Genetics & Molecular biology and, Crop & Pasture Science

This thesis is presented for the degree Research Masters with Training of Murdoch University

2021

Perth, Western Australia

Supported by the Australian National Phenome Centre (ANPC)

Declaration

I, Courtney Hall, verify that in submitting this thesis;

the thesis is my own account of the research conducted by me, except where other sources are fully acknowledged in the appropriate format;

the extent to which the work of others has been used is documented by a percent allocation of work and signed by myself and my Principal Supervisor;

that any editing and proof-reading by professional editors comply with the standards approved by the Research Degrees and Scholarship Committee;

that all necessary ethics and safety approvals were obtained, including their relevant approval or permit numbers, as appropriate.

Courtney Amber Hall

Abstract

The outbreak of the SARS-CoV-2 virus has brought prominence to the concept of immune health for individuals. A common means of attempting to do so is by incorporating immune supplements into everyday life. While immune supplements generally contain well-documented traditional herbs, knowledge about the quality and safety of these commercial products is minimal. In Australia, the Therapeutic Goods Administration (TGA) regulates and enforces advertising, labelling and compositional consistency of immune supplements; however, minimal pre-market assessment omits the potential harm and adulteration regularly cited in the literature.

A multifaceted approach to these products' overall safety and quality is essential in safeguarding human health. Following TGA guidelines, seventeen immune supplements were investigated for their labelling compliance with the *Therapeutic Goods Order No. 92* for non-prescription medicines. Although systemic labelling non-compliance was observed throughout the products, this was not associated with their potential to cause harm. Thus, stringency in this area is not necessarily applicable to protecting consumers. More focus should be put on high throughput pharmacovigilance methods that examine immune supplements' compositional integrity and consistency. For this study, the composition of immune supplements was analysed via nuclear magnetic resonance (NMR) spectroscopy using metabolomics. NMR provides detailed 'snap shots' into the chemical profile of immune supplements that can be interpreted via multivariate statistics to indicate the consistency of products across numerous batches. Therefore, this thesis aims to provide an overview of the quality and safety of Australian immune supplements. At the same time, it is recognising the place of metabolomics in regulatory environments as a high throughput mechanism of quality assurance.

Acknowledgements

At the beginning of this project, I had no idea the amount of dedication, time, and learning needed to complete this thesis in a year. Luckily enough, I had the support of two supervisors who helped me in many ways and at all hours.

To Garth Maker, my primary supervisor. A full-time advisor and a part-time therapist, your understanding of academics, natural medicines and psychology was the foundation of my success.

Secondly, Torben Kimhofer who pushed me in ways that accelerated my learning years beyond what I thought capable. Thank you for teaching me to code, do statistics, and most importantly, not give up on the first few hundred errors I received while learning python.

You have both taught me skills far beyond the classroom, and I am so grateful for your commitment.

To Kyle and Novia, my coding partners in crime. Thank you for always being there when a for loop went wrong, or I needed to work through a concept aloud. They always say that you're in the wrong room if you are the smartest in the room. I was definitely in the right room and am continuously inspired by you both.

Finally, this project would not have been done without my housemates, parents and friends, who had no idea about my project, but let me turn blue in the face talking about NMR and statistics. Thank you for sharing the cleaning load, dinners and wine nights that kept me sane enough to keep going. I owe you guys.

Table of Contents

Table of Tables

Chapter 1

Chapter 2

Chapter 3

Table of Figures

Chapter 1

Chapter 3

Abbreviations table

Abbreviation	Explanation
ARTG	Australian Register for therapeutic goods
CAM	Complementary and Alternative Medicine
COSY	^{1}H-^{1}H correlation spectroscopy
CV	Coefficient of Variation
GC	Gas Chromatography
HDS	Herbal Dietary Supplements
HILI	Herb induced liver injury
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High Performance Liquid Chromatography
HPTLC	High Performance Thin-Layer Chromatography
HQSC	Heteronuclear Single Quantum Coherence
HRMAS	High-resolution Magic Angle Spinning
LC	Liquid Chromatography
MS	Mass Spectrometry
NIPALS	Nonlinear iterative partial asymmetric least squares
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
q-NMR	quantitative Nuclear Magnetic Resonance
QTOF-MS	Quadruple-time-of-flight Mass Spectrometry
TCM	Traditional Chinese Medicine
TGA	Therapeutic Goods Administration
TGO 92	Therapeutic Goods Order 92
TLC	Thin-Layer Chromatography
UHPLC	Ultra High-performance Liquid Chromatography

Background: Immune supplements in Australia: Safety, efficacy, and quality control in a regulatory context.

Herbal supplements are defined by the World Health Organisation (WHO) to include herbs and herbal preparations as any product that may "contain, as active ingredients, parts of the plants, other plant material or combinations thereof" (1). They have been used worldwide for thousands of years, with the earliest record of herbal medicinal use being Chinese Emperor Shen Nong in 220 BC, known as the Shennong Materia medica)(2). Their ceremonial, social, and community uses are highly geography-specific, and modern regulatory frameworks mimic local area perspectives. Like allopathic medicines, the diversity in herbal supplement regulation prevents the formation of global bodies that could objectively examine the quality, safety, and efficacy of these products.

A 2019 statement from the WHO highlighted that although herbal medicine use is widely accepted throughout the world, with 88% of member states recognising the use of traditional herbal medicines, product quality assurance is often inadequate, posing a potential health risk to consumers (1). A 2017 study found that 43.2% of adults had used at least one dietary supplement in the last two weeks in Australia (3). Furthermore, the outbreak of the SARS-CoV-2 virus has dramatically increased the sales of supplements globally (4-6). An American case study indicated sales of elderberry and zinc products increased by 255% and 475% in the first week of March 2020 (7). The spotlight has now been put on individual health, particularly the immune system, and has urged consumers to include largely untested or quality assured products into their lifestyle. In this research, immune supplements, although often denoted as herbal and dietary supplements (HDS), are defined as any nutraceutical product advertised under generalised categories, including "immune-boosting" and "everyday health support" or other related immune and general well-being products. These

products are known to contain common immune stimulating herbs such as *Echinacea purpurea*, *Astralagus membranaceus, Andrographis paniculata,* and *Sambucus nigra,* but also contain organic compounds including vitamin C (ascorbic acid). While not officially classified by regulatory bodies, their increased usage has been noted since the beginning of the SARS-CoV-2 pandemic (4, 5). Thus, calls for a greater understanding of *immune boosters* have been presented in numerous literature reviews published in 2020-2021 (4, 6, 8-10).

Therefore, the scientific community's simultaneous call for greater regulation and pharmacovigilance of herbal supplements (11-15) and global regulatory bodies increased interest in controlling these highly unmanaged products (1) stimulates the need for a thorough examination into the safety and quality of immune supplements in Australia. The Therapeutic Goods Administration (TGA) is responsible for approving and registering therapeutic and cosmetic goods, including immune supplement products. The primary legislation that provides this power is the *Therapeutic Goods Act 1989.* This act establishes the regulation of all therapeutic goods in Australia, regardless of their manufacturing origins. Furthermore, it attempts to provide a unified approach to consumer health by states, territories and pharmacists (11) by signposting the management and overriding enforcement under one unified body, the TGA. The regulation of immune supplements falls under the umbrella term ‘complementary medicines’; these include herbs, nutritional supplements, minerals, vitamins, and homoeopathic remedies (12).

Product labels of approved substances are required to state the way it is regulated, either registered (AUST R) or listed (AUST L) on the Australian Register of Therapeutic Goods (ARTG). The ARTG is a public record of all therapeutic, complementary, and alternative medicines (CAM) legally sold on the Australian market. Complementary therapies are

primarily listed (L) on the ARTG because they are considered low-*risk* (unlikely to cause harm) to consumer health (13). In contrast, prescription medicines are registered (R) and more tightly regulated. The TGA determines the potential risk of a new product, considering factors such as the potential for toxicity or side effects and the likelihood of harmful events during long term-consumption. Unlike registered prescription and over-the-counter medicines, the TGA neither assesses the efficacy or quality of listed substances nor ensures their correct labelling before being sold on the Australian market (14).

Furthermore, the importation of herbal supplements into Australia does not require independent quality control (QC), and the TGA only performs QC after several significant adverse effects are reported (15). The TGA penalises advertising and compositional non-compliance in many forms, ranging from criminal convictions to warning letters. In 2021, two-sport supplement manufacturers were fined over A$12 million and 12 months in prison for undisclosed misconduct under the Act(16). Therefore, both scientific literature and reporting by the TGA indicate high levels of non-compliance that needs to be addressed through new pharmacovigilance strategies combining the technical knowledge of scientists with the regulatory capacity of the Government. This merger between research and regulation is observed in China, where the Chinese Pharmacopeia applies rigorous metabolomic processes, specifically Liquid Chromatography Mass Spectrometry (LC-MS), to assure the quality and consistency of herbal products (17).

Furthermore, regulatory bodies around the world do not have a consolidated definition of herbal products or a globalized approach to regulation, as each country has its legislation or guidance on herbal products. Much of the inherent ‘safety’ of natural products by regulators and consumers, which partly explain the immense industry growth, may be

false. For example, recent toxicity reports of herbal supplements, highlight adverse reactions ranging from mild to severe and sometimes lethal (18). The organs most affected are the liver, kidney, and brain (19); however, HDS adverse reactions are inadequately reported, and a thorough understanding of their toxicity is limited (20). The most comprehensive HDS toxicity study in 2020 analysed 413 patients with acute/chronic liver pain. They found that over 90% incurred herb-induced liver injury (HILI) and 79 patients required a liver transplant (21).

In most cases, the biochemical pathways by which herbs may induce toxicity have not been fully elucidated, with the most likely mechanisms being oxidative and mitochondrial stress and dysfunction (22). A more comprehensive understanding of the mechanisms of toxicity is required, especially in the context of dosage ambiguity (23, 24). Notably, in Australia, listed medicines do not need to comply with the TGA's dosage uniformity agreements. Additionally, contaminants such as microbes and heavy metals have been identified to cause organ damage. A recent investigation using the standards set by the Chinese Pharmacopeia found extracts of *Andrographis paniculata* and *Curcuma longa* to be above the Provisional Tolerable Daily Intake levels for arsenic and lead (25). An examination of 6,712 adult women in the United States found women who used various supplements had lead levels 20% higher than those who used no herbal supplements (26).

Furthermore, current research using DNA barcoding for pharmacovigilance surveillance found 27% of herbal products from 37 countries had some form of adulteration (27). Other studies have also highlighted significant degrees of adulteration from fillers, contaminants, and blatant product substitution (28). These adulterants may give the appearance of clinical effectiveness, increase the weight or quantity of the supplement, or dye the product to align with consumer beliefs about the nature of the product. This is

observed in the case of *Curcuma longa* (turmeric) where yellow dye is often added to deceive consumers about the turmeric content of the supplement (29).

Therefore, the widespread belief that herbal supplements are inherently low risk contradicts the increased reporting of the 'dark side' of dietary supplements. The present analysis calls for methodological and analytical mechanisms' that can assure the compositional integrity of herbal supplements—at the same time, creating new avenues for regulatory relationships that blend modern technological advancements with Government enforcement to protect Australian consumers.

Research Aim & Objectives

In the context of the increased prevalence of immune supplement usage, this analysis aims to determine the quality and safety of immune supplements available on the Australian market. This is done by providing insight into better methods of quality control and supplement safety by aiming to:

1. Determine if product labels comply with the respective regulatory guidelines released by the Therapeutic Goods Administration. This aim will compare products to determine if labelling compliance has been upheld and is a valid form of quality assurance.
2. Perform compositional analysis with 1 and 2D NMR experiments to test the application of high throughput NMR for the investigation of the chemical profile of supplements, including active ingredients, excipients, and potential contaminants.
3. Characterise batch-to-batch variation of supplements using NMR signal quantifications of active ingredients and principal components analysis as a multivariate statistical analysis method operating on full resolution spectra. Three

different product batches of seven immune supplements will be compared to examine batch variations in chemical profiles.

Significance and rationale

The advent of the COVID-19 pandemic has seen natural supplements increasingly marketed with protective properties against SARS-CoV-2 infections. However, the scientific evidence for these health claims is inadequate with critical issues in study design relating to small sample size, the omission of intention-to-treat analysis, lack of objective outcomes, such as, hypothesis-free clinical trials, and lack of adverse effect reporting (14, 30). Furthermore, there is conflicting data about the biochemical pathways and mechanisms of action by which these dietary supplements act, and studies that verify or validate efficacy are missing. Besides the lack of sufficient scientific evidence of efficacies, a body of literature reports product contamination and adulteration as explained in the introduction (25, 27, 28, 31-33).

At present, the composition of immune supplements available on the Australian market is provided by manufacturers on the product label, mainly reporting herbs and micronutrients and not listing all not all the ingredients or their concentration units.

In Australia, there has been no published assurance of the batch quality consistency of immune supplements. Many variables contribute to batch-to-batch discrepancies, including growth conditions, plant species and varieties, harvest timing and geographical origins of each ingredient, and manufacturing process. For example, one study found 10% of *Echinacea pallida* sample preparations contained no measurable active ingredient when DNA barcoding was performed (34). Reliable data on immune supplements' composition and phytochemical variation is lacking, with insufficient information regarding product composition being provided to the consumers.

Outline of thesis

Throughout recent years there has been an increased investment in research regarding the safety of all CAM products. In 2017, Byard et al., posed the question, “What are the risks to the Australian community from herbal medicines?” (35). This thesis represents part of a more extensive study into CAM products on the Australian market following more recent findings using DNA barcoding (27) and mass spectrometry (36). Therefore, it is only appropriate to explore the current regulatory framework for labelling and advertising of immune supplements in Australia and the potential application of high-throughput nuclear magnetic resonance (NMR) spectroscopy to examine the compositional integrity of supplements. The thesis is structured into two parts:

Part A: Compliance of Australian immune supplements/CAM products with Australian regulations for product labelling. This section centralises the current regulatory landscape for labelling compliance, comparing current guidelines and requirements with actual labelling practices of CAM manufactures that release products on the Australian market. Furthermore, this thesis section explored how labelling non-compliance might cause and its potential to cause harm to consumers.

Part B: Examination of immune supplements' compositional integrity, including the presence of active ingredients and quantification of production batch variability of 7 selected immune health supplements. This section examines the compositional integrity and batch quality consistency of products based on proton (^{1}H) NMR by comparing product spectra with reference standards and database entries of active ingredients proposed on product labels. Three different batches of products were used to examine compositional consistency and to detect potential manufacturing discrepancies across different manufactured batches.

Chapter 1: Metabolomics for botanical fingerprinting of herbal supplements: A review of immune specific herbs.

Introduction

A growing trend throughout medical research centralises the need for quality control of unregulated products, regardless of debate about efficacy (37-40). A 2019 study found that paediatricians in Spain prescribed immune supplements alongside allopathic medicines in 89% cases (41). Additionally, the SARS-CoV-2 pandemic has shifted focus towards HDS designed to benefit the immune system. Khabour and Hassanein highlighted that increased consumption of *immune-boosting* product*s* directly correlated with participants fear of the virus (4) despite reports of inadequate support for their purported health benefits and lack of rigorous quality assurance. These findings illustrate the requirement to ensure compositional integrity in an effective quality assurance framework to safeguard consumer health (9).

Traditional authentication and identification of herbal medicines using Pharmacopeia (documents pretaining to herbs and their uses) which primarily relied on macroscopic and microscopic plant descriptors for taxonomical identification (28). This required expert knowledge and relied on morphological features, typically examining whole-plant appearance and leaf features characteristic for a certain herb. However, HDS are generally formulated as multi-component products that have no defined plant species characteristics due to modern manufacturing. Now product authentication is less reliant on plant species identification but focuses on chemical composition analysis.

Recent technological advancements allow a more untargeted compositional analysis for the identification of plant cellular components with method from the field of systems biology and functional genomics providing new avenues for genetic and chemical analysis.

For example, genetic fingerprinting is nowadays widely adopted for food authentication, population biodiversity, and monitoring herbal exports. Parveen et al. (2013) highlighted that the use of polymerase chain reaction (PCR) for DNA-based identifcation of species is generally excellent. However this technology has limitations relating to DNA quality, primer choice and amplicon length, and an inability to generate information related to environmental factors such as climate and soil (42). In contrast to DNA barcoding, chemical compositional analysis is less capable of plant species identification but can characterise environmental patterns, trends and potential outliers relating to natural and unnatural variability from paddock to pill. Furthermore, immune supplement active components are often polyphenolic compounds that cannot be measured with DNA-based methods but are easily measured with spectroscopic methods (43).

Untargeted chemical analysis methods such as liquid or gas chromatography-mass spectrometry (LC-MS and GC-MS, respectively) and nuclear magnetic resonance (NMR) spectroscopy have provided the most significant advancements in assessing herbal composition. These technologies often combine separation with spectroscopic techniques to efficiently derive structural and quantitiative information from complex mixtures, for example using calibration and reference methods. These methods are now known as dereplication strategies(44). Between 2007 and 2015, citations containing ‘dereplication’ or ‘de-replication’ more than doubled (45). The most effective techniques were dereplication strategies that incorporated various complementary and supplementary technologies for small molecule (2 kDa) detection in natural products.

There are discrepancies found in the literature used to describe small molecule research, referring to phonetically similar terms: *Metabolomic*s *or metabonomic*s. Some root

metabonomics with genomics, particularly in toxicology works (46). Others believe it is a technical discrepancy between NMR and MS, although this is not consistent throughout the literature. Thus, for the sake of this review, the term *metabolomics* will be used to describe a systems-based approach to small molecule identification (<2kDa). Metabolomics has proven successful in quality control of natural products (50) and identifcation of characteristic compounds for herbal identification (33) using both targeted and untargeted analysis strategies. In this context, the terms untargeted and targeted are used to describe the measurement of compounds, with the former aiming to measure as many compounds as possible, while the latter aims to absolutely quantify specific compounds or class of compounds (42, 45, 47).

Most commonly, in natural product identifcation both strategies, untargeted and targeted analysis of samples are eqaully applied in the metabolomics for quality control (48). Untargeted methods provide a means of metabolomic discovery that can be used to compare control and test groups to identify differences in the composition and concentration of chemicals. Targeted methods quantify the absolute concentration value of known compounds, or from a specific group of compounds. Selecting the most appropriate analysis technique often relies on technical considerations, such as the intention to identify unexpected compounds or absolute quantification of known compounds.

As outlined above, targeted techniques can pinpoint compounds unique to each species, for instance to quantify a specific molecule in a given product. However, the targeted chemical examination is currently limited by the lack of commercially available standards, due to the extensive number of plant compounds of which not all have been isolated and made available in abundance (49). For example, biomarkers for *Curcuma Longa* (Turmeric) are

relatively well understood. However, their interaction and other supplementary compounds are unknown (50). Therefore, current trends in the identification and quantification of complex mixtures primarily rely on hypothesis-free untargeted metabolomics to provide a more detailed overview of the entire chemical profile of herbs.

Metabolomics for botanical profiling

The rise in -omics science has relied upon technologies from other branches of science to birth new disciplines, like metabolomics. Omics sciences are a variety of disciplines found in systems medicine and are generally structured according to the biological flow of information from the gene (genomics) to RNA (transcriptomics) to protein (proteomics). In the quest for small molecule (<2kDa) identification and quantification, two main analytical techniques are used to determine a sample's chemical composition: Mass spectrometry (MS), used either in direct injection mode or in combination with pre-separation techniques, such as ultra-performance liquid chromatography (UPLC), and nuclear magnetic resonance spectroscopy (NMR). While each has its own costs and benefits, no single analytical workflow can characterise all compounds in complex mixtures, and both technologies are often considered to provide complementary information.

Furthermore, metabolic discovery is a relatively new advancement of systems biology resulting in major terminology ambiguities in the literature when describing the chemical composition of samples, including chemical profiling, structural elucidation, and fingerprinting. While the use of these techniques is well known, their application is often confusing and unclear. Fingerprinting is a semi-quantitative analysis applied to quality control in botanicals, and is based on the overall evaluation of spectral or chromatographic data using reference compounds and chemical markers (17). This differs from characteristic profiles that

provide a less comprehensive composition overview, comparing reference peaks with those found in the mixture. Instead of using many compounds, it relies on a single chemical marker, one compound from one herb (51).

Structural elucidation refers to the interpretation of the spectroscopic data, whilst chemical profiling seeks to understand the definition and distribution of botanical compounds within a specific herb or supplement. In this thesis, the term metabolic profiling will refer to the systematic detection and qualitative characteristics of compounds in HDS. An essential tool used in metabolic profiling are spectral databases typically accessible in the public domain with examples given below. Additionally, comparative analysis of spectra is done by utilising interdisciplinary techniques of multivariate statistics, data mining using software and scripting languages such as TopSpin©, or open source libraries in R, or Python, respectively (like those used in dereplication strategies) and graph analytical approaches, for example biochemical reaction network modelling (27, 33, 52). Continuous increases in publications regarding metabolomics or dereplication strategies in herbal mixture studies (Figure 1.1) highlight the progressive use of these techniques in herb and plant species identification and differentiation. As can be seen in Figure 1.1, there is a clear increasing time trend for publications containing key words ‘metabolomics’ and ‘dereplication’ strategies in natural products that undertake characteristic profiling and fingerprinting (45).

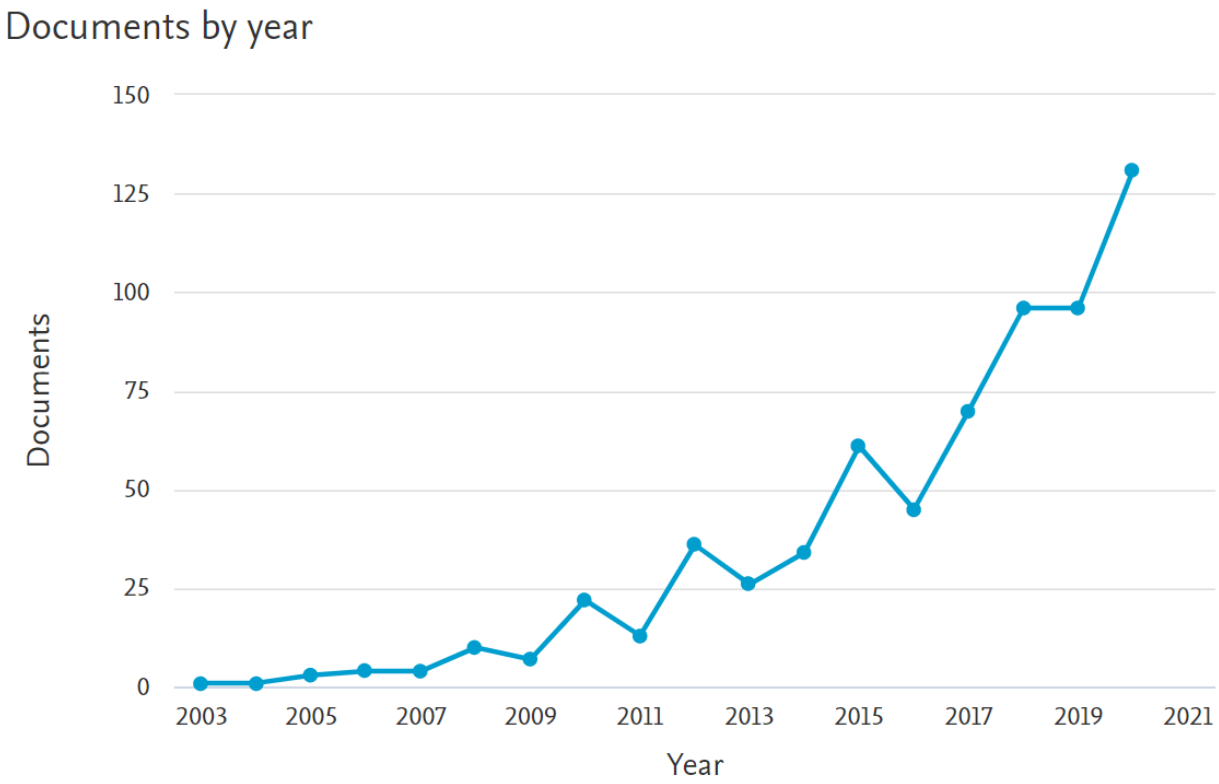


Figure 1. 1 Increasing trend of "metabolomics" or "dereplication" in "herbal mixtures" or "nutraceuticals" or "dietary supplements" publications obtained from Scopus (https://www.scopus.com). At time of submission, the document count for 2021 was 81.

This is due to the concept of phytoequivalence, which posits that the chemical profile of any herb or botanical can be compared with a certified reference sample (53), to assure that particular herb is contained in the product. The use of biomarkers for herbal identification is well known and much of the literature surrounding characteristic profiling relies on the development of consistent biomarkers. For example, A*ndrographis paniculata* contains andrographolide as a major chemical constituent, and it has been shown that this can be used as both a quantitative and qualitative biomarker for quality control of *A. paniculate (54)*.

For metabolite identification it is common practice to compare signals derived from chemical reference standards with signals found in product spectra. Alternatively, where no

reference standards are available, signals are compared with spectral database entries. However, the current incompleteness of databases, and lack of standard experimental procedures limit the comparability and applicability of database comparison for identification (55).

It is clear from the literature that authentication of herbal products is not synonymous with the complete identification of herb/mixture chemical profile. Nonetheless, a comprehensive review by Heyman and Meyer (2012) articulated the success of chemometrics to observe molecular relationships in different dimensions, and presenting these in an easily understood graphical manner (40). The true identity of a product for quality control does not necessarily rely on every phytochemical constituent but certain biomarkers that can reliably indicate the qualitative and semi-quantitative presence of the herb (56-60). Although more holistic development of biomarkers for herbs/herbal mixtures that generally have many active compounds working together (40) could provide a greater understanding of the synergistic effects of phytochemicals. In a typical setting a combination of spectroscopic methods, separation science and database searching are applied to identify extracts in high-throughput commercial or regulatory setting for CAM.

Application of ^{1}H NMR

Recent years have seen a movement towards the use of proton (^{1}H) NMR for natural product quality control. ^{1}H NMR provides information about the structure of a chemical profile by observing magnetic relaxation properties, chemical shift and intensity to reflect the exact environment of protons. Recent work by Kim et al. highlighted that separation techniques (eg. HPLC, UPLC, GC) can be used in complement to NMR; however, quantitiative NMR alone can be used for targeted chemical fingerprinting and untargeted analysis of

impurities, particularly in complex mixtures (61). While NMR is commonly used complementary to Mass Spectometry (MS) techniques, it carries certain benefits. Firstly, the natural abundance of hydrogen atoms within organic molecules enables lower concentrations to be recognised via ^{1}H NMR. Secondly, there is a large range of detection schemes possible, providing many ways of answering metabolic questions (40). Finally, NMR is non-destructive and enables samples to be used for multiple experiments and analyses (62).

However, the use of NMR is not trivial, as sample processing and set-up have a significant influence on the quality of data produced. Some parameters that must be considered are sample salt type and concentration, dissolved oxygen content, choice of solvent, sample pH, temperature and temperature gradient, spectrometer field homogeneity and primary magnetic field strength (63). Furthermore, experimental reproducibility and replication rely on decisions about appropriate data acquisition methods such as pulse sequence, the length of the 90º pulse, quality of shimming and tuning, number of acquisitions and data points. Nonetheless, ^{1}H NMR remains a convenient way of examining organic molecules such as phenolic compounds, alkamides, flavonoids and saponins which are thought to give therapeutic benefits in immune supplements(61).

Data acquisition using NMR is well understood as spectra are produced according to the contributing hydrogens contained within a product. These observable signals are produced via the Fourier transform free induction decay (FID). The FID is a time domain signal that is Fourier transformed to obtain readable spectra in the frequency domain. The intensity of signal is directly proportional to the contributing number of protons residing in an electrochemical environment. This is problematic in complex or multi-component mixtures due to homonuclear (^{1}H-^{1}H) coupling interactions and closely related chemical shifts that lead

to overlapping signals that lead to non-lorentzian peak shapes, and the potential convolution of minor signals (60).

Two discrete post-processing approaches of NMR metabolomics have been established: Peak deconvolution and statistical approaches operating on the full resolution spectrum. Both are viable methods to compare ^{1}H NMR spectra that identify compounds of interest. Depending on the complexity and signal presentation of compounds, different methods are more applicable than others. For instance, quantitative comparison of individual compounds requires integration of free-standing signals, while characteristic chemical profiles are typically generated using multivariate statistical analysis(64). Principal Component Analysis (PCA) is an unsupervised multivariate statistical method that reduces the data dimension by forming principal components that summarises original variables as a composite index(59).

On the other hand, deconvolution methods seek to integrate the signals while accommodating for overlap and unresolved peaks commonly found in proton NMR (59). Neither of these methods is faultless and recent approaches focus on process automation using machine learning methods for the compositional analysis of complex mixtures (65-67). Accordingly, the success of dereplication strategies in metabolomics relies on comprehensive databases that enable cross-institutional validation of chemical structures. The Human Metabolome Database (HMDB) and Natural Products Magnetic Resonance Database (NP-MRD) are currently the most comprehensive for natural product identification. However these are still limited in their experimental specificity and application to complex mixtures in context of plant research. Other databases that focus on chemical structure and mechanistic information (e.g., based on biochemical pathways) include Chemical Entities of Biological Interest (ChEBI)(68) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (69). Therefore,

scientific community efforts in building comprehensive natural product databases for ^{1}H NMR spectroscopy will enable higher confidence in assigning compositional features for natural products.

Table 1. 1 . Review of current methodologies used to characterise common herbal and inorganic ingredients found in Australian immune supplements.

***All abbreviations are explained in the abbreviations table above.**

Herb	Purported therapeutic effects	Known biochemical interactions	Biomarker	Kegg ID	Method of biomarker identification*	Reference
***Echinacea purpurea* (Purple Coneflower)**	– Strengthening immunity – anti-inflammatory	– Stimulation of interleukin-10 (IL-10), IL-1, and IL-6 – Inhibition of tumour necrosis factor-alpha(TNF-a) – Decreased pro-inflammatory cytokines(70, 71)	**Chichoric acid**	**C10437**	HPLC with TLC scanner system	Zolgharnein et al., 2020(72)
					^{1}H-NMR	Frederich et al., 2009(34)
					Various non-targeted MS-hyphenation strategies	Waidyantha(66) et al., 2020
***Astralagus membranaceus* (Mongolian milkvetch)**	– Increasing athletic performance – improvement of exercise-related fatigue – supporting weight loss	– Prevents production of ROS. – Caps endogenous Cu ions to form CA-Cu complex – Reduces phosphorylation of c-Jun N-terminal kinases (JNK-1) (73)	**Caffeic acid**	C01197	UPLC-MS	Li et al., 2019(74)
	– Immune mechanism stimulation – Antioxidant properties – Aids in hepatoprotection – Weight loss	– Inhibits NF-kB activation in times of inflammation – Inhibits acetylcholinesterase.	**Chlorogenic acid**	C00852	^{1}H NMR	Marchetti et al., 2021(59)

		– Activates GABA receptor by binding to benzodiazepine site. (75)				
Andrographis paniculata	– Anti-pyretic – Immunostimulant – Anti-protozoan – Antioxidant – Anti-tumorigenesis.	– Inhibits NF-kB, TNF-a, IL-6, IL-12 activation in times of inflammation (54)	**Andrographolide**	C20214	^{1}H-NMR	Srirameneni et al., 2019(76)
					^{1}H-NMR	Tajidin et al., 2019(77)
Sambucus nigra (Elderflower)	– Immune mechanism stimulation – Antioxidant – Liver protection – Anti-inflammation.	– Decreases cytokine-induced inflammation. –antiproliferative effects (30, 78)	**cyanidin-3-glucoside**	C08604	NIR spectroscopy	Stuppner et al., 2020(79)
					UHPLC-MS/MS	Appenteng et al., 2021(80)
					GC-MS	Salvador et al., 2016(81)
Allium sativum **(Garlic)**	– Anti-viral – blood pressure regulation – lowers cholesterol – Anti-tumour – Immunomodulatory	– Decreases lipopolysaccharide inflammation – Increases mast cell production during infection. – Decreases pro-inflammatory cytokines by inhibiting ERK1/2 signalling phosphorylation(82, 83)	**Alliin (S-allyl-L-cysteine)**	C08265	LC-MS/MS	Molina-Calle et al., 2017(84)
					HRMAS-NMR	Riota et al.,(85)
					HPLC-MS/MS	Liu et al., 2015(86)
Olea europaea **(Olive)**	– Antioxidant. – Antimicrobial activity – Anti-inflammatory	– Decrease in copper-sulphate induced oxidation	**Oleuropein**	C09794	UHPLC-HRMS	Lemonakis et al.,2013 (56)

		– Increases protection of nitrous oxide in macrophages(87)				
Curcuma longa **(Turmeric)**	– Relieves abdominal pain – Anti-inflammatory – Anti-arthritic – Anti-anxiety	– Reductions in CDAI & inflammatory indices – Increased vitamin C and E levels – Reduces glutathione S-transferase activity, PGE2 production and TNG-a. – Induction of p53 expression(4, 29)	**Curcumin**	C10443	DNA barcoding	Parvathy et al., 2015(32)
					^{1}H-NMR and HPTLC	Chatzinasi et al., 2019(29)
					qHNMR	Kim et al., 2021 (61)
Ascorbic acid **(Vitamin C)**	– Treatment of scurvy – Aid in recovery from RTI. –Anti-inflammatory	– Activation Nuclear factor-kB (NF-kB). – Increased expression of catecholamines, vasopressin & collagen neurotransmitters (88, 89)	**Ascorbate**	C00072	^{1}H-NMR	Reid et al., 1989(90)
					UHPLC-MS	Baenas et al., 2019(91)

Botanical fingerprinting of immune supplements

Most herbal supplements contain a combination of active and excipient ingredients. Active ingredients are responsible for the claimed remedial purposes. Immune supplements in Australia commonly include traditionally therapeutic herbs such as *Echinacea purpurea*, *Astragalus membranaceus, Sambucus nigra, Olea europea, Allium sativum, Andrographis paniculata* and *Curcuma longa* and organic compounds including vitamin C (ascorbic acid) and zinc amino acid chelates. A single set of compounds often characterises the herbs purported therapeutic properties, which generally includes antioxidant behaviour via polyphenols (Table 1)(50). Polyphenols are a large group of organic molecules derived from the phenylpropanoid and polyketide pathway in plants, containing more than one phenolic unit. There are immense structural differences both among polyphenols and throughout plant species. This complexity forces the use of advanced technologies such as MS and NMR for qualitative and quantitative assessment of herbal materials(92).

In the case of immune supplements, the proposed activity of polyphenols is attributed to the "biochemical scavenger theory". Plant polyphenols are called 'scavengers' of different oxygen species, binding their aromatic rings to free radicals such as HO^- (93). These free radicals are involved in human pathological events via hydrogen and lipid peroxidation(94, 95). For example, SARS-CoV-2 infection increases pro-inflammatory cytokine storms resulting in hyperinflammation, which could cause the progressive lymphopenia observed in seriously ill COVID-19 patients(96). The correlation between the polyphenolic content of common immune stimulating herbs and cytokine-inducing SARS-CoV-2 could be the reason for the dramatic increases in the use of immune supplements. There has also been a change in the behaviour of supplement users to forefront immune function and overall wellness(8): Despite

the lack of significant results from trials using immune supplements in the treatment of COVID-19(87). A recent review of randomized control trials published on COVID-19 found only twelve studies that could be included, all of which had low methodological quality and high risk of performance and selection bias, using the Cochrane risk assessment(10). Therefore, this review does not centralise the efficacy of these products but states some background into why they are chosen to be used in commercial supplements. Nonetheless, priority is given to the elucidation of the polyphenolic content of these supplements using metabolomic techniques, particularly NMR, in the belief this will enable greater control of the compositional integrity of these products in a time of increased community prevalence and absence of accurate efficacy reporting.

Specific components of immune supplements

Vitamin C is the generic name of *L*-ascorbic acid; this active isomer protects proteins, lipids, and nucleotides from oxidative damage by acting as an antioxidant(84) and is a prooxidant as various signal cascades result in the production and depletion of oxidative species and inflammatory mediators(97). Numerous mechanisms are proposed to explain the therapeutic potential of vitamin C supplementation in respiratory illness, including improving immune cells phagocytic properties, antioxidant and cytotoxic effects in inflammation, and antihistamine response(97-102). This function highlights its possible role in combatting viral infections, with some research indicating a direct inhibitory effect of ascorbic acid on pathogens(97). However, vitamin C has exhibited no real curative or preventative properties in benefitting recovery or decreasing hospitalisations in influenza-like illnesses (100).

The use of spectroscopy and spectrometry techniques in vitamin C extraction from complex mixtures does not generate the same quality of data as other work regarding its

efficacy and healing properties (97, 100-103). Many MS methods have been applied to separating and detecting water-soluble vitamins, including ascorbate, in body fluids, including blood, plasma, cerebrospinal fluid, and leukocytes (98, 104). More recently, separation techniques such as UHPLC-MS have been applied to solid and liquid vegetables, indicating robust validation methods for vitamin C/ascorbate (the base form) in food products (91). However, high-throughput methods, such as ^{1}H-NMR, have not been applied to identify ascorbate in complex supplement mixtures. Therefore, all studies suffer from a focus on the use of ascorbate in health and well-being without registering its safety or presence in common immune supplements.

Echinacea purpurea (purple *Echinacea*) is endemic to North America, where Native Americans established its therapeutic application in fighting upper respiratory tract infections. It is one of the most widely used dietary supplements in America and Europe, with over 800 *Echinacea*-containing products available on the market (71). More than 600 papers have examined the immunopharmacology, biochemistry, and clinical application of *E. purpurea*. However, there is little consensus on the specific mode of action by which *E. purpurea* may provide protective coverage over the host when infected with a pathogen(70). All *Echinacea* species contain phenylpropanoid derivatives of caffeic acid, primarily chicoric acid. Caffeic acid is the base hydroxycinnamic acid that is commonly found amongst various plant speceis(73). A 2009 study used ^{1}H-NMR to differentiate between medicinal strains of *Echinacea* and, while the metabolic profile was very similar, species could be discriminated using the aromatic rings of chicoric acid and echinoids(34). This was similar to the findings of Waidyanatha et al. (2020), who performed untargeted MS analysis on 24 different *Echinacea* products (finished and unfinished). Of the finished products, chicoric acid was the primary constituent of *E. purpurea* products, corroborated by Zolgharnein and colleagues (72). Thus,

chicoric acid is an essential biomarker for quality control of *E. purpurea* due to its degradation and dependency on harvesting and manufacturing processes (72). While characteristic phytoconstituents are well understood, the strategies attempted are time-consuming and labour intensive. ^{1}H NMR has not currently been applied and more complex mixtures containing *Echinacea* have not been examined using metabolomic techniques.

Astragalus membranaceus (Qang Hi) is an adaptogenic herb with significant cultural importance in traditional Chinese medicine (105). The traditional primary use of *A. membranaceus* in the treatment of immunodeficiency disorders and a wide variety of infections that cause "insufficient *qi*" (life energy), but there is minimal research into its therapeutic effects in a clinical setting (106). The primary therapeutic effect of *A.Membranaceus* in the treatment of respiratory illness is due to its stimulation of cell-mediated immune mechanisms and antioxidant behaviour that aids in hepatoprotection; chlorogenic acid polyphenols are the reported facilitators of these mechanisms(70). ^{1}H NMR has been used to elucidate chlorogenic acid from herbal preparations, and predictive models were built to account for the complexity of 1D-NOESY using a PLS regression model(59). However, this has not been done in the case of *A. membranaceus*, even though UPLC-MS methods identified chlorogenic acid in run times as short as 2.17 min(74). Therefore, further research is needed to confirm the potential of chlorogenic acid as a biomarker for *A. membranaceus* in mixtures, particularly when using ^{1}H NMR.

Also referred to as the "king of bitters", *Andrographis panicualta* has been used throughout Asia and Europe to stimulate the immune system against infections. Andrographolide is the principal constituent, and Dai and colleagues presented a good overview of its pharmacological activities in 2018(54). Interestingly in 2011, the TGA

requested adverse event reporting from sponsors using *Andrographis* in their products. They found fifty previously unreported adverse events of anaphylaxis and allergic reactions. In 2012, a further 72 adverse events included hypersensitivity, swelling, throat tightness, wheezing, palpitations, and pruritic rash(107). Thus, greater warning conditions have been placed on *A. paniculata* in Australia. However, it remains on the permissible ingredients list. Therefore, it is evident that the determination of biomarkers for the identification and quantification of *A. paniculata* is needed in the interest of consumer health. A 2019 study successfully used ^{1}H NMR to identify andrographolide in *A. paniculata* by comparing chemical shifts in the presence and absence of chloroform extracts(76). The elucidation of andrographolide in complex mixtures has not yet been attempted; although, current literature suggests its potential as a meaningful biomarker for quality control purposes(108, 109).

In the case of *Sambucus nigra* (elderflower), cyanidin-3-glucoside and -sambubioside are the primary anthocyanins (polyphenols) found in all parts of the plant material used in supplements. Currently, no study has applied NMR spectroscopy to confirm the presence of anthocyanins in *S. nigra*; however, previous works using near-infrared (NIR) spectroscopy and LC-MS suggest it could be used for chemical referencing in optimised experimental conditions(79). Anthocyanidins have been elucidated using 1D ^{1}H NMR. However, Sriramaneni and colleagues found heteronuclear multiple bond correlation (HMBC) and heteronuclear multiple quantum coherence (HQMC) experiments essential to understanding the long range coupling between ^{1}H and ^{13}C for ultimate anthocyanidin count(92). Another study isolated similar cyanidin flavonoids from different species of red chicory using both 1- and 2D NMR techniques(51). Furthermore, a 2019 review indicated that cyanidin-3-glucoside in plasma and urine of healthy humans is a potential biomarker of anthocyanin-rich berry

intake (110). These anthocyanins are believed to promote various health effects, ranging from antiproliferation of tumours (78) to viral mitigation (111). Therefore, as elderflower continues to be applied to herbal remedies, it is essential to find a high-throughput means of identifying them in complex mixtures, outside of the scientific community.

Olea europaea, more commonly known as olive, was traditionally distributed around the Mediterranean, where decoctions of dried fruit and leaves were used to relieve respiratory infections. Olive phenolics such as oleuropein have been found to exhibit antioxidant and antimicrobial activity. In 2020, an exploratory study used both targeted and untargeted analysis of HRT-MS spectra of 13 olive-based dietary supplements to suggest the presence of 378 compounds. They quantified 26 of these, including oleuropein, to find that only six reached the recommended therapeutic hydroxytyrosol levels when using oleuropein complexes(37). Much of the metabolomic literature focuses on olive oil mixtures, thus indicating that oleuropein can be used as a biomarker for *O.europaea*, particularly if olive leaves are used for preparation, where oleuropein concentration is highest (112).

Recent investigations have specified the mechanisms of action that have given turmeric *(Curcuma longa)* its place in traditional healing for thousands of years. Turmeric is found in many supplements that claim to have benefits on the immune system and everyday well-being. The demand for turmeric has increased the prevalence of adulterated turmeric as fillers and substitutes decrease production costs and give the impression of high-quality turmeric based on bright yellow colours(29, 32, 113). Parvathy et al. used DNA barcoding to highlight 30% of turmeric powders tested were adulterated with cassava (32). In addition, metabolomic studies of 56 commercial products used NMR to infer the presence of *C. longa.* The researchers concluded that the extent of compositional change exhibited was likely

caused by manufacturing differences and lack of clarity surrounding the constituents, which give the purported therapeutic effect(29). These studies into the quality of turmeric are important as curcumin (the main active ingredient) is considered to be the active anti-inflammatory agent (6), and studies in healthy people have highlighted its potential in muscle recovery(114) and anti-anxiety properties(115). However, it is difficult to draw conclusive results from these studies due to the wide range of dosages and definitions of healthy people(116). Therefore, in the case of *C.Longa*, more is understood about the quality diversity than efficacy of active constituents.

Garlic (*Allium sativum*) is a common functional food that is known to contain compounds linked to immunity. An animal study on supplementation with alliin (S-allyl-L-cysteine sulfoxides), the primary purported therapeutic compound for garlic, improved glucose tolerance and decreased inflammation via IL-6 and stabilization of MCP-1 and TNF-alpha serum levels(83). It was previously believed that alliin reverted to allicin and that allicin gave the characteristic benefits of garlic. In contrast, Amagase (2006) insisted that allicin is not necessary for the biological activities of garlic to take place (117). Nonetheless, there is currently no metabolomic approach to the botanical fingerprinting of garlic. Dereplication techniques have been used in studies characterising the traceability of Italian garlic varieties (85) and differentiating between fresh and black garlic(84). Although, this has not been applied to complex mixtures. This is problematic as no coherance on appropriate biomarkers for identification is cited in the literature.

Meanwhile, there are evident discrepancies between biochemical knowledge about the therapeutic constituents of garlic and the use of garlic as a healing plant, and there are currently no studies into the compositional integrity of garlic supplementation. However, a

recent randomized, double-blind, placebo-controlled study found significant decreases in inflammation in an intervention group of 35 women with rheumatoid arthritis(82). Nonetheless, the intervention group received 1000 mg/day for eight weeks. This dose is much higher than those found in commercial supplements and highlights the challenges when efficacy and application to everyday life are not appropriately assessed.

Conclusion

At present, the exact composition of immune supplements available on the Australian market relies heavily on manufacturers disclosing ingredients on the product label. The relative paucity of metabolomics literature encompassing immune supplements or herbal mixtures in general means that the above herbs have been reviewed separately and not as the complex mixtures they present in commercially. This enabled better identifcation of biomarkers that could be used as chemical reference standards to aid in identification. Previous publications that used dereplication strategies forefronted the phytochemical elucidation of raw plant materials and provided a foundation for high throughput identification of chemical markers that indicate the presence of a herb (28, 45, 72, 109, 110, 112). Therefore, future research needs to focus on applying these methods to complex herbal products, such as immune health supplements.

Given the widespread awareness of SARS_CoV_2 and the subsequent demand for symptomatic relief and immune-boosting regimes, the quality of these products must be more rigorously assessed and understood before their sale on the Australian market. Methods for such analysis should be optimised for complex mixtures of herbal and inorganic material. In consequence, the rise in metabolomic analysis presents a new methodology for examining complex mixtures for herbal authentication. In particular, targeted analysis via ^{1}H

NMR can be applied for a high throughput process that serves both the community and regulatory bodies.

Attribution Statement- Chapter 2: The prevalence of labelling non-compliance in Australian immune supplements.

STATEMENT FOR MANUSCRIPT – Should precede all chapters written in manuscript form.

The current manuscript is prepared for submission.

The following authors contributed to this manuscript as outlined below.

Authorship order	Contribution (%)	Concept Development	Data Collection	Data Analyses	Drafting of manuscript
Courtney Hall	75%	X	X	x	X
Garth Maker	20%	X	-	X	X
Emily Davies	5%	X	-	X	-

Contribution indicates the total involvement the author has had in this project. Placing an 'X' in the remaining boxes indicates what aspect(s) of the project each author engaged in.

By signing this document, the Candidate and Principal Supervisor acknowledge that the above information is accurate and has been agreed to by all other authors.

Courtney Hall **Garth Maker**

Candidate **Principal Supervisor**

Chapter 2: The prevalence of labelling non-compliance in Australian immune supplements.

Chapter Link

The composition of Immune supplements relies on the primary label to convey messages about active ingredients and provide insight into their therapeutic value. If this information is not factually correct, consumers may be misled into buying products without full knowledge of their safety. Therefore, a review of the current compliance with labelling standards is an integral part of examining the quality of products sold in pharmacies across Australia.

Introduction

The label of a commercial supplement is the primary source of information for consumers(118). Approximately 95% of people acknowledge reading the label of a health food product, even if they do not have the technical skills to critique its nutritional value (119). However, little attention is placed on the reliability of labels to convey consumer health messages nor certify the ingredients contained within. Another study further enforced this, as 68% of participants decision-making in supermarkets was influenced by the label information. Advertising techniques used in labelling are often not overt, and their subtle nature can significantly affect consumer perspectives. Increasing research into the psychology of labels in the natural health space indicates that labels play a crucial role in providing health information at the scale of population-level intervention (120, 121).

In Australia, advertising and labelling compliance is primarily enforced via the Therapeutic Goods Order 92 (TGO92), which dictates the standard for non-prescription

medicine labels, including advertising claims and permitted indications. As of March 2021, the four-year transition period for sponsors of CAM and over the counter medicines ended, and TGO92 came into full effect(122). The TGA allows only classified low-risk ingredients and low-level claims such as “relieves throat irritation’ for non-prescription medicine (10). The dominant product evaluation technique is a post-market analysis, using random sampling or targeting flagged substances (14). From 2020 to 2021, numerous warning letters were sent to supplement advertisers for false claims about the prevention and treatment of viruses, specifically COVID-19, for products listed on the Australian Register of Therapeutic Goods (ARTG)(123). Furthermore, the *Therapeutic Goods Act 1989* defines mandatory monitoring responsibilities for sponsors. Current complaint management relies on six key compliance aspects: advertising, labelling, the information provided on the ARTG entry, manufacturing, quality and formulation, safety, and evidence (124). These responsibilities are divided across numerous legislative and regulatory departments. Proper enforcement of legislation is therefore essential to maintaining quality assurances.

The TGA 2019-2020 annual report highlighted that the most significant act of non-compliance found among CAM products was labelling adherence. In both random and targeted auditing of product labels, 32% and 29% non-compliance were found, respectively. Within these categories, 163 targeted reviews and five random reviews found that 74% of medicines had verifiable compliance breaches (125). The same report from the TGA indicated that no complementary and alternative products had been rejected from being approved for sale on the market. Similar work by Musgrave et al. (2014) examined 121 products for labelling adherence and highlighted systemic non-compliance. However, criticism from the TGA focused on the small sample size (121 out of 11604 products listed on the ARTG) and the stringency used in determining non-compliance (126). For example, 90% of products were

deemed non-compliant, although many of these discrepancies were minor and rectified almost immediately (e.g., missing full stops). Nonetheless, other non-compliance issues such as lack of registration on the ARTG, differing concentrations from what was listed on the ARTG, and incorrect concentration or dosage instructions are concerning due to their potential to cause harm (127).

This study aimed to produce a refreshed in-house assessment of a small number of immune supplements to better understand the prevalence of labelling non-compliance and its association with potentially harmful outcomes. As COVID-19 has increased the number of false or misleading advertisements, the TGA has placed this as its number one priority when addressing advertising compliance (128). Thus, this examination into labelling compliance of immune supplements paints a well-timed picture of the effects of labelling on consumer decision-making and health, notably, in response to the increased use of immune-stimulating products.

Methods

Seventeen immune supplements were purchased online from a prominent Australian chemist at the end of 2020. These products were chosen based on advertising claims such as "immune-boosting," "everyday vitality," or similar, in an attempt to mimic the selection choices of consumers during a global pandemic of a respiratory virus. No conscious thought to brands, ingredients, or dosages was given. Nine different nutraceutical brands were included, with all products manufactured in Australia. Once products were documented, their labels were compared with TGO92, which outlines the labelling standards for non-prescription medicines in Australia. Products were also screened for listing on the ARTG. Additional TGA regulations, such as the Poisons Standard (129) and TGA Permissible

Ingredient List(130), were used to determine appropriate warning labels for particular active ingredients such as zinc and *Andrographis paniculata*. A product's public summaries published by the TGA was compared to current physical labels to confirm similarity. Finally, databases including Google Scholar, Scopus, and PubMed were screened for individual product names to identify any published literature about the safety or efficacy of the product.

Analysis of labels

Compliance analysis was performed using a Microsoft Excel spreadsheet to divide the various compliance conditions into more defined variables that could be used for statistical analysis. Compliance data were managed using "yes/no" conclusions according to the requirements set out by the TGA. Raw data is presented in Appendix 1. One primary reviewer was used for all non-compliance; however, when specific ambiguity arose, two external reviewers were included in determining the compliance status. TGO92 is a relatively comprehensive document that defines the labelling for numerous non-prescription therapeutic products such as sunscreens, homoeopathic preparations, and various medicinal packaging. Thus, much of the Order was not relevant in the case of immune supplements, and sections were merged and blended for a clearer CAM-specific perspective on labelling compliance. An overview of the sections that are directly related to immune supplements are provided in Table 2.1

Table 2. 1 Overview of TGO 92 sections that specifically apply to the labelling requirements of immune supplements.

Section	Overview of requirements listed	Sections of interest
7	**"general requirements, including label presentation"** provided an overview of the information needing to be contained on main labels and that the text and background colour must contrast strongly	All
8	**"Information to be included on the label"** provides a relatively general overview of necessary label requirements such as, medicine name, active ingredients list, dosage and quantity of medicine, batch number, expiry date, storage conditions, sponsor contact details, relevant warning statements, directions for use and statement of purpose	All (excluding homeopathic remedies)
9	**"Information to be included on the main label"** and provides greater insights into the presentation of information, including information about naming and compliance requirements in regards to active ingredients	9.2, 9.5,
10	**"Qualifications and special requirements"** provided the specifics for other medicine packagings such as strips, dials and blister packs. As well as small containers.	10.9(a)
11	**"How information is expressed"** indicates the particular dosage units and metrics required for compliant products.	11.1, 11.2,11.5

In some cases, reviewers had to derive their own interpretation for legislation due to ambiguous phrasing (this interpretation is specified in the table of Appendix A). For example, S9.2 states that "the name of the medicine on the main label must be presented in a continuous, uninterrupted manner and not broken up by additional information or background text" (122). This was interpreted as the name listed on the ARTG and defined uninterrupted as the complete non-inclusion of any symbols, pictures, or logo that could confuse consumers. Furthermore, immune supplements are commonly multi-component

mixtures, many required compliance checks with S9.5 that referred to products containing two or more active ingredients that are a vitamin, mineral or herbal preparation.

S11.2.d indicates that text must not be <1.5 mm unless otherwise specified (active ingredients must be >2.5 mm). Measurement of font size was done to the specifications outlined on the TGA website, measuring only capital letters for consistency (122). No product was classified as a small container (<60 ml), and thus, all followed the same compliance protocols. Two products with strip packaging had secondary packaging independently assessed under S10.9 for "individually wrapped medicines." According to TGO92, S10.9(a), compliance for "strip dial and blister packs" relies on compliance with sections 8 and 9 for the main label. Both products in the sample population were compared with their appropriate compliance conditions and deemed compliant or non-compliant accordingly. Moreover, comparison between labels and public summaries primarily examined permitted warning statements, ingredient dosages, and permitted indications.

Discussion and Results

It was quickly determined that all products had some form of non-compliance under TGO92, limiting the understanding of the systemic effects of labelling non-compliance in a small sample size. Furthermore, previous studies into labelling non-compliance found that many of the requirements did not safeguard consumer health but were instead highlighting too rigorous legislation when the priority should be given to ingredient safety and efficacy studies. The previous misdemeanours examined in other labelling reporting (127) and negative industry response (13) suggested it was necessary to forefront issues on non-compliance that could potentially lead to harmful outcomes.

Twenty compliance components were measured and sorted according to the previously outlined definitions of potential to cause harm (Table 1). Potentially harmful non-compliance is defined as any form of non-compliance that could significantly impact human health due to labelling confusion or mislabelling. Hence, non-harmful non-compliance is non-compliance that is very unlikely to potentiate harm and may result from unnecessary stringency of the TGA. This process is limited by subjectivity; however, relevant risk-based regulation literature was assessed to ascertain potential analysis mechanisms (23, 35, 131, 132). It was determined that harm-based approaches are commonly used throughout the industry (133). Numerous statistical techniques were employed to add validity to the results. Chi-squared analysis of independence was applied to highlight non-harmful and potentially harmful non-compliance when examining individual compliance components of TGO92. Both a null and alternative hypothesis was developed to analyse the association between non-compliance and potential to cause harm:

H_0: Systemic non-compliance of immune supplements are independent of their potential to cause harm;

H_1: Systemic non-compliance of immune supplements are associated with their potential to cause harm.

Table 2. 2 Investigation into the labelling non-compliance of products and their potential to cause harm.

Potentially harmful	**N=**	**Non-harmful**	**N=**	
Wrong active ingredient presentation (S11)	0 (0%)	Non-cohesive title (S9.2)	17 (100%)	
Incorrect/missing listing status (S7.2(d)(i))	2 (11.76%)	Active ingredient font <2.5mm (S9)	17 (100%)	
Incorrect/missing warning statement (S8.1(k))	0 (0%)	Wrong/missing quantity (S9.1)	0 (0%)	
Warning statement <1.5mm (S7.2(d))	2 (11.76%)	Quantity and dosage form displayed <1.5mm (S7.2(d))	1 (5.88%)	
Missing/Incorrect batch details (S8.1(f))	0 (0%)	Listing status <1mm but still legible (S7.2(d)(i))	0 (0%)	
Missing/incorrect expiry date (S8.1(g))	0 (0%)	Batch prefix <1.5mm but still legible (S7.2(d))	1 (5.88%)	
Incorrect/missing directions of use (S8.1(l))	0 (0%)	Expiry prefix <1.5mm but still legible (S7.2(d))	1 (5.88%)	
Directions of use <1.5mm (S7.2(d))	1 (5.88%)	Missing/incorrect sponsor details and below 1.5mm but still legible (S8.1(i))	2 (11.76%)	
Wrong/missing storage instructions (S11.4)	0 (0%)	Storage instructions still legible but <1.5mm (S7.2(d))	1 (5.88%)	
Differing ingredient concentrations, permitted ingredients, or permitted indications on ARTG and physical label	2 (11.76%)	Incorrect/missing dosage form	0 (0%)	
	3 (17.64%)		**17 (100%)**	**Total products non-compliant**

The two variables, compliance and potential to cause harm, were not entirely independent and creating two-way tables of association was impractical in this small sample

size and low level of harmful non-compliance observed. Therefore, other methods were assessed for validity, such as Fischer's exact test, Barnard's, and Boschloo's statistical tests. Doing so revealed that adding statistical reasoning to the relationship between the potential to cause harm and non-compliance would not benefit the interpretation of results. As stated by Andrade (2019), just because something has statistical significance (*P*-value < 0.05) does not mean that it is valid nor relevant to the outcome (134). Therefore, the researchers kept the same hypotheses; however, centralised deeper examination of the legislation to understand the nature of non-compliance in immune supplement labelling in Australia. This aligned with the American Statistical Association who affirm that P-values for statistical significance are not exemplary to scientific reasoning(135).

Therefore, while the TGA believes it takes a risk-based approach (124), the declaration that CAM products are generally 'low harm' prevents necessary investigations into safety and efficacy. The ambiguity of legal jargon presented in the TGO92 made it challenging to determine the necessary stringency to apply when determining compliance, such as the lack of guidance on the expression of continuous titles in S9.2. It is a requirement for registration on the ARTG that each product contains a unique name. However, that name may be either proprietary or non-proprietary(136). Thus, many use their logo or brand name to register the product. This caused widespread non-compliance as logos used trademark branding, whereas extensions to the title such as "Immunoforce" or "Echinacea Plus" were provided in a different font, text size, colour, or border. While this does not suggest risk to consumer health, such products must be deemed non-compliant with TGO92 and not be sold in Australia.

Moreover, Information must be displayed in a font no smaller than 1.5 mm unless otherwise specified. Drawing on the work of Fuchs and colleagues, the font size has minimal

effects on individual comprehension and the ability to locate information (137). Hence, it was deemed to have no potential to cause harm when the information did not include important safety messages. While much non-compliance was deemed to be a direct result of unnecessary stringency of the TGA, the analysis was ambivalent about which stakeholder is at fault regarding active ingredient font size. TGO92 requires all active ingredients in a container with a volume capacity greater than 60 ml to have a font size greater than or equal to 2.5 mm, and 100% of products were deemed non-compliant on this variable.

Nonetheless, active ingredient lists were still legible and in only two cases went below 2 mm. Additionally, no discrepancies were found in the presentation of the active ingredients according to S11 of TGO92, "how information is to be expressed". Therefore, this was deemed unlikely to cause harm. Neither of the two products that included strip packaging had readily removable dosage units and thus were compliant with the conditions of S10. However, as S10 depends on compliance in sections 8 and 9, these strip packets were also non-compliant, as both products had issues with the active ingredient text size (>2.5 mm) and cohesive title components of compliance. Furthermore, one of these products claimed to "relieve symptom severity by 30%". This is not on the TGA's "permitted indications" for the product, and further review of the literature highlighted poor clinical testing due to reliance on *in vitro* studies and a higher administrated dose than recommended on the product (138). Nonetheless, this was the only product included in the study to be mentioned in peer-reviewed journals (139). Thus, further research into labelling and advertising of HDS should investigate the scope and quality of research published on specific products. A recent investigation into a Herbalife-brand associated with severe hepatotoxicity in 54-year-old women (140) highlights the importance of continued research. Increased case reporting by

clinicians will give greater insight into specific instances of toxicity and may force regulatory bodies to focus on the safety of these products.

Therefore, while this labelling compliance assessment is specific to immune supplements, these findings re-enforce the high level of non-compliance across the CAM industry in Australia found in earlier works (127) and the 2019-2020 TGA annual report (141). This unique study design attempted to calculate the association between compliance and risk and found that only three products (17.64%) had significant compliance issues that could directly cause harm. The reasons for non-compliance differed immensely. One product was registered under a different name on the ARTG, decreasing the traceability and accountability of the product, sponsor and manufacturer. Another product with no listing number also had lower *Echinacea purpurea* and calcium pantothenate content listed on the ARTG's public summary than on the label. Interestingly, these products came from the same major sponsor. The final non-compliant product with the potential to cause harm had a wide range of non-compliance in font size and non-harmful non-compliance in different sponsor names.

The font size was determined to be harmful in the case of critical safety information such as health warning statements (HWS) and directions of use. A recent review article highlighted the significant impact HWS had on selecting food and drink (142). Without appropriate legibility, warning statements could be overlooked. According to Gabriels and Lambert (2013), warning statements and use directions significantly influenced purchasing decisions (143). Furthermore, a limited presentation of warning statements, such as those required for *Andrographis paniculata* or high levels of zinc, increases consumer assumptions that natural products are safe and risk-free (131).

Under current TGA guidelines, all products included in the study should not be for sale on the Australian market. Thus, this paper joins the call from other toxicologists and pharmacologists (127) in the need for better enforcement and validation of herbal ingredients without focusing on unnecessary bureaucratic details that do not impact consumer health. An amelioration of current compliance legislation is essential and must attempt to meet the reality of manufacturers and regulators while centralising human health (38, 118).

Attribution Statements- Chapter 3: The composition and batch consistency of Australian immune supplements: A metabolomics approach.

STATEMENT FOR MANUSCRIPT – Should precede all chapters written in manuscript form.

The current manuscript is prepared for submission.

The following authors contributed to this manuscript as outlined below.

Authorship order	Contribution (%)	Concept Development	Data Collection	Data Analyses	Drafting of manuscript
Courtney Hall	75%	X	X	X	X
Garth Maker	5%	X	-	-	X
Torben Kimhofer	17.5%	-	-	X	X
Emily Davies	7.5%	X	X	-	-

Contribution indicates the total involvement the author has had in this project. Placing an 'X' in the remaining boxes indicates what aspect(s) of the project each author engaged in.

By signing this document, the Candidate and Principal Supervisor acknowledge that the above information is accurate and has been agreed to by all other authors.

Courtney Hall

Candidate

Garth Maker

Principal Supervisor

Chapter 3: The Composition and Batch consistency of Australian immune supplements: A Metabolomics Approach.

Chapter link

The previous analysis found that all immune supplements were non-compliant according to the TGA regulations. Nonetheless, the majority of this non-compliance was unlikely to cause harm. Therefore, a further look at the composition of these products will provide a more insightful examination into the safety and quality of immune supplements, particularly in reference to their active ingredients.

Introduction

The macroscopic features (leaf type, flower, root system) of plant varieties belonging to the same herb species are generally very similar. However, their molecular phenotype and chemical composition can vary substantially. Consequently, this variation is reflected in plant-derived products, such as herbal supplements, where the overall composition of products can vary due to two fundamental aspects (92, 144, 145). Firstly, the natural environment variability in which plants are grown, including geographic aberrations relating to climate or soil discrepancies. Studies investigating *Echinacea purpurea* have highlighted various phytochemical differences during growth stages and throughout growth regions (146). Mei *et al*. indicated much higher yields in tetraploid *E. purpurea* than diploid individuals of the same species (144). This study also found that the number of phenolic compounds for chicoric and chlorogenic acid were highest during the blooming time (144). Thus, fluctuations in the natural environment can significantly impact the biochemical composition of plant material. These fluctuations may cause variations in the therapeutic nature of immune supplements and decrease the reliability of beneficial components. Secondly, manufacturing and herb

processing procedures can affect levels of bioactive herbal compounds. A 2014 study employing HPLC-MS and NMR spectroscopy observed the variability of turmeric (C*urcuma longa)* along its commercial supply chain and found that high temperature processing of both extracts and raw product had drastic effects on the presence of therapeutic curcuminoids (113).

Additionally, herbal supplement production often relies on freeze drying and long storage periods, which can affect the phytostability of active therapeutic components (147). Therefore, a mix of uncontrollable factors such as genetic ploidy, weather conditions, production processes, and storage variables have caused some manufacturers to substitute and adulterate herbal supplements to maintain production output (148). Thus, increased attention to the inclusion of off-label products by the scientific community has highlighted the need for greater oversight of herbal products (28, 32, 36, 49).

In Australia, the Therapeutic Goods Administration (TGA) is the primary regulator of herbal supplements, with Good Manufacturing Practice (GMP) guidelines defined in the *Therapeutic Goods (Manufacturing Principles) Determination 2020* as the primary quality assurance strategy for CAM batch quality (149). A batch is defined as a uniquely manufactured quantity of a product that is expected to be homogenous. One study found that 69% of university students regularly took dietary supplements; however, minimal attention is given to the compositional quality and integrity of supplement batches in Australia (99). A significant increase in metabolomic literature regarding herbal supplements (Figure 1.1 of review article) suggest that this methodology warrants consideration for analysing the safety and quality of these products. Furthermore, previously unknown small molecules and their pathways are now profiled into databases, such as the Human Metabolome Database

(HMDB)(11), which is the parent database of the Natural Product Magnetic Resonance Database (NP-MRD)(150), specifically for plant metabolomics and DrugBank (151), for pharmaceuticals. The NP-MRD lists over 40,000 compound entries to aid in the identification and structural elucidation of phytochemicals when using NMR spectroscopy and should be used as a tool in compositional identification of plant species in herbal products.

NMR spectroscopy has become a popular technique that…..

Furthermore, technological developments in multivariate statistics have provided a suitable path for similarity analysis of complex botanical mixtures. In NMR, multivariate datasets are generated via the integrations of chemical shifts and peaks values into a spectrum. Xiong and colleagues (2013) demonstrated that principal component analysis (PCA) formulates a well-rounded inference about the statistical significance of variance found in botanical products(152). However, similarity analysis can often overfit the model when the coefficient of determination (r^2) is too low and must be accounted for during the interpretation of the analysis. Therefore, it is imperative to approach bioinformatic pipeline creation cautiously. The use of automated softwares such as SIMCA, .. and has lead researchers to often overlook the intricacies of applying statistical validation to complex mixtures such as immune supplements- including inherent difficulties such as metal ion strength effect on chemical shift, and herbal degradation. By using open-access development environments and database depositories it is possible to maintain a raw data fact checking rigour, enabling new and improved pre-processing pipelines to be achieved for better statistical validation and modelling.

Thus, this chapter investigates the batch variation of seven immune supplement products with three different batches per product, using ^{1}H NMR spectroscopy while

developing characteristic compositional profiles of products. Two multivariate strategies, principal component analysis (PCA) and statistical total correlation spectroscopy (STOCSY), were employed to examine the inter-batch variances. In instances where chemical structures were difficult to elucidate due to chemical shift and peak overlap, 2D J-resolved experiments were incorporated to gain structural information about the compositional identity and similarities shared between different batches of commercial immune supplements sold in Australia.

Methods

Product Sampling

From the original selection of seventeen products (Table 2.1, see chapter 2), ten products were randomly selected by a random number online generator(153) that used equal sample probabilities. Selected products were purchased at local West Australian chemists, where three different manufacturing batches were acquired for each product (identified by different batch numbers on the packaging). Different production batches were available for seven out of ten products in August 2021. Products were chosen when they held a unique batch number and the same listing number according to the ARTG. All samples (with brand name anonymised), their product identification, form, and active ingredients are listed in Table 3.1. The product form included tablets and capsules.

Table 3. 1 Overview of selected immune supplements. Listed are the product's physical form and active ingredients declared on the main label used to purchase analytical standards.

Product ID	Form	Active Ingredient(s)*
A	Tablet	*Andrographis paniculata, Echinacea Purpurea*, olive leaf, turmeric, vitamin C, zinc.
E1	Capsule	*Echinacea Purpurea*
E2	Capsule	*Echinacea Purpurea*
V1	Tablet	Vitamin C, Echinacea Purpurea, cholecalciferol, zinc.
V2	Tablet	Vitamin C, *Echinacea Purpurea*, zinc, garlic.
V3	Tablet	Vitamin C, zinc, citrus bioflavonoid extract.
S	Capsule	*Sambucus Nigra, v*itamin C, zinc.

* **According to product label**

Chemical References

Nine chemical standards were purchased from Sigma-Aldrich (Melbourne, Australia): curcumin from *Curcuma longa*, chlorogenic acid, chichoric acid, caffeic acid, andrographolide, oleuropein, L-alliin, and kuromanin chloride. The standard for ascorbic acid was from an in-house collection ((+)-Sodium L-ascorbate). These compounds are active polyphenols found within the supplements active herbal or inorganic material (ascorbate) (Table 1.1).

Sample processing

Solid tablets were crushed with a mortar and pestle, and 10mg of the CAM powder or capsule contents were transferred into a microcentrifuge tube and mixed with 1 mL of

98% deuterated methanol (CD_3OD) solution (10% w/v). Analytical standards were prepared according to their molecular weight in a mixture of 1mL CD_3OD, with a final concentration of 10Mm.

These mixtures were vortexed using an MS1 minishaker at 1.8 x 10^5 rpm until completely dissolved. The solutions were then centrifuged for 5 minutes at 7378 x g. A volume of 585 μL supernatant was transferred to a fresh microcentrifuge tube where 65 μL phosphate-buffered sodium trimethylsilyl-[2,2,3,3-2H4]-propionate (TSP) buffer (7.4pH) was added and mixed with deuterated methanol at a ratio of 1mg of TSP to 1ml of CD_3OD. 650μL of the final mixture for each sample was pipetted into an NMR tube (5mm) with a lid and placed into a 96 well rack to be immediately loaded into the ambient autosampler of the NMR spectrometer.

Proton NMR Spectroscopy

1D and 2D ^1H NMR experiments were performed on Bruker Avance IVDr 600 MHz spectrometer. Standard 90˚ 1D 1H NMR spectra were acquired using a standard NMR pulse sequence (noesygppr1d), with the receiver gain (RG) set to 90.5 AU and 64 scans. These standard Bruker parameters were used for cross-institutional replicability and implementation into high throughput regulatory environments. In its most basic form, this pulse sequence can be described as:

$$RD - 90° - D_{MIX} - ACQ$$

The relaxation delay (RD) amounted to 4.0s, followed by a standard $90°$ radio frequency (RF) pulse, a mixing time (D_{mix}) of 0.01 s and free induction decay (FID) acquisition time (ACQ) of 2.7s. The FID was collected with a spectral width of 20 ppm and 131,072 data points expressed as δ chemical shift (ppm). J-resolved experiments (jresgpprqf) were

performed with RG of 90.5. The number of datapoints collected for the direct (f2) and indirect (f1) dimensions were 8192 and, 40 respectively.

The following additional 2D experiments were performed for chemical reference standards: Heteronuclear Multiple Bond Correlation (HMBC), Heteronuclear Single Quantum Coherence (HSQC), and ^{1}H-^{1}H Correlation Spectroscopy (COSY) 2D using standard experimental parameters for future analysis.

Spectral data pre-Processing

NMR experiment data were imported into Python workspace, with 1D spectra represented as a two-dimensional matrix *X* with dimension *n x m*, where *n* and *m* represent the number of rows (spectra, *n*=72) and columns (chemical shift variables, *m*=131,072), respectively. 1D ^{1}H NMR spectra were pre-processed using Python software libraries provided by the Australian National Phenome Centre (ANPC). The pre-processing pipeline included chemical shift calibration (TSP apex centred at 0.0 ppm), excision of solvent signals (CD_3OD: 3.23-3.35 ppm and 4.7-4.9 ppm) and signal-free sections up/downfield of 0.5 and 9 ppm. Baseline correction (asymmetric least squares) was then performed, where the optimal parameters were determined graphically ($\lambda = 10^{6}$, $p = 10^{-4}$). The pre-processed data matrix, *X* (*n*=72, *m*=55,260), was the input for the statistical analysis as outlined below.

Multivariate Statistical Analysis

In contrast to univariate statistical methods, where a single variable is investigated at a time, multivariate statistics enable the simultaneous analysis of multiple variables. Principal Components Analysis (PCA) was performed as an unsupervised multivariate statistical method, which has previously been applied in the field of metabolomics to comparative herbal supplement analysis (11-14). PCA reduces the dimensionality of large

NMR matrices $(X_{n,m})$ which allows interpretation of systematic variation in few components. PCA generates scores (*T*) and loadings (*P*) according to:

$$X = T * P + E$$

With T and P of dimensions $n\ x\ p$ and $p\ x\ m$ respectively, with p representing the number of principal components and the residual matrix $E_{n,m}$ accounting for unstructured variation, such as noise intensities. In PCA, a principal component (*PC*) is a linear composite of the original variables, which produces one scores value for a single spectrum on each *PC*. Each *PC* describes systematic variances that are observed throughout the spectral data with variable weightings across each PC being orthogonal. Prior to PCA, data were autoscaled according to:

$$\hat{x} = \frac{x - \bar{x}}{\sigma_x}$$

Where x is a column vector of X, $\bar{x}$ is the statistical mean and σ_x is the standard deviation. $\hat{x}$ is the autoscaled version of x which was used to calculate PCA model, to prevent effects of magnitude dominating the PCA model.

PCA model interpretation in metabolomics is performed using a scores scatter plot with t_i on the x-axis and t_{i+1} on the y-axis, with i representing the index for PC. Thus, the Matrix T and P are derived from multiplying the elements of $X_{n,m}$ using nonlinear iterative partial least squares (NIPALS) which includes the calculation of a weighting's matrix (W) that is typically not interpreted. In addition to scores plotting, Hotelling's T^2 is calculated for PC i and $i + 1$ and presents as an ellipse in the scores plot. The Hotelling's T^2 can be interpreted as a multivariate confidence interval (95%) that can be used for the identification of outliers.

PCA model loadings for PC_i were statistically reconstructed using Pearson's correlation calculated from autoscaled X with t_i, according to:

$$r = \frac{1}{n}(t_i^T, X)$$

Further details on how this formulates as a correlation matrix is available in Appendix B. The statistical reconstructions of loadings were visualised in a line plot, that resembled an NMR spectra to ease interpretation. The x-axis represents the chemical shift variable (ppm), the y-axis shows the covariance of $T_{n,i}$ and $X_{n,m}$, encoding magnitude information of the original chemical shift variables. Whereas the correlation is denoted as r which is represented as the continuous line colour gradient, ranging from 0 to 1.0 (absolute values), indicating the variables model importance with the sign of covariance representing the characteristic peaks according to the scores values of that PC. These PCA model plots were set side by side with pre-processed spectra to aid in interpretation of reconstructed loadings and to interpret individual signal patterns in relation to the PCA model.

Metabolite identification

Statistical Total Correlation Spectroscopy

Statistical total correlation spectroscopy (STOCSY) is commonly used to observe the correlation between different peaks present in a full resolution spectrum, providing an avenue for directly identifying peaks from the same molecule. STOCSY calculates the Pearson's correlation between variables associated with a user-defined driver peak. The driver peak is selected manually, depending on PCA model importance and generally are peaks of high intensity. Similar to statistical reconstructions of PCA loadings, the STOCSY plot is constructed with the chemical shift (ppm) on the abscissa (x-axis), covariance

between driver and all columns of $X_{n,m}$ on the ordinate (y-axis) and absolute Pearson's correlation as a continuous line colour gradient. Therefore, the reconstructed spectrum highlights how resonance intensities of a single molecular species change with sample concentration in a constant pattern.

Chemical reference standards and spectral overlay plots

To identify compounds in the product spectra, chemical references representing potential biomarkers for active ingredients were used to match peak patterns with signals in product spectra. These biomarkers were identified through literature searching based on the phytoequivalence theory, which states that references of compounds found in plants can be used for comparison in complex mixtures.

Peak shifts in reference spectra were compared with spectral information obtained from three databases, HMDB(11), NP-MRD(150), and DrugBank(151), to validate peak similarities between in house chemical reference standards and online sources. Hence, chemical reference standards were overlayed with each product spectrum to devise the similarity of peak shape and position, thus indicating if a standard was present in the respective product. Where there were instances of immense peak overlap in product spectra, 2D J-resolved spectra were used to extract information about peak shape and deconvolve multiple signals via visual assessment.

To identify the product's excipient ingredients listed in the Australian Register of Therapeutic Goods (ARTG), chemical shifts were compared with database entries where available. NMR spectra of compounds gelatin (ChEBI ID 5291) and hypromellose (ChEBI ID 30618) were not found on public databases and simulated/predictive spectral data were

interpreted with caution. To improve visualisation, reference spectra were scaled in intensity to match proton NMR spectral intensities of products.

Compound relative quantification

In proton NMR, the area of a peak is directly proportional to the respective compound concentration in the sample. Therefore, peak integration was performed on selected signals to obtain a relative measure of compound concentration. Chemical shift sections used for integration were manually selected, focusing on non-overlapping major peaks and characteristic for respective compounds identified by spectral comparison with reference standards or database entries. Integrals were estimated using the sum of intensities for a particular peak. Integral estimates are tabulated in the respective results section. The coefficient of variation (CV) for each integration was calculated accordingly:

$$CV_x = \left(\frac{\sigma_x}{\bar{x}}\right) * 100$$

Where the standard deviation (σ_x) of the peak integral of a compound was divided by its mean $(\bar{x})$ and multiplied by one hundred to obtain a percentage value.

In order to provide a reference for future identification of chemical standards, all major peak shift positions were recorded together with relative integrals displayed in a Table 3.2 of the results section.

Results

PCA Analysis of all products

PCA was performed using spectra of all the different batches to obtain an overview of the main spectral variation patterns. A two-component model was calculated where the percentage of total variation expressed in PC 1 and PC2 amounted to 28.6% and 11.3%,

respectively (Figure 3.1). According to products, the distribution of the scores showed a clear clustering trend, with most batches clustering on the t1 axis. As the majority of products had a negative scores value (t1<0) on PC 1 scores plots, some systematic coherence between immune supplements, in general, was highlighted. Excluding product A, which had a positive score value (t1>0), was classified as an outlier on the t^2 hotelling ellipse. This is significantly reflected in the PC1 loadings, where a strong correlation is expressed positively and highlights multiple systematic differences in the raw spectrum. Spectral signals with high positive variable importance on PC1 include signals in the aliphatic region, 2.25ppm(t) 1.28ppm(s), 0.88ppm(t) and are most likely relate to stearic acid (a common excipient ingredient). However, high model importance was also indicated for peaks in the aromatic region and indicated the presence of polyphenols in supplements. These signals were found in high intensity in product spectra A, E1 and E2, which all declare *E.purpurea* as an active ingredient.

Firstly, peaks arising from product A share extremely high collinearity exhibited by the peak at 1.20ppm. Further covariance is a result of chemical shift, highlighted by the singlet at 1.27-1.28ppm. Furthermore, there is apparent variation observed within the individual batches of A that needs to be independently reviewed to understand the batch effects of individual products. Additionally, t2 expressed less overall variance as product A was not an outlier. Interestingly, there is a clear distinction between the echinacea products (E1 & E2), which were positively correlated, while the rest were negatively skewed. There is some inter batch discrepancies exhibited within products E2, S, and A primarily.

Investigation into individual batches is required for more detailed analysis. The loadings for PC2 indicate both positive and negative correlations. At the same time, the

distribution of the scores of t2 indicates that these correlation differences are primarily derived from products S and E2. As E2 is closer to being classified as an outlier, this will have a more significant weighting in the covariance matrix and thus have the most prominent effects in the PCA model. There are two key variance trends accounted for in PC2. Firstly, the different peaks associated with E2 and S. For example, the singlet for positively correlated products at 4.56ppm shows a high covariance. This is caused by a systematic difference in peak shift and intensity differences related to correlation. In this instance, both factors are likely to play a role. Three distinct peaks are negatively correlated and share systematic differences. A singlet at 1.28ppm, a doublet at 3.66ppm and a multiplet at 3.88ppm are compared between the PC2 loadings and raw spectra to indicate that these peaks are not found in all compounds. From the distribution of the t2 scores, it is evident that this covariance primarily results from product S. Further insight is provided by the t3 plane, expressing 8% of the total variation. Products V1, V3, E1, E2, and A, are clustered around zero and minimally contribute to the loadings model of PC3. Importantly, products S and V2 have immense differences and replicates found outside the 95% confidence ellipse. The loadings for PC3 are indicative of the immense impact that V2 has on the covariance matrix. Nonetheless, further comparison between the scores and spectra presents the large chemical shift observed in the positive covariance of the reconstructed spectra between 3.5-3.9ppm of PC3. This is directly related to peak shift, and a closer examination of the spectra highlights no difference in peak shape, only chemical shift.

Interestingly, the minor negatively correlated peaks between regions 7.0ppm and 7.5ppm showed high covariance (r^2>0.8) due to the influence of peak shape caused by dilution. Therefore, the chemical shift relating to product V2 and the concentration differences of product S is accounted for in the PC3 model. Although products A, V2, and S

present outside the 95% t^2 hotelling ellipse, no outliers were excluded due to the nature of the unsupervised analysis, small sample size, and exploratory nature of the hypothesis. Thus, relying on the combination of scores, loadings, and raw spectra analysis can indicate the systemic variation more accurately than relying on their information separately. Hence using this method presented specific common trends that were observed throughout all the batches.

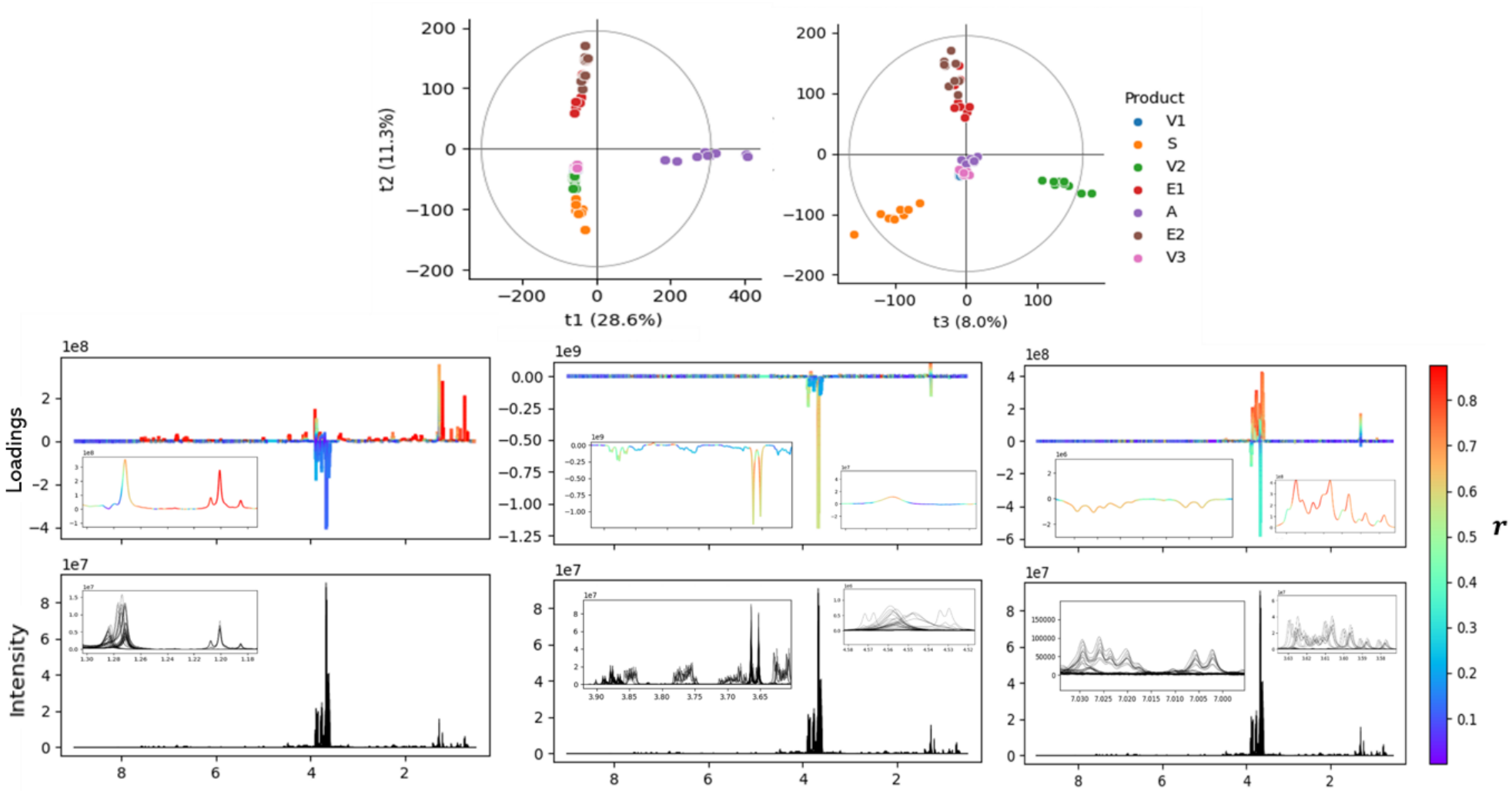


Figure 3. 1 3 Principal component model of all products. Top: Scores plots for t1 (28.6%), t2 (11.3%) and t3 (8.0%). Bottom: Reconstructed loadings comparison (PC 1, PC 2, PC 3 (left to right)) of all products, highlighting the common variance trends such as chemical shift and compositional differences.

Reference integration

Table 3. 2 Chemical reference integrations according to peak positions with Chemical Entity of Biological interest identification number (ChEBI).

Reference standard (ChEBI ID)	Signal position(ppm)	Integration	P*
Chicoric acid (3594)	5.78(s)	0.978	A, E1, E2, V1, V2
	6.35(d)	0.981	
	6.77(d)	0.947	
	6.98(dd)	1.000	
	7.06(d)	0.986	
	7.64(d)	0.955	
Chlorogenic Acid (16112)	2.04(m)	0.983	A, E1, E2, V1, V2
	2.18(m)	1.000	
	3.70(dd)	0.557	
	4.15(m)	0.555	
	5.32(m)	0.534	
	6.24(d)	0.549	
	6.75(s)	0.538	
	6.93(dd)	0.571	
	7.03(d)	0.562	
	7.54(d)	0.542	
Caffeic acid (16433)	6.20(d)	0.561	A, E1, E2, V1, V2
	6.76(d)	0.896	
	6.99(dd)	1.000	
	7.01(d)	0.948	
	7.50(d)	0.922	
Ascorbic acid (38290)	3.19(m)	0.011	A, V1,
	3.43(m)	0.011	
	3.68(m)	1.000	V
	3.85(td)	0.487	V
	4.35(d)	0.463	S
Kuromanin (71661)	2.14(s)	0.026	S
	3.67(m)	1.000	
	3.89(dd)	0.519	
	5.28(d)	0.472	
	6.64(d)	0.286	
	7.01(d)	0.546	
	8.02(d)	0.411	
	8.25(dd)	0.498	
	9.01(s)	0.510	
Andrographolide (65408)	0.76(s)	0.459	A
	1.22(s)	1.000	
	4.12(d)	0.321	
	4.16(dd)	0.153	
	4.47(m)	0.147	
	5.013(d)	0.147	
	6.85(td)	0.149	
Alliin (2596)	1.96(s)	0.027	A
	3.04(m)	0.217	
	4.10(dd)	0.293	
	5.48(m)	1.000	
	5.94(m)	0.432	
Oleuropein (7747)	1.18(t)	0.093	A
	1.67(dd)	1.000	
	2.44(m)	0.326	
	2.70(dd)	0.330	
	2.77(t)	0.010	
	3.71(s)	0.971	
	3.97(m)	0.343	
	4.11(m)	0.332	
	4.21(m)	0.334	
	5.91(s)	0.345	
	6.08(m)	0.322	
	6.55(dd)	0.341	
	6.68(m)	0.628	
	7.51(s)	0.279	
Curcumin (3962)	3.914(s)	1.000	A
	5.96(m)	0.031	
	6.62(d)	0.341	
	7.57(d)	0.363	
	7.49(d)	0.085	
	7.22(d)	0.336	
	7.11(dd)	0.336	
	6.82(m)	0.404	

***Expected product according to main label**

Composition & Batch variation of Product A

^{1}H NMR spectra of product A were graphically compared with reference spectra of the five active ingredients listed on the product label: Andrographolide, ascorbic acid, chicoric acid, curcumin, oleuropein (Table 3.1). In all cases, the full signals trace of the reference spectrum could not be identified in any of the product A spectra due to peak overlap between compounds (Figure 3.2). High-moderate intensity signals of the reference spectra were identified for curcumin (3.90 ppm (d), 7.56 ppm (d), 7.31 ppm(s), and 7.09 ppm (dd)), oleuropein (2.75ppm(t), 2.69pmm(d), 2.67ppm(d), 1.64(dd), 6.53ppm(dd), 6.81(d), andrographolide (0.73ppm(s), 4.44ppm(m), 6.832ppm(td)) and ascorbate (2.75ppm(t), 3.87ppm(m), 3.17(m)). These characteristic chemical shift patterns were confirmed by J-resolved spectroscopy (Figure 3.1.1, for example, of curcumin).

The comparison of NMR signals for andrographolide reference with product A spectra were inconclusive due to the peak overlap of the complex mixture. For example, the andrographolide reference standard has 13 peaks; however, only three were represented in the spectra. Therefore, relative concentrations could only be derived from free-standing peaks with no contribution from other compounds. Interestingly, there were minimal concentration differences across batches for characteristic peak 0.73ppm (s) for andrographolide. Contrarily, curcumin had the most variability with the highest coefficient of variation (42.6%) and concentrations doubled in some products compared with others. No relative integration for chicoric acid could be deduced due to the lack of free-standing peaks. Therefore, it is not possible to confirm the presence of chicoric acid within product A. Nonetheless, batch variation was consistent between batches as batch S.B had the highest concentration and batch S.A had the lowest throughout all chemical reference standards.

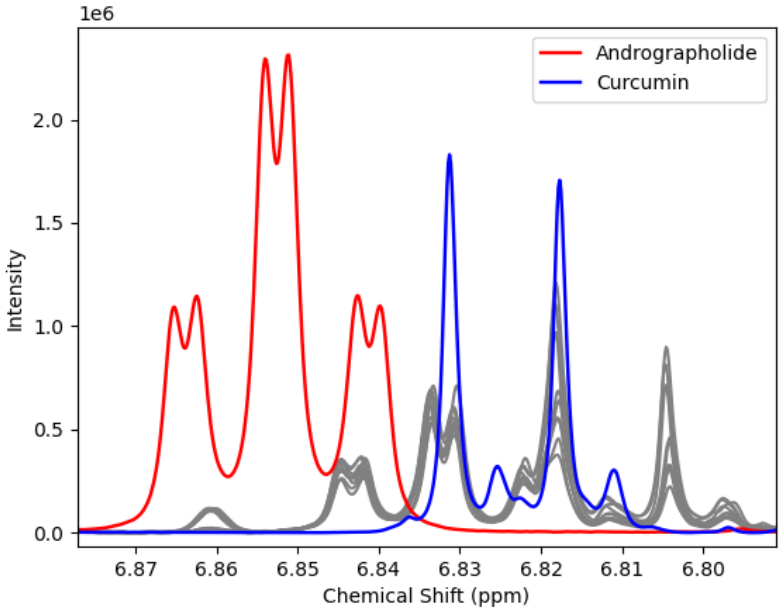


Figure 3. 2 **1H NMR Spectral comparison of chemical shift region 3.80-3.88ppm for andrographolide and curcumin reference standards with product A. Highlights peak overlap of complex mixtures, as one of the fundamental problems of 1H analysis of complex mixtures.**

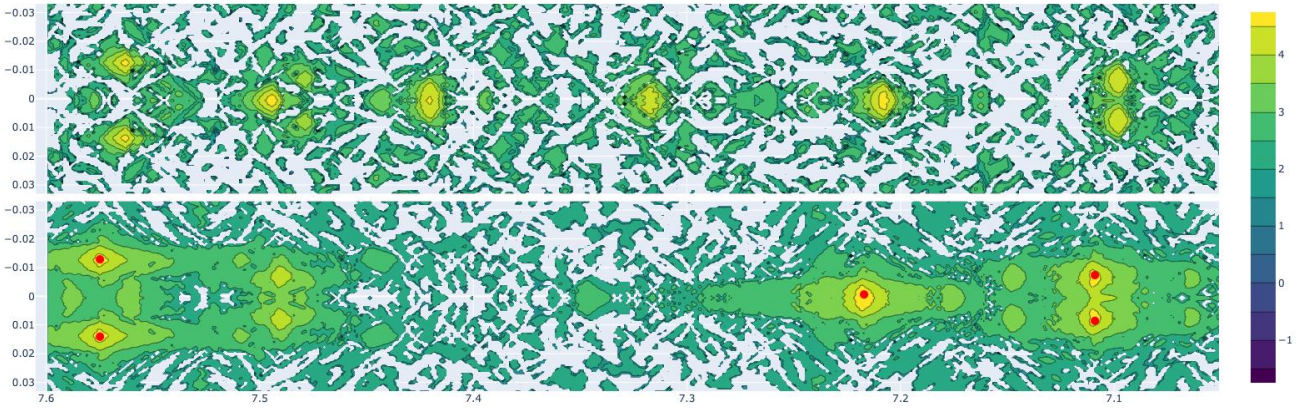


Figure 3. 3 **Comparison the chemical shift region 7.05-7.60ppm for product A (top) and curcumin standard using J-resolved spectroscopy. Red coloured points indicate local peak maxima determined by zero crossing of the first derivative.**

The comparison of NMR signals for andrographolide reference with product A spectra were inconclusive due to the peak overlap of the complex mixture. Therefore, relative concentrations could only be derived from free-standing peaks with no contribution from other compounds. Interestingly, there were minimal concentration differences across batches for characteristic peak 0.73ppm (s) for andrographolide. Contrarily, curcumin had

the most variability with the highest coefficient of variation (42.6%) and concentrations doubled in some products compared with others (Figure 3.3). No relative integration for chicoric acid could be deduced due to the lack of free-standing peaks. Therefore, it is not possible to confirm the presence of chicoric acid within product A. Nonetheless, batch variation was consistent between batches as batch S.B had the highest concentration and batch S.A had the lowest throughout all chemical reference standards. Further statistical analysis, including PCA and STOCSY, was required to observe particular peaks contributing to the overall batch variance.

Table 3. 3 Relative integration for peaks relating to active ingredients found in product A.

Compound	**Signal position**	**Relative integration**	
		Mean (range)	**Coefficient of Variation (%)**
Andrographolide	0.73ppm(s)	$1.2x10^8$($1.0x10^8$-$1.4x10^8$)	10.7%
Oleuropein	2.74ppm (t)	$2.6x10^7$($2.1x10^7$-$3.4x10^7$)	21.0%
Curcumin	7.2ppm(s)	$1.9x10^7$($1.0x10^7$-$3.4x10^7$)	42.6%
Ascorbate	3.87ppm (td)	$3.0x10^8$($2.4x10^8$-$3.9x10^8$)	15.7%
Chicoric Acid	NA	NA	NA

In order to assess batch variations for product A, a two-component PCA model was calculated where the total variation explained by PC 1 and PC 2 amounted to *36.8*% and *19.7*%, respectively. The scores (t) distribution for PC1 shows a clustering trend of batch B spectra, with positive t1 values, and batch A and C spectra, with negative t1 values (Figure 3.4). High-intensity NMR signals contributing to this scores distribution are mainly located in the aliphatic chemical shift region for batch C spectra (PC 1), with STOCSY analysis indicating peaks at position 0.7*3* ppm (s), 1.24 ppm (s), 1.58 ppm (s), 3.86ppm (m) and 3.66 ppm (d) being structurally related (driver peak = 0.8821 ppm)(Figure 3.5).

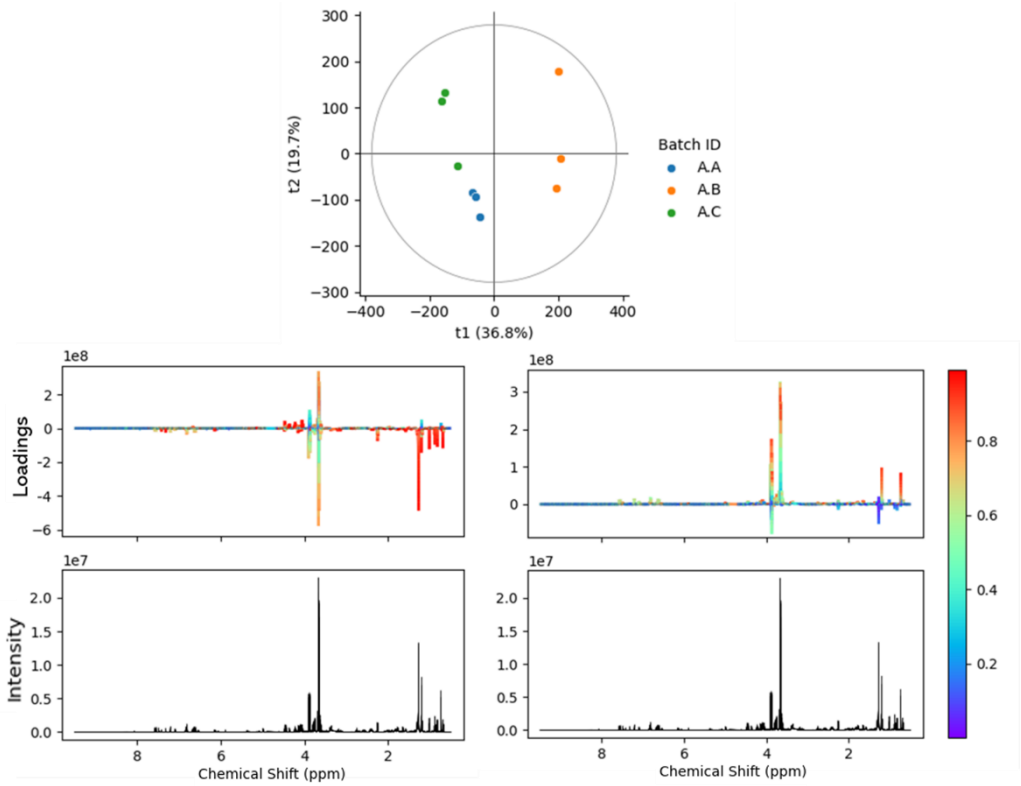


Figure 3. 4 **2 principal component model of product A. Top: Scores plot of t1 and t2 axis. Bottom: PC 1 (left) reconstructed loading comparison with raw spectra and PC 2 statistical comparison (right).**

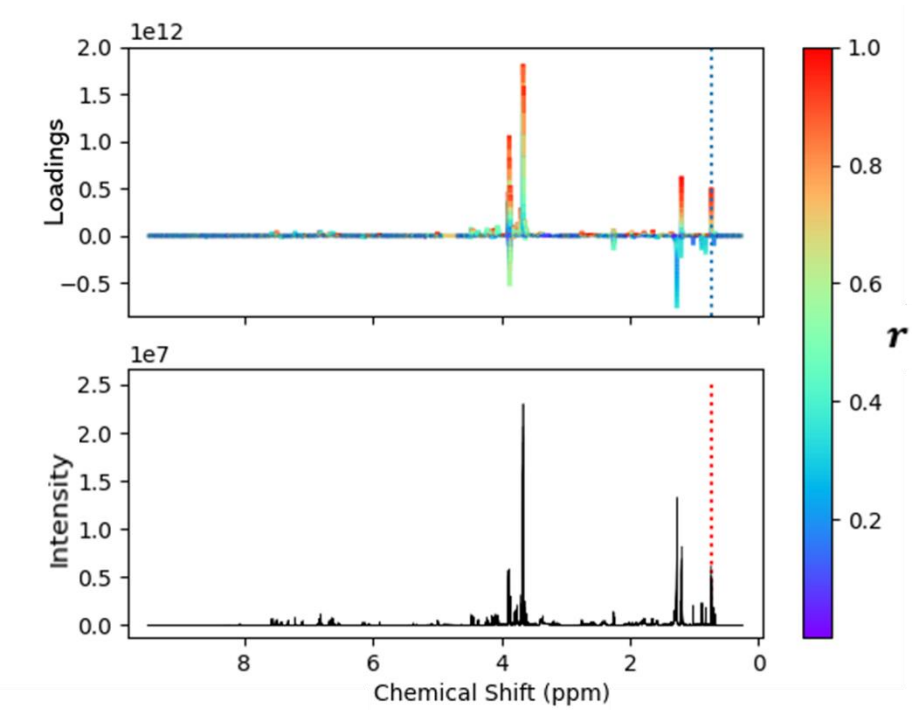


Figure 3. 5 STOCSY comparison with pre-processed spectra of product A (d=0.8821ppm). Highlights four peaks that express systematic collinearity and are likely contributing to the same compound in the mixture.

The STOCSY-derived peak patterns did not match the pure compound-spectrum of reference spectra or databases (Figure 3.1.3). However, a chemical shift comparison using an online database search on (HMDB) ranked lauryl sulphate as a potential candidate (0.859ppm(t), 1.248ppm(s),1.488ppm(t)).

This is a common excipient that was not listed in the product public summary. Characteristic NMR signals for batch A and B are located in chemical shift regions 4.0 to 4.5 ppm. The STOCSY analysis for these signals did not produce coherent information related to structural correlation since the low-intensity signals produced high correlations coefficients for the entire chemical shift range. Signals with high PCA model importance on PC2 include 0.73(s), 1.201(s), 3.66(d) and 3.87 ppm (all characteristic for batch C). It is noteworthy that significant chemical shift variations were observed across all spectra of product A (Figure 3.1.2). These signal variations are visible in PCA loadings of principal component (PC) 1 and 2, with strong peak shift variations observed for signals in the chemical shift regions of major ascorbate-like signals 3.66ppm(d) and 3.87ppm(m) for PC1, respectively (Figure 3.1.2). A detailed evaluation regarding the impact of chemical shift variation on the effectiveness of PCA for the comparison of batch variations is provided in the discussion section.

Echinacea Products

Composition and Batch analysis of Product E1

Product E1 and E2 are high strength Echinacea supplements, and ^{1}H NMR spectra of both products were graphically compared with reference standards of chicoric acid, caffeic acid and chlorogenic acid. These compounds represent the active ingredients listed on the product label and are reported in the literature to be distinct components of *Echinacea*

Purpurea. All of the reference standards gave rise to NMR signal patterns in the aromatic chemical shift region, with minor differences observed in chemical shift positions and relative signal concentrations (Table 3.3). The 1D ^{1}H NMR spectra of product E1 showed comparable signals to all reference compound spectra (6.4ppm(d), 6.76ppm (d), 6.94ppn(dd), 7.06ppm(d),7.61ppm(d)).

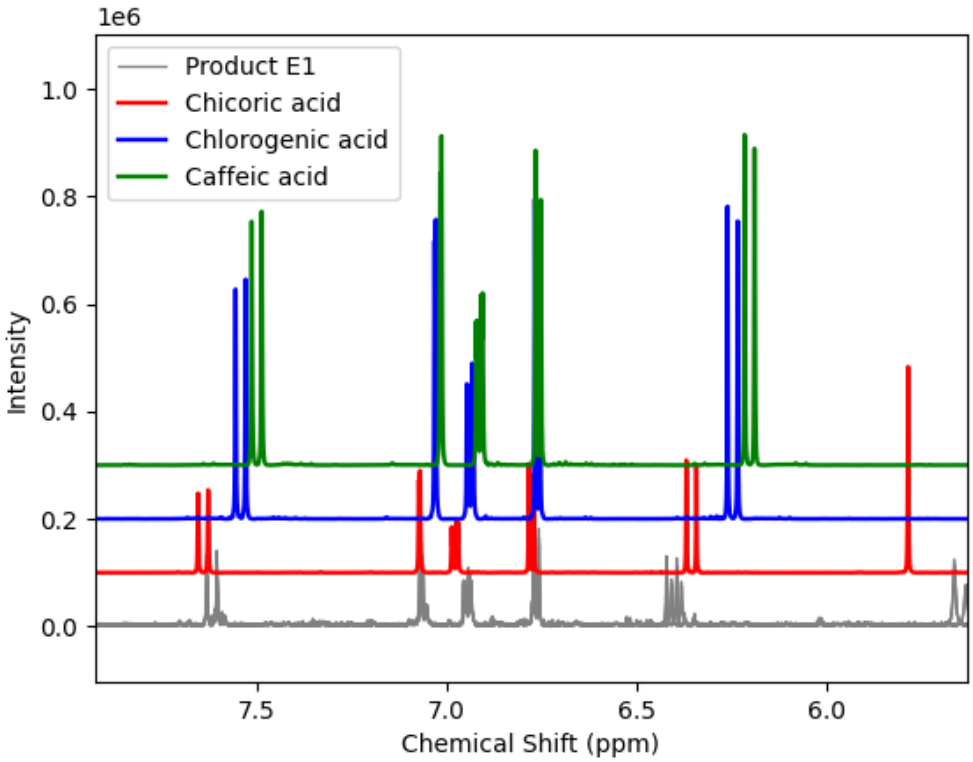


Figure 3. 6 1H NMR spectral comparison of the aromatic region of all samples for product E2, compared with reference standards: Chicoric acid, Chlorogenic acid, Caffeic acid.

Additional signals with high intensity were observed in the aliphatic chemical shift region, particularly the singlet at 1.28ppm. Although, the singlet peak identity could not be confirmed with chemical reference standards, STOCSY, or database searching. NMR signal patterns of product E1 did not completely overlap with signals observed for reference standards, and a conclusive compound assignment was not possible using 1D NMR

experiments. NMR signal difference among the product spectra included one additional singlet at 5.75 ppm for chicoric acid (Figure 3.6).

To improve compound identification in regions of peak overlap, a further comparison was performed using J-resolved experimental data. The signal traces of chicoric acid located downfield of 5.7 ppm are visible in product spectra, although with minor chemical peak shift variations (Figure 3.6). Due to the peak similarities of chemical reference standards, it was not possible to determine which reference standard was contained in the product. All peaks found in the aromatic region of the product, besides the singlet at 5.75ppm, shared chemical shift similarities with all three reference standards, this resulted in peak overlap, and peak intensity increases that were directly related to the inclusion of numerous compounds in one peak.

Thus, relative peak concentrations were needed to decipher their presentation in the product. While chicoric acid was definitively found in the product due to the observed full signal trace (figure 3.2.1). The peak pattern ratio found in the product did not match chicoric acid and is most likely a result of overlap with signals relating to chemical reference standards chlorogenic acid and/or caffeic acid. The relative intensities were observed in the product are 5.63ppm (s, 1.0), 6.39ppm (d, 1.0), 6.76 ppm(d,~1.2), 6.94ppm(dd, 1.0),

7.06ppm(d, ~1.3), 7.61ppm(d, ~1.1). Therefore, it is evident that either chlorogenic acid/caffeic acid or both contribute to the signals observed in product spectra.

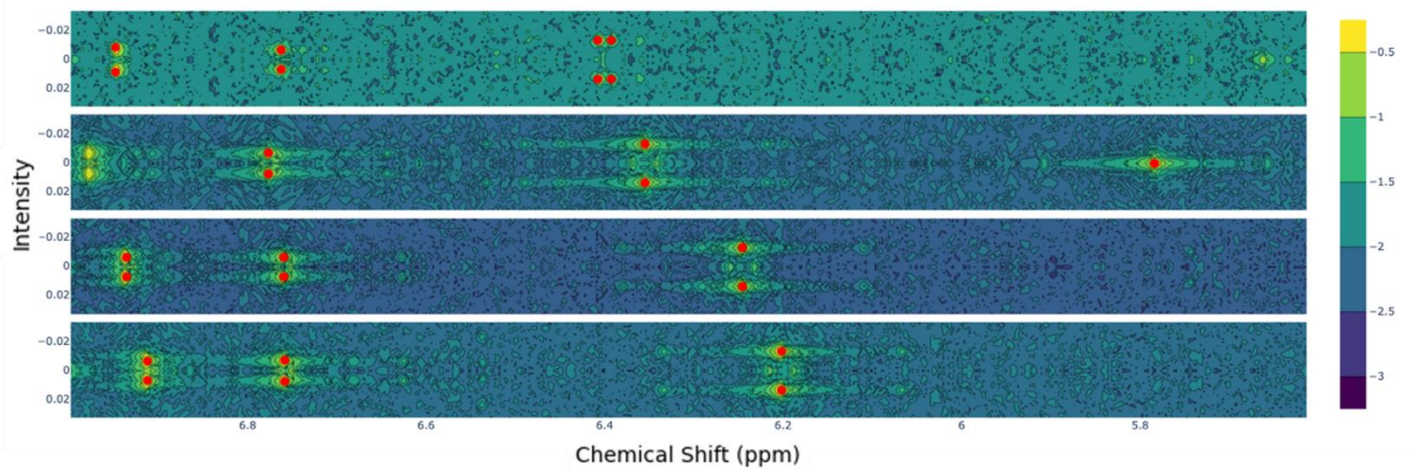


Figure 3. 7 J-resolved spectral comparison for the of product E2. The min-max scaled mean of all samples of product E2 (top) compared to the min-max scaled reference standards in descending order: chicoric acid (row 2), chlorogenic acid (row 3), caffeic acid (row 4) for the aromatic chemical shift region (5.6 - 7.0 ppm). The product doublets at 6.39ppm is a shifted doublet across spectra and present as two discrete doublets in the statistical mean spectrum. Red coloured points indicate local peak maxima determined by zero crossing of the first derivative.

PCA was performed better to understand compositional variation across batches of product E1. PC 1 and 2 explain a percentage of the total variation (r^2), 27.4% and 19.9%, respectively. The t1 values highlighted a clear cluster of batches E2.B and E2.C. with low t1 values (t1<0) whereas E2.A. was formed separately with a high t1 value (t1>0). The clustering trend exhibited in t1 is primarily a result of chemical shift (Figure 3.8). This is observed in signals of high model importance such as 1.40ppm (s) and 1.28ppm (s), where peak shift was systematic across batches (Figure 3.8). Peaks relating to the chemical reference standards (Table 3.3) are also of high PCA model importance. However, this also relates to systematic peak shift as illustrated in Figure 3.9. These systematic peak shifts are likely related to ionic strength differences. For example, produced by the inclusion of varying amounts of magnesium stearate as an excipient ingredient (see also discussion).

High loadings for PC 2 were observed for signals showing both, batch related systematic peak shift mainly in regions with major peak overlap, and intensity differences of signals located in the aromatic region where characteristic signals of chemical standards are found.

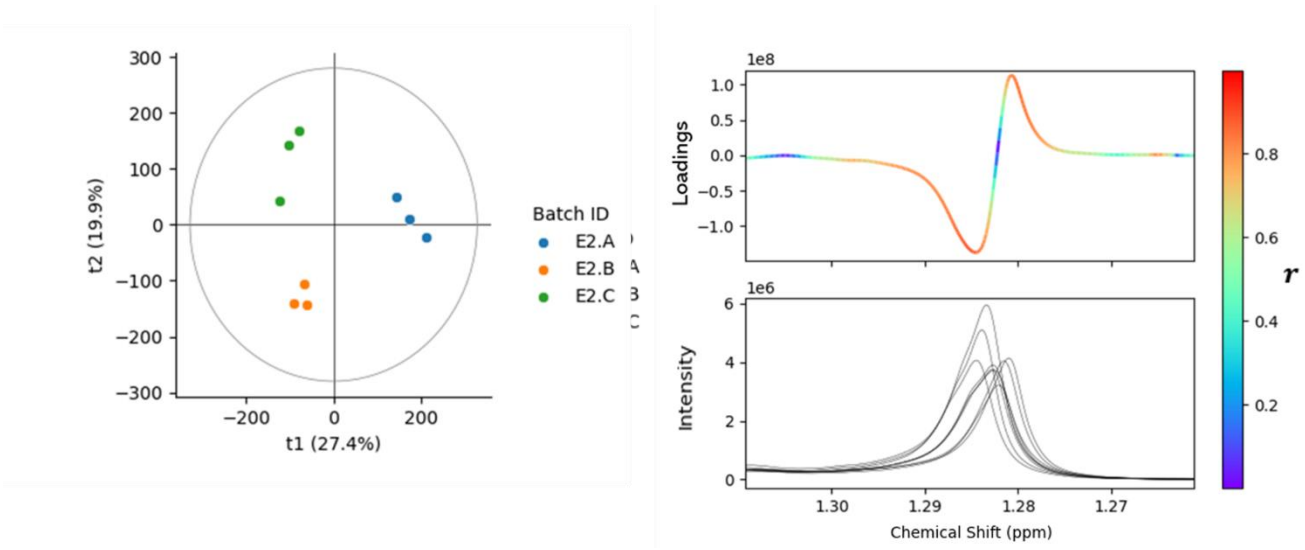


Figure 3. 8. PCA analysis of product E2. A) Scores of PC1 and PC2. B) reconstructed loadings of chemical shift region 1.24-1.34ppm with major signal split with high negative and positive model importance on PC 1 (upper panel). This is related to systematic peak shift for the singlet at 1.28ppm across batch spectra (lower panel).

To further elucidate the concentration differences of active ingredients in product E1, the spectra of all batch samples were integrated for the selected signals: Singlet at 5.63ppm, and a doublet of doublet at 6.94ppm, which could be related to chicoric acid, chlorogenic acid and/or caffeic acid, respectively. These integrals showed significant differences as the maximal differences approximately doubled , with the singlet's (chicoric acid) mean integral amounting to $7.1x10^6$ and ranging from $4.2x10^6$-$9.2x10^6$ (standard error, CV=22%) and the mean of the doublet of doublets (chlorogenic acid and/or caffeic acid) being $8.5x10^6$ with a range of $5.2x10^6$-$1.1x10^7$(CV=24.0%) (Table 3.3) Highest concentration differences for chicoric acid were observed between Batch E1.B and E1.A, with relative concentration in the latter being 16% less. Chicoric acid concentration in batch C was 41%

less than batch B. Similar batch concentration differences were found for the doublet of doublets (6.94ppm) representing a mixture of two or more chemical references (chicoric acid and chlorogenic acid or caffeic acid).

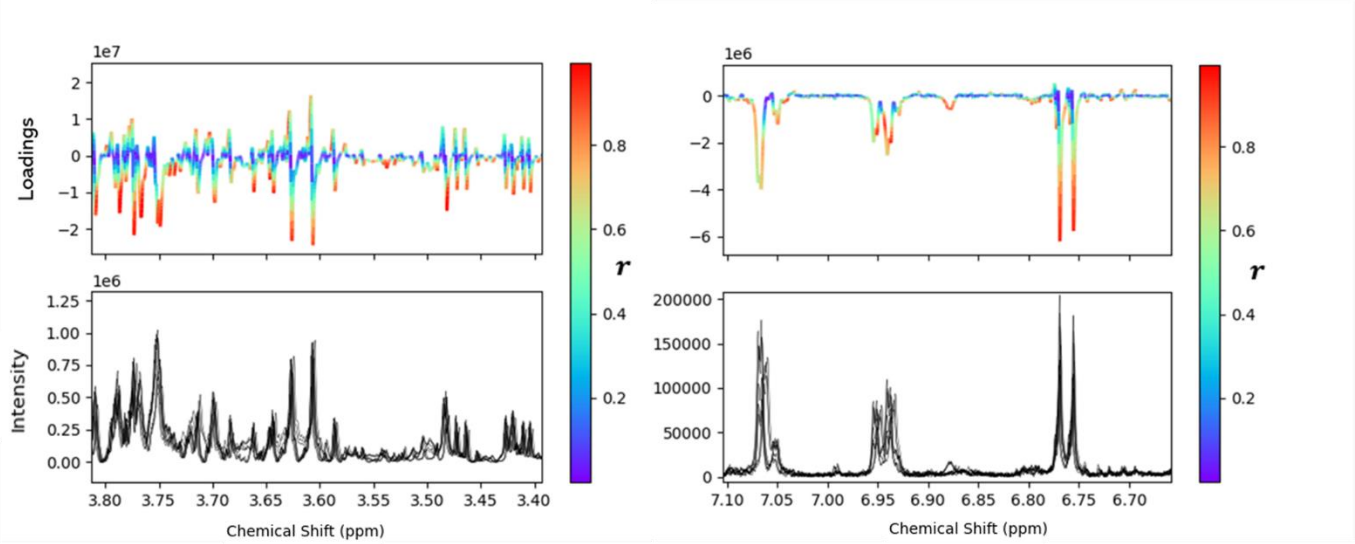


Figure 3. 9 PCA principal component 2 reconstructed loadings of selected chemical shift regions for E1.Top left panel: PCA loadings of chemical shift region 3.4-3.8ppm, Bottom left panel: Product spectra. high model importances showing extensive chemical shift and with peak overlapped. Top Right:PCA loadings for spectral region 6.7-7.1ppm; highlighting three characteristic peaks also observed in reference standards for E.purpurea. Bottom right: Product spectra indicating concentration differences across product batches.

Further examination of the characteristic chemical profile of product E1 included a comparison with listed excipient ingredients: Calcium hydrogen phosphate dihydrate, gelatin, magnesium stearate, purified water, silicon dioxide and sodium lauryl sulfate. All attached salt components will dissociate in solution (i.e., Magnesium stearate: Magnesium dissociates from stearate), and all inorganic ions will not be observed as peaks in NMR spectra. However, inorganic ions contribute to the ionic strength observed in spectra and can lead to peak shift. Products such as gelatin will appear as broad peaks that were not identifiable in the complex mixtures. Due to the peak overlap of the aliphatic region, it was not possible to determine the presence of stearate. Interestingly, product spectra contained

signal patterns that matched to lauryl sulfate (0.882(t), 1.28(s), 1.488(t)) an excipient that was listed on the public summary for the product.

Table 3. 4 Relative integration for peaks relating to active ingredients found in product E2.

Compound	Signal position (ppm)	Relative integration	
		Mean (range)	**Coefficient of Variation (%)**
Chicoric Acid	5.64ppm(d)	$7.1x10^6(4.2x10^6\text{-}9.2x10^6)$	23.7%
Standard mixture (Chicoric acid with chlorogenic or caffeic acid)	6.94ppm(dd)	$8.5x10^6(5.2x10^6\text{-}1.1x10^7)$	24%

STOCSY was performed to identify peaks with structural correlation (d=1.284ppm). However, STOCSY trace did not recognise unique structural patterns due to major chemical shift. This expressed similar results to the PC1 loadings and did not provide greater insight into peaks correlated to the same compound. STOCSY analysis was performed on a peak that had low peak variability between batches in the aromatic region (d=6.77) of E1. Similar to previous findings, there was too much within batch chemical shift to accurately attribute STOCSY profiles to individual compounds, as found in the supplementary information.

Composition and Batch analysis E2

Similar to product E1, product E2 also contains E. purpurea extracts and graphical comparison of E1 product spectra focused on the same reference compounds for product E2: Chicoric acid, chlorogenic acid, caffeic acid. Comparable with previous findings, the reference standard chlorogenic acid and caffeic acid's chemical shift patterns were incompletely found in the product spectra. The spectral comparison between product E1 and E2 showed remarkable similarities, with all NMR signals in E2 spectra being also present in E1 spectra, showing only minor chemical shift differences (figure 3.2.5). Although these

products are marketed under different product names, the present analysis indicated little to no compositional differences between these products. Further integration of the same chicoric acid peak (5.64ppm(s) as product E2 highlighted concentration differences between the two very similar products. E1 had a lower mean of $5.3x10^6$ compared with E2 ($7.1 x10^6$). This aligns with product labels which state that E2 has a higher echinacea content (222.22mg) than E1 (111.11mg). Chicoric acid, chlorogenic acid and caffeic acid have high-intensity peaks found from 5.75-7.65ppm (Table 3.2). When these peaks were compared with product E1, it was difficult to discern correlated peaks due to the large concentration differences between product and standard. Thus, chemical standards and product E1 were min-max scaled for ease of interpretation.

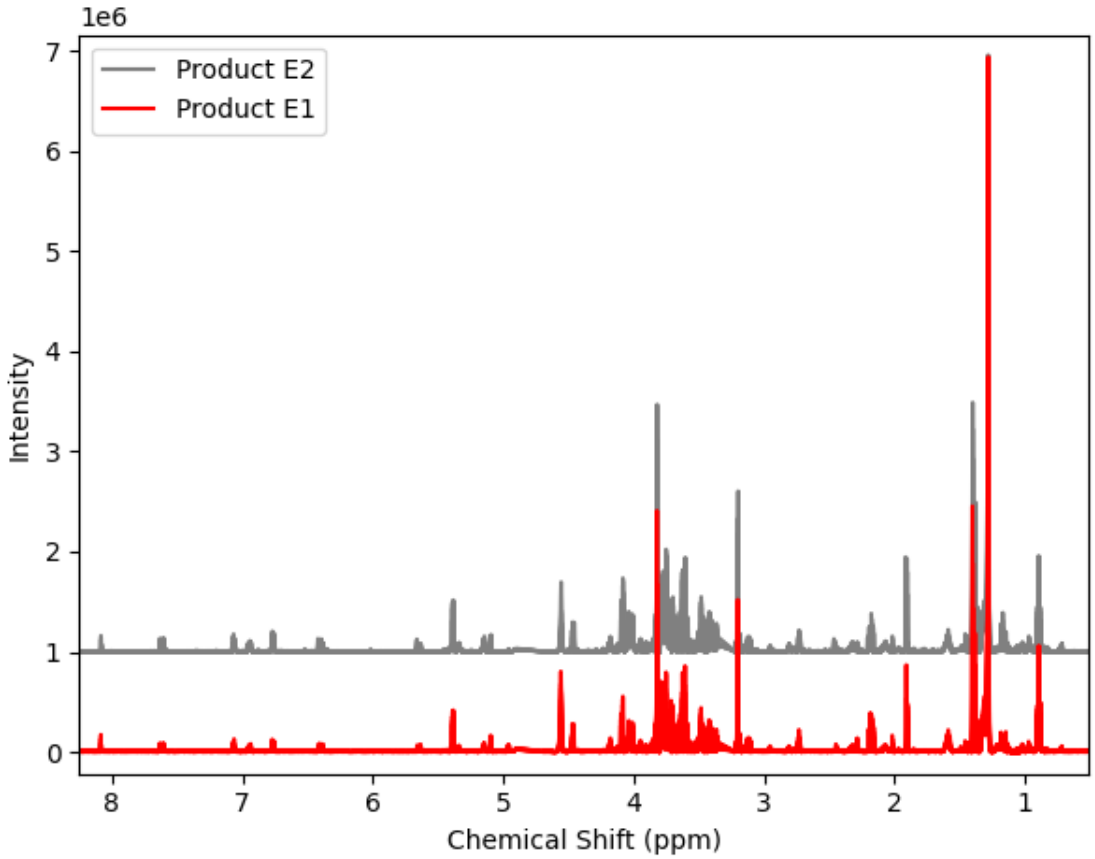


Figure 3. 10 1H Spectral overlay comparing products E1 and E2. This highlights minimal peak differences using a y-axis offset (+1x105 to provide E2) for improved comparison.

This provides clarity of peak shape identification and can be clearly observed in Figure 3.10 where all peaks from 6.4ppm to 7.6ppm are of a similar position and shape. Nonetheless, peak shift was observed in signals with chemical shifts similar to the five peaks relating to chicoric acid, chlorogenic acid and/or caffeic acid. It is likely that these signals relate to the active ingredients listed for *E.purpurea* in E1, since there were distinct minor peaks found in the region from 1.5ppm to 5.0ppm that also matched with signals in the reference spectra and that could not be accounted for by the active ingredients list of the product (Appendix B) . A review of the excipient ingredients from the public summary of produce E1 found gelatin, magnesium stearate, silicon lauryl sulfate, silicon dioxide and calcium hydrogen phosphate dihydrate.

Further signal comparisons with DrugBank database entries indicated sodium lauryl sulfate's presence, as three out of the four characteristic peaks for this compound were recognised in the product spectra. Lauryl sulfate peaks that could not be identified in the product spectra (three singlets at 3.7ppm), were located in chemical shift regions with high peak overlap. Other integrations compared peaks resulting from the polyphenols mixture, such as the doublet of doublets at 6.94ppm. This highlighted significant variation in the relative concentrations found within the products but did not indicate the inter-batch variation. Therefore, it was necessary to examine batch variations. Due to the similarity between product E1 and E2, STOCSY was not performed as it did not provide any insight into the structural relationship between peaks for these products.

Table 3. 5 1H NMR shifts of reference compounds for product E2 and signals observed in spectra

Compound	**Peak used for integration**	**Relative integration**	
		Mean (range)	Coefficient of Variation (%)
Chicoric Acid	5.64ppm(s)	$5.3x10^6$ ($4.4x10^6$-$7.1x10^6$)	16.2%
Chicoric acid with chlorogenic acid/caffeic acid	6.94ppm(dd)	$5.9x10^6$($4.5x10^6$-8.610^6)	22%

PCA was used to identify batch variations, where the scores plot for E2 expressed 25.8% and 20.6% of total variation across the t1 and t2 axis, respectively (Figure 3.11). Examination of t1 axis indicated showed negative scores for both batches E2.A and E2.C and increasing positive score values for batch E2.B. The loadings plot for PC1 indicated chemical shift variations across the entire spectral range, in particular for the replicate from E2.C that was close to the 95% confidence ellipse. The t2 distribution showed clusters for each batch. Batch 21.A replicates clustered the closest to the origin (t2 ~0)—Whilst t2 values for batches E2.A and E2.C have high negative (E2.A) and strong positive (E2.C) values, respectively.

The reconstructed loadings of PC2 showed significant peak variations across the ppm range, particularly for positively correlated variables. Within the loadings plot, the singlet at 4.56ppm (Figure 3.11) is characteristic for batch A, with no major peak shift observed. This indicates concentration differences, in particular for batch A and batch C and highlights the influence of dilution on peak shape. Integration of peaks was performed to quantify relative concentration differences for the same peaks integrated for product E1. These signals are also found to be of high PCA model importance in PC 1 for Batch B. Batch B had the highest average concentration for the two integrated peaks. The most significant difference was observed for the doublet of doublets, where the relative concentration of Batch B was

$7.3x10^6$ (range=$6.0x10^6$- $8.9x10^6$, CV=15.9%). Compared to batch B, the mean intensity for this signal was reduced by 29% and 26% for batch A and C, respectively (Table 3.5). Therefore, clear between batch differences were observed. However, batch E2.B was the most different from the other two, due to concentration differences.

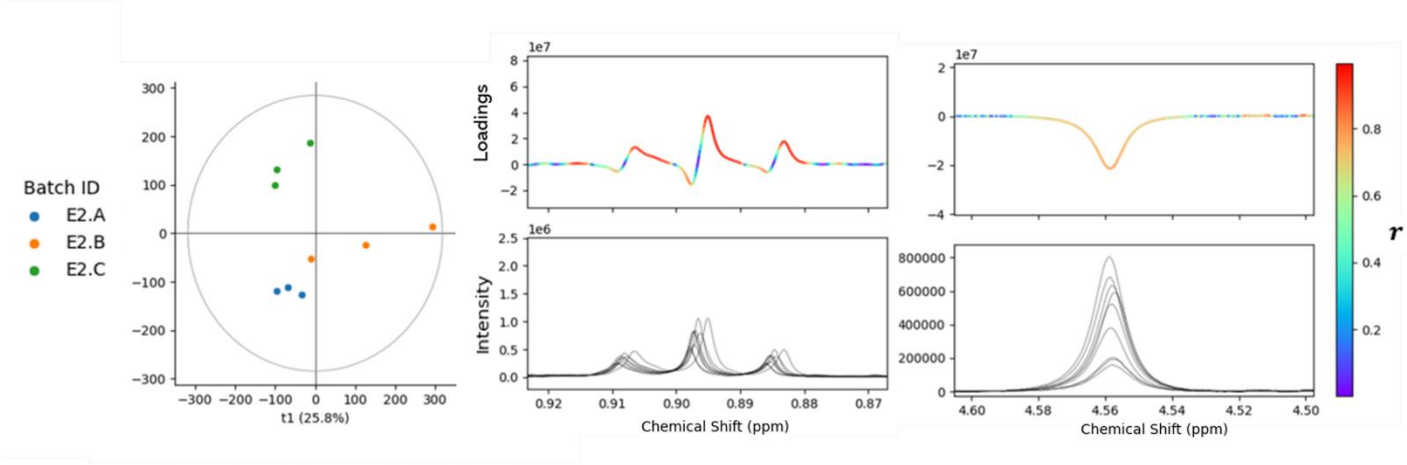


Figure 3. 11 Product E2 principal components analysis highlighting inter-batch discrepancies using scores and reconstructed loadings plots using a two-component model. Left: Scores plot for t1 (r2= 25.8%) and t2 (r2=20.6%). Middle: PC 1 inset highlights how the model's signals of high importance are characterised by chemical shift. Right: PC 2 of singlet (4.56ppm) indicating peak intensity differences potentially related to sample concentration

Vitamin C products

Composition and Batch analysis of product V1

Product V1 was compared with reference spectra of active ingredients listed on the main label (*Allium Sativum* (Alliin), Ascorbate, and *Echinacea Purpurea* (Chicoric acid, chlorogenic acid, and caffeic acid) resulted in no compounds being fully elucidated. Firstly, the active component of Allium Sativum, alliin, was not identified in the product and only shared peak similarities with two minor singlets found at 1.97ppm and 8.06ppm. However, no major characteristic peak patterns were observed. Ascorbate reference also shared one major

peak with product V1 (3.8772ppm(m)) and two minor peaks (3.405ppm(m) and 3.172ppm(m)). Although, no ascorbate was definitively identified. However, it is probable an ascorbate-like substance is used in the manufacturing of V1 due to the peak position similarities of the reference multiplet (3.679ppm) and product spectra doublet at 3.658ppm.

Lastly, *Echinacea Purpurea* reference standards (chicoric acid, chlorogenic acid, caffeic acid) found no indication of *E.Purpurea* in product V1 (Figure 3.12). J-resolved consolidated this in the aromatic chemical shift region. One doublet was found in the J-resolved spectra at 6.37ppm, which was most closely related to chicoric acid reference (6.354 (d)). However, this cannot confirm the presence of *E.Purpurea* and could not be identified in the J-resolved NMR spectra due to peaks being found too close to the noise region (Figure 3.12). Cholecalciferol reference spectra from HMDB was compared with product spectra, and no cholecalciferol signals were found. Therefore, only one comparison of the relative concentration of a major characteristic peak between batches could only be performed on the ascorbic acid-associated multiplet at 3.87ppm. Integration occurred from 3.853-3.904ppm, where the average concentration was $1.1x10^9$ with a minimum of $8.4x10^8$ and maximum of $1.3x10^9$ (CV=11.5%). High range was observed for batch A, which had contained the maximum and minimal peak integral across all samples (mean=$1.1x10^9$,

range= $9.7x10^9$-$1.0x10^9$, CV=2.5%) and $1.2x10^9$ (range= $1.1x10^9$-1.210^9, CV=4.3%), respectively.

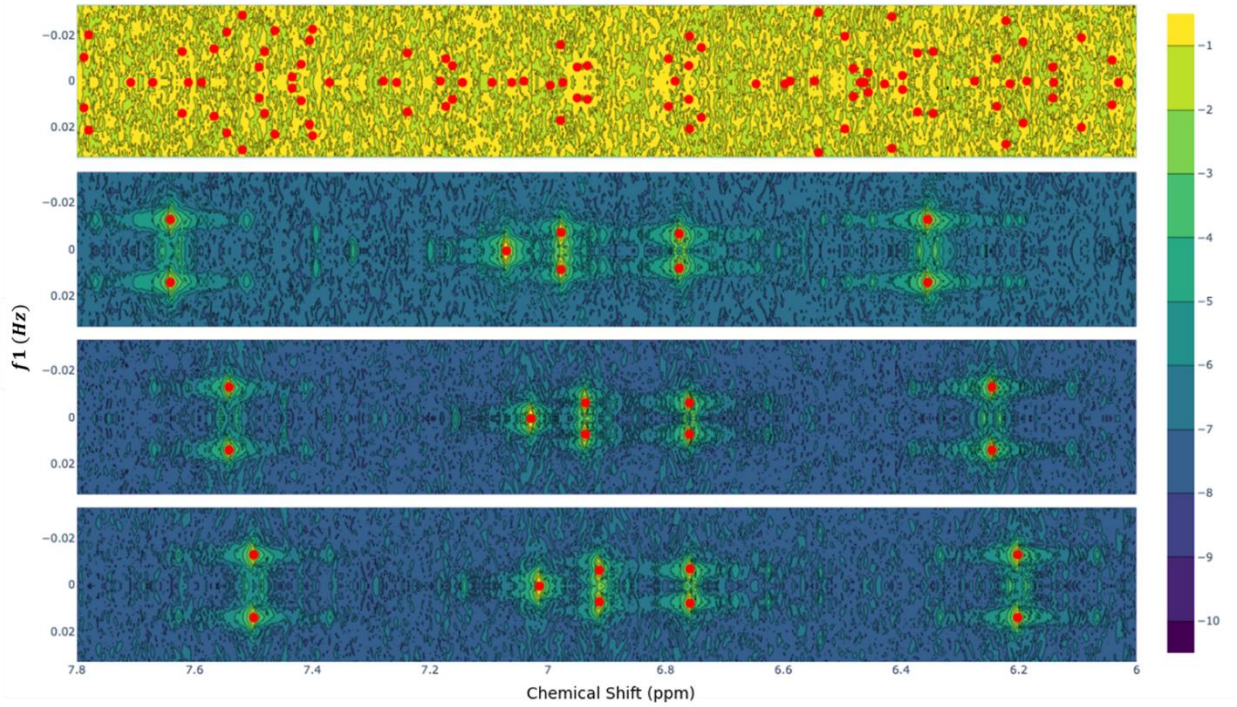


Figure 3. 12 J-resolved spectral comparison of mean-scaled product V1 with E.purpurea biomarkers. Top panel: mean scaled spectrum of V1. Lower panel in order: chicoric acid, chlorogenic acid and caffeic acid in the aromatic region (6-7.8ppm). Red coloured points indicate local peak maxima determined by zero crossing of the first derivative. Product spectra show small peak intensities close to noise signals.

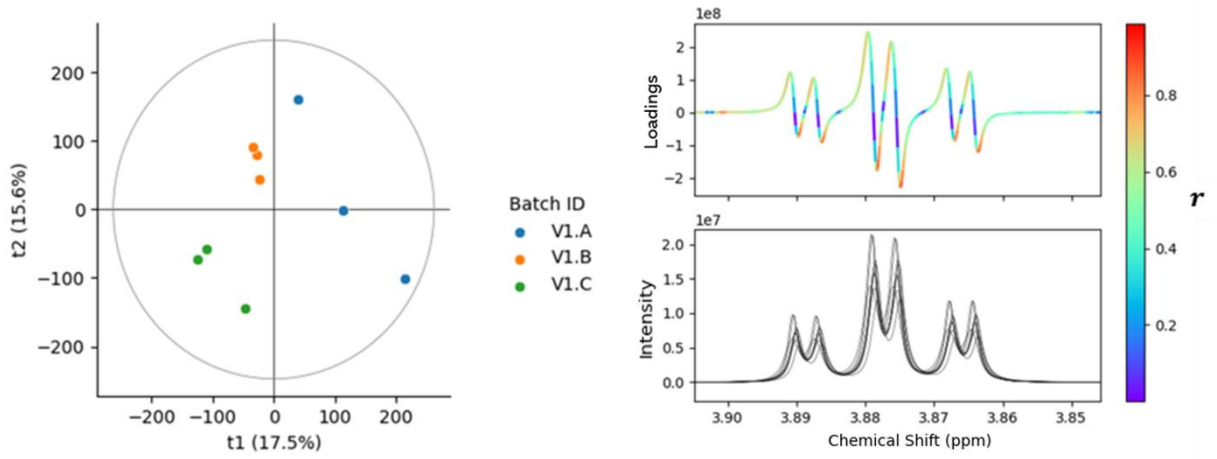


Figure 3. 13 Figure 3.3.2. PCA analysis of product V1, highlighting batch variation. Left: Scores plot for t1 and t2. Right: Comparison of PC 1 loadings with pre-processed spectra of triplet at 3.877ppm, highlighting intensity differences and shift perturbation.

The integral peak disparity for V1.A was highlighted on the t1and t2 axis (r^2=17.5% and r^2=15.6%) (Figure 3.13) where batch A had a positive clustering trend in PC 1 and immense variation in PC 2. Moreover, the scores plot for V1, indicated significant inter-batch variation with minimal clustering along t1 and t2. Interestingly, V1.B and V1.C shared scores similarities, whereas V1.A had one replicate as almost outside the t^2 ellipse, and all samples were positively distributed along t1. Signals with high positive and negative loadings for PC1 were observed over the same spectral region(3.2-4.0ppm). Interestingly, high covariance (r>0.8) was primarily found in small peak regions accounting for broad resonances from macromolecules (3.0-4.5ppm). Interestingly, the ascorbate-like multiplet at 3.87 had high model importance on PC1, which showed minor chemical shift variation as well as concentration differences in product spectra (Figure 3.13). Further comparison between PC1 and PC2 signals with high model importance discriminates between the chemical shift variance found in PC1 and compositional variance that is more accurately inferred via PC2.

By observing the multiplet at 3.87ppm in both PC1 and PC2 loadings, the high covariance that is negatively correlated is due to the chemical shift in PC1, is not shown in PC2. PC2 loadings highlight the collinearity between the multiplet at 3.87ppm and 3.66ppm. To test if these signals are structurally related, STOCSY was performed using the driver peak at 3.664 ppm (Figure 3.14). STOCSY indicated structural correlation where the multiplet at 3.87ppm was highly correlated with the driver peak. At the same time, these peaks did not match with ascorbate reference standard signals(Figure 3.14)

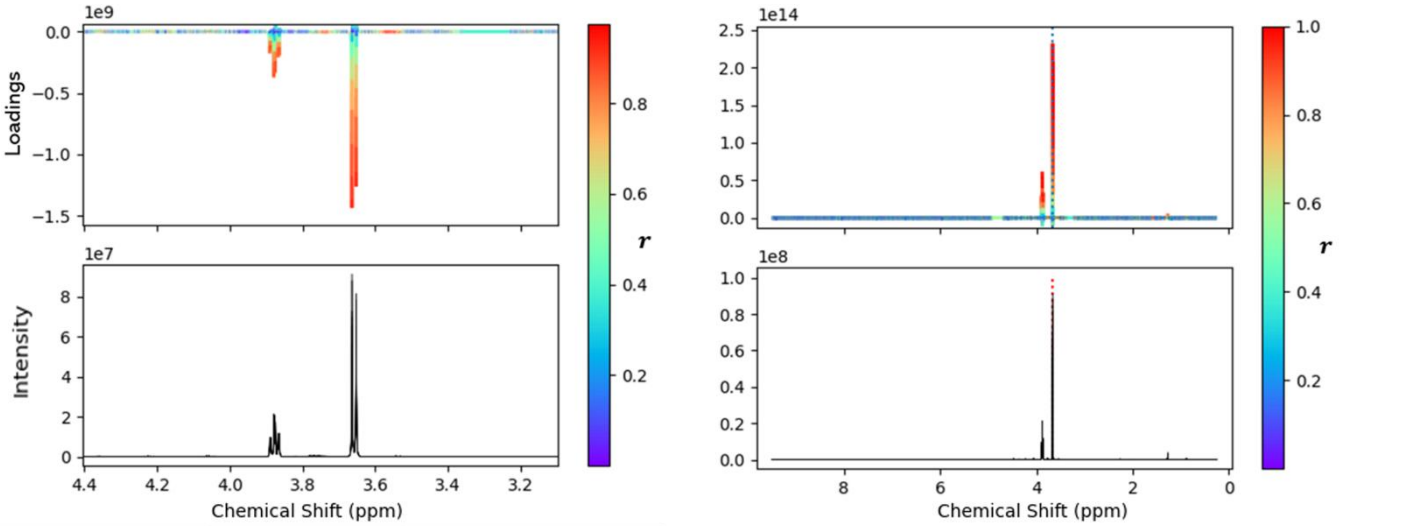


Figure 3. 14 Comparison of PC 2 reconstructed loadings (left) and STOCSY analysis (right) to highlight similar collinearity effects in both models.

Composition and Batch analysis of product V2

Product V2 was the only product advertised as a specific children product. ^{1}H NMR spectra of product V2 were graphically compared with reference spectra of the main active ingredient listed in the product: Ascorbate (ascorbic acid from calcium ascorbate dihydrate, ascorbic acid from sodium ascorbate, ascorbic acid). As there were three different forms of ascorbic used and NMR cannot show dissociative salts, it was impossible to elucidate the various forms of ascorbate in the 1H NMR, resulting in a significant chemical shift from the differing ionic strengths of the included salts. Although the product advertised ascorbate, no complete metabolite identification could be found. The major peaks for V2 were found between 3.5-4.0ppm, with significant peak overlap. Therefore, J-resolved images were used to deconvolve this area (Figure 3.15). However, no prominent peak of ascorbate was

represented in V2, as the triplet of doublets at 3.87ppm presented as a doublet of doublets in product V2, and other major peaks were missing.

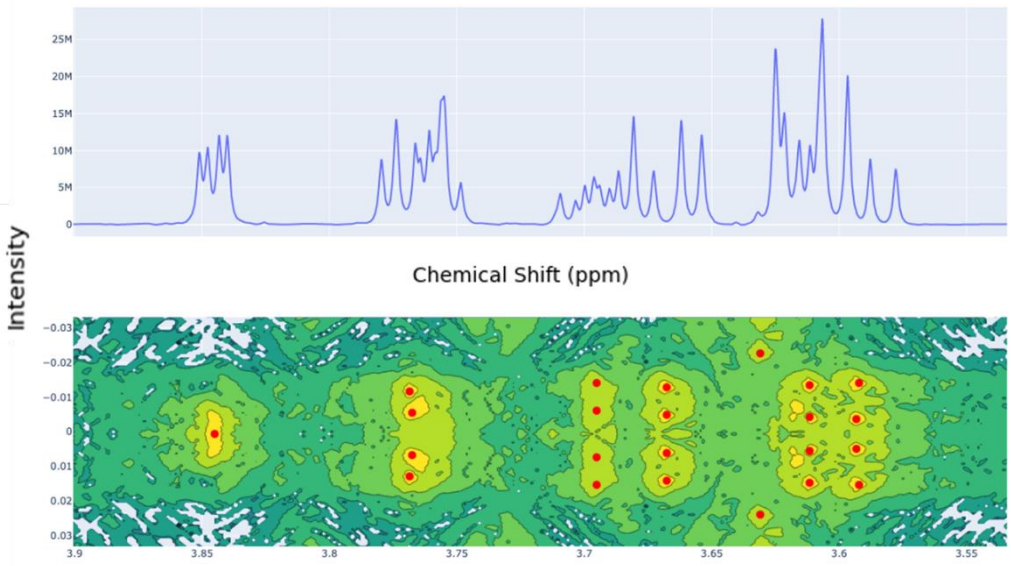


Figure 3. 15 Comparison of 1H NMR spectra of heavily overlapped region (3.55-3.85ppm) of product V2 with J-resolved image. Red dots indicate automated peak picking process. Due to decreased resolution of J-res it was not possible to use this in identification of overlapping products.

Further investigation into the excipient ingredients observed HMDB reference spectra for sucralose, stearic acid, and sorbitol. Stearic acid could be characterised from reference spectra as peaks at 0.88ppm(t), 1.272ppm(s), 2.23ppm(t) are most likely attributed to the excipient ingredient. However, peak shift, dilution effects, and the influence of other compounds on minor signals make definite determination difficult. Similarly, sorbitol has major peaks in the same region as ascorbate and most likely contributes to the major multiplets found between 3.5-4.0ppm. Mainly, similarities were found between the multiplets at 3.76ppm and 3.60ppm. However, peak overlap in this region made identification impossible without further experimentation and acquiring sorbitol reference standard in the same experimental conditions as products.

Moreover, sucralose spectra from HMDB did not share any peak similarities with product V2. Thus, it was not contained in the product. Therefore, no peak integration could be performed as no free-standing peak sufficiently related to reference standards.

Therefore, due to the difficulty of integration, the PCA model for product V2 was generated for similarity analysis. The scores plot for t1 and t2, infer relatively high numbers of total variance at 22.4% and 14.9%, respectively (Figure 3.16).

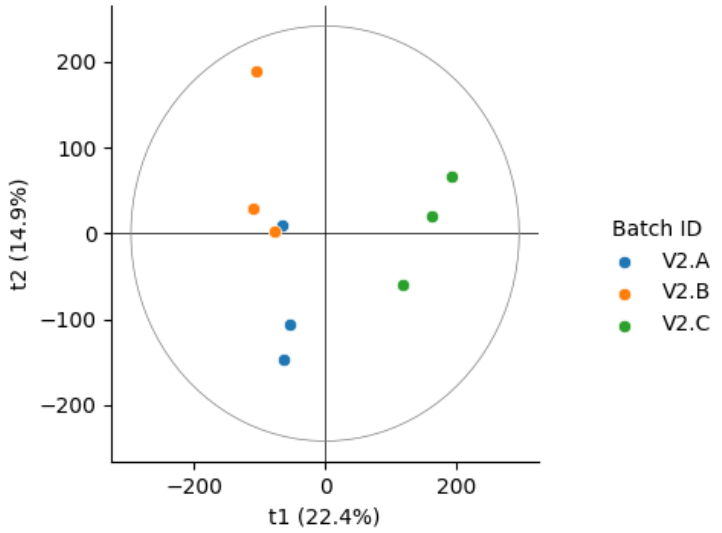


Figure 3. 16 Scores plotting presenting the t1 (r2= 22.4%) and t2 (r2=14.9%) axis of product V2. Products only had a minor tendency to cluster according to both scores axis.

There is a clear split across the t1 axis where batches V2.A and V2.B have negative scores values and relatively similar. Contrarily, the other three replicates of batch V2.C have positive scores values. Moreover, PC1 loadings highlight that the significant positive and negative scores plots for t1 are a result of immense chemical shift (Figure 3.17). The t2 plane had minimal batch clustering for V2.A and V2.B, where two replicates from either batch were very similar according to PCA. V2.C had the biggest batch clustering around the origin.

PC2 loadings, still exhibited some variance caused by shift. However, peak shape differences caused from either different molecules or concentration is exhibited from 4.52-4.58ppm, where some peaks appear as a singlet, while others are doublets. This is a result of dilution and concentration effects on peak shape. The only major peak of interest was from the multiplet at 3.76ppm, although this is related to shift perturbations (Figure 3.17). Due to peak overlap in regions with major peaks, STOCSY analysis received similar results to PC 1 and therefore, did not aid in identification nor batch variation effects.

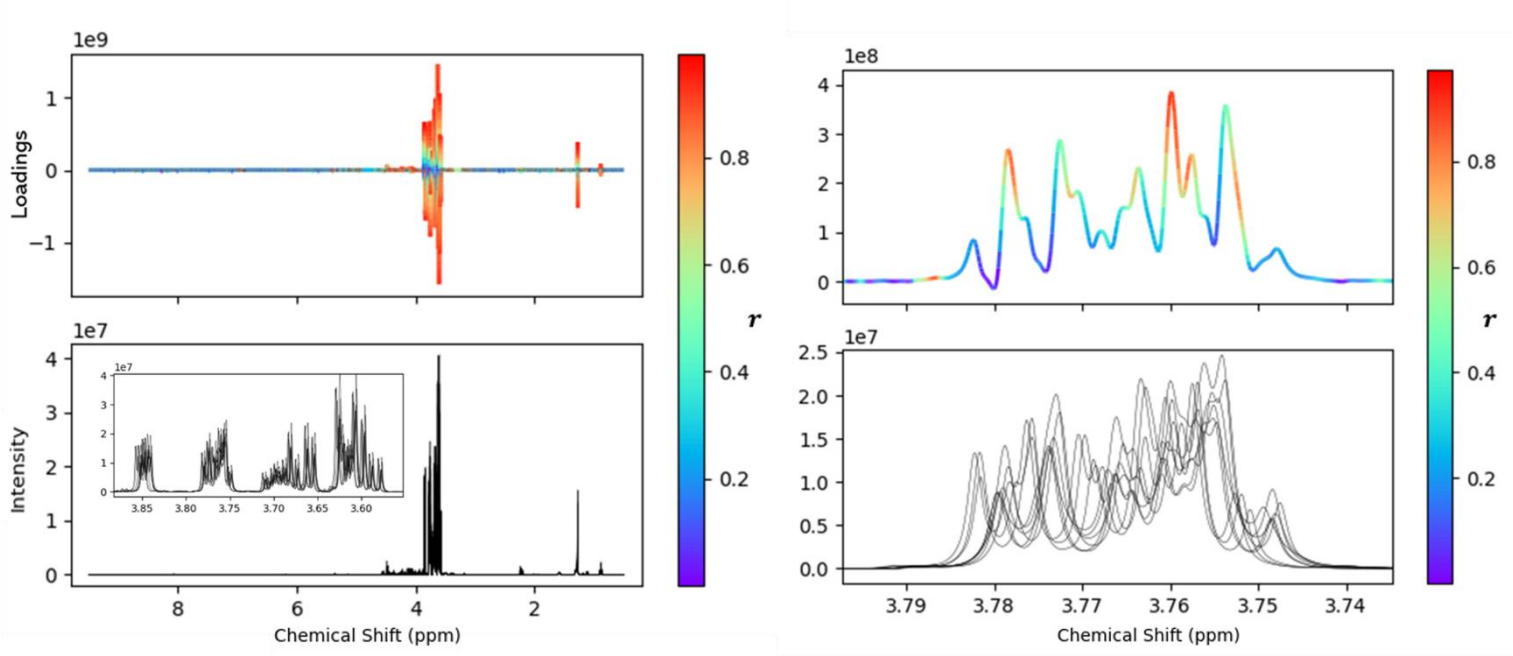


Figure 3. 17 Statistical reconstructions of the loadings for product V2. Left: PC 1 indicating immense chemical shift perturbations that effect the model. Right: PC 2 of spectral range 3.74-3.80ppm, presenting the effect that dilution and chemical shift have on peak shape and consistency.

Composition and Batch analysis Product V3

The vitamin C product V3 declared the following active ingredients on the product label: Ascorbic acid, cholecalciferol and *E.purpurea* (chicoric acid, chlorogenic acid, and

caffeic acid). Therefore, reference standards were compared with ^{1}H NMR spectra of three different batches from V3 to develop a characteristic profile of its contents. The two major peaks found in the product spectra are likely to relate to ascorbate-like compounds, these are a triplet of doublets (3.87ppm) that correlated with another triplet of doublets (3.85ppm) and these were also found in the ascorbate reference spectrum. Similar to previous observations, the major doublet at 3.65ppm does not have the same peak shape nor ratio to other signals found in the ascorbate reference standard spectrum. Similarly, the major ascorbate doublet (4.32ppm) is similar to the minor doublet found at 4.36ppm in the product. Due to that not all ascorbate reference signals were found in product spectra, no definite confirmation of the presence of ascorbate in product samples was possible.

Comparison between the three active compounds of *E.purpurea* (chichoric acid, chlorogenic acid, caffeic acid) was based on J-resolved spectroscopy due to the peak shift in the aromatic region of the product spectra (Figure 3.18). Due to the low concentration of reference standards found in the product, the automated peak picking function did not provide insight into *E.purpurea* specific compounds. Thus a mixture of ^{1}H NMR spectra and 2D J-resolved spectra were required for identification (Figure 3.19). Three minor peaks could be attributed to chicoric acid (7.05ppm(d), 6.76ppm(d) and 6.37ppm(d)). However, integration for peaks related to *E.purpurea* showed close their proximity to noise, and it is concluded that these compounds are not or only in minor concentrations present in product V3 (Figure 3.19).

Excipient ingredients likely to present in NMR spectra include crospovidone, stearate, alpha-tocopherol, and sucrose, due to their chemical structure and ability to be seen in NMR spectroscopy. ^{1}H NMR signals of these compounds were extracted from HMDB

reference spectra. No crospovidone peaks were identified, nor sucrose or alpha-tocopherol. However, sterate was observed in the three characteristic peaks at 0.88ppm(t), 1.27ppm(s), and 2.25ppm (t).

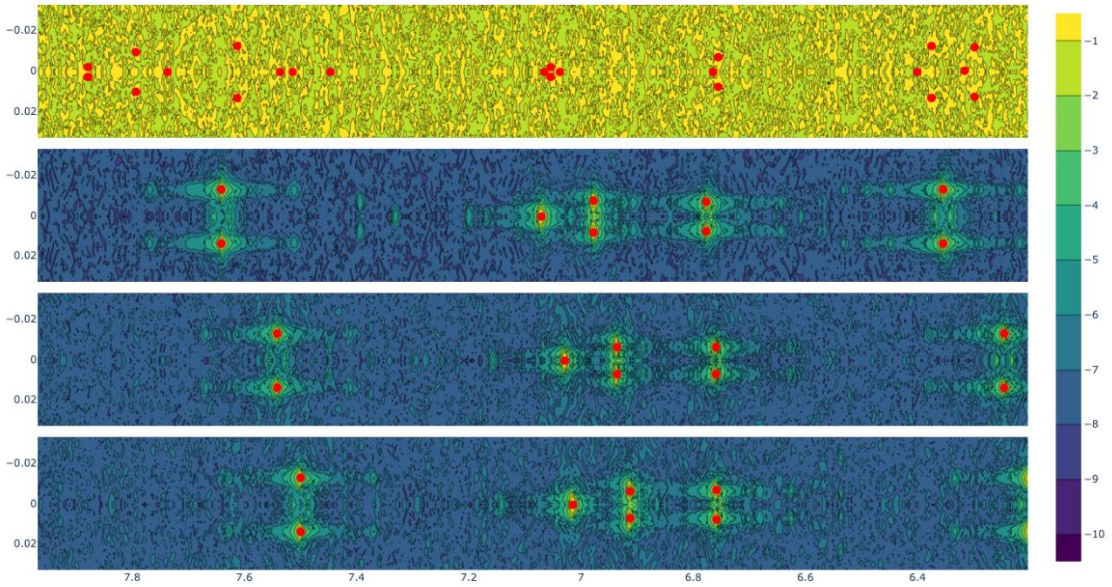


Figure 3. 18 2D J-resolved spectral comparison of product V3 (top) with three active compounds found in E.purpurea (chicoric acid, chlorogenic acid, caffeic acid (bottom))

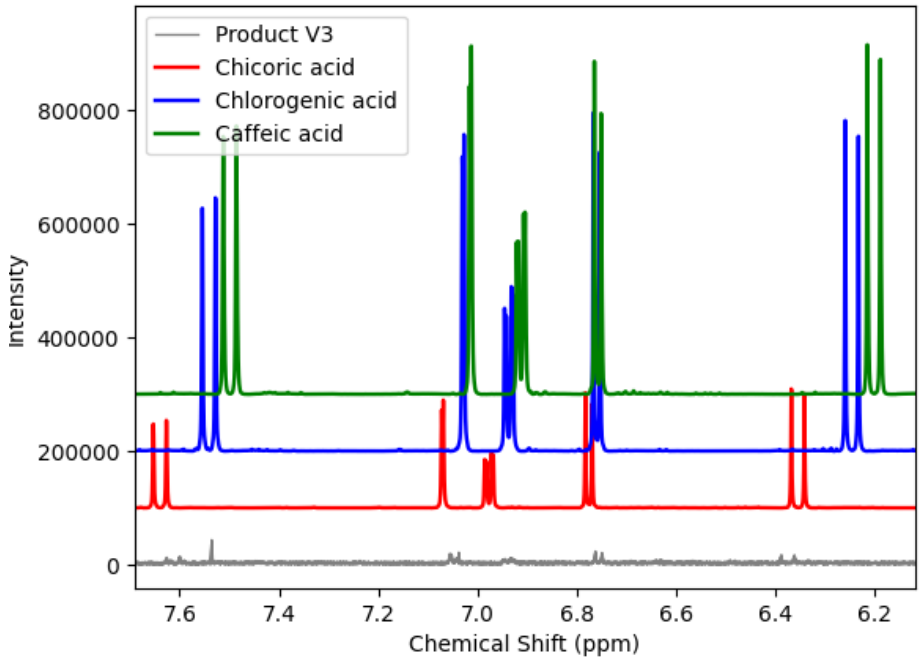


Figure 3. 19 1H NMR spectral overlay of product V3 with three active components of E.purpurea using a $1x10^5$ offset for each reference spectra to indicate the complexity of peak pattern matching for similar compounds.

The only peak where integration was possible was the triplet of doublets similar to the ascorbate reference signal at 3.87ppm (integrated from 3.842-3.913ppm). Nonetheless, Inter-batch variation for this signal was observed with V3.B showing the lowest relative concentration and highest coefficient of variation(mean =3.8x10^8, range 2.9 x10^8-4.6 x10^8, CV=18.9%). Batch V3.C had the highest average integration for ascorbate-like multiplet (mean=5.2x10^8, range=5.0 x10^8-5.4 x10^8, CV=2.75%), and batch V3.A had an integration mean of 4.9x10^8 (range= 4.2 x10^8-5.5 x10^8, CV=10%).

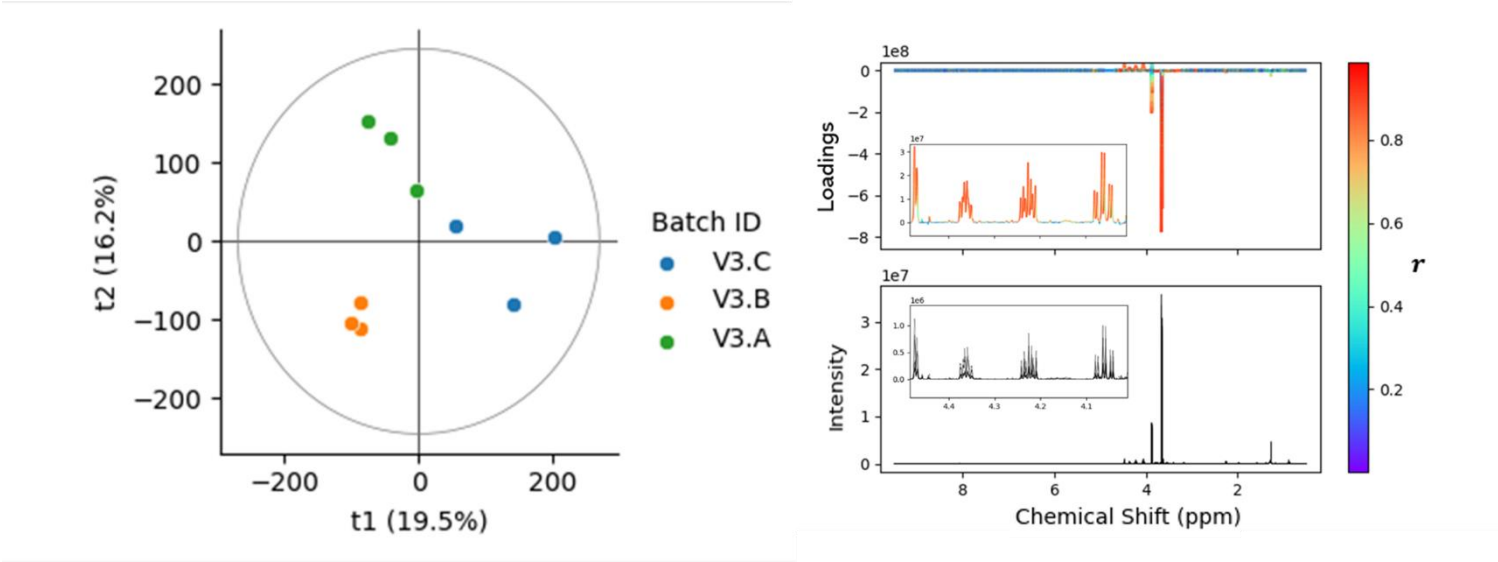


Figure 3. 20 PCA model of product V3. Left: Scores plot of t1 (r2= 19.5%) and t2 (r2=16.2%) axis. Right: PC 1 reconstructed loadings highlighting collinearity between two ascorbate-like peaks and inset shows four minor peaks likely relating to the same compound.

Further batch variation effects were explored using PCA analysis using the first two principal components. The PC1 and PC2 axis highlighted 19.5% and 16.2% of the total variance, respectively (Figure 3.20). There was batch clustering observed on the PC1 axis, with the least tendency to cluster being observed for V3.B, although all scores were positive. PC 1 loadings highlighted the high model importance of two major ascorbate-like peaks and high collinearity between four minor peaks (4.061ppm, 4.224ppm, 4.362ppm, and 4.473ppm). Analysis on the t2 axis, indicated similar clustering trends to t1 (Figure 3.20).

However, batch V3. A had positive scores values, and V3.C had the most negative scores values.

Nonetheless, a comparison between the reconstructed loadings of PC 2 and pre-processed spectra highlights that the main cause of variance is the chemical shift of the two major ascorbate-like peaks (Figure 3.21). The peaks with the highest model importance in PC 1 were used to select STOCSY driver peaks. Interestingly, STOCSY provided similar results to PC 1 for collinearity effects of peaks related to driver peak 4.475ppm (Figure 3.21). This showed that all four peaks are likely related to the same compound.

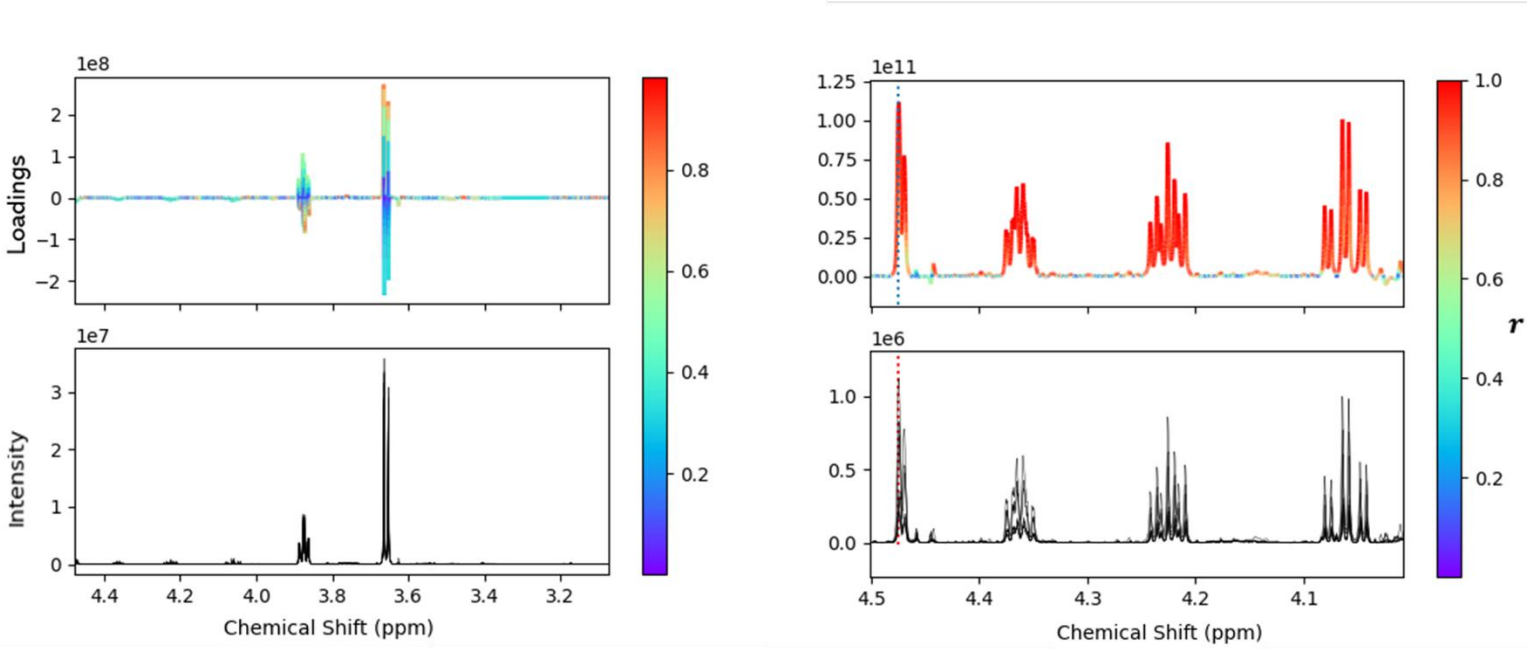


Figure 3. 21 Statistical analysis of product V3. Left: PC 2 model of two major peaks that are important to the model due to chemical shift perturbations. Right: STOCSY analysis with driver peak at 4.475ppm; highlighting similar collinearity effects to PC 1.

Compositional analysis of product S

Product S was the only immune supplement analysed to contain the herbal ingredient *Sambucus Nigra*. ^{1}H NMR spectra of product S were initially compared with ascorbate Vitamin C) and kuromanin reference standards for characterisation of the active ingredient listed on the product label: Vitamin C and *Sambucus Nigra* (reference chemical kuromanin as biomarker of herbal product).

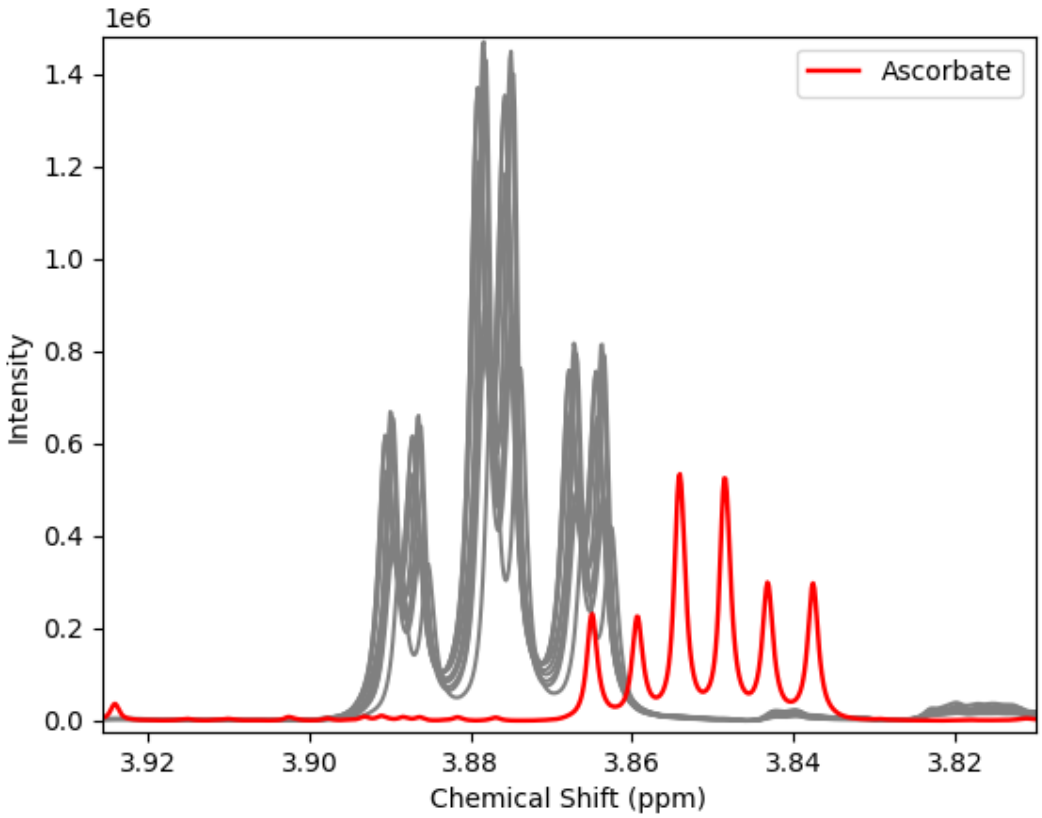


Figure 3. 22 Spectral comparison of scaled ascorbate reference spectra with product S, highlighting the peak similarity of the one major peak at 3.87ppm.

Similar to the comparison of product E2, the full signal traces of the reference spectra could not be identified in any of the samples of product S. High-intensity signals of kuromanin (Appendix B) were not present in the product spectra. Particularly in the region of 3.4-4.0ppm, where the product has immense peak overlap. However, low-intensity peaks in the aromatic region of kuromanin reference standard were identified in the product

spectra: (6.48ppm(s), 6.53ppm(s), 6.70ppm(m)). Comparison of ascorbate reference signals with product S signals highlighted one major signal that was common to both sample types, located at (3.87ppm(m), and two minor multiplets located at 3.17ppm and 3.40ppm (Figure 3.22).

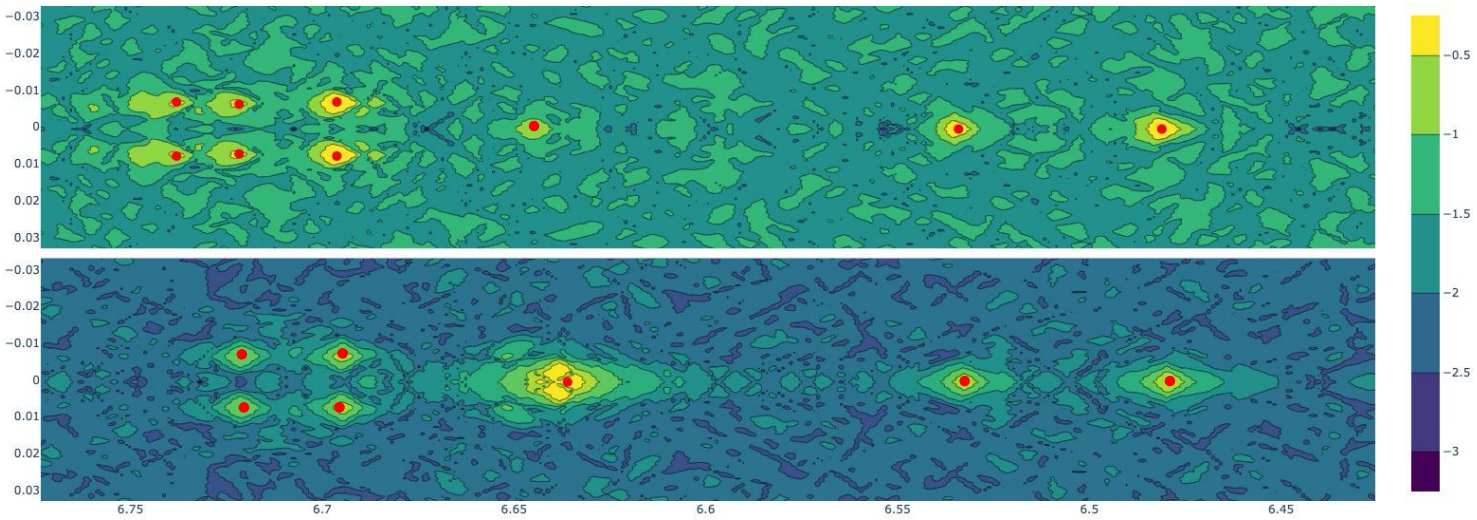


Figure 3. 23 2D J-resolved spectra comparison of chemical shift region 6.45-6.8ppm of product S (top) and kuromanin reference standard (bottom), indicating matching minor peaks. Red coloured points indicate local peak maxima determined by zero crossing of the first derivative.

These findings were confirmed using J-resolved spectroscopic data (Figure 3.23). Relative concentration was calculated from one characteristic peak for each reference standard. Within batch, concentration differences were found significant for both identified products as the range for kuromanin minor peaks was very large with a high standard error (table 3.4.1). In the case of kuromanin, batch S.A had the lowest concentration and was 22% less that batch S.C. Contrastingly, a trend was observed for the ascorbate peak at 3.876ppm, where the highest concentration was $9.7 \times 10^7$for S.A.3 and lowest (4.9×10^7)for S.C.3. Nonetheless, the dispersion of concentrations remained between batches and minimal within batch discrepancies were noted (Table 3.6).

Table 3. 6 Comparison of relative integration for minor free-standing peaks of kuromanin between batches of product S.

Signal position of kuromanin	Batch	Mean (range)	Coefficient of variation (%)
6.48ppm	**S.A**	$1.9x10^6$($1.8x10^6$-$2.0x10^6$)	4.75%
	S.B	$2.2x10^6$($1.4x10^6$-$3.1x10^6$)	30.9%
	S.C	$2.4x10^6$($2.0x10^6$-$2.8x10^6$)	12.63%
6.53ppm	**S.A**	$1.6x10^6$($1.4x10^6$-$1.8x10^6$)	9.37%
	S.B	$1.8x10^6$($1.2x10^6$-$2.7x10^6$)	33.8%
	S.C	$2.1x10^6$($1.8x10^6$-$2.4x10^6$)	9.68%

The PCA model for product S produced a scores plot with PC1 and PC2 explaining 24.3% and 15.9% of the total spectral data variation, respectively. A clear batch clustering trend was visible; however, batch S.A had the greatest clustering tendency on PC1 (Figure 3.24). Contrarily, batch S.B showed immense variation in t1 and t2, with one replicates being outside the Hotelling's T^2 ellipse. Signals of high model importance on PC1 included a broad signal at 1.27ppm, representing an overlap of multiple signals/broad signals that produced considerable peak shape variations across product spectra.

Other contributing signals for PC1 include smaller broad peaks with chemical shift of 0.88(t), 1,57ppm(t)and 2.25ppm(t) . Signals of high importance in the PC 2 loadings include a triplet of doublets (3.87ppm) and minor broad peaks of multiplets at 4.06ppm, 4.22ppm and 3.36ppm. Similar to product V3, this peak pattern could not be identified as a particular compound using HMDB databases. Interestingly, the positive covariance with high correlation values found, were primarily a result of dilution factors affecting peak shape, as

presented in figure 3.7.2. The major peaks of high importance in PCA model, were derived from peaks that aren't identified in the aliphatic chemical shift region for batch C and the outlier of batch S.B (0.8822ppm(t),1.273ppm(s), 2.2563ppm(t), 5.084ppm(d)).

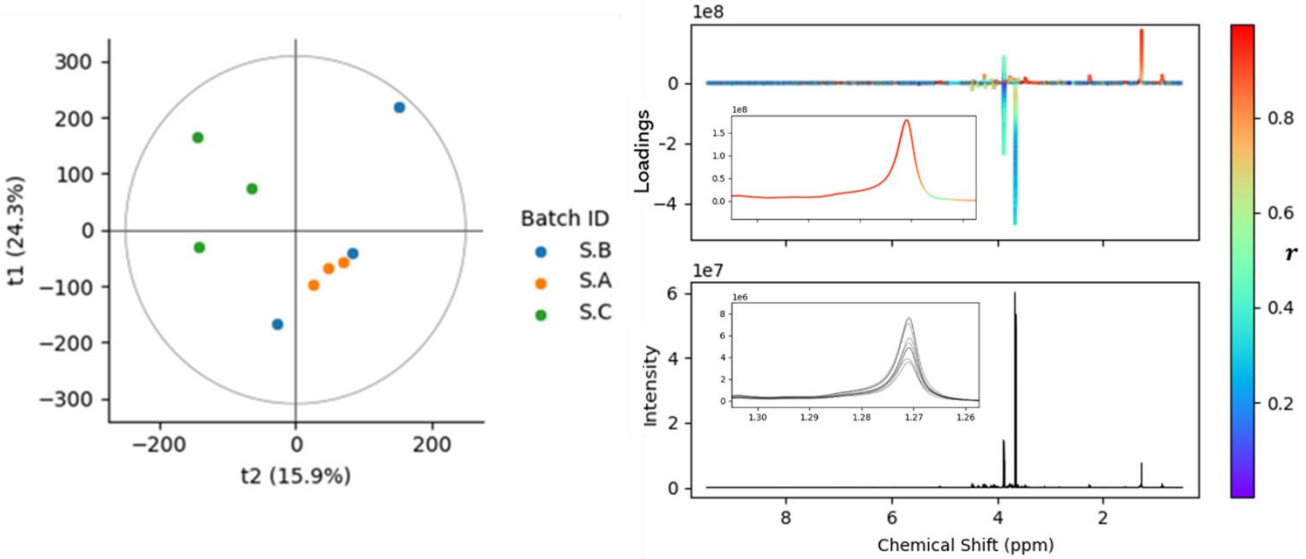


Figure 3. 24 Figure 3.6.3 PCA model for product S. Left: Scores plot for t1 and t2 highlighting minimal batch clustering. Right: PC 1 reconstructed loadings primarily indicating between-batch chemical shift and peak shape dilution.

STOCSY analysis (d=1.2714) indicated that peaks at positions 0.88ppm(t), 1.27ppm (s), 2.25ppm (t), and 5.08ppm(d) are correlated (Figure 3.25). However, the chemical shift database search did not provide any conclusive metabolite ID matches. Moreover, the t2 axis is highly variable as batch S.C disperses positively. The reconstructed loadings spectra for PC2 were influenced by both the negative and positive t2 values. Nonetheless, these were also a mixture of chemical shift (4.12ppm) and intensity effects on peak shape (4.06ppm). Furthermore, STOCSY performed with d=4.4745ppm, highlighted structurally related signals at 4.063ppm(m), 4.228ppm(m), 4.359ppm(m), 4.459ppm(m). However, these signal patterns did not match any of the acquired pure compound spectra, nor were these coherent with online database annotations (Figure 3.25).

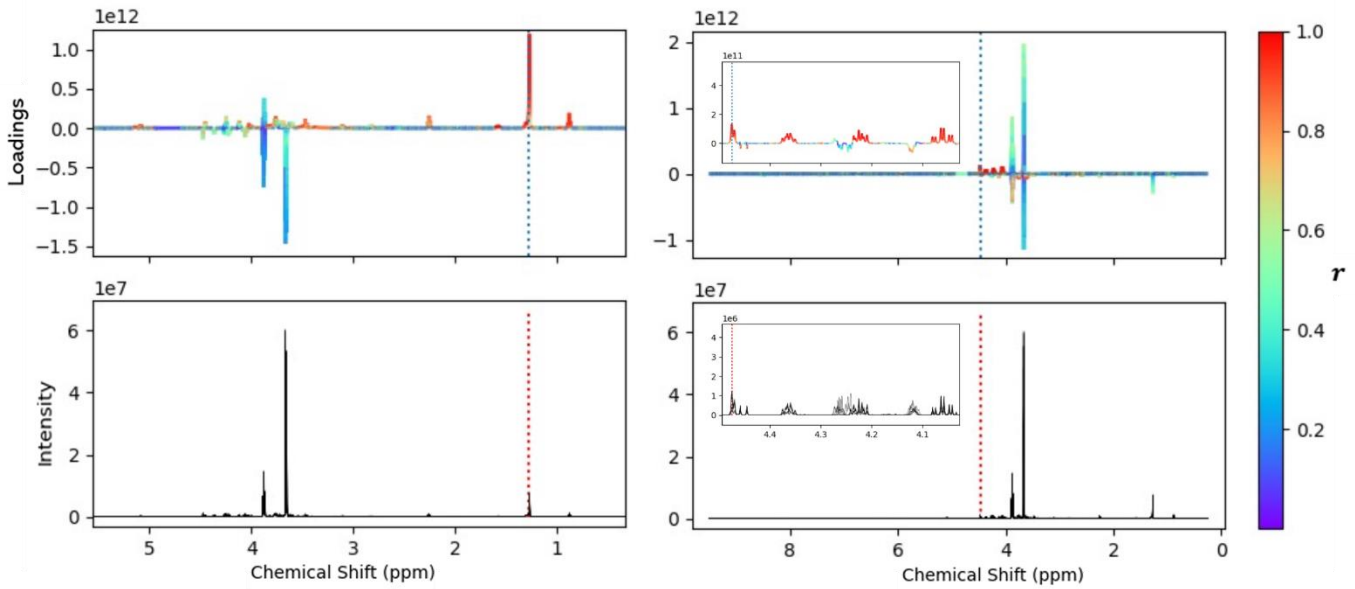


Figure 3.25 STOCSY analysis on product S using two different driver peaks (d= 1.2414 and d=4.4746ppm) to identify peaks that contribute to the same compounds. These collinearity effects were compared with databases and produced no conclusive results.

Discussion

Chemical analysis

All immune supplements investigated in this analysis showed comparable spectral signals and patterns within batches of products, as determined by spectral overlay plots. Significant peak shift variations were observed across all spectra. These were most likely related to pH variations and ionic strength as discussed below. Further analysis focused on the identification of signals in product spectra by means of comparative analysis with reference spectra (active ingredients). This was done graphically. Conclusive identification of active ingredient *E.purpurea,* using chicoric acid biomarker, was found in E1 and E2. Where the complete spectral peak patterns were identified. Nonetheless, challenges in identification of reference compound signals in product spectra included significant peak shift, peak overlap, and similarity of signal traces related to different chemical reference standards, such as chlorogenic acid and caffeic acid.

Additionally, batch disparity of relative integrals of biomarker compounds was present for all products, as assessed by relative concentrations. The Food and Drug Administration (USA) recommends a threshold in coefficient of variation below 10% for drug bioequivalence and 20% for food label bioequivalence(154). For the free-standing peaks that were integrated it was found that all products had batch relative concentration differences CV>10%. This was the lowest in Product S (11.3%) and highest in product A curcumin relative concentration of 40.2%. Hence, no product fulfilled the drug bioequivalence requirements set out by the FDA. However, within batch CV was as low as 2.3% in product A curcumin peaks and as high as 30.3% in batch S.B ascorbate-like peak. High within batch coefficient of variation is likely to be related to the peak intensity of chosen signal to calculate CV values where higher peak magnitudes show high variances and vice versa: this data characteristic is commonly referred to as heteroscedasticity. These figures are based on relative concentrations, and it would be desirable to have absolute quantification based on pill or product unit. Therefore, further experimentation would need to occur, prior to absolute quantification of products and is discussed in detail in the chemometric analysis section.

Compositional identification relied on NMR spectra of reference compounds being graphically compared to signals in spectra acquired from herbal products. The comparison was primarily performed in the aromatic spectral region (~6.5-8.0 ppm), where NMR signals of plant products are commonly observed due to the presence of polyphenols (92). Due to the complex nature of supplements, NMR signal identification was limited by the crowding and peak overlap, particularly in regions 3.0-4.0ppm, and alongside peak shift variations across the entire spectral range(155). This peak shift was likely to be a result of different ionic strengths in the samples (156), introduced inorganic ions such as Zn^{2+}, Ca^{2+}, or Mg^{2+}, which are common constituents of immune supplements and purported health benefits such as zinc

with known antiviral properties (157). These introduced major systematic shift patterns that negatively impacted on compound identification as well as, the evaluation of quantitative differences across batches and is further described in chemometric analysis.

Moreover, NMR spectroscopy is less sensitive and has a high limit of detection when compared with MS (158). NMR is dilution sensitive meaning that products with low concentrations will produce spectral signals which can differ in chemical shift as well as peak pattern signals(159). For instance, a triplet appearing in a high concentration sample may present as a singlet in a very low concentration sample (Appendix 7.B). This may promote spurious correlations, notably when experimental details differ in solvent concentration and the number of scans. Immune supplements are heavily diluted substances with some active ingredients reported in quantities as low as 0.5 mg/pill. Optimally the experimental set up should account for this, such as increasing the number of scans for ^{1}H NMR from 64 to 256 scans for weakly concentrated samples (39). However, due to the aim of this analysis in considering batch variability, all experiments would need to be run under the same experimental conditions increasing run time significantly, as NMR signal to noise (which is a measure of sensitivity) and number of scans have a square root relationship. For example, to double the signal to noise, the number of scans must increase by a factor of four, making this method less high-throughput, thus less applicable for purity assessment in regulatory environments.

As previously stated, there was significant peak overlap, particularly in regions between 3.0 and 4.5 ppm, where major peaks were most commonly found within complex mixtures. This made the identification and integration of compounds resonating in this region (glucose, sucrose, sorbitol, ascorbate, and other carbon-based excipient ingredients) difficult

due to the lack of free-standing peaks, as seen in Products V3 and S. In these cases, J-resolved spectroscopy did provide some benefit in deconvolving complex signals into more easily observable scalar couplings. This did not wholly aid in identification due to the lack of sensitivity as a result of acquisition performed with low number of scans (ns=4)(160).

Interestingly, all products and some reference standards (ascorbate, curcumin, etc) contained a peak group containing at least three signals at 1.28 ppm. This peak group is likely related to a methyl (-CH2-) alkyl functional group. However, these signals could not be attributed to any single compound and is likely a result of the same electrochemical environment in structurally related compounds. Thus, it was not possible to quantify nor reliably use this peak as a driver in STOCSY analysis due to the compositional nature of the signal group. This major signal group showed high correlation with both active and excipient ingredients (lauryl sulfate and magnesium stearate) and was not further characterised. In order to simplify this significant peak overlap, future work should consider sample processing methods from the field of separation science such as chromatographic techniques.

Moreover, the signal traces in products advertised to contain ascorbate as an active ingredient were very similar to signals observed by the reference standard; however these traces did not match completely. Thus, Products S, V1, V2 and V3 contained compounds that are structurally related to ascorbate such as precursor molecules or stereoisomers (the replacement of the multiplet at 3.73 ppm with a doublet and exclusion of doublet at 4.66 ppm). Further investigation into the commercial production of vitamin C highlighted numerous precursor compounds: sorbitol, L-sorbose, L-sorbosone, 2-keto-L-gluonic acid, and L-ascorbic acid (88). However, database searching did not clarify which form of vitamin C was

present in the supplements, even in cases where sorbitol was listed as an excipient ingredient (Product V2).

In most cases, peaks relating to chemical reference standards declared as active or excipient ingredients by the manufacturer were not found in product spectra. This could result from compositional changes introduced during harvest, manufacturing processes, or supplement preparation; such as soaking time, temperature(161), and degradation during storage(2). Gimenez and colleagues studied the potential degradation of ascorbic acid in dietary supplements and found a mean loss of 16% throughout 25 samples. They attributed this degradation to pre-processing conditions and light sensitivity(162). Additionally, the concentration of non-reducing sugars, such as sucrose (commonly used as an excipient ingredient) directly relates to the rate of degradation and oxidation of ascorbic acid (163). Hence, more research into the interactions between excipient and active ingredients is required to better understand the degradation and bioavailability of purported therapeutic compounds derived from immune supplements.

The selected biomarkers further challenged the characteristic profiling of active ingredients. Chicoric acid(66), curcumin (29), kuromanin (30) and andrographolide (164) are well understood characteristic compounds of herbs *E. purpurea*, *C. longa*, *S. nigra* and *A. paniculata,* respectively. In contrast, chlorogenic and caffeic acids are not exclusively found in many *E.purpurea* and could originate from different herbs (73). Spectra of E*chinacea* product*s (E1* and *E2)* showed low-intensity peaks consistent with chicoric acid, chlorogenic acid, and/or caffeic acid. Due to ^{1}H NMR signal similarity of reference chemicals for chlorogenic acid and caffeic acid (Appendix B), the exact product composition relating to these compounds could not be determined. Chlorogenic acid contains an ester bond of caffeic

acid and quinic acid. The quinic acid protons give rise to the multiplet at 2.18 ppm and doublet of doublets at 3.70 ppm in the chlorogenic acid ^{1}H NMR spectrum. These are the only measurable differences between caffeic and chlorogenic acids in the complex herbal supplement mixtures, and due to the high concentrations of excipient ingredients in this chemical shift region, it was impossible to derive signals specifically relating to the quinic acid ester of chlorogenic acid. Thus, the low concentration of active ingredients such as chicoric, chlorogenic and caffeic acids, oleuropein, kuromanin and alliin influenced spectral output via dilution sensitivity. This was clearly observed in Product V1, where no peaks for *E. purpurea* could be observed due to the low signal to noise ratio.

Furthermore, there is debate over the active compounds of *Allium sativum.* Since 1999, researchers have recognised chemical differences between fresh garlic and garlic powder, most likely related to the enzymatic activity of alliinase and drying processes (165). Allinase converts alliin to allicin when broken down, meaning that allicin could be a more appropriate biomarker for identifying *A. sativum* when found in heavily processed herbal supplements (166). However this has not been investigated in complex mixtures nor using proton NMR. Future experiments should consider different standards to identify *A.sativum* in product samples.

No oleuropein was found in Product A. This is consistent with previous examination into oleuropein content in 8 herbal products using LC-MS. They found that only one product contained the correct quantity of oleuropein, as stated on the product label (56). While kuromanin is a well-known biomarker for *Sambucus nigra,* previous investigation into the loss of polyphenols such as kuromanin have highlighted significant photodegradation and instability in anthocyanin content in elderberry juice (147). Therefore, regardless of the initial

inclusion of herbal and inorganic material in immune supplements, it is likely that active ingredients underwent some degradation, even though all were tested before their expiry dates. Further testing of the phytostability of these products is required to understand the potential changes that occur during manufacturing and shelf-life.

Overall, compositional investigation relied on manual processes combining product comparison with active ingredient reference standards and database searching for excipient ingredients or unassigned peaks. Database comparison was limited as ^{1}H NMR spectra of natural compounds are not always available or have been derived by mathematical simulations (155, 167), which do not always produce accurate or complete NMR signal patterns. In the case of ascorbate, ^{1}H NMR spectra listed on HMDB and NP-MRD were dissimilar to the spectrum of the reference standards acquired for this study (Appendix B) or found in the literature (168). Additionally, incorrect identification of natural compounds is common (169-171) due to the similarity of carbon skeletons, as indicated by Burns and Reynolds (172). Thus, there is a major need for validating database reference spectra. For instance by using cross-institutional comparison that enables greater rigour in spectral quality and consistency that is uploaded to a central repository. Increased data quality would improve meaningful comparison of reference standard equivalence when identifying natural products, particularly if free induction decay (FID), and details regarding sample preparation and pre-processing parameters were uploaded supplementary to spectral information.

The interpretation of 2D-NMR experiments that are often used in developing an analytical reference library, such as HMBC (Heteronuclear Multiple Bond Correlation) and HSQC (Heteronuclear Single Quantum Coherence) can provide a more comprehensive overview of peak correlations, specifically between the carbons of quinic acid from

chlorogenic acid and shared protons of caffeic and chlorogenic acids to improve compound identification and differentiate between these two acids. Burns and Reynolds (2019) suggested using ^{1}H-^{1}H correlation spectroscopy (COSY) and HSQC as a starting point and using HMBC to tie deprotonated fragments together (172). Furthermore, numerous recent retractions of incorrectly assigned natural compounds indicate the need for further experimentation (172).

Chemometric analysis

PCA was limited by extensive peak chemical shift variations observed across samples and posed a challenge in data analysis and interpretation(173). Moreover, the low concentration of active ingredients found within products meant peaks had a low signal to noise ratio, affecting PCA modelling, such as in the case of product S and *E.purpurea* (174). Therefore, in all PCA models, PC-1 reconstructed loadings expressed high levels of variance introduced by peak shift, as reviewed in the pre-processed spectra. This variance was a significant feature in batch-to-batch comparisons. The same shortcoming was observed in STOCSY analysis, where major peak shifts were observed for most product spectra (excluding the four unidentifiable peaks found in Products V3 and S). Therefore, structural correlation patterns could not be established, and neither PCA nor STOCSY were deemed appropriate for the analysis of batch-to-batch variations or metabolite identification due to their reliance on peak alignment.

To resolve peak shift due to differences in ionic strength, future experiments should consider sample processing techniques that are able to remove salts prior to examination with NMR spectroscopy. For example, electrodialysis (ED) or solid-phase extraction (SPE) could be used to remove ionic compounds before performing NMR spectroscopy (4). ED

removes salts that have a high separation efficiency. However, it is more costly than SPE (175), and neither method has been used on divalent metals such as Ca^{2+} and Mg^{2+} in complex mixtures prior to ^{1}H NMR spectroscopy. A workflow developed by Jiang et al. (2012) for the binding of metal ions in urine using ethylenediaminetetraacetic acid (EDTA) and potassium fluoride (PF) provided uniformity in chemical shift regions and enabled quantification of Ca-EDTA and Mg-EDTA complexes(176). Furthermore, pre-processing techniques have been refined, such as spectral binning(5), wavelet-based deconvolution(177) (Appendix B), and automated alignment mechanism(6). However, these are difficult when the amount of chemical shift is unknown and may produce artefacts in the presence of peak overlap(63), particularly wavelet deconvolution (Appendix B).

In the present analysis, relative peak integration provided the greatest insight into the batch-to-batch variation of selected compounds, with relative signal concentrations across samples highlighting inter- and intra-batch differences. Although relative concentration could indicate product trends, no absolute concentration could be determined, and these results cannot be used in a clinical setting. This was due to the experimental design and thus reduced the comparability across studies(178). These findings are consistent with other metabolomics literature highlighting the difficulty of using NMR alone in absolute quantification(158).

Furthermore, identification relating to entire peak patterns of chemical reference standards must be taken provisionally. The sensitivity threshold of NMR is relatively high (in the low μM range) compared with MS (in the low pM range)(179). Active ingredients are often found in extremely low concentrations, mainly when using the sample processing methods of this analysis. When taking a subset of a pill's heterogeneous mixture, all ingredients in the complex mixture may not appear. Therefore, determining the absolute concentration per

product unit (e.g., tablet) would better infer consumer dosages. This should be done by taking the average weight of a pill and capsule and calculating the expected concentration of compounds, alongside a higher concentration of mixture being included in the sample.

Future applications

Future studies should include complementary techniques with low detection limits to consolidate absolute quantification findings. Various researchers throughout the world have developed complementary methods that include MS and NMR. Additionally, high performance liquid chromatography (HPLC) is noted as the gold standard for validation and batch quality consistency evaluation as multiple biomarkers may be measured simultaneously and similarity analysis can be performed relatively easily (180). For herbal supplement quantification, various stock mixtures that contain similar heterogeneity to herbal supplement products could be used as references for MS quantification and spiking experimentation in NMR. Quantitative NMR (q-NMR) has significantly evolved throughout recent years in drug quality control, including new developments such as the inclusion of cryoprobes and dynamic nuclear polarisation (179). Currently, q-NMR can certify drug product purity with a precision of +/- 1% with an uncertainty measure <1.0%, however this has not yet been applied to herbal supplements (179).

Such approaches are inherently limited because excipient ingredient quantities are not provided in Australia, meaning that active ingredient dosages cannot be reliably measured relative to these compounds. Therefore, targeted methods could be used to quantify these defined excipients. Furthermore, database searching for excipient ingredients highlighted the presence of lauryl sulfate, stearic acid, and sorbitol. Other excipient

ingredients could not be identified due to the lack of matching database or literature entries of ^{1}H NMR spectra. Th us, numerous factors are required to enable dosage calculations to provide a stronger link between immune supplement quality, safety, and efficacy. Furthermore, the complexity of pill mixtures could apply further experimentation with stock mixtures, which could enable faster experimental and model variation while presenting the contribution of different compounds to the same peak(56).

The present experimental design, the use of ^{1}H NMR for the complete identification of compounds contributing to the complex mixture of immune supplements is not possible. However, characteristic profiles of reference compounds could be produced for supplements that contain active ingredients. Therefore, the unequivocal determination of active ingredients in herbal supplements using high throughput NMR is a complex task that requires a cross-institutional repository for spectral trace comparison.

Conclusion

The development of characteristic profiles of immune supplements using ^{1}H NMR spectroscopy provided an avenue for understanding the presence of active ingredients declared by product manufacturers. The nature of immune supplements (containment of inorganic ions and heterogenous mixture) resulted in study limitations when using ^{1}H NMR, related to chemical shift perturbations and peak overlap. These limitations might be overcome by the tailoring sample processing and separation techniques to improve spectral data output. Nonetheless, apparent relative concentration differences of active ingredients observed between batches highlighted the need for greater control over the batch consistency of immune supplements. These findings illustrate the need for future research for confirmatory analysis and to further improve the use of quantitative NMR in herbal

supplement analysis for quality control. This would enable further development of robust methods that could be used by regulatory bodies to enforce meaningful changes to the complementary and alternative medicine sector in the scope of public safety.

Conclusion and future applications

The current analysis attempted to determine the use of labelling compliance and ^{1}H NMR as a means of quality assurance for immune supplements in Australia. While, it was found that all products were non-compliant in their labelling format, it was not clear that this indicated the harm potential of the supplement. Furthermore, the sole use of NMR in ion rich complex mixtures, such as Immune supplements proved difficult without a suitable bioinformatic pipeline and further pre-sample processing. This work presents a multidisciplinary (regulatory and metabolomic) examination of IS product quality based on regulatory guidelines published by the TGA. This examination into the regulatory oversight of labelling compliance and compositional integrity of immune supplements has demonstrated the need for greater control over the quality of these largely unregulated products. The transition period to new regulatory frameworks for labelling compliance meant the TGA spent considerable resources and time on labelling and advertising. Moreover, there was no alarming safety concerns reported in the NMR analysis such as toxic contaminants or adulterants. The largest concerns being the lack of active ingredient and between batch concentration differences. Nonetheless, this cannot be only attributed to product design and could be a result of sample pre-processing and NMR experimentation.

Package labelling

Analysis of immune supplements found all products non-compliant with the TGO92, defining minimum pre-market requirements/criteria for complementary medicines in Australia. However, further risk-based investigation highlighted that labelling incompliance often did not impede product safety due to minor discrepancies, particularly in font size. Moreover, there was no alarming safety concerns reported in the NMR analysis such as toxic contaminants or adulterants. The largest concerns being the lack of active ingredient and between batch concentration differences. Nonetheless, this cannot be only attributed to product design and could be a result of sample pre-processing and NMR experimentation. Furthermore, the small sample size of products used is not a complete representation of consumer habits when using immune supplements and requires a greater sampling base. Therefore, this thesis fore fronted the development of an open access bioinformatic pipeline that could be used by data depositories for cross-institutional and regulatory comparison or quality assurance. Data will be deposited to metabolights to enable other researchers to identify false database entries for compounds.

^{1}H NMR Analysis

With this knowledge, the proposed sample analysis pipeline, including a combination of 1D NMR experiments and PCA as multivariate analysis methods, was not able to to generate quantitative information for the complete set of constituents of herbal supplement mixtures. Certain inherent disadvantages from sample processing, NMR experimentation set up, data pre-processing and bioinformatics make the standardisation and comparison of herbal supplements difficult without complementary techniques including Liquid Chromatography or solid phase extraction (Discussed in detail throughout Chapter 3).

For example, the Chinese Pharmacopoeia (ChP) sets out the enforcement of traditional Chinese medicine (TCM) in China while unifying with the standards recommended in the European Pharmacopeia. This is the only pharmacopoeia to enforce batch consistency analysis, requiring at least ten batch chromatographs to be used for comparison(57, 58, 180). While there are some criticisms to the quality control used by the ChP, TCM researchers globally, develop improved methodologies for quality consistency while being supported by government regulators, often using hyphenated mass spectrometry systems alongside NMR (152, 181-183). This provides greater validation and understanding of ionic effects on composition. There is already a common understanding of the role of complementary metabolomics in quality control of TCM(183). Such as those, recognising the limitations of databases in small molecule identification for NMR, the use of pre-separation techniques for the removal of metal ions, and the simultaneous use of MS and NMR spectroscopy for a completer and more robust structural picture of herbal supplements. Therefore, proton NMR spectroscopy will be more useful, in the application of pre-developed quality markers (Q-markers) that have been standardised using complementary practices and robust sample collection/preparation procedures for efficacy-orientated quality assurance(152).

Complementary methods that involve either greater pre-processing or the combination of spectroscopic and spectrometry techniques will enable an enriched sample processing pipeline that relieves the requirement for additional pre-processing during data analysis. Specifically, as machine learning processing such as binning, and alignment have been known to generate artefacts that interfere with statistical analysis. Moreover, the integration of toxicity assays, bioavailability testing, and active constituent identification, providing a holistic quality and safety assurance approach.

Numerous questions remain about the compositional integrity of immune supplements in this pilot study, further indicating the need for innovative and cost-effect quality assurance measures that should be further investigated in a validation study. the relative concentration differences exhibited between the batches spectral output did not enable an understanding into the 'why there are concentration differences?'. Therefore, future validation studies will optimise quantitative features of NMR (particularly 2D JRes NMR) for absolute concentration metrics. In contrast, greater forethought into experimental procedures will provide absolute concentration determination from proton NMR and will not rely on further 2D experimentation which decreases the high throughput aims of this analysis. By being able to apply the ^{1}H NMR spectra as botanical vouchers in scientific research and quality assurance consistency in species identification, and chemical composition, can be achieved (184). Due to their long-term low cost and time effectiveness once developed, such an approach could facilitate pre-manufacturing regulation and provide accurate taxonomic and chemical identification for comparison with processed material and extracts. For example, the difference in chemical shift when comparing Ascorbic acid found in products has led to further questioning about the changes that occur to the chemical structure of herbal and inorganic materials from pasture to market. This could result in decreased physiological bioavailability or prove potentially harmful in high dose ascorbic acid processes.

Future Applications

Further research into this would enable a better understanding of therapeutic compound degradation and allow for evidence-based legislation, as the TGA currently has no shelf-life evaluation for listed goods(185). While the present analysis does not indicate any significant adulterations these quality assurance practices should be implemented pre-

market as the assumption that CAM products are inherently low risk does not reflect the current scientific literature (23, 27-29, 39, 49, 61, 127).

In 2008, Australian researchers provided a list of consumer-safety related solutions for complementary medicine regulation, suggesting greater scrutiny of clinical trial data and product identification that should be accessible to the public. Moreover, they highlighted that the “100% cost-recovery” strategy of the TGA causes prioritization of economic stability over safeguarding consumers(185). Based on the analysis presented in this thesis, we agree with the proposed solutions. Meaningful change will not occur until scientific literature, technological innovation, and regulators unify to produce an evidence-based workflow, including, sample processing, data processing, and compositional identification to ensure the quality and safety of CAM products. The use of ^{1}H NMR demonstrated the potential for high throughput mechanisms of analysing compositional integrity and its further use in batch quality consistency using machine learning programs. Although, there is a requirement for further development of experimental design, the creation of an open access pre-processing pipeline using Python highlights the need for standardised platforms that can be cross-institutionally applied and consolidated for NMR to enable greater returns on the spectral data that is acquired. Further automation, such as those seen in the mass spectrometry space needs to be applied to NMR. Including, the creation of a wavelet theory-based identification probability automation as indicated in appendix B. While further pre-experimentation processes would need to be included to remove significant between sample chemical shift. This includes solid-phase extraction of heavily concentrated metal ions in immune supplements and would provide greater accessibility to wavelet-based automations and alignment algorithms due to the removal of significant chemical shift

It is unlikely that consumers will decrease their use of immune supplements, particularly with the global pandemic continuing. Therefore, it is the onus of researchers and regulators to unify under one international body to protect consumers from the potential harm caused by the misuse of resources in the quality assurance of herbal medicines.

References

1. Organisation WH. WHO Global Report on Traditional and Complementary Medicine 2019. Geneva: World Health Organisation 2019.
2. Pan S-Y, Litscher G, Gao S-H, Zhou S-F, Yu Z-L, Chen H-Q, et al. Historical Perspective of Traditional Indigenous Medical Practices: The Current Renaissance and Conservation of Herbal Resources. Evidence-Based Complementary and Alternative Medicine. 2014;2014:525340.
3. O'Brien SK, Malacova E, Sherriff JL, Black LJ. The Prevalence and Predictors of Dietary Supplement Use in the Australian Population. Nutrients. 2017;9(10):1154.
4. Khabour OF, Hassanein SFM. Use of vitamin/zinc supplements, medicinal plants, and immune boosting drinks during COVID-19 pandemic: A pilot study from Benha city, Egypt. Heliyon. 2021;7(3):e06538.
5. Wagner D MA, & Caulfield T. "Immune Boosting" in the time of COVID: selling immunity on Instagram. Allergy, Asthma & Clinical Immunology. 2020;16(1).
6. Kunnumakkara AB, Rana V, Parama D, Banik K, Girisa S, Sahu H, et al. COVID-19, cytokines, inflammation, and spices: How are they related? Life Sci. 2021:119201-.
7. Lockyer S. Effects of diets, foods and nutrients on immunity: Implications for COVID-19? . Nutrition Bulletin. 2020;45:456-73.
8. Zupo R, Castellana F, Sardone R, Sila A, Giagulli VA, Triggiani V, et al. Preliminary Trajectories in Dietary Behaviors during the COVID-19 Pandemic: A Public Health Call to Action to Face Obesity. Int J Environ Res Public Health. 2020;17(19):7073.
9. Adams KK, Baker, W. L., & Sobieraj, D. M. Myth Busters : Dietary Supplements and COVID-19. Ann Pharmacother. 2020.
10. Feng Z, Yang J, Xu M, Lin R, Yang H, Lai L, et al. Dietary supplements and herbal medicine for COVID-19: A systematic review of randomized control trials. Clinical Nutrition ESPEN. 2021;44:50-60.
11. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Research. 2017;46(D1):D608-D17.
12. Administration TG. Complementary medicines overview In: Administration TG, editor. 2019.
13. Therapeutic Goods Administration: complementary medicines In: Office ANA, editor. Canberra2011.
14. Industry Audit and Trends. Canberra: Complementary Medicines Australia; 2020.
15. Administration. TG. How we regulate medicines 2019.
16. TGA. Court action. In: Deparment H, editor. Canberra: Australian Government; 2021.
17. Shen M-R, He Y, Shi S-M. Development of chromatographic technologies for the quality control of Traditional Chinese Medicine in the Chinese Pharmacopoeia. Journal of Pharmaceutical Analysis. 2021;11(2):155-62.
18. Schoepfer AM, Engel A, Fattinger K, Marbet UA, Criblez D, Reichen J, et al. Herbal does not mean innocuous: ten cases of severe hepatotoxicity associated with dietary supplements from Herbalife products. J Hepatol. 2007;47(4):521-6.
19. Nortier JL, Vanherweghem JL. Renal interstitial fibrosis and urothelial carcinoma associated with the use of a Chinese herb (Aristolochia fangchi). Toxicology. 2002;181-182:577-80.
20. Woo JJ. Adverse event monitoring and multivitamin-multimineral dietary supplements. Am J Clin Nutr. 2007;85(1):323s-4s.
21. Chen Y, Wang C, Yang H, Huang P, Shi J, Tong Y, et al. Epidemiology of Drug- and Herb-Induced Liver Injury Assessed for Causality Using the Updated RUCAM in Two Hospitals from China. BioMed Research International. 2021;2021:8894498.
22. Liang Q, Xie M. [Analysis of toxicity of traditional Chinese herbal medicine and its connotation]. Zhong Xi Yi Jie He Xue Bao. 2009;7(2):101-4.
23. Avigan MI, Mozersky RP, Seeff LB. Scientific and Regulatory Perspectives in Herbal and Dietary Supplement Associated Hepatotoxicity in the United States. Int J Mol Sci. 2016;17(3):331.
24. Organisation WH. WHO guidelines for assessing quality of herbal medicines with reference to contaminants and residues. Geneva; 2007.
25. Luo L, Wang B, Jiang J, Fitzgerald M, Huang Q, Yu Z, et al. Heavy Metal Contaminations in Herbal Medicines: Determination, Comprehensive Risk Assessments, and Solutions. Frontiers in pharmacology. 2021;11:595335-.
26. Buettner C, Mukamal KJ, Gardiner P, Davis RB, Phillips RS, Mittleman MA. Herbal supplement use and blood lead levels of United States adults. J Gen Intern Med. 2009;24(11):1175-82.

27. Ichim MC. The DNA-Based Authentication of Commercial Herbal Products Reveals Their Globally Widespread Adulteration. Frontiers in Pharmacology. 2019;10(1227).
28. Pawar RS, Handy SM, Cheng R, Shyong N, Grundel E. Assessment of the Authenticity of Herbal Dietary Supplements: Comparison of Chemical and DNA Barcoding Methods. Planta Med. 2017;83(11):921-36.
29. Chatzinasiou L, Booker A, MacLennan E, Mackonochie M, Heinrich M. Turmeric (Curcuma longa L.) products: What quality differences exist? Journal of Herbal Medicine. 2019;17-18:100281.
30. Ulbricht C, Basch E, Cheung L, Goldberg H, Hammerness P, Isaac R, et al. An evidence-based systematic review of elderberry and elderflower (Sambucus nigra) by the Natural Standard Research Collaboration. J Diet Suppl. 2014;11(1):80-120.
31. Crighton E CM, Farrington R, et al. . Toxicological screening and DNA sequencing detects contamination and adulteration in regulated herbal medicines and supplements for diet, weight loss and cardiovascular health. Journal of Pharmaceutical and Biomedical Analysis. 2019;176.
32. Parvathy VA, Swetha VP, Sheeja TE, Sasikumar B. Detection of plant-based adulterants in turmeric powder using DNA barcoding. Pharm Biol. 2015;53(12):1774-9.
33. Ichim MC, Booker A. Chemical Authentication of Botanical Ingredients: A Review of Commercial Herbal Products. Front Pharmacol. 2021;12:666850.
34. Frédérich M, Jansen C, de Tullio P, Tits M, Demoulin V, Angenot L. Metabolomic analysis of Echinacea spp. by 1H nuclear magnetic resonance spectrometry and multivariate data analysis technique. Phytochemical Analysis. 2010;21(1):61-5.
35. Byard R, Maker G, Bunce M. What risks do herbal products pose to the Australian community? The Medical Journal of Australia. 2017;207:134.
36. Crighton E, Weisenseel J, Bunce M, Musgrave I, Trengove R, Maker G. Exploring the Application of the DSA-TOF, a Direct, High-resolution Time-of-Flight Mass Spectrometry Technique for the Screening of Potential Adulterated and Contaminated Herbal Medicines. Journal of The American Society for Mass Spectrometry. 2019;30.
37. Garcia-Aloy M, Groff N, Masuero D, Nisi M, Franco A, Battelini F, et al. Exploratory Analysis of Commercial Olive-Based Dietary Supplements Using Untargeted and Targeted Metabolomics. Metabolites. 2020;10(12).
38. Garg V, Dhar VJ, Sharma A, Dutt R. Facts about standardization of herbal medicine: a review. Zhong Xi Yi Jie He Xue Bao. 2012;10(10):1077-83.
39. Gilard V, Balayssac S, Tinaugus A, Martins N, Martino R, Malet-Martino M. Detection, identification and quantification by 1H NMR of adulterants in 150 herbal dietary supplements marketed for improving sexual performance. Journal of Pharmaceutical and Biomedical Analysis. 2015;102:476-93.
40. Heyman HM, Meyer JJM. NMR-based metabolomics as a quality control tool for herbal products. South African Journal of Botany. 2012;82:21-32.
41. Güemes Heras I, Santamaría-Orleans A, Colinas Herrero JF, Gómez Sorrigueta P, Ortiz González L, de la Iglesia-Arnaez R, et al. Use of Dietary Supplements among Spanish Pediatricians in Daily Practice: A Cross-Sectional Survey Study. Journal of Nutrition and Metabolism. 2019;2019:5819305.
42. Sarker SD, Nahar L. An introduction to natural products isolation. Methods Mol Biol. 2012;864:1-25.
43. Nybom H, Weising K, Rotter B. DNA fingerprinting in botany: past, present, future. Investigative Genetics. 2014;5(1):1.
44. Hubert J, Nuzillard J-M, Renault J-H. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? Phytochemistry Reviews. 2017;16(1):55-95.
45. Tawfike AF, Viegelmann C, Edrada-Ebel R. Metabolomics and dereplication strategies in natural products. Methods Mol Biol. 2013;1055:227-44.
46. Sumner LW, Mendes P, Dixon RA. Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry. 2003;62(6):817-36.
47. Beutler JA, Alvarado AB, Schaufelberger DE, Andrews P, McCloud TG. Dereplication of phorbol bioactives: Lyngbya majuscula and Croton cuneatus. J Nat Prod. 1990;53(4):867-74.
48. Hachem R, Assemat G, Balayssac S, Martins-Froment N, Gilard V, Martino R, et al. Comparative Chemical Profiling and Monacolins Quantification in Red Yeast Rice Dietary Supplements by (1)H-NMR and UHPLC-DAD-MS. Molecules. 2020;25(2).
49. Crighton E, Coghlan ML, Farrington R, Hoban CL, Power MWP, Nash C, et al. Toxicological screening and DNA sequencing detects contamination and adulteration in regulated herbal medicines and supplements for diet, weight loss and cardiovascular health. J Pharm Biomed Anal. 2019;176:112834.
50. Ding S, Jiang H, Fang J. Regulation of Immune Function by Polyphenols. J Immunol Res. 2018;2018:1264074-.

51. D'Evoli L, Lucarini M, Potenza A, Ritota M, Sequi P, Lombardi-Boccia G. Anthocyanin Profile of Two Italian Cichorium Intybus L. Cultivars. Acta horticulturae. 2012;939:337-40.
52. Avula B, Cohen PA, Wang YH, Sagi S, Feng W, Wang M, et al. Chemical profiling and quantification of monacolins and citrinin in red yeast rice commercial raw materials and dietary supplements using liquid chromatography-accurate QToF mass spectrometry: Chemometrics application. J Pharm Biomed Anal. 2014;100:243-53.
53. Yang J, Wang A, Li X-j, Fan X, Yin S-S, Lan K. A chemical profiling strategy for semi-quantitative analysis of flavonoids in Ginkgo extracts. Journal of Pharmaceutical and Biomedical Analysis. 2016;123.
54. Dai Y, Chen S-R, Chai L, Zhao J, Wang Y, Wang Y. Overview of pharmacological activities of Andrographis paniculata and its major compound andrographolide. Critical Reviews in Food Science and Nutrition. 2019;59(sup1):S17-S29.
55. Sorokina M, Steinbeck C. Review on natural products databases: where to find data in 2020. Journal of Cheminformatics. 2020;12(1):20.
56. Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL. Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products--application to oleuropein preparations. PLoS One. 2013;8(10):e78277.
57. Li S, Han Q, Qiao C, Song J, Lung Cheng C, Xu H. Chemical markers for the quality control of herbal medicines: an overview. Chinese medicine. 2008;3:7-.
58. Li Y, Guo S, Zhu Y, Yan H, Qian D-w, Wang H-q, et al. Comparative analysis of twenty-five compounds in different parts of Astragalus membranaceus var. mongholicus and Astragalus membranaceus by UPLC-MS/MS. Journal of Pharmaceutical Analysis. 2019;9.
59. Marchetti L, Rossi MC, Pellati F, Benvenuti S, Bertelli D. HR-1H NMR spectroscopy and multivariate statistical analysis to determine the composition of herbal mixtures for infusions. Phytochemical Analysis. 2021;32(4):544-53.
60. Pacholczyk-Sienicka B, Ciepielowski G, Albrecht Ł. The Application of NMR Spectroscopy and Chemometrics in Authentication of Spices. Molecules. 2021;26(2):382.
61. Kim SB, Bisson J, Friesen JB, Bucchini L, Gafner S, Lankin DC, et al. The Untargeted Capability of NMR Helps Recognizing Nefarious Adulteration in Natural Products. Journal of Natural Products. 2021;84(3):846-56.
62. Krishnan P, Kruger NJ, Ratcliffe RG. Metabolite fingerprinting and profiling in plants using NMR. Journal of Experimental Botany. 2004;56(410):255-65.
63. Emwas A-H, Saccenti E, Gao X, McKay RT, dos Santos VAPM, Roy R, et al. Recommended strategies for spectral processing and post-processing of 1D 1H-NMR data of biofluids with a particular focus on urine. Metabolomics. 2018;14(3):31.
64. Bartel J, Krumsiek J, Theis FJ. STATISTICAL METHODS FOR THE ANALYSIS OF HIGH-THROUGHPUT METABOLOMICS DATA. Computational and Structural Biotechnology Journal. 2013;4(5):e201301009.
65. Takis PG, Schäfer H, Spraul M, Luchinat C. Deconvoluting interrelationships between concentrations and chemical shifts in urine provides a powerful analysis tool. Nature Communications. 2017;8(1):1662.
66. Waidyanatha S, Pierfelice J, Cristy T, Mutlu E, Burback B, Rider CV, et al. A strategy for test article selection and phytochemical characterization of Echinacea purpurea extract for safety testing. Food Chem Toxicol. 2020;137:111125-.
67. Wolfender J-L, Marti G, Thomas A, Bertrand S. Current approaches and challenges for the metabolite profiling of complex natural extracts. Journal of Chromatography A. 2015;1382:136-64.
68. Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, et al. ChEBI in 2016: Improved services and an expanding collection of metabolites. Nucleic Acids Res. 2016;44(D1):D1214-9.
69. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.
70. Block K MM. Immune System Effects of Echinacea, Ginseng, and Astragalus: A Review. Cancer Therapies 2003;2(3):247-67.
71. Saeidnia S MA, Vazirian M. . Echinacea purpurea: Pharmacology, phytochemistry and analysis methods. Pharmacognosy Reviews 2015;9(17):63.
72. Zolgharnein J, Niazi A, Afiuni S, Zamani K. Determination of Cichoric Acid as a Biomarker in Echinacea Purpurea Cultivated in Iran Using High Performance Liquid Chromatography Open Access. Chinese Medicine. 2010;1:23-7.
73. Cao X WC, Tian Y, Guo P. The caffeic acid moiety plays an essential role in attenuating lipid accumulation by chlorogenic acid and its analogues. RSC Advances. 2019;9(22):12247-54.

74. Li Y, Guo S, Zhu Y, Yan H, Qian D-W, Wang H-Q, et al. Comparative analysis of twenty-five compounds in different parts of Astragalus membranaceus var. mongholicus and Astragalus membranaceus by UPLC-MS/MS. Journal of Pharmaceutical Analysis. 2019;9(6):392-9.
75. Muhammad Naveeda b, VH, MA, AAK, , Khane GJ, et al. Chlorogenic acid (CGA): A pharmacological review and call for further

research. Biomedicine & Pharmacotherapy. 2018;97:67-70.
76. Sriramaneni RN, Sadikun A, Kumar Hs N, Shivashekaregowda H, Venkatappa A, Asmawi M. Identification of Bioactive Diterpenoid Lactones using Nuclear Magnetic Resonance from Andrographis paniculata. Pharmacognosy Research. 2019;11:121-6.
77. Tajidin NE, Shaari K, Maulidiani M, Salleh NS, Ketaren BR, Mohamad M. Metabolite profiling of Andrographis paniculata (Burm. f.) Nees. young and mature leaves at different harvest ages using 1H NMR-based metabolomics approach. Scientific Reports. 2019;9(1):16766.
78. Lamy S, Muhire E, Annabi B. Antiproliferative efficacy of elderberries and elderflowers (Sambucus canadensis) on glioma and brain endothelial cells under normoxic and hypoxic conditions. Journal of Functional Foods. 2018;40:164-79.
79. Stuppner S, Mayr S, Beganovic A, Beć K, Grabska J, Aufschnaiter U, et al. Near-Infrared Spectroscopy as a Rapid Screening Method for the Determination of Total Anthocyanin Content in Sambucus Fructus. Sensors (Basel). 2020;20(17):4983.
80. Appenteng MK, Krueger R, Johnson MC, Ingold H, Bell R, Thomas AL, et al. Cyanogenic Glycoside Analysis in American Elderberry. Molecules. 2021;26(5):1384.
81. Salvador ÂC, Rudnitskaya A, Silvestre AJD, Rocha SM. Metabolomic-Based Strategy for Fingerprinting of Sambucus nigra L. Berry Volatile Terpenoids and Norisoprenoids: Influence of Ripening and Cultivar. Journal of Agricultural and Food Chemistry. 2016;64(26):5428-38.
82. Moosavian SP, Paknahad Z, Habibagahi Z, Maracy M. The effects of garlic (Allium sativum) supplementation on inflammatory biomarkers, fatigue, and clinical symptoms in patients with active rheumatoid arthritis: A randomized, double-blind, placebo-controlled trial. Phytotherapy Research. 2020;34(11):2953-62.
83. Sánchez-Sánchez MA, Zepeda-Morales ASM, Carrera-Quintanar L, Viveros-Paredes JM, Franco-Arroyo NN, Godínez-Rubí M, et al. Alliin, An Allium sativum Nutraceutical, Reduces Metaflammation Markers in DIO Mice. Nutrients. 2020;12(3):624.
84. Molina-Calle M, de Medina VS, Priego-Capote F, de Castro MDL. Establishing compositional differences between fresh and black garlic by a metabolomics approach based on LC–QTOF MS/MS analysis. Journal of Food Composition and Analysis. 2017;62:155-63.
85. Ritota M, Casciani L, Han BZ, Cozzolino S, Leita L, Sequi P, et al. Traceability of Italian garlic (Allium sativum L.) by means of HRMAS-NMR spectroscopy and multivariate data analysis. Food Chem. 2012;135(2):684-93.
86. Liu L, Chen L-J, Zhang Y, Bing B, Ji S. Isolation determination of garlic allyl sulfides and their antioxidant activity. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering. 2015;31:268-74.
87. Silveira D, Prieto-Garcia JM, Boylan F, Estrada O, Fonseca-Bazzo YM, Jamal CM, et al. COVID-19: Is There Evidence for the Use of Herbal Medicines as Adjuvant Symptomatic Therapy? Frontiers in Pharmacology. 2020;11(1479).
88. Pappenberger G, Hohmann HP. Industrial production of L-ascorbic Acid (vitamin C) and D-isoascorbic acid. Adv Biochem Eng Biotechnol. 2014;143:143-88.
89. Shakoor H, Feehan J, Al Dhaheri AS, Ali HI, Platat C, Ismail LC, et al. Immune-boosting role of vitamins D, C, E, zinc, selenium and omega-3 fatty acids: Could they help against COVID-19? Maturitas. 2021;143:1-9.
90. Reid RS. The proton NMR spectrum of ascorbic acid: A relevant example of deceptively simple second-order behavior. Journal of Chemical Education. 1989;66(4):344.
91. Baenas N, Salar FJ, Domínguez-Perles R, García-Viguera C. New UHPLC-QqQ-MS/MS Method for the Rapid and Sensitive Analysis of Ascorbic and Dehydroascorbic Acids in Plant Foods. Molecules. 2019;24(8):1632.
92. Tzima K, Brunton NP, Rai DK. Qualitative and Quantitative Analysis of Polyphenols in Lamiaceae Plants-A Review. Plants (Basel). 2018;7(2):25.
93. Sroka Z, Cisowski W. Hydrogen peroxide scavenging, antioxidant and anti-radical activity of some phenolic acids. Food and Chemical Toxicology. 2003;41(6):753-8.

94. Spiteller G. Peroxidation of linoleic acid and its relation to aging and age dependent diseases. Mechanisms of Ageing and Development. 2001;122(7):617-57.
95. Kirkinezos IG, Moraes CT. Reactive oxygen species and mitochondrial diseases. Seminars in Cell & Developmental Biology. 2001;12(6):449-57.
96. Bohn MK, Hall A, Sepiashvili L, Jung B, Steele S, Adeli K. Pathophysiology of COVID-19: Mechanisms Underlying Disease Severity and Progression. Physiology. 2020;35(5):288-301.
97. H.Hemilä. Vitamin C and Infections. Nutrients. 2017.
98. Attila S, Vancea S, Kiss I, Donáth-Nagy G. Quantification of Plasma and Leukocyte Vitamin C by High Performance Liquid Chromatography with Mass Spectrometric Detection. Journal of Analytical Chemistry. 2020;75(9):1168-76.
99. Barnes K, Ball L, Desbrow B, Alsharairi N, Ahmed F. Consumption and reasons for use of dietary supplements in an Australian university population. Nutrition. 2016;32(5):524-30.
100. Hemilä H LP. Vitamin C for preventing and treating pneumonia. Cochrane Database of Systematic Reviews 2013.
101. Hoang B SG, Fang W, Han B. Possible application of high-dose vitamin C in the prevention and therapy of coronavirus infection. Journal of Global Antimicrobial Resistance. 2020:256-62.
102. Pehlivan F. Vitamin C: An Antioxidant Agent. IntechOpen. 2017.
103. Carr A MS. Vitamin C and Immune Function. Nutrients. 2017.
104. Khaksari M, Mazzoleni LR, Ruan C, Kennedy RT, Minerick AR. Data representing two separate LC-MS methods for detection and quantification of water-soluble and fat-soluble vitamins in tears and blood serum. Data in Brief. 2017;11:316-30.
105. Zou C SG, Wu Y, Lu F, Mao W, Liu X. . Astragalus in the Prevention of Upper Respiratory Tract Infection in Children with Nephrotic Syndrome: Evidence-Based Clinical Practice. Evidence-Based Complementary and Alternative Medicine 2013:1-10.
106. Hong F XW, Ragupathi G, Lau C, Leung P, Yeung K et al. . The Known Immunologically Active Components of Astragalus account for Only a Small Proportion of the Immunological Adjuvant Activity When Combined with Conjugate Vaccines. Planta Medica 2010;77(8):817-24.
107. Safety review of Andrographis paniculata and anaphylactic allergic reactions. Canberra: Therapeutic Goods Administration, Department H; 2015.
108. Sagineedu S, Kong K, Phang S, Toh B, Chu V, Hamzah A, et al. Quantitative analysis of andrographolide, a biomarker of Andrographis paniculata, using qNMR spectroscopy. Frontiers in Pharmacology. 2019;10.
109. Burgos RA, Alarcón P, Quiroga J, Manosalva C, Hancke J. Andrographolide, an Anti-Inflammatory Multitarget Drug: All Roads Lead to Cellular Metabolism. Molecules. 2020;26(1).
110. Sandoval-Ramírez BA, Catalán Ú, Fernández-Castillejo S, Pedret A, Llauradó E, Solà R. Cyanidin-3-glucoside as a possible biomarker of anthocyanin-rich berry intake in body fluids of healthy humans: a systematic review of clinical trials. Nutrition Reviews. 2019;78(7):597-610.
111. Mohammadi Pour P, Fakhri S, Asgary S, Farzaei MH, Echeverría J. The Signaling Pathways, and Therapeutic Targets of Antiviral Agents: Focusing on the Antiviral Approaches and Clinical Perspectives of Anthocyanins in the Management of Viral Diseases. Frontiers in pharmacology. 2019;10:1207-.
112. Kritikou E, Kalogiouri NP, Kolyvira L, Thomaidis NS. Target and Suspect HRMS Metabolomics for the Determination of Functional Ingredients in 13 Varieties of Olive Leaves and Drupes from Greece. Molecules (Basel, Switzerland). 2020;25(21):4889.
113. Anthony Booker DF, Deborah Johnston, Chinenye Umealajekwu, Eike Reich, Michael Heinrichad. Chemical variability along the value chains of turmeric (Curcuma longa): A comparison of nuclear magnetic resonance spectroscopy and high performance thin layer chromatography. Journal of Ethnopharmacology. 2014;152(2).
114. Drobnic F, Riera J, Appendino G, Togni S, Franceschi F, Valle X, et al. Reduction of delayed onset muscle soreness by a novel curcumin delivery system (Meriva®): a randomised, placebo-controlled trial. J Int Soc Sports Nutr. 2014;11:31.
115. Esmaily H, Sahebkar A, Iranshahi M, Ganjali S, Mohammadi A, Ferns G, et al. An investigation of the effects of curcumin on anxiety and depression in obese individuals: A randomized controlled trial. Chinese journal of integrative medicine. 2015;21(5):332-8.
116. Hewlings SJ, Kalman DS. Curcumin: A Review of Its Effects on Human Health. Foods. 2017;6(10):92.
117. Amagase H. Clarifying the Real Bioactive Constituents of Garlic. The Journal of Nutrition. 2006;136(3):716S-25S.

118. Sachan AK, Vishnoi G, Kumar R. Need of standardization of herbal medicines in Modern era. International Journal of Phytomedicine. 2016;8:300-7.
119. Viola GCV, Bianchi F, Croce E, Ceretti E. Are Food Labels Effective as a Means of Health Prevention? J Public Health Res. 2016;5(3):768-.
120. Campos S, Doxey J, Hammond D. Nutrition labels on pre-packaged foods: a systematic review. Public Health Nutrition. 2011;14(8):1496-506.
121. Oliveira D, Machín L, Deliza R, Rosenthal A, Walter EH, Giménez A, et al. Consumers' attention to functional food labels: Insights from eye-tracking and change detection in a case study with probiotic milk. LWT - Food Science and Technology. 2016;68:160-7.
122. Administration TG. Medicine labels: Guidance on TGO 91 and TGO 92. In: Health Do, editor. Canberra: Australian Government; 2021.
123. TGA. Advertising Complaints and Investigations Information. In: Deparment H, editor. Canberra2021.
124. Guidance R. The TGA's risk management approach. In: Administration TG, editor. Canberra: Australian Government; 2021.
125. Therapeutic goods advertising compliance: 2019-20 Annual Report. Canberra: Therapeautic Goods Administration, Deparment H; 2020.
126. Full responses from the TGA and MJA authors, Byard et al2017; (4). Available from: https://insightplus.mja.com.au/2017/4/full-responses-from-the-tga-and-mja-authors-byard-et-al/.
127. Blacksell L, Byard RW, Musgrave IF. Forensic problems with the composition and content of herbal medicines. J Forensic Leg Med. 2014;23:19-21.
128. Administration TG. How we manage advertising compliance. In: Deparment H, editor. Canberra: Australian Government; 2020.
129. The Standard for the Uniform Scheduling of Medicines and Poison, (2003).
130. Therapeutic Goods (Permissible Ingredients) Determination (2021).
131. Boon HS, Kachan N. Natural health product labels: is more information always better? Patient Educ Couns. 2007;68(2):193-9.
132. Harris PJ, Shaw IC. Solving New Zealand's Complementary and Alternative Medicine (CAM) product crisis: A risk-based proposal: University of Canterbury; 2018. 35–60 p.
133. Esayas S, Mahler T. Modelling compliance risk: a structured approach. Artificial Intelligence and Law. 2015;23:271-300.
134. Andrade C. The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives. Indian J Psychol Med. 2019;41(3):210-5.
135. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. The American Statistician. 2016;70(2):129-33.
136. TGA. Australian regulatory guideline for

over-the-counter medicines-

Appendix 3: Guidelines on presentation aspects

of OTC applications. In: Deparment H, editor. Canberra: Australian Government; 2012.
137. Fuchs D, Heyer T, Langenhan D, Hippius PDM, editors. Influence of Font Sizes on the Readability and Comprehensibility of Package Inserts2008.
138. Vlachojannis JE, Cameron M, Chrubasik S. A systematic review on the sambuci fructus effect and efficacy profiles. Phytotherapy Research. 2010;24(1):1-8.
139. Tiralongo E, Wee SS, Lea RA. Elderberry Supplementation Reduces Cold Duration and Symptoms in Air-Travellers: A Randomized, Double-Blind Placebo-Controlled Clinical Trial. Nutrients. 2016;8(4):182-.
140. Jurčić D, Gabrić M, Troskot Perić R, Liberati Pršo AM, Mirat J, Včev A, et al. HERBALIFE® asscoiated severe hepatotoxicity in previously healthy woman. Acta Clin Croat. 2019;58(4):771-6.
141. TGA. Therapeutic goods advertising compliance. In: Health D, editor. Canberra: Australian Government 2020.
142. Clarke N, Pechey E, Kosīte D, König LM, Mantzari E, Blackwell AKM, et al. Impact of health warning labels on selection and consumption of food and alcohol products: systematic review with meta-analysis. Health Psychology Review. 2020:1-24.
143. Gabriels G, Lambert M. Nutritional supplement products: does the label information influence purchasing decisions for the physically active? Nutrition Journal. 2013;12(1):133.

144. Mei B, Xie H, Xing H, Kong D, Pan X, Li Y, et al. Changes of phenolic acids and antioxidant activities in diploid and tetraploid Echinacea purpurea at different growth stages. Revista Brasileira de Farmacognosia. 2020;30(4):510-8.
145. Zaushintsena A MI, Babich O, Noskova S, Kiseleva T, Popova D et al. . Quantitative and qualitative profile of biologically active substances extracted from purple echinacea (Echinacea Purpurea L.) growing in the Kemerovo region: functional foods application. . Foods and Raw Materials 2019:84-92.
146. Kreft S. Cichoric Acid Content and Biomass Production of Echinacea purpurea. Plants Cultivated in Slovenia. Pharmaceutical Biology. 2005;43(8):662-5.
147. Johnson MC, Thomas AL, Greenlief CM. Impact of Frozen Storage on the Anthocyanin and Polyphenol Contents of American Elderberry Fruit Juice. Journal of Agricultural and Food Chemistry. 2015;63(23):5653-9.
148. Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S. DNA barcoding detects contamination and substitution in North American herbal products. BMC medicine. 2013;11(1):1-13.
149. Therapeutic Goods Act 1989.
150. Wishart DS, Sayeeda Z, Budinski Z, Guo A, Lee BL, Berjanskii M, et al. NP-MRD: the Natural Products Magnetic Resonance Database. Nucleic Acids Research. 2021.
151. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46(D1):D1074-d82.
152. Xiong Y, Hu Y, Li F, Chen L, Dong Q, Wang J, et al. Promotion of quality standard of Chinese herbal medicine by the integrated and efficacy-oriented quality marker of Effect-constituent Index. Phytomedicine. 2018;45:26-35.
153. Haahr M. Random number generator Ireland: Randomness and Integrity Services Ltd.; 2021 [
154. Bioequivalence Guidance for Industry. Rockville: Food and Drug Administration, Services USdoHaH; 2006.
155. Yesiltepe Y, Nuñez JR, Colby SM, Thomas DG, Borkum MI, Reardon PN, et al. An automated framework for NMR chemical shift calculations of small organic molecules. Journal of Cheminformatics. 2018;10(1):52.
156. Tredwell GD, Bundy JG, De Iorio M, Ebbels TMD. Modelling the acid/base 1H NMR chemical shift limits of metabolites in human urine. Metabolomics. 2016;12(10):152.
157. Pal A SR, Picozza M, Pawar A, Rongioletti M, Dutta A et al. . Zinc and COVID-19: Basis of Current Clinical Trials. Biological Trace Element Research 2020.
158. Crook AA, Powers R. Quantitative NMR-Based Biomedical Metabolomics: Current Status and Applications. Molecules. 2020;25(21):5128.
159. Kretschmer A, Backer M. Concentration dependence of the 1H- and 29Si-NMR chemical shifts of tetramethyldisiloxanediol in deuterated chloroform. Journal of Organometallic Chemistry. 2001;628(2):233-40.
160. McKenzie JS, Donarski JA, Wilson JC, Charlton AJ. Analysis of complex mixtures using high-resolution nuclear magnetic resonance spectroscopy and chemometrics. Progress in Nuclear Magnetic Resonance Spectroscopy. 2011;59(4):336-59.
161. didi ismanto S, Desri Rahmi I, Febrian A. The Influence of Drying Temperature on Chemical Components of Herbal Tea Leaves (Spondias Dulcis, Soland). IOP Conference Series: Earth and Environmental Science. 2016;583.
162. Giménez R, Cabrera C, Olalla M, Ruiz MD, López MC. Ascorbic acid in diet supplements: loss in the manufacturing process and storage. International Journal of Food Sciences and Nutrition. 2002;53(6):509-18.
163. Clydesdale FM, Ho CT, Lee CY, Mondy NI, Shewfelt RL, Lee K. The effects of postharvest treatment and chemical interactions on the bioavailability of ascorbic acid, thiamin, vitamin a, carotenoids, and minerals. Critical Reviews in Food Science and Nutrition. 1991;30(6):599-638.
164. Sagineedu SR, Kong K, Phang SY, Toh BL, Chu VS, Hamzah AS, et al. Quantitative analysis of andrographolide, a biomarker of Andrographis paniculata, using qNMR spectroscopy. Frontiers in Pharmacology. 2019;10.
165. Krest I, Keusgen M. Quality of herbal remedies from Allium sativum: differences between alliinase from garlic powder and fresh garlic. Planta Med. 1999;65(2):139-43.
166. Reyes BAS, Dufourt EC, Ross J, Warner MJ, Tanquilut NC, Leung AB. Chapter 4 - Selected Phyto and Marine Bioactive Compounds: Alternatives for the Treatment of Type 2 Diabetes. In: Atta ur R, editor. Studies in Natural Products Chemistry. 55: Elsevier; 2018. p. 111-43.
167. Aires-de-Sousa J, Hemmer MC, Gasteiger J. Prediction of 1H NMR Chemical Shifts Using Neural Networks. Analytical Chemistry. 2002;74(1):80-90.
168. Tian X, Tian D, Wang Z-Y, Mo F. Synthesis and Evaluation of Chitosan-Vitamin C Complex. Indian journal of pharmaceutical sciences. 2009;71:371-6.

169. Liu Y, Saurí J, Mevers E, Peczuh MW, Hiemstra H, Clardy J, et al. Unequivocal determination of complex molecular structures using anisotropic NMR measurements. Science. 2017;356(6333):eaam5349.
170. Maier ME. Structural revisions of natural products by total synthesis. Nat Prod Rep. 2009;26(9):1105-24.
171. Brown PD, Lawrence AL. The importance of asking "how and why?" in natural product structure elucidation. Nat Prod Rep. 2017;34(10):1193-202.
172. Burns DC, Reynolds WF. Minimizing the risk of deducing wrong natural product structures from NMR data. Magnetic Resonance in Chemistry. 2021;59(5):500-33.
173. Cañueto D, Salek RM, Correig X, Cañellas N. Improving sample classification by harnessing the potential of (1)H-NMR signal chemical shifts. Sci Rep. 2018;8(1):11886-.
174. Halouska S, Powers R. Negative impact of noise on the principal component analysis of NMR data. J Magn Reson. 2006;178(1):88-95.
175. Gurreri L, Tamburini A, Cipollina A, Micale G. Electrodialysis Applications in Wastewater Treatment for Environmental Protection and Resources Recovery: A Systematic Review on Progress and Perspectives. Membranes (Basel). 2020;10(7):146.
176. Jiang L, Huang J, Wang Y, Tang H. Eliminating the dication-induced intersample chemical-shift variations for NMR-based biofluid metabonomic analysis. Analyst. 2012;137(18):4209-19.
177. Tong X, Zhang Z, Zeng F, Fu C, Ma P, Peng Y, et al. Recursive Wavelet Peak Detection of Analytical Signals. Chromatographia. 2016;79(19):1247-55.
178. Fei Q, Wang D, Jasbi P, Zhang P, Nagana Gowda GA, Raftery D, et al. Combining NMR and MS with Chemical Derivatization for Absolute Quantification with Reduced Matrix Effects. Analytical Chemistry. 2019;91(6):4055-62.
179. Simmler C, Napolitano JG, McAlpine JB, Chen S-N, Pauli GF. Universal quantitative NMR analysis of complex natural samples. Curr Opin Biotechnol. 2014;25:51-9.
180. Li Y, Wu T, Zhu J, Wan L, Yu Q, Li X, et al. Combinative method using HPLC fingerprint and quantitative analyses for quality consistency evaluation of an herbal medicinal preparation produced by different manufacturers. Journal of Pharmaceutical and Biomedical Analysis. 2010;52(4):597-602.
181. van Beek TA, Montoro P. Chemical analysis and quality control of Ginkgo biloba leaves, extracts, and phytopharmaceuticals. Journal of Chromatography A. 2009;1216(11):2002-32.
182. Xu X-R, Zhang T, Li P, Xu Q, Feng B, Tan P, et al. Overview and prospects of traditional Chinese medicine blending technology oriented by quality consistency. Zhongguo Zhong yao za zhi= Zhongguo zhongyao zazhi= China journal of Chinese materia medica. 2019;44(22):4786-92.
183. Wei X-C, Cao B, Luo C-H, Huang H-Z, Tan P, Xu X-R, et al. Recent advances of novel technologies for quality consistency assessment of natural herbal medicines and preparations. Chinese Medicine. 2020;15(1):56.
184. Culley TM. Why vouchers matter in botanical research. Applications in Plant Sciences. 2013;1(11):1300076.
185. Harvey KJ, Korczak VS, Marron LJ, Newgreen DB. Commercialism, choice and consumer protection: regulation of complementary medicines in Australia. Medical Journal of Australia. 2008;188(1):21.

Appendix

Appendix A

Appendix table A.1 Example of contingency tables for statistical analysis.

	non-compliant	compliant
harmful	17	0
non-harmful	3	17

Contingency tables were created in numerous ways to attempt a statistical analysis of independence. Nonetheless, the chosen variables were not entirely independent and always had a case of zero as labels could never be harmful and compliant at the same time (without further compositional analysis). Therefore, this statistical analysis could be performed after characteristic profiles were developed. Although, this would have indicated 100% harmful non-compliance due to the lack of complete peak patterns found. Additionally, a single 1 could have been added to all values to make the model valid. However, due to the small sample size, this would have added spurious significance. Consequently, this analysis did not benefit for bivariate and independent statistical testing commonly used in compliance testing.

Appendix table A. Product compliance according to legislation set out in TGO 92.

An overview of the different areas of non-compliance and each products compliance status. A tick (✓)highlights instances of non-compliance. A cross (✗) indicates compliance.

Supplement	Non-cohesive title (S9.2)	Wrong active ingredient presentation (S11)	Active ingredient font <2.5 mm	Incorrect/missing listing status (S7.2(d)(i))	Incorrect/missing warning statement (S8.1(k))	Warning statement <1.5 mm (S7.2(d))	Missing/incorrect expiry date (S8.1(g))	Quantity and dosage form displayed <1.5 mm (S7.2(d))	Incorrect/missing directions of use (S8.1(l))	Directions of use <1.5 mm (S7.2(d))	Wrong/missing storage instructions (S11.4)	Differing ingredient concentrations, permitted ingredients, or permitted indications on ARTG and physical label	Incorrect/missing dosage form	Storage instructions still legible but <1.5 mm (S7.2(d))	Missing/incorrect sponsor details and below 1.5mm but still legible (S8.1(i))	Expiry prefix <1.5mm but still legible (S7.2(d))	Batch prefix <1.5mm but still legible (S7.2(d))	Listing status <1mm but still legible (S7.2(d)(i))
2020-01	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-02	✓	✗	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-03	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗
2020-04	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-05	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗

2020-06	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-07	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-08	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✗	✗	✓	✓	✓	✓	✗
2020-09	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-10	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-11	✓	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗
2020-12	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-13	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-14	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-15	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-29	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2020-30	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗

Appendix B.

For this analysis covariance and correlation were expressed in matrix notation. However, these are commonly examined as mathematical formulas, including the following:

To determine the covariance values for each coordinate of the covariance matrix is done by highlighting how the variation found between and within the sample batches can be given in terms of the spectra (x) and ppm (y):

$$Cov(x,y) = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{n}$$

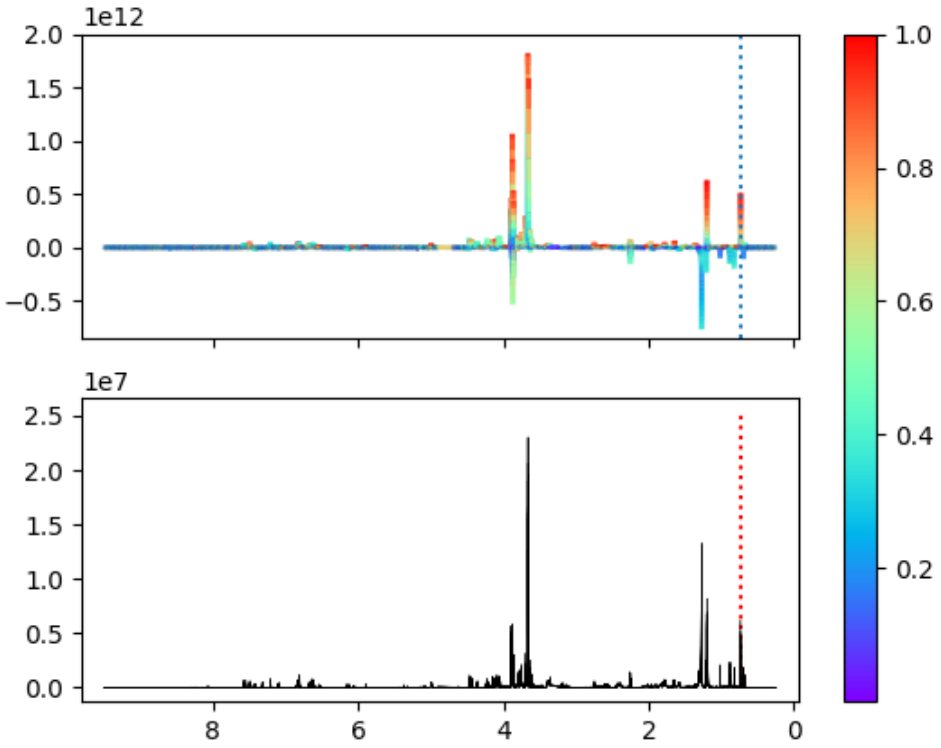
The 3-dimensional matrix was generated using a symmetric $(p \times p)$ matrix for variable x (spectra), y (ppm), and z (standardised scores) to form a 3x3 matrix:

$$\begin{matrix} Cov(x,x) & Cov(x,y) & Cov(x,z) \\ Cov(y,x) & Cov(y,y) & Cov(y,z) \\ Cov(z,x) & Cov(z,y) & Cov(z,z) \end{matrix}$$

The correlation is derived from covariance but is not affected by scale and measures the strength of two variables when all other conditions remain constant. This is formula is expressed as:

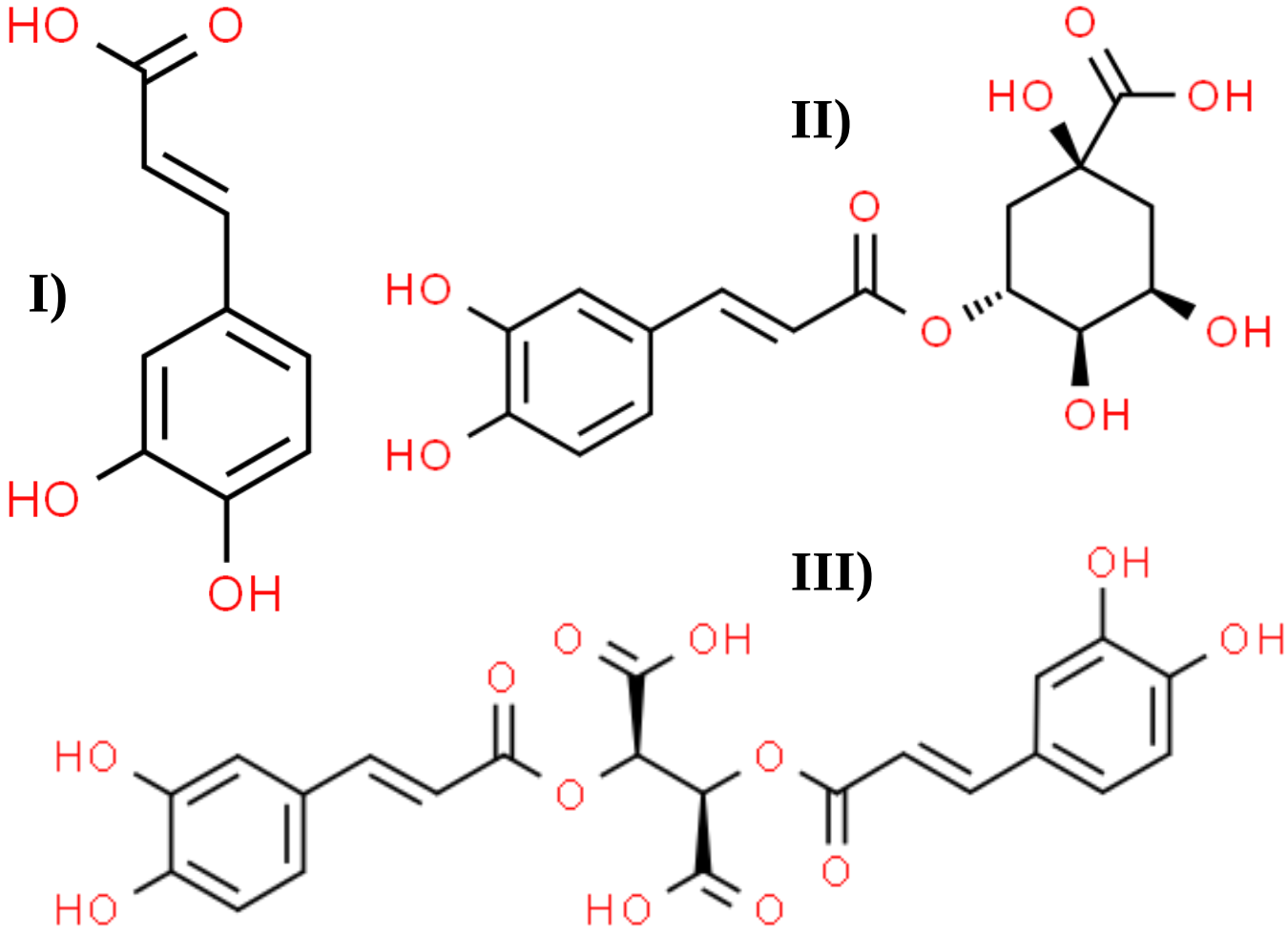
$$Cor(x,y) = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Figure 1.A.B. STOCSY results of product A.



STOCSY was performed with the driver peak at 0.88ppm. The only result from the HMDB database was 3-phosphoglyceric acid. This is unlikely to be the correct compound due a 50% match. Furthermore, the peak position and shape did not match simulated reference spectra. Therefore, STOCSY proved ineffective at highlighting collinearity of peaks derived from the same compounds. Database searching is also limited and requires greater experimental results that are applicable to users over predictive models.

Figure 2.A.B Comparison of active polyphenols found in *E.purpurea*: Chicoric acid, chlorogenic acid, caffeic acid.

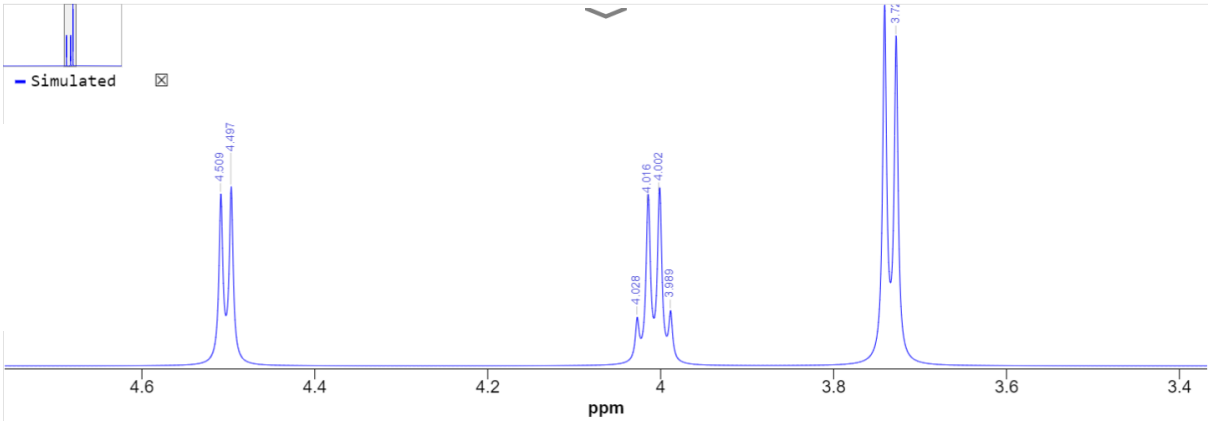


Three compounds were overlaid to determine the presence of *E.purpurea*. These compounds are all hydroxycinnamic acids and share very similar chemical structures which explains their spectral similarity and difficulty in determination using ^{1}H NMR. I) Caffeic acid is the root chemical structure of all three compounds. This phenyl ring differs from cinnamic acids due to the substitution of hydroxy groups at positions 3 and 4. Caffeic acid makes up both II) Chlorogenic acid and III) Chicoric acid, which contain additional links and explain slight differences in NMR spectra. Chlorogenic acid also has an ester link with quinic acid.

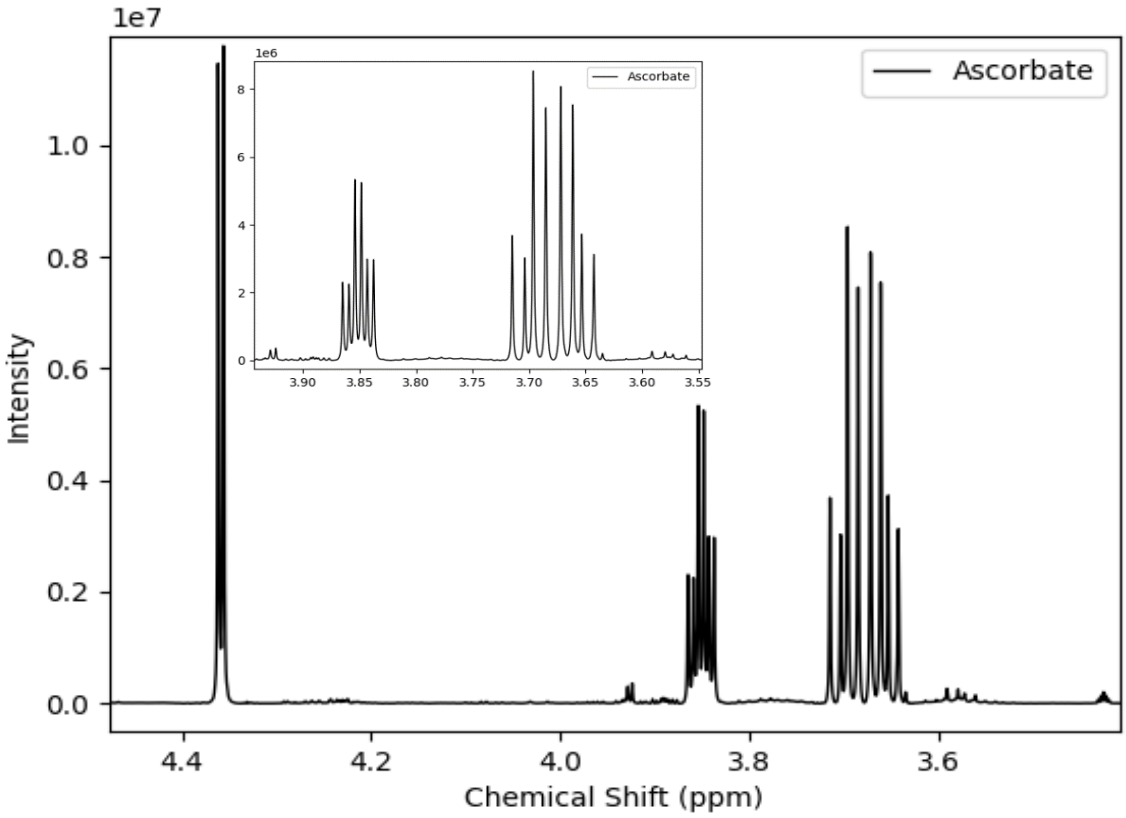
Whereas chicoric acid has two caffeic acids joined by tartaric acid. Nonetheless, their aromatic regions are very similar which explains peak pattern similarity from 6.0-8.0ppm. The quinic acid found on chlorogenic results in the split multiplets located between 2.0 and 2.5ppm.

Figure 3.A.B Comparison of simulated database reference spectra (NP-MRD), in-house chemical reference and literature.

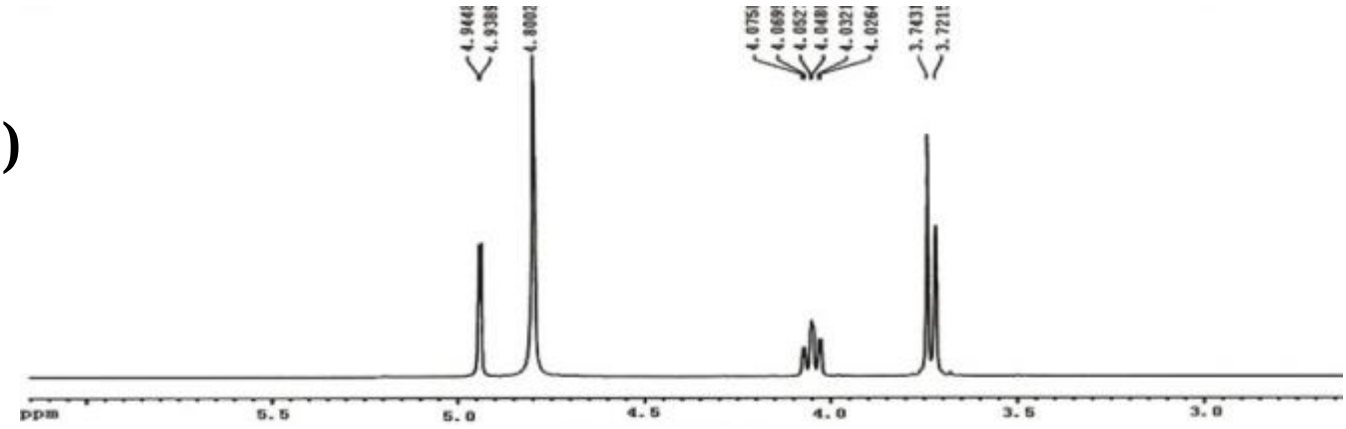
I)



II)



III)



Three reference spectra for ascorbate were compared. I) ascorbate simulated reference spectra generated from NP-MRD and HMDB. This did not match II) in-house reference spectra nor the work by III) Tian et al (168). Peak positions, line shape, and peak patterns were not comparable and therefore posed a problem in identification.

Figure 5.A.B Comparison of kuromanin reference spectra with product S, specifically identifying major peaks in areas of significant peak overlap.

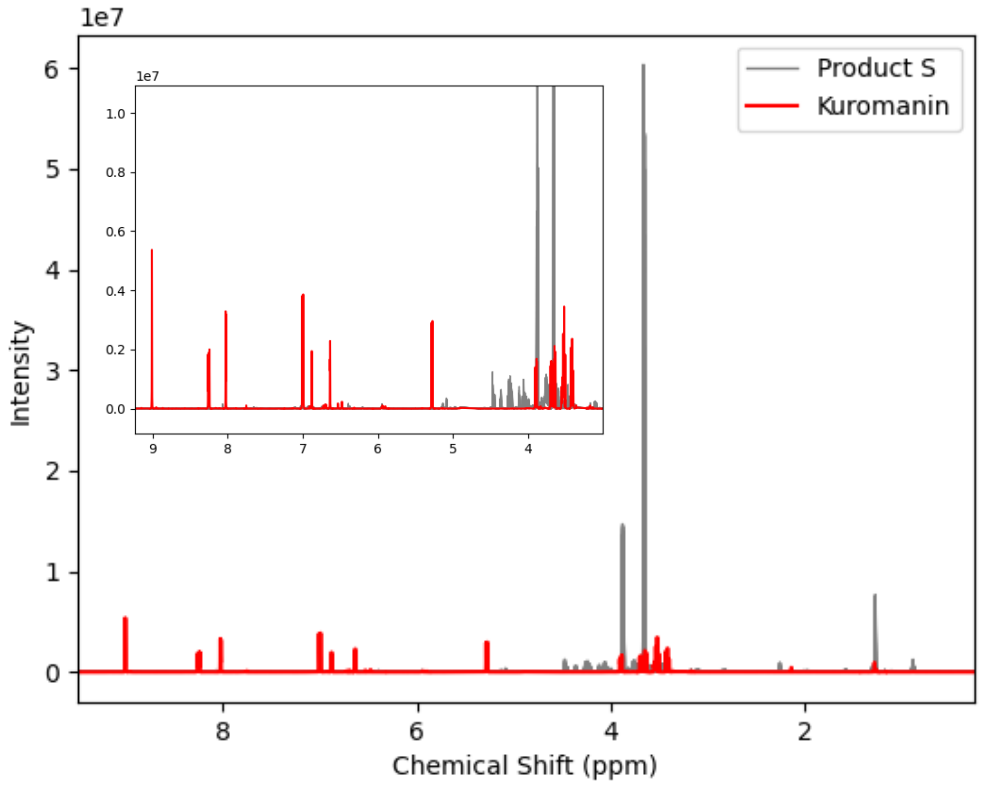
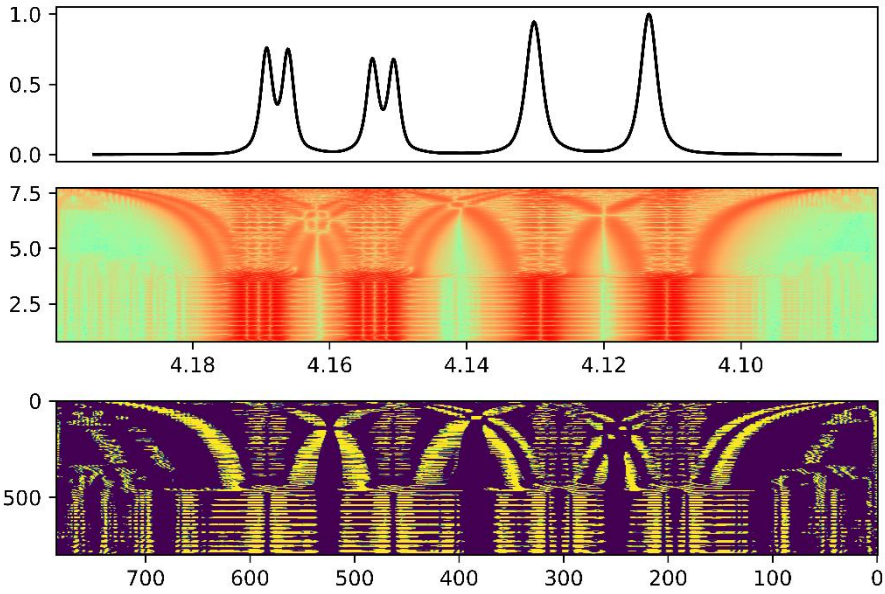


Figure B.6. Deconvolution of spectral waves using the Continuous wave transform and Shannon wavelet.



Two different examples of andrographolide reference standard being deconvolved via the continuous wave transform using the morlet wavelet (top) and shannon wavelet (bottom). This can be applied to a machine learning algorithm that will automatically scan the full spectral range for analytical reference standard. Unfortunately, due to time constraints and the chemical shift perturbations observed in product spectra, this was not applicable for this project.

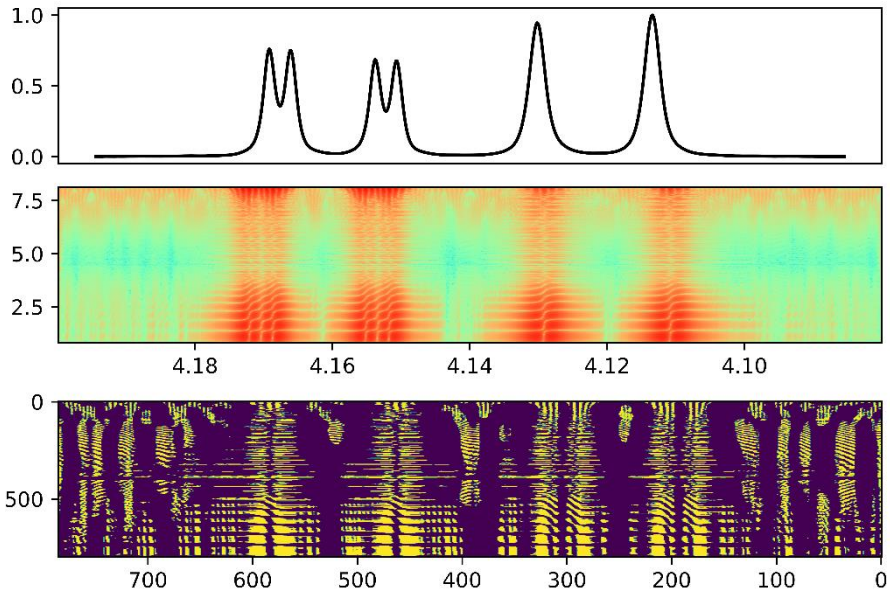
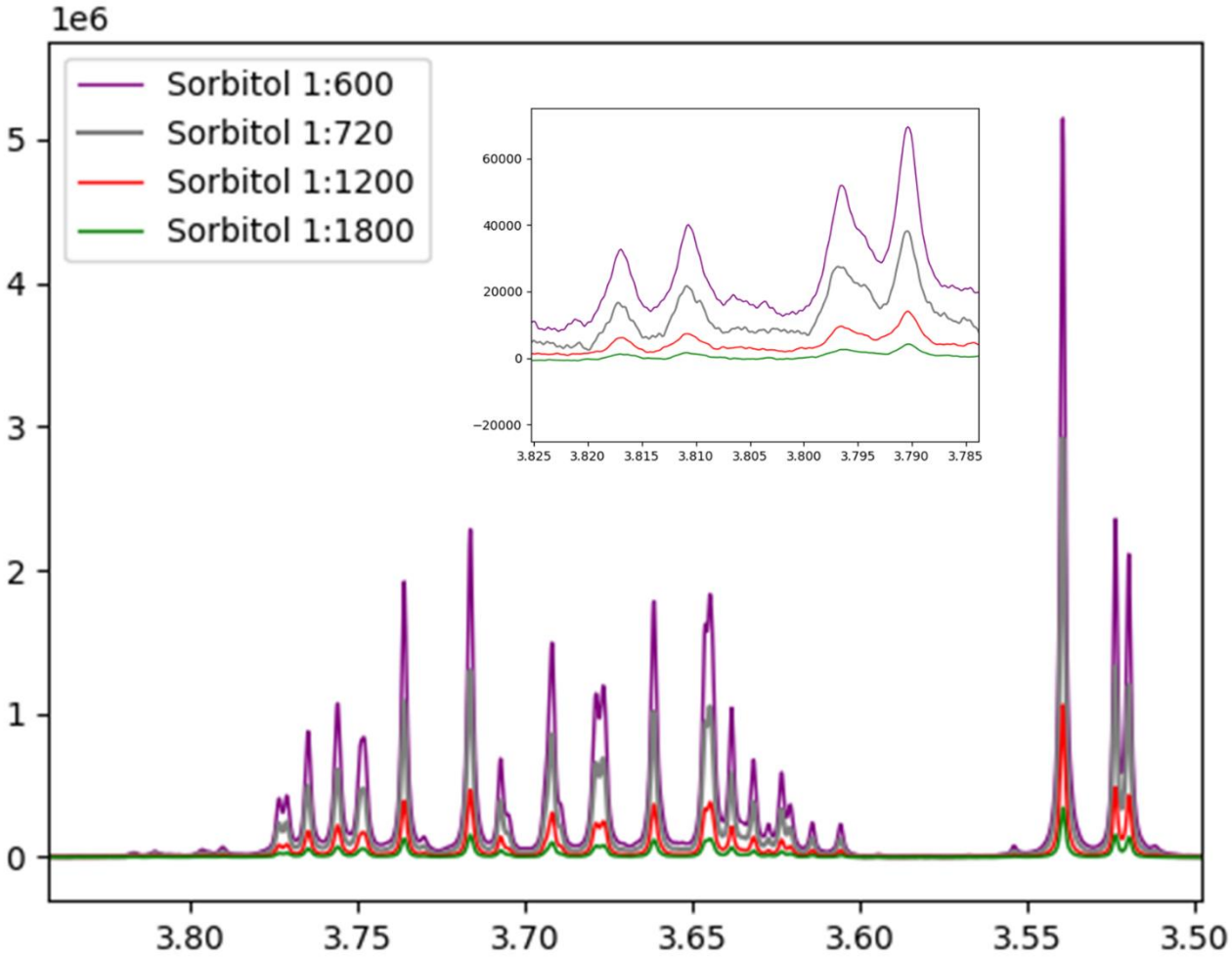


Figure B.7. The effect of dilution on peak shape pattern in sorbitol.



In house chemical reference sorbitol was used to highlight the impact that dilution has on peak shape. Sorbitol was diluted by 20%, 100%, and 200% using ^{1}H NMR (10 scans each). This showed there is some resilience to peak shape in major signals. Although, minor signals easily become distorted and are difficult to determine. This consolidated the effects of sample dilution in chapter 3 and assisted in determining alternative sample processing methods for future experimentation.