NEW ADVANCES IN CAPILLARY ELECTROPHORESIS FOR BIOMONITORING IN POPULATION HEALTH AND NEWBORN SCREENING OF CYSTIC FIBROSIS

NEW ADVANCES IN CAPILLARY ELECTROPHORESIS FOR BIOMONITORING IN POPULATION HEALTH AND NEWBORN SCREENING OF CYSTIC FIBROSIS

By STELLENA MATHIAPARANAM, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements of the Degree

Doctor of Philosophy

June 2022

DOCTOR OF PHILOSOPHY (2022) McMaster University

(Chemistry and Chemical Biology) Hamilton, ON

TITLE: New Advances in Capillary Electrophoresis for Biomonitoring in Population Health and Newborn Screening of Cystic Fibrosis

AUTHOR: Stellena Mathiaparanam, B.Sc. (McMaster University)

SUPERVISOR: Professor Philip Britz-McKibbin

PAGES: xviii, 183

Abstract

Biological markers (*i.e.,* biomarkers) are essential in clinical and epidemiological studies as they may provide mechanistic insights into the developmental origins of disease, as well as improve diagnostic testing and risk assessment for disease prevention. However, major challenges remain due to the lack of rapid yet selective analytical methods for high throughput screening that are also amenable to volume-restricted specimens. This thesis includes two major research themes that take advantage of capillary electrophoresis (CE) separations, including (1) the targeted analysis of urinary iodide and thiocyanate for assessment of nutritional adequacy and tobacco smoke exposures in the population, and (2) the discovery of new biomarkers in sweat specimens that may improve universal newborn screening programs for cystic fibrosis (CF) infants beyond impaired chloride transport. *Chapter II* examines the prevalence and risk factors associated with iodine deficiency in 24 h urine samples collected from 800 participants across four clinical sites in Canada as part of the Prospective Urban and Rural Epidemiological (PURE) study when using CE with UV detection in conjunction with sample self-stacking. Importantly, regional variations in iodine status were revealed with participants from Quebec City and Vancouver at greater risk for iodine deficiency than Hamilton and Ottawa. Overall, iodine supplement use, thyroxine prescription, urinary sodium excretion, and self-reported dairy intake were found to be protective factors against iodine deficiency. *Chapter III* applied a validated CE assay to measure urinary thiocyanate as a biomarker of tobacco smoke and dietary exposures in an international cohort of 1000 participants from the PURE study spanning 14 countries with varied income status, smoking habits, and diet quality. Current smokers residing in high-income countries had the highest extent of cyanide exposure indicative of greater harms from tobacco smoke compared to middle- and low-income countries after adjusting for smoking intensity and other covariates. *Chapter IV* introduces a rapid CE method with indirect UV detection to simultaneously measure sweat chloride and bicarbonate

from presumptive CF infants' residual sweat samples. Although bicarbonate did not provide clinical value in neonatal CF diagnosis, sweat chloride testing by CE may reduce test failure rates due to insufficient volumes from infants in a clinical setting. Lastly, *Chapter V* applied an untargeted strategy to characterize the sweat metabolome from presumptive CF infants when using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS). A panel of sweat metabolites were found to discriminate CF from non-CF (*i.e.*, unaffected carriers) infants, including aspartic acid, glutamine, oxoproline, and pilocarpic acid, which also correlated with sweat chloride. The clinical utility of these sweat metabolites to prognosticate late-onset CF infants from indeterminate sweat chloride test results was also explored. In summary, this thesis contributes innovative separation methods for biomarker screening and discovery in clinical and epidemiological studies for the prevention and early treatment of human diseases that benefit from optimal nutrition.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Philip Britz-McKibbin, for giving me the opportunity to join his research group and for his continuous support and encouragement during my undergraduate and graduate studies. My admiration and passion for research associated with bioanalytical chemistry initially emerged when I started in his group as an undergraduate student and has only exponentially increased since then. I truly appreciate his time and feedback during numerous discussions that have made this Ph.D. productive and worth it. Dr. Britz's guidance and enthusiasm helped shape me into the proactive and independent researcher I am today.

Next, I would like to thank my committee members, Dr. Stephen Hill, Dr. Giuseppe Melacini, and Dr. Jeremy Hirota. Their expertise, critical feedback, and beneficial suggestions have helped and motivated me during my graduate career. A special thank you to Dr. Stephen Hill, whose immense contribution played a significant role in my clinical projects. He went above and beyond by submitting REB applications, storing and de-identifying biological specimens, and taking the time to find missing metadata that I urgently needed. In addition to his contribution, Dr. Stephen Hill continuously provided me with words of encouragement (even through anecdotes) that will always stay close to my heart and has motivated me to reach the end of my Ph.D. degree.

Furthermore, I sincerely thank all the collaborators involved in the studies described in this thesis for their time and effort. In especial, I thank Dr. Linda Pedder and Lauren Brick (at McMaster Children's Hospital) for being incredibly attentive by helping me trace patient samples for my clinical studies and for being patient with me by answering my numerous follow-up participant and genomics related questions. I would also like to thank our collaborators at The Hospital for Sick Children in Toronto, Dr. Tanja Gonska and Katherine Keenan, who also provided valuable sweat samples. Thank you to our collaborators at Population

Health Research Institute for allowing me take part in their large-scale epidemiological studies; Dr. Andrew Mente for providing helpful statistical input.

I would like to thank past members of the Britz research group, including Adriana Nori de Macedo, Alicia DiBattista, Jennifer Wild, Karen Lam, Nadine Wellington, Mai Yamamoto, Michelle Saoi, and Sandi Azab. A special thank you to Adriana for being the most knowledgeable and patient mentor. I would also like to thank the current members of the Britz lab: Meera Shanmuganathan, Ritchie Ly, Biban Gill, Erick Helmeczi, Jordan Holzschuher, Megan Magee, and Zach Kroezen. Thank you all for the continuous support, words of wisdom, and laughs throughout this journey. An additional sincere thank you to laboratory technicians, Paul Chindemi and Patricia Martin for letting me borrow equipment that helped my experiments run smoothly.

Lastly, I would like to thank my family. To my younger sisters: Pamela, Rohada, and Chrisaka, words cannot describe how much I love you guys and how lucky I am to have all of you in my life. Thank you for all the love, laughs, support, and for being the most resilient people I know. You guys taught me how to love and appreciate life, which made this journey achievable. Importantly, I would like to express my deepest gratitude to my mom, Rivanna Ratnasingam, for her endless love, support, and comfort throughout my entire life. Thank you for always being there for me from staying up with me until I finished my homework, to helping me pay for my university applications, to motivating me through undergrad, to being my biggest supporter when I wanted to pursue graduate school, to teaching me how to cook, so I could sustain life away from home, and to finally reaching the finish line together. Thank you for everything and for being the best role model to me and my sisters. Thank you for giving me the room to grow and always encouraging me to fulfill my passions and dreams. I love you very much and owe all my success to you. This one is for you mommy, we did it!

Table of Contents

List of Figures

Supporting Figures

List of Tables

Supporting Tables

List of Abbreviations and Symbols

$[M+H]^-$	Deprotonated molecular ion
$[M+H]^+$	Protonated molecular ion
1-OHP	1-hydroxypyrene
AGM	Adjusted geometric means
AHEI	Alternative healthy eating index
ANCOVA	Analysis of covariance
ANOVA	Analysis of Variance
ATP	Adenosine 5'-triphosphate
AUC	Area under the curve
BGE	Background electrolyte
BMI	Body mass index
CAWG	Chemical Analysis Working Group
CE	Capillary electrophoresis
CE-ESI-MS	Capillary electrophoresis-electrospray ionization-mass spectrometry
CE-ESI-TOF-MS	Multisegment injection-capillary electrophoresis-electrospray ionization-time-of-flight-mass spectrometry
CE-iUV	Capillary electrophoresis-indirect ultraviolet
CE-MS	Capillary electrophoresis-mass spectrometry
CE-UV	Capillary electrophoresis-ultraviolet
CF	Cystic Fibrosis
CF-SPID	Cystic fibrosis-screen positive inconclusive diagnosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CFTR-RD	CFTR-related disease
CI	Confidence interval
Cl-Tyr	3-chlorotyrosine
COVID	Coronavirus disease
CPD	Cigarettes per day
CV	Coefficient of variation
Da	Dalton
DI-MS	Direct infusion-mass spectrometry
EAR	Estimated average requirement
EIC	Extracted-ion chromatogram
EK	Electrokinetic

EOF	Electroosmotic flow
ESI	Electrospray ionization
ESI-MS	Electrospray ionization-mass spectrometry
ETS	Environmental tobacco smoke
FC	Fold change
FDR	False discovery rate
F-Phe	Fluorophenylalanine
FWHM	Full width at half maximum
GC-MS	Gas chromatography-mass spectrometry
HIC	High-income
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
IDD	Iodine deficiency disorders
IQR	Interquartile range
IRT	Immunoreactive trypsinogen
IS	Internal standard
ITP	Isotachophoresis
iUV	Indirect ultraviolet
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LIC	Low-income
LOD	Limit of detection
m/z	Mass-to-charge ratio
MA	Mexican American
MEHP	Mono-2-ethylhexyl phthalate
MIC	Middle-income
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSI	Metabolomics Standards Initiative
MSI-CE-MS	Multisegment injection-capillary electrophoresis-mass spectrometry
NBS	Newborn screening
NDS	1,5-naphthalene disulfonic acid
NETS	Not exposed to environmental tobacco smoke
NHANES	National Health and Nutrition Examination Survey
NHB	Non-Hispanic Black
NHW	Non-Hispanic White
NMR	Nuclear magnetic resonance
NMS	Naphthalene sulfonic acid
NSM	Non-smoker

NSO	Newborn screening Ontario
OR	Odds ratio
OTH	Other unclassified ethnicities
PAH	Polycyclic aromatic hydrocarbons
PCA	Principal component analysis
PEG	Polyethylene glycol
Phe	Phenylalanine
PKU	Phenylketonuria
PLS-DA	Partial least squares-discriminant analysis
$PM_{2.5}$	Fine particulate matter
PON	Paraoxonase
PURE	Prospective Urban and Rural Epidemiological
QC	Quality control
QNS	Quantity not sufficient
RMT	Relative migration time
ROC	Receiver operating characteristic
RPA	Relative peak area
RS	Recovery standard
RSD	Relative standard deviation
RT	Retention time
SDS	Sodium dodecyl sulfate
SHS	Second-hand smoke
SM	Smoker
SNR	Signal-to-noise ratio
T4	Thyroxine
TGP	Total goiter prevalence
TNE-7	Total nicotine equivalents
TOF	Time-of-flight
TR	Transfer ratio
UI	Urine iodine
UIC	Urinary iodine concentration
UIC	Urinary iodine concentration
UIE	Urinary iodine excretion
UL	Upper limit
UV-Vis	Ultraviolet-visible
VIP	Variable importance in projection
WHO	World Health Organization

Declaration of Academic Achievement

The following material has been previously published and is reprinted with written permission:

Chapter II. Reprinted and adapted from Mathiaparanam, S.; Macedo, A. N.; Mente, A.; Poirier, P.; Lear S. A.; Wielgosz, A.; Teo, K. K; Yusuf, S.; Britz-McKibbin, P. The Prevalence and Risk Factors Associated with Iodine Deficiency in Canadian Adults. *Nutrients* **2022**, *14* (13), 1-17. Copyright (2022) American Chemical Society.

Chapter IV. Reprinted and adapted from Mathiaparanam, S.; Macedo, A. N.; Gill, B.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P. Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-Limited Sweat Specimens. *J. Cyst. Fibros.* **2022**, https://doi.org/10.1016/j.jcf.2022.05.004. Copyright (2022) American Chemical Society.

The following material has been submitted for publication:

Chapter III. Mathiaparanam, S.; Gill, B.; Thirunavukkarasu, S.; Guillaume, P.; Teo, K. K.; Yusuf, S.; Britz-McKibbin, P. Validation of Urinary Thiocyanate as a Robust Biomarker of Active Tobacco Smoking in the Prospective Urban and Rural Epidemiological Study. *Nicotine & Tobacco Research*, **2022**. American Chemical Society.

Chapter I

Introduction to Biomarkers and Their Use in Clinical and Epidemiological Studies

Chapter I: Introduction to Biomarkers and Their Use in Clinical and Epidemiological Studies

1.1 Introduction to Biomarkers in Clinical and Epidemiological Research

The term *biomarker* is an abbreviation of the word *biological marker*.[1] Biomarkers are discrete molecules measured in biological samples that may provide deeper insights into the origin, progression and classification of human diseases, which are widely used in epidemiological studies and clinical medicine. Clinical applications of biomarkers include the screening, diagnosis and prognosis of chronic or inherited disorders, as well as treatment responses to therapy that can better inform clinical decision making on an individual level.[2] Initially, biomarkers were defined to be biochemical or molecular alterations in various biological specimens that could lead to improved classification of a disease (*e.g.*, disease-causing gene mutations) that are associated with preventative and risk factors.[2] However, the official US National Institutes of Health (NIH) working group had expanded the definition of a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacologic responses to therapeutic intervention".[3] In most cases, biomarkers are quantified by various bioassays and instrumental methods when using standardized protocols ranging from colorimetric immunoassays, fluorescent-based enzyme kinetic assays to liquid chromatography-tandem mass spectrometry (LC-MS/MS); however, validating a biomarker for clinical use in diverse populations can be challenging. There are three conditions that are required to be satisfied: (i) the extent to which a biomarker explains the biological circumstance that is occurring; (ii) correlating compatible and significant phenotypic traits of the disease; (iii) determining how the biomarker relates to the disease as compared to existing methods.[4]

The utilization of biomarkers in immunological and genetic disorders, acute infections and chronic diseases, such as cancer are widely acknowledged.[2,5]

For example, Wild *et al*.[6] had reported urinary phenylalanine (Phe) and various other catabolites for non-invasive screening, confirmatory diagnosis and treatment monitoring of phenylketonuria (PKU) in support of universal newborn screening (NBS) programs. Similarly, Nori de Macedo *et al*.[7] demonstrated that endogenous (asparagine and glutamine) and exogenous (pilocarpic acid and mono-2-ethylhexyl phthalate, MEHP) metabolites in sweat samples enable differentiation of cystic fibrosis (CF) from non-affected non-CF infants, which provided new insights into CF disease mechanisms beyond impaired chloride transport. Furthermore, biomarkers may also function as objective indicators of dietary habits, lifestyle and/or environmental exposures.[2,8] For example, Rafiq *et al*.[9] summarized evidence for supporting specific dietary biomarkers of food intake as an alternative to self-reports from questionnaires that are prone to bias. In this case, Wellington *et al*.[10] described a contrasting dietary intervention to further validate biomarkers associated with a health-promoting Prudent diet, such as proline betaine, that is a robust biomarker of citrus food intake in urine or plasma. Lastly, urinary 1-hydroxypyrene (1-OHP) is a biomarker associated with recent smoke exposure to polycyclic aromatic hydrocarbons (PAH), which can be applied for the biomonitoring of firefighters' exposure to air pollution when combatting major wildfires as a way to improve occupational health and risk assessment.[11] The clinical relevance of three distinct biomarkers associated with diet and health outcomes (iodide), tobacco smoking (thiocyanate), and confirmatory testing of cystic fibrosis (chloride) will be further discussed below.

1.1.2 Monitoring of Iodine Nutritional Status for Population Health

Iodine, an essential trace micronutrient, represents a widely measured biomarker for evaluating nutritional status on a population level.[8] Iodine deficiency remains a global public health issue that is associated with a wide variety of deleterious health outcomes, including hypothyroidism, impaired cognitive development,

Table 1.1. The spectrum of iodine deficiency disorders across the life span. Adapted from de Benoist *et al*.,[16] Hetzel,[17] Laurberg *et al*.[18] and Stanbury *et al*.[19]

Stage in human life span	Adverse health outcomes
Fetus	Abortions
	Stillbirths
	Congenital anomalies
	Increased perinatal mortality
	Endemic cretinism
	Deaf mutism
Neonate	Neonatal goiter
	Neonatal hypothyroidism
	Endemic mental development
	Greater susceptibility of thyroid gland to nuclear radiation
Child and adolescent	Goiter
	(Subclinical) hypothyroidism
	(Subclinical) hyperthyroidism
	Impaired mental function
	Impaired physical development
	Greater susceptibility of thyroid gland to nuclear radiation
Adult	Goiter with complications
	Hypothyroidism
	Impaired mental function
	Spontaneous hyperthyroidism in the elderly
	Iodine-induced hyperthyroidism
	Greater susceptibility of thyroid gland to nuclear radiation

depression, cardiovascular diseases and cancer.[12,13] Optimal iodine intake from specific foods (*e.g.*, seafood, dairy), iodized table salt and/or supplements is required for the biosynthesis of thyroid hormones, which are important in regulating cellular metabolism, normal growth and cognitive development.[14,15] In fact, most countries worldwide have adopted table salt iodization as a public health policy to ensure a consistent source of iodine to the population for the prevention of iodine deficiency disorders (IDDs)[16] as summarized in **Table 1.1**. Iodine nutritional adequacy is particularly important for pregnant women during critical stages of fetal development to prevent adverse health outcomes in their offspring later in life.

Despite the early success of public health measures to prevent IDDs, changing dietary patterns in contemporary societies have contributed to increased ingestion of processed foods lacking iodized salt.[20,21] Furthermore, a prompt increase to dietary iodine is not a feasible solution since excess iodine intake is also correlated to negative health outcomes, such as hyperthyroidism, tachycardia and arrhythmia.[22] Thus, continuous monitoring of iodine nutrition in the population is crucial to implement revisions to public health policies, such as iodization programs.[23–25] In fact, total goiter prevalence (TGP) was suggested to be the gold standard indicator to evaluate IDD prevalence until the 1990s.[16] However, TGP was determined to be a poor indicator of IDD prevalence, as it only reflected a history of severe iodine deficiency as opposed to recent iodine sufficiency.[16] As a result, the median urinary iodine concentration (UIC) remains the most sensitive indicator of recent iodine nutritional status on a population level. The predominant species of iodine excreted in urine is iodide, which reflects more than 90% of recently ingested iodine from foods or supplements.[26] According to the World Health Organization (WHO), iodine nutrition can be assessed by monitoring UIC from single-spot random urine (μg/L or μM) or 24 h urine (μg/day) samples as shown in **Table 1.2**. Currently, the recommended dietary intake to ensure iodine nutritional adequacy in adults and pregnant women is 150 μg/day and 250 μg/day, respectively.[26]

1.1.3 Urinary Thiocyanate as a Biomarker of Tobacco Smoke Exposure

In addition to iodine serving as a biomarker in assessing recent dietary exposures, thiocyanate has also shown to be a promising biomarker in population health. For instance, thiocyanate functions as an iodine uptake inhibitor that can increase risk for IDDs due to overconsumption of isothiocyanate derived glucosinolates from cruciferous plants,[28] and it has been strongly correlated to tobacco smoke exposure due to inhalation of hydrogen cyanide.[29] Importantly, cyanide is an acute toxin that is associated with tobacco smoking and cardiorespiratory disease

Table 1.2. Categories of iodine nutrition for adults (excluding pregnant and breastfeeding women) based on median urinary iodine concentration (UIC) and daily urine iodine (UI) excretion.[26,27]

Category of iodine nutrition	Median UIC (µg/L)	Median UI excretion (µg/day)
Severe deficiency	< 20	< 30
Moderate deficiency	20-49	30-74
Mild deficiency	50-99	75-149
Adequate intake	100-199	150-299
More than adequate intake	200-299	300-449
Excessive intake	>299	>449

risk.[30] It mainly enters the human body via inhalation of hydrogen cyanide gas produced during the combustion process of nitrogen containing compounds.[30] There are two major pathways (detoxified and acute toxicity) associated with cyanide exposure through tobacco smoking.[31] The human body is able to detoxify about 0.017 mg of cyanide per kg of body weight per minute without the aid of therapeutic measures.[31] The main detoxification pathway involving low concentrations of cyanide to thiocyanate occurs via a sulfuration process with the presence of thiosulfate by mitochondrial rhodanese.[30,31] Although this mechanism is highly effective, it is incapable of converting high concentrations of cyanide to thiocyanate due to the reduction of sulfur donors.[31] In the event of acute poisoning, cyanide metabolism via endogenous pathways are immediately affected with the abundant levels of cyanide exposure. Consequently, cyanide concentrations accumulate and this leads to the advancement of toxic effects due to the deficiency of sulfur donors. Through this pathway, cyanide is known to inhibit the activity of cytochrome c oxidase with complex IV containing a/a3-domain, which is a proton pump responsible for producing energy in the form of adenosine 5'-triphosphate (ATP).[32] This process leads to an inhibition of the electron transport chain, ultimately resulting in loss of aerobic metabolism and reduced production of ATP.[31] Towards the final steps of the energy production process, cytochrome c oxidase oxidizes the converting reactive species, H^+ and O_2, to water.[32] However, the inhibition of cytochrome c oxidase can lead to

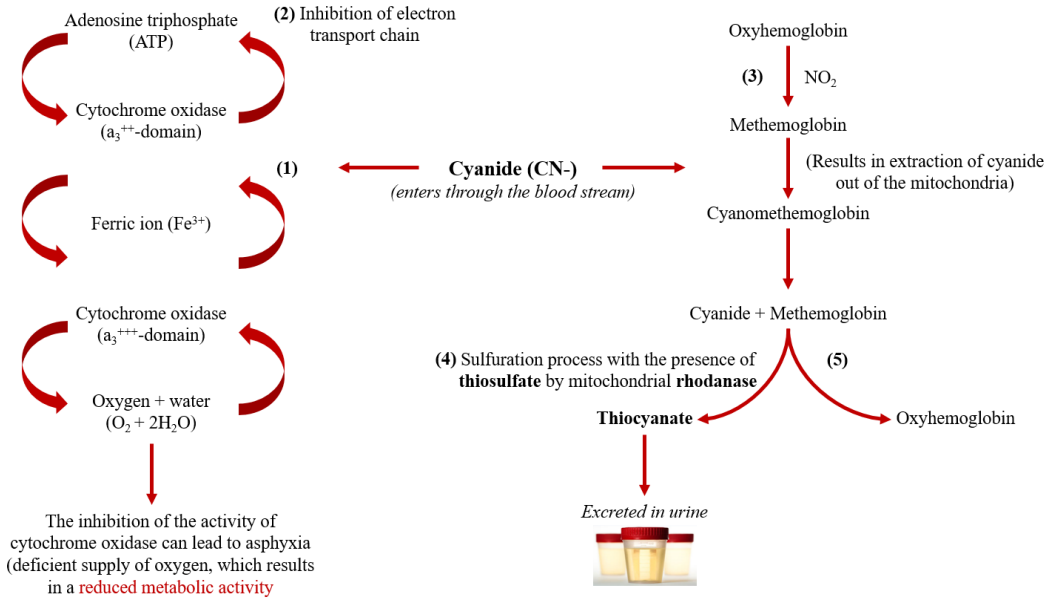


Figure 1.1. Cellular respiration suddenly comes to an end when the cytochrome oxidase initiates its binding pattern within the mitochondria. **(1)** Step 1 illustrates the initial binding of cyanide in the bloodstream to the ferric ion (Fe^{3+}) in the cytochrome a_{3+} complex. **(2)** Step 2 displays the inhibition of the electron transport chain, resulting in loss of aerobic metabolism and a reduction of the formation of adenosine triphosphate (ATP). **(3)** Step 3 shows nitrites (NO_2) reacting with oxyhemoglobin to produce methemoglobin, which results in the extraction of cyanide out of the mitochondria, ultimately forming cyanomethemoglobin. **(4)** Step 4 illustrates the occurrence of the sulfuration process through the presence of the enzyme rhodanase that combines thiosulfate with cyanomethemoglobin to produce a less toxic version of thiocyanate that is excreted through the urine. **(5)** Step 5 displays methemoglobin essentially converting back to oxyhemoglobin by the methemoglobin reductase enzyme. Adapted from Gracia *et al.*[31]

asphyxia, or also known as a deficient supply of oxygen to the human body, which results in a reduction in metabolic activity and possibly apoptosis.[33] An overview of the major pathways involved in cyanide detoxification and intoxication are shown in **Figure 1.1**. As a result, the quantification of thiocyanate is important for the biomonitoring of current tobacco use, including passive smoke exposures among never smokers.

Table 1.3. The adjusted geometric means (AGM) with 95% confidence interval (95% CI) for urinary thiocyanate concentrations (μg/L) and categorized by age, sex, ethnicity and smoking status. The data was obtained from the National Health and Nutrition Examination Survey 2005-2012.[36]

Urinary SCN-	Age 12-19 years		Age ≥ 20 years	
	GM (95% CI)	Sig.	GM (95% CI)	Sig.
M	1174.4 (1011.3-1363.8)		1024.2 (950-1104.2)	M<F (p<0.01)
F	1061.4 (949.5-1186.4)		1251.4 (1174.6-1333.3)	
NHW	1118 (966.7-1292.9)	NHW<NHB (p=0.047)	1305.1 (1208.9-1408.8)	NHW>NHB (p=0.04); HNW<MA (p<0.01); NHW<OTH (p<0.01)
NHB	1339.8 (1173.6-1529.6)	NHB>MA (p<0.01)	1217.9 (1116-1329.1)	NHB>MS (p<0.01); HNB>OTH (p<0.01)
MA	998.7 (821.8-1213.7)		988 (92.4-1055.8)	
OTH	1038.5 (822.9-1310.6)		1046.2 (955.6-1145.4)	
SM	1122.9 (1086.5-1160.5)		1277.7 (1200.1-1360.2)	SM>NSM (p=0.01)
NSM	1110 (845.6-1457.1)		1003.2 (866-1162.2)	
ETS	1042 (844.8-1285.3)		1215.6 (1086.2-1360.6)	ETS>NETS (p=0.02)
NETS	1196.2 (1039.3-1376.7)		1054.4 (991.2-1121.6)	

M = male; F = female; NHW = Non-Hispanic White; NHB = Non-Hispanic Black; MA = Mexican American; OTH = Other unclassified ethnicities; SM = Smoker; NSM = Non-smoker; ETS = Exposed to environmental tobacco smoke; NETS = Not exposed to environmental tobacco smoke; Sig. = Statistical significance

Thiocyanate can be measured in various biological specimens, including blood, urine, sweat or saliva.[29] However, the sensitivity and stability of thiocyanate is dependent upon the specific biofluid, as well as smoking intensity for active smokers or passive exposure from second-hand smoke (SHS) exposure.[29] Various type of tobacco smoke, including mainstream (arises from the mouth after exhaling the cigarette puff) and sidestream smoke (smoke that enters the environment from the end of the cigarette) produce environmental tobacco smoke (ETS) or SHS.[34,35] A study performed by Jain[36] investigated the trends and variability of many inorganic anions, including urinary thiocyanate collected from the National Health and Nutrition Examination Survey (NHANES) that spanned

from 2005-2012. The trends were analyzed based on age, sex, ethnicity, smoking status and exposure to ETS as summarized in **Table 1.3**.[36] A total of 17,220 participants from the United States were examined, including 4,034 participants aged 12-19 years of age and 13,186 were aged ≥ 20 years.[36] Also, urinary thiocyanate concentrations were reported for different ethnicities, including non-Hispanic White (NHW), non-Hispanic Black (NHB), Mexican American (MA) and other ethnicity (OTH).[36] Overall, current smokers had significantly higher levels of urinary thiocyanate compared to non-smokers in the ≥20 years old category (p = 0.01).[36] Furthermore, males ≥20 years of age have lower thiocyanate levels compared to females (p < 0.01),[36] and ETS exposure was associated with increased levels of urinary thiocyanate (p = 0.02).[36]

1.1.4 Sweat Chloride for Confirmatory Diagnosis of Cystic Fibrosis

Cystic fibrosis is one of the most common inheritable childhood-onset and life-threatening autosomal recessive disease in the Caucasian population impacting about 1 in 3600 live births in Canada.[37–39] The disease is caused by a mutation in a gene located on the long arm of human chromosome 7, specifically in the q31-q32 region.[40] Cystic fibrosis is most widely associated with an impairment in electrolyte transport/permeability across the cystic fibrosis transmembrane conductance regulator (CFTR) channel.[41,42] The CFTR protein functions as a channel that conducts chloride ions and other anions across the epithelial cell membranes of the sweat glands, pancreas and lungs,[43] which are responsible for producing sweat, digestive enzymes and mucus. Although more than 2,000 CFTR mutations have been reported, CF is often associated with delF508, protein deficiency and loss in CFTR function resulting in impaired chloride transport and increased viscosity and stickiness of mucus.[43] Confirmed CF patients show increased levels of chloride in their sweat.[44] In this case, a defect in the CFTR channel leads to an inhibition of the reuptake of chloride in sweat glands, which results in an abnormal elevation in chloride levels in sweat.[44] Elevated sweat

chloride is a characteristic physiological measure of CF, which is used for confirmatory testing and diagnosis of presumptive CF patients.[45] A severe loss in anion permeability across the CFTR channel can result in various health complications that manifest early in life, such as malnutrition, acute or persistent respiratory symptoms, pancreatic exocrine insufficiency and intestinal obstruction.[41,42] However, early detection and prompt intervention of CF affected infants can lead to better growth and development (*e.g.*, nutritional status), improved lung function and a longer average life expectancy.[46]

For these reasons, newborn screening (NBS) for CF consists of a two-tiered screening algorithm followed by a confirmatory sweat chloride test, which represents the gold standard for CF diagnosis, as illustrated in **Figure 1.2**.[47] The first tier for CF screening of neonates consists of an immunoreactive trypsinogen (IRT) test, which contains poor specificity with an extremely low positive predictive value (PPV) ranging from 3-10%.[48] Infants with elevated IRT levels (>99.8%), which may or may not be specific to CF, proceed to the second tier screen comprising a panel of CFTR disease-causing mutations that differ between juridications.[49] Although the most common CFTR mutation in CF, delF508 is prevalent in about 70% of all CF cases, there is a significant difference in the frequency of specific homozygous or heterozygous mutations across ethnicities.[48] In general, presumptive screen-positive infants with elevated IRT along with detection of one or two CFTR mutations then proceed to a sweat chloride confirmatory test at a regional pediatric hospital. The sweat chloride test was first introduced in the 1950s which was conducted via pilocarpine-stimulated iontophoresis as a standardized protocol to collect small volumes of sweat (> 20 μL) topically on the forearm of infants.[51] Measurements of sweat chloride concentrations are performed using a chloridometer via conductivity detection or coulometric titration as a confirmatory test on all presumptive screen-positive CF infants. In this case, elevated chloride concentrations (≥ 60 mmol/L) is indicative

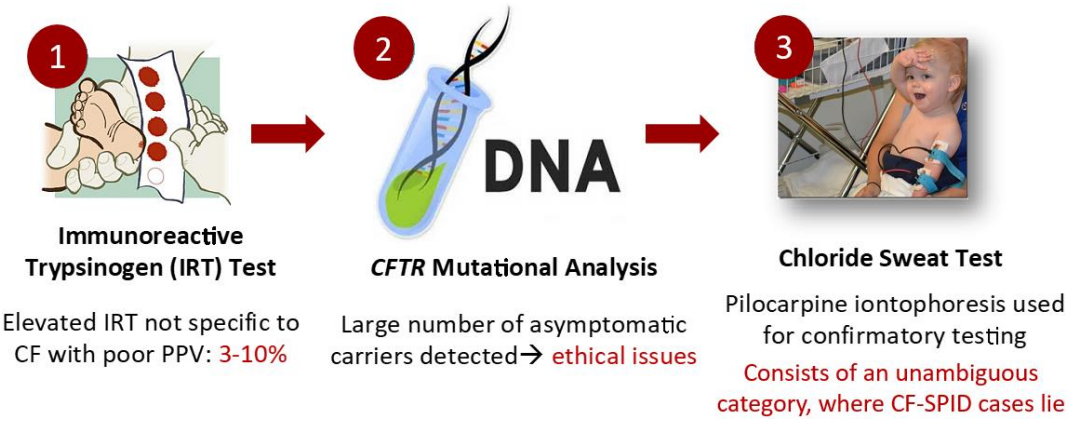


Figure 1.2 An overview of the newborn screening algorithm for cystic fibrosis (CF). The two-tiered screen consists of an immunoreactive trypsinogen (IRT) immunoassay screen, followed by a CFTR mutational analysis panel. If the infant has elevated IRT (> 96th percentile) and 1 to 2 disease-causing CFTR mutations, they proceed for a third-tier confirmatory sweat chloride test via pilocarpine-stimulated iontophoresis. Adapted from Wilcken *et al.*[48] and Munck *et al.*[50]

of CF, whereas chloride concentrations < 30 mmol/L suggest the chance of being diagnosed with CF is unlikely (*e.g.,* unaffected carrier or false-positive).[52,53] However, infants with sweat chloride concentrations ranging from 30-59 mmol/L have an ambiguous diagnosis and are referred to as cystic fibrosis-screen positive inconclusive diagnosis (CF-SPID) cases.[50] A subset of CF-SPID infants may have a CF diagnosis later into childhood (CF-SPID → CF), which contributes to significant parental anxiety and uncertainty.[54,55] Recently, Gonska *et al.*[56] performed a longitudinal study over 7.7 years (95% CI of 7.1-8.4 years) demonstrating that 24 out of 115 (≈ 21%) children, who were initially classified as CF-SPID (≥ 40 mmol/L), were later reported to having elevated sweat chloride levels (≥ 60 mmol/L). As a result, new prognostic biomarkers that can may better predict the risk for late onset CF among ambiguous screen-positive infants are urgently needed.

1.2 Capillary Electrophoresis Separation Theory and Application in Clinical and Epidemiological Studies

Capillary electrophoresis (CE) represents a versatile microseparation platform, which is applicable to the resolution of a wide range of ions in aqueous medium.[57] Moreover, CE requires minimal sample preparation due to its high tolerance for complex matrices, which is ideal for the analysis of small volumes (in the nanoliter range) of samples.[58] The application of CE to volume-limited samples may be beneficial in clinical settings. For example, when performing a sweat chloride test for CF using a Macroduct device, a minimum sweat volume of 15 μL is required.[59] However, samples with smaller volumes are labeled as "quantity not sufficient" (QNS) and are prevalent in infants younger than three months who have experienced conventional sweat tests.[60,61] More advantages of CE include rapid analysis, high separation efficiency, low operating costs and prediction of migration behaviour.[57,62] Additionally, it is possible to perform in-capillary sample enrichment procedures without instrument configuration to increase sensitivity.[63] For example, online sample preconcentration methods enable a sample self-stacking process, where analytes consisting of low concentrations in a long injected sample zone are concentrated into a short zone (stack).[64,65] The stacked analytes are then separated and individual zones are measured.[65] For example, transient isotachophoresis (ITP) stacking is frequently used in practice and is usually created by a bulk sample analyte (*e.g.*, chloride present in urine) that plays the role of a transient leader or terminator.[63,64] The stacking of the analytes is temporary and disappears as the bulk sample undergoes electromigration dispersion.[63] This results in a lower migration velocity of the analyte of interest and a decrease in the length of its zone (shorter zone leads to a more concentrated analyte of interest).[63] The advantages of transient ITP stacking include the possibility of injecting large sample volumes without the requirement of low conductivity of the sample zone, resulting in sub-micromolar analyte concentration enhancements, even in a macrocomponent matrix.[63,64] Therefore,

sample self-stacking provides improved separation efficiency and detection sensitivity, while maintaining low ion suppression especially in high-salt samples (*e.g.*, urine).[65,66] However, disadvantages of CE include poor concentration sensitivity, variation in migration time and limited inter-laboratory comparisons.[57,67,68]

1.2.1 Brief Overview of Capillary Electrophoresis Theory

In CE, each ion in a standard solution or sample is separated based on its effective charge and hydrodynamic radius (*e.g.*, size) when an electric potential is applied across a silica-fused capillary filled with background electrolyte (BGE).[69] The two electrodes located at each end of the capillary that are immersed in the buffer reservoirs/vials are connected to the power supply that applies a voltage across the capillary, which initiates the separation process.[70] The two electrodes are referred to as either the cathode or anode, which can switch depending on what polarity (normal or reversed) is assigned to the separation. A sample is commonly injected via a hydrodynamic pressure through the capillary for a fixed time or ions in the sample are introduced electrokinetically after a voltage has been applied across the capillary. The electrophoretic mobility is a fundamental physicochemical property of an ion in a specific BGE that is independent of electric field strength.[70] Faster mobility occurs for ions having a higher charge and smaller hydrodynamic radius.[70] The charge of the ion is fixed for fully dissociated ions, such as strong acids or small ions, whereas for weak acids or bases, the effective charge is dependent on the buffer pH.[70] Furthermore, the ionic strength of the BGE impacts the apparent electrophoretic mobility by forming an electric double layer surrounding an ion in solution, where higher ionic strength reduces ion mobility. The electrophoretic mobility of an ion is also dependent on the temperature or viscosity of the solution as described below for a spherical/uniformly charged ion:

$$\mu_{ep} = \frac{q}{6\pi\eta r} \tag{1.1}$$

Where q represents the charge of the ion, η is the velocity of the buffer and r refers to the radius of the ion.[70]

Another important factor to consider in CE is the electroosmotic flow (EOF), which represents the bulk flow of the BGE through the capillary.[70] The surface of the uncoated silica-fused capillary consists of ionisable/weakly acidic silanol groups, which are in contact with the BGE.[70] The negatively charged inner wall of the capillary attracts positively charged ions from the BGE and ultimately produces an electric double layer (rigid and diffuse) and a potential difference in close proximity to the inner capillary wall (*e.g.*, zeta potential).[70] Assuming a flat charged surface, the magnitude of this EOF can be expressed using the Smoluchowski equation:

$$\mu_{app} = -\frac{\varepsilon\zeta}{\eta} \tag{1.2}$$

Where ε represents the dielectric constant, η is the viscosity of the buffer and ζ refers to the zeta potential.[71] The main advantage of the EOF includes its flat flow profile characteristic, which leads to higher peak efficiency, better resolution and reduces band broadening from mass transfer compared to hydrodynamic flow systems, such as high-performance liquid chromatography (HPLC).[72] Furthermore, migration times in CE are inversely proportional to the apparent mobility, which is the vector sum of electrophoretic mobility and EOF.[73] Disproportionate shifts of migration times are a result of an alteration to the EOF as described in **Equation 1.2**.[71,73] Cations and anions are separated based on variations in their apparent mobilities.[74] The bulk of the BGE is carried with the cations located in the diffuse layer, which are free to migrate towards the cathode.[70] The result is a net flow of the cations in the diffuse layer and the bulk solution of the BGE migrate towards the cathode.[70] Cations migrate in the same direction as the EOF, whereas anions migrate in the opposite direction of

electroosmosis.[74] The neutral solutes migrate in the same direction and velocity as the EOF and do not undergo separation.[74]

1.2.2 Capillary Electrophoresis with UV Absorbance Detection

On-line photometric detection methods are frequently used in CE separations to quantify the concentration of an ion with the support from a calibration curve. There are many types of detection methods that are commonly coupled to CE, including direct ultraviolet (UV) absorbance and indirect ultraviolet (iUV) detection, as well as laser-induced fluorescence and electrospray ionization-mass spectrometry (ESI-MS).[75] Direct UV detection is used when UV absorbing analytes are measured corresponding to their peak absorbance wavelength.[76] Although CE with direct UV absorbance when performed using a photodiode array detector provides spectral information about the putative identity and purity of an analyte, it suffers from poor concentration sensitivity in due to its narrow optical pathlength and small sample injection volumes with a limit of detection (LOD) ranging from 10^{-5}-10^{-7} mol/L depending on molar absorptivity.[75] Regardless of the poor concentration sensitivity of classical direct UV absorbance detection, most studies are performed with this type of detector mainly due to its low-cost, robustness and compatibility with CE instrumentation with little to no modications.[75] As a result, studies with analyte concentrations ranging from μmol/L to mmol/L are generally performed using this detector.[75] [75] In-capillary sample concentration techniques in CE may be used to further lower detection limits for analysis of dilute analytes, such as transient isotachophoresis, dynamic pH junction, sweeping or sample self-stacking.[64,77] Furthermore, indirect photometric detection is a universal, inexpensive and simple alternative for the determination of non-absorbing UV analytes.[76,78] It requires the addition of a UV absorbing co-ion (an ion with the same charge sign as the analyte of interest) to the BGE that undergoes a displacement by a non-absorbing analyte to generate a negative signal or peak.[78] The absorbing co-ion is referred to as a probe or

chromophore.[78] The mobility and concentration of the probe have important effects on the resolution and peak symmetry of CE separations when using indirect UV detection.[78] The molar absorptivity of the probe impacts concentration sensitivity and should be kept as high as possible at an optimal wavelength.[78] The LOD, which is generally 1-2 orders of magnitude higher than direct UV absrobance[75] is dependent on the baseline noise (N_{BL}), transfer ratio (TR), which corresponds to the number of probe ions displaced by a single analyte ion, the molar absorptivity of the probe (ε), and the effective pathlength (l) of the detection cell:[78]

$$\mathrm{LOD} = \frac{N_{BL}}{TR\varepsilon l} \tag{1.3}$$

The LOD is independent of the concentration of the probe. Thus, the lowest LOD can be achieved by minimizing N_{BL} and maximizing TR, ε and l. Additionally, the sharpest and most symmetrical peaks are achieved when the value of the electrophoretic mobility of the probe matches the analyte and the difference between the analyte concentration in the migrating zone and the probe concentration in the BGE is maximized (*e.g.*, lowest possible sample injected and highest possible concentration of the probe is used in the BGE).[78] However, increasing the concentration of the probe in the BGE increases the background absorbance, which results in an increase in the baseline noise leading to spectral interferences or system peaks.[78]

1.2.3 Capillary Electrophoresis and Mass Spectrometry

The analysis of complex sample mixtures can be achieved by combining the high separation efficiency of CE with the sensitivity and selectivity of mass spectrometry (MS).[75] In this case, the analytical separation performance is enhanced, as CE enables the separation of ions that cannot be mass-differentiated (*e.g.*, isomers and isobars), while high resolution MS provides a second separation dimension for ions that may co-migrate based on differences in their mass-to-

charge ratio (*m/z*) in the gas-phase.[57] However, the analyses are limited by the choice of ion source and its operating conditions required to generate gas-phase ions for subsequent mass resolution.[75] Unlike on-line photometric detection techniques, MS requires a compatible ion source with the most commonly used interface being electrospray ionization (ESI).[57] ESI-MS is a *soft* ionization technique that produces a minimum number of fragments and therefore is suitable for the analysis of labile molecules.[79] ESI utilizes a voltage application to aid the transfer of solvated ions from solution into the gas-phase for MS analysis.[80,81] During solute ionization, the sample solution is infused into the ESI interface by a heated capillary.[79] With a high voltage applied to the solution, a spray of charged droplets will initially be produced at the tip of the capillary located in the ESI interface.[79] As the ion source temperature exceeds the boiling point of the solvent, the solvent evaporates steadily into smaller charged droplets to facilitate solute desorption as a gas-phase ion for subsequent mass resolution (*e.g.*, time-of-flight, TOF).[79] Additionally, the stability of the spray formation can be improved by incorporating a flow enhancing sheath liquid to the BGE as soon as it exits the capillary.[82] The electrical contact between the electrode and CE effluent can be achieved when the sheath liquid is mechanically (*e.g.*, through an isocratic pump) or electrokinetically pumped.[82] Both the analyte ionization process and mass compatibility of the separation buffer benefit by this application.[82]

Capillary electrophoresis-mass spectrometry (CE-MS) frequently uses a coaxial sheath-flow liquid ESI interface design,[82,83] where the capillary is inserted into two larger diameter stainless steel tubes configured coaxially and electrically grounded as shown in **Figure 1.3**.[83,84] The flow rate of the sheath liquid solution located in the inner coaxial tube ranges from 1 to 10 μL/min range, which provides a constant solvent composition for solute ionization during the separation.[83] The outer coaxial tube is used for a nebulizer gas to further facilitate droplet formation in ESI-MS.[83] Although this type of interface can decrease sensitivity by diluting the concentration of the analyte, it contributes to overall

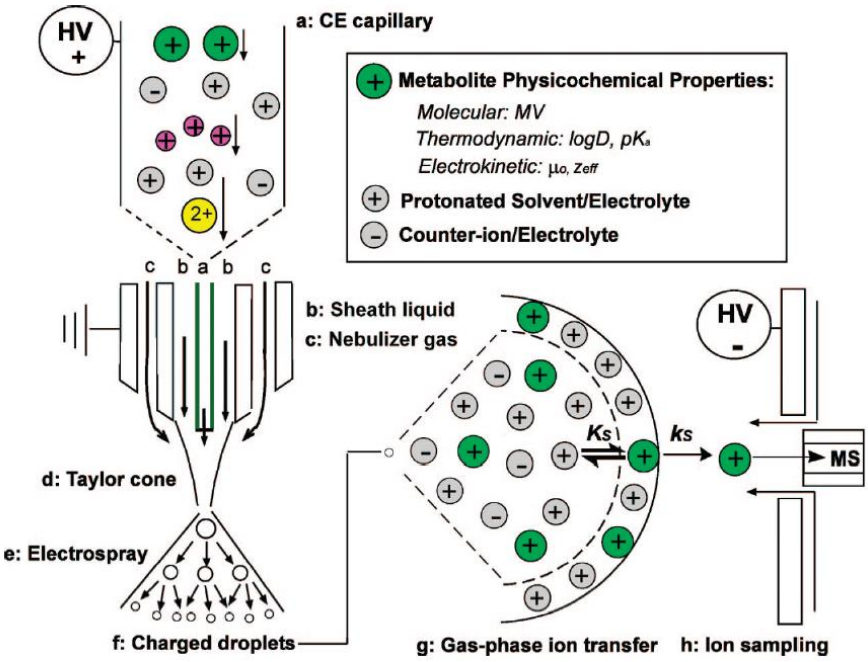


Figure 1.3. Schematic of a coaxial sheath liquid interface for CE-MS that provides a stable spray formation for ion desorption into the gas-phase without significant matrix-induced ion suppression effects. Various thermodynamic and electrokinetic factors influence solute ionization efficiency in ESI that results in wide variations in sensitivity in ESI-MS when analyzing different classes of polar metabolites. Adapted from Chalcraft *et al*.[84]

robustness by providing a more stable spray that minimizes potential ion suppression or enhancement effects from the sample matrix.[84] Furthermore, coaxial sheath-flow ESI interface often provide more robust and reproducible analyses in comparison to low-flow/sheathless nanospray devices and liquid junction interfaces,[82] as well as sheathless interfaces that exhibit unstable spray formation in CE due to their very low flow rates.[85] However, recent advancements in low-flow/sheathless interfaces coupled with pre-column chemical derivatization techniques can greatly enhance sensitivity for expanding metabolome coverage.[86] Kuehnbaum *et al*.[86] introduced multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) as a high throughput platform for metabolomics with high data fidelity. The method offers a

multiplexed separation platform that increases sample throughput by an order of magnitude, while implementing coaxial sheath-flow ESI interface that provides a homogenous solution for stable solute ionization under steady-state conditions.[84,86] Therefore, MSI-CE-MS with a coaxial sheath-liquid flow interface offers a promising high throughput platform for targeted and/or non-targeted analysis of a diverse range of polar/ionic metabolites in complex biological samples.

1.3 Overview of Biomarker Discovery in Metabolomics

1.3.1 Study Design of Metabolic Workflows

There is a growing concern of a crisis in reproducibility with many scientific publications reporting false discoveries.[87] The chance that a research finding is true may depend on two key factors: study power and bias.[87] However, current metabolomic studies have advantageously gained from the advancements made in study design that account for sources of bias.[88] Thus, the need for an optimal study design for metabolomic workflows, standardized protocols, and analysis of quality control (QC) samples using validated analytical methods is extremely crucial.[88] An effective study design of metabolomic workflows will aid in authenticating biomarker discovery for clinical and epidemiological research.[88] For example, **Figure 1.4** illustrates an overview of an accelerated workflow for biomarker discovery for metabolomics involving cystic fibrosis (CF). The workflow is as follows: (1) Experimental design, which is comprised of diseased *vs.* healthy patients; (2) Dilution trend filter of a pooled sample, which is helpful for untargeted metabolite characterization and rejection of spurious peaks, adducts and background ions;[89] (3) High-throughput analysis and targeted metabolite profiling using MSI-CE-MS; (4) Statistical analysis and data visualization; (5) Biological interpretation following unambiguous metabolite identification and quantification.

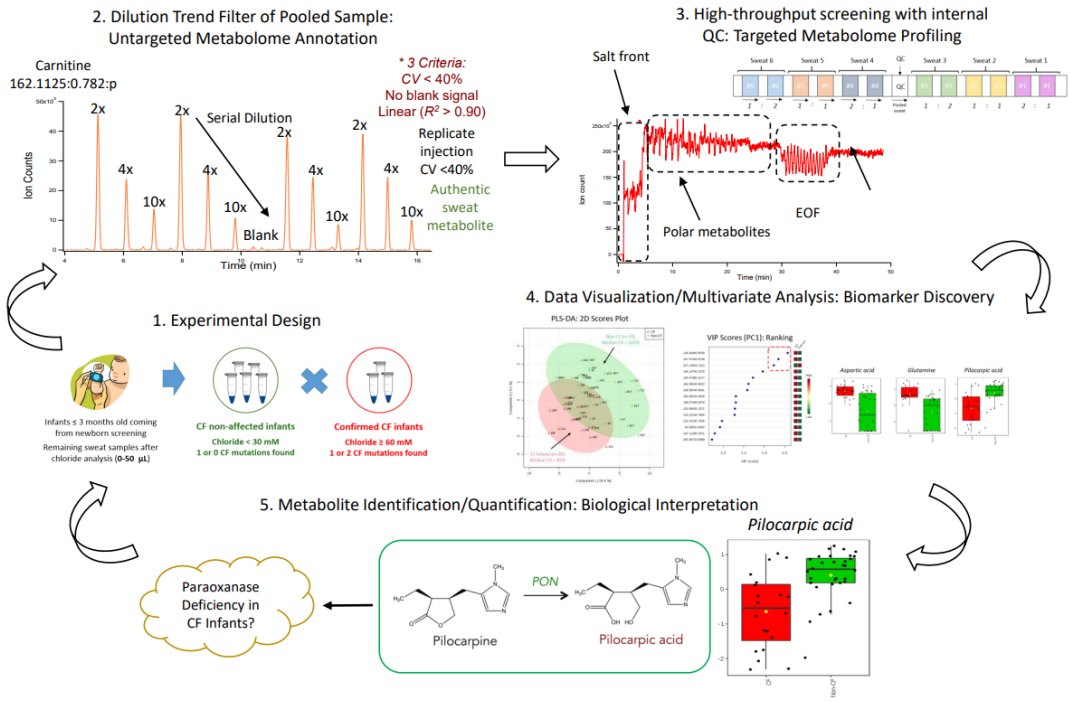


Figure 1.4. An overview of an accelerated data workflow for biomarker discovery in metabolomics for improved screening and/or diagnosis of CF infants.

1.3.2 Sample Collection, Preparation and Storage

There are many sources of variation in metabolomic studies that are direct consequences of the pre-analytical processes.[90] Pre-analytical steps are important in metabolomics study designs, including sample collection, sample workup, and storage, which can have a significant impact on the integrity of the sample, such as metabolite stability.[90] The most frequent sources of variation in pre-analytical processes of a biological specimen (*e.g.*, urine, blood, feces, tissue and cells) have three commonalities.[90] Firstly, volume or weight is important to consider prior to sample collection.[90] Secondly, biological samples should be processed and stored at freezing temperatures (-80 °C) shortly after collection to prevent chemical degradation and enzymatic activity.[90] Lastly, aliquoting samples from their original vials are highly recommended upon collection into separate sample vials to avoid repeat freeze-thaw cycles that could lead to a loss of sample integrity.[90]

1.3.3 Targeted Metabolite Characterization

Metabolomics can be characterized as a comprehensive approach in the analysis of all low-molecular weight compounds (< 1500 Da) present in a biological sample.[91] Metabolomics research has become an increasingly promising tool in clinical medicine for biomarker discovery and elucidating metabolic perturbations due to internal and external stimuli or exposures.[92] Although the study of metabolomics lies at the end of the "-omics" cascade in molecular biology (*e.g.*, genomics → transcriptomics → proteomics), it has shown to be sensitive to dynamic changes from both gene expression and environmental factors.[93] Thus, aberrant metabolism is closely associated with both phenotype changes and clinical outcomes.[94] Targeted metabolite analyses consists of extracting known metabolites and have been proven to be thoroughly useful for evaluating response with respect to disease pathology,[95–97] dietary patterns,[10,98,99] xenobiotic exposure,[7,100] environmental factors,[101] and drug discovery and precision medicine.[102,103] In other words, the chemical structure of metabolites are known that can be readily measured with high accuracy and precision using available standards.[110] Metabolomics is defined as the global analysis of both known metabolites and unknown compounds in complex biological samples, which typically uses high resolution separation and detection platforms, such as nuclear magnetic resonance (NMR) spectroscopy, direct infusion-mass spectrometry (DI-MS), liquid chromatography-mass spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS) and CE-MS. The following strengths and limitations of each instrumental platform are outlined in **Table 1.4**. No single analytical platform is capable of analyzing a diverse range of metabolites present in biological samples that vary in their fundamental physicochemical properties (*e.g.*, polarity, molecular weight, stereochemistry).[91] However, the use of complementary techniques is required for comprehensive analysis of the human metabolome (if feasible) depending on several study limitations, such as total sample volume, sample throughput, and operating costs.[111]

Table 1.4. Various analytical platforms used in metabolomics based studies with their corresponding strengths and limitations.

Analytical Platform	Strengths	Limitations
Nuclear magnetic resonance (NMR)[91,104,105,106]	- Robust platform - Non-destructive to sample - Largely automated - Absolute quantification - Chemical identification	- Low/poor sensitivity - Limited 1D resolution - Poor metabolome coverage - Larger sample volumes
Direct infusion-mass spectrometry (DI-MS)[107]	- Rapid and high-throughput - Robust platform - Simpler data processing without separation	- Susceptible to ion suppression - Isomers/isobars not resolved - Complex sample processing required
Liquid chromatography-mass spectrometry (LC-MS)[91,107]	- Broad metabolome coverage - High resolution for resolving isomers/isobars - Reduced ion suppression - High sensitivity - Metabolite identification by MS/MS and retention time	- Lower sample throughput - Different separation mechanism and column types (RP/HILIC) - Susceptible to retention time drift requiring time alignment - Extra time for column conditioning and equilibration
Ultra Performance Liquid Chromatography (UPLC-MS)[91]	- High peak capacity, enhanced specificity and sensitivity - Broader range due to optimum linear velocity - Rapid analysis without loss of resolution - Accurate mass measurement levels of less than 5 ppm	- Increased pressure requires high maintenance - High pressure analysis can result in early column damage and reduced column life
Gas chromatography-mass spectrometry (GC-MS)[91,108]	- High efficiency separations - Reproducible retention times - Extensive EI-MS database for structure identification	- Limited to volatile and thermally stable compounds - Complex sample processing required
Capillary electrophoresis-mass spectrometry (CE-MS)[7,86,91,109]	- High separation efficiency - Ideal for ionic metabolites - Low sample volumes - Simple sample preparation - Multiplexed separations to improve throughput and QC	- Variability of migration times - Poor sensitivity - Limited metabolome coverage if nonaqueous not used for lipids - Limited inter-laboratory comparisons

1.3.4 Data Processing and Statistical Analysis in Metabolomics

Once data acquisition has been completed for samples that were analyzed in a random manner with the inclusion of QC and blank runs, a vast amount of data needs to be properly handled and converted for data analysis. However, a major bottleneck in the "-omics" field, (*e.g.*, metabolomics), is the time-consuming data processing step due to the large amount of data generated.[112,113] In general, data handling can be split into two steps, which include data pre-processing and statistical data analysis. The data pre-processing step consists of transforming the acquired raw data with a peak integration processing software/website (*e.g.*, Agilent MassHunter Qualitative Analysis, mzMine, and Skyline).[114] The key goal of the data processing step is to extract important information from the observed raw signals from the raw data files of each measured ion. This information includes mass-to-charge ratio (*m/z*), retention/migration time, area, height, width, resolution measured by full width at half maximum (FWHM), and signal-to-noise ratio (SNR). The general data processing schematic involves multiple steps, including 1. Filtering (*e.g.*, the removal of background noise), 2. Feature detection (*e.g.*, peak picking authentic signals while rejecting spurious signals and background ions), 3. Time alignment and data smoothing (*e.g.*, required for correcting differences in retention times between runs by clustering peaks), and 4. Normalization (*e.g.*, eliminates undesired systematic variation within runs and can be implemented by incorporating multiple or a single internal standard, IS).[115] Additionally, the running of quality control (QC) and blank samples are important to implement during metabolomics analyses to ensure high quality data, system stability, analytical reproducibility, and monitor sample carry-over/contamination.[116,117] In this case, QC samples are strategically placed in a random manner among the samples per run. The precision of the retention/migration time and the signal intensity of metabolites are recommended to have a relative standard deviation (%RSD) of < 2% for retention/migration times, < 20% for the area of abundant peaks, and < 30% for extremely low

abundant peaks with low signal-to-noise ratio.[117] Peaks that fail these criteria are excluded from the metabolomics dataset to reduce false discoveries.[117]

In metabolomic studies, there is no consensus on the optimal normalization technique to account for concentration variability, notably when analyzing single-spot urine samples.[118] Normalization processes adjust the data for both analytical and biological variability.[118] Concentration normalization techniques may be applied prior to or post data acquisition.[119] Concentration normalization methods implemented prior to data acquisition include, diluting the samples to the same concentration based on osmolality or specific gravity parameters.[118–121] Also, urinary creatinine is commonly used to correct for variations in hydration status, where all the peak intensities of the other metabolites within a sample are normalized to the peak intensity of creatinine.[122] However, creatinine excretion levels are affected by several other factors, such as age, sex, weight, diet, exercise, muscle mass, body mass index (BMI), ethnicity and health conditions involving renal function.[118,123–132] Since creatinine excretion levels vary based on several factors, it is not an ideal normalization method. Prior to multivariate analysis, additional data pretreatment methods may be required to improve biological interpretability by correcting for conditions that impede the biological meanings/significances of metabolomics datasets, such as drastic differences in metabolite concentrations.[133] Examples of data pretreatment techniques include log transformation, power transformation, centering, autoscaling, Pareto scaling, range scaling and vast scaling.[133] Each pretreatment step contains its own strengths and limitations as each step or a combination of steps highlight various portions of the data.[133]

The statistical analyses chosen to be conducted is dependent on the complexity of the dataset. The type and distribution of the original dataset dictates which statistical test is best suited for univariate analysis. Generally, for non-normally distributed data, non-parametric tests are applied to the dataset,

including Wilcoxon Rank sum test (compares two dependent samples), Mann-Whitney U test (compares two independent samples), Spearman correlation (measures the degree of association between two variables) and Kruskal-Wallis test (compares three or more groups on a dependent factor). Equivalently, if the data is normally distributed, parametric tests are applied to the dataset, such as Paired t-test, Unpaired t-test, Pearson correlation, One way Analysis of variance (ANOVA) and Analysis of covariance (ANCOVA) using covariate adjustments to correct for variable confounding. For ANOVA and ANCOVA analysis, *post hoc* tests are required if a significance is determined in the overall analysis. Multiple hypothesis testing correction methods is important for metabolomics datasets, as they filter out false positive cases. The simplest and most widely used test for multiple testing correction in metabolomics is the Bonferroni adjustment.[134] The Bonferroni correction uses the significance threshold (usually at $p = 0.05$ or 95% level) and divides it by the number of statistical tests conducted.[135] Although the Bonferroni adjustment is widely used, it is considered to be a conservative approach.[135] Bonferroni minimizes false positives (Type I error); however, it magnifies the risk of false negatives (Type II error).[135] That being said, controlling the false positive rate can lead to inconsistencies that can negatively impact significant findings.[136] As an alternative, the false discovery rate (FDR) estimation has shown to be less conservative and can control false positive error rates.[134,136] The application of FDR have been implemented in multivariate analysis involving metabolomics related clinical and epidemiological research.[7,10,137] Furthermore, the most commonly used classification techniques in metabolomics data processing include principal component analysis (PCA, unsupervised approach) and partial least squares-discriminant analysis (PLS-DA, supervised analysis) or orthogonal partial least squares-discriminant analysis (OPLS-DA).[138] These types of multivariate analysis techniques explore the clustering and variability between groups.[138] In summary, univariate and/or

multivariate statistical techniques are routinely utilized in metabolomics research for biomarker discovery, including classification and discrimination.[139]

Although metabolomics is increasingly being studied and reported for disease diagnosis, prognosis and risk prediction, many metabolomic studies show minimal consistency on how the significant biomarkers are evaluated.[140] Additionally, only a few studies address the sensitivity, specificity or provide receiver-operating characteristic (ROC) curves with corresponding confidence intervals to support the potential and utility of significantly determined metabolites.[140] A ROC curve is a plot that illustrates the diagnostic ability of a binary classification system (e.g., predictive positive: disease vs. predictive negative: healthy) as its discrimination threshold is varied.[140] The true positive rate (sensitivity) is plotted against the *y*-axis and the false positive rate (1-specificity) is plotted against the *x*-axis and are calculated for each concentration value of the metabolite.[140] Each concentration value equal or above the threshold are considered to be positive (diseased) or below the threshold are considered to be negative (healthy control).[140] Moreover, ROC curves generate a value for the area under the curve (AUC), which represents the degree or measure of discrimination. A simple guide to evaluating the utility of a biomarker according to its AUC is as follows: 0.9-1.0 excellent, 0.8-0.9 → good, 0.7-0.8 → fair, 0.6-0.7 → poor and 0.5-0.6 → fail.[140] Receiver-operating characteristic curves are an adequate tool to assess the diagnostic validity of potential biomarkers required in clinical settings that may result in a positive contribution to improved health outcome and/or quality of life.

1.3.5 Untargeted Metabolite Characterization and Identification

Untargeted metabolite analysis is implemented for hypothesis generating research, since targeted methods may overlook potential novel and significant biomarkers of clinical significance. Although untargeted metabolomics is challenging, as it requires extensive data pre-processing and unambiguous

structural elucidation, it has proven to be a powerful discovery approach that leads to a deeper understanding of cellular metabolism and its interaction with complex exposures.[141] When utilizing DI, LC or CE with ESI-MS, untargeted metabolomics can be performed under both positive and negative ionization modes with full-scan data acquisition. In other words, unknown metabolites are required to be able to generate positively and/or negatively charged species(*e.g.*, $[M+H]^+$, $[M+H]^-$) among other combinations, such as in-source fragment ions or salt adducts.[141] It is also encouraged that certain chemicals be avoided when preparing samples when using ESI-MS, as they can cause an interference by behaving like a contaminant.[142] Also, if these compounds ionize, they can cause ion suppression, which can lead to an increase in background noise.[141] Examples of these compounds include plasticizers and surfactants, such as sodium dodecyl sulfate (SDS) and polyethylene glycol (PEG).[141] In MS-based metabolomic studies, established algorithms are implemented when screening for potential compounds during an untargeted analysis, such as putative metabolite identification based on accurate mass search in a public database.[143] Peak redundancy, where a metabolite generates multiple MS signals, such as adducts, dimers, isotopes, or in-source fragments can contribute to data overfitting,[144] as well as potential false annotations.[145] However, these unwanted features comprising redundant co-migrating ions can be filtered out when using a temporal signal pattern recognition in MSI-CE-MS.[89]

In 2007, the Chemical Analysis Working Group (CAWG) of the Metabolomics Standards Initiative (MSI) had suggested a guide of minimum reporting standards in metabolomics based on four levels of confidence for the identification of metabolites.[146] However, Schymanski *et al.*[147] recommended an alternative tiered guide, which included a fifth level in the minimum reporting standards in 2014. Consequently, the Metabolite Identification task group of the Metabolomics Society proposed that it was "time to reassess the current reporting standards for metabolite identification and identify any changes required because

of recent developments".[148] Levels of metabolite identification confidence using MS include: Level 1 dictates an identified metabolite with a confirmed chemical structure; level 2 signifies a putative identified metabolite with a probable structure; level 3 represents a putatively annotated metabolite class with a tentative candidate structure; level 4 dictates an unknown metabolite with an unequivocal molecular formula; level 5 represents a unique molecular feature with an exact mass of interest together with retention or migration time.[147,149] Levels 1-3 can require the use of a spectral library, database or software search as confirmation, which can be performed on multiple online platforms, including KEGG, HMDB, ChemSpider, and PubChem.[149] Levels 1-3 can also require the utilization of tandem mass spectrometry (MS/MS) with collisional-induced dissociation when confirming the level of metabolite identification confidence.

1.4 Thesis motivations and objectives: Biomarker Analysis by Capillary Electrophoresis for Clinical and Epidemiological Studies

Research in biomarker discovery have been essential in clinical medicine and population health studies. However, metabolomics still has significant challenges related to biomarker validation and successful clinical translation. Furthermore, large-scale epidemiological studies are rarely conducted to accurately assess nutritional status and biomonitoring of toxic compounds in diverse populations worldwide. Therefore, the work in this thesis includes an accurate, sensitive, and robust CE-UV for the targeted analysis of specific inorganic anions in complex biological samples. The first project includes the analysis of inorganic anions present in urine that are relevant to human health and dietary patterns across Canada (n = 800), which include iodide and two environmental iodine uptake inhibitors, nitrate and thiocyanate (*Chapter II*). A modified CE-UV technique with improved resolution for thiocyanate was next performed for the biomonitoring of tobacco smoke exposure in adults (n = 1000) in 14 countries at different socioeconomic status (*Chapter III*). Moreover, a novel and rapid targeted

CE-iUV assay had been developed and validated for the determination of sweat chloride and bicarbonate from presumptive CF infants (*Chapter IV*). Lastly, an untargeted approach using MSI-CE-MS is utilized to identify sweat biomarkers that could discriminate between CF affected and unaffected infants, which may provide new insights into disease progression in indeterminant cases.

1.4.1 The Prevalence and Risk Factors Associated with Iodine Deficiency in Canadian Adults

The consequences of iodine deficiency have been widely perceived for many decades as it is a modifiable risk factor to a plethora of iodine deficiency disorders, including endemic goiter, intellectual and growth impairment, and hypothyroidism.[17,150,151] Although the universal salt iodization program has been implemented in Canada since the 1920s,[150] continuous monitoring of iodine nutrition is important due to changes in environmental chemical exposures and dietary patterns on a population level. *Chapter II* of this thesis focuses on risk and protective factors associated with iodine deficiency in 24 h urine samples collected from 800 participants as a part of the Prospective Urban and Rural Epidemiological (PURE) study. This large-scale study takes advantage of the sample self-stacking performance offered by CE-UV that allows for sub-micromolar detection of urinary iodide, as well as more abundant iodine uptake inhibitors, nitrate and thiocyanate.[152] Our main findings indicate regional variations in iodine status were revealed with participants from Quebec City and Vancouver at greater risk for iodine deficiency compared to Hamilton and Ottawa due to specific dietary patterns and chemical exposures. Overall, supplements containing iodine, thyroxine (T4) prescription, urinary sodium excretion, and dairy intake were determined to be protective factors against iodine deficiency.

1.4.2 Validation of Urinary Thiocyanate as a Robust Biomarker of Active Tobacco Smoking in the Prospective Urban and Rural Epidemiological Study

Tobacco smoking is the leading preventable cause of death that results in more than 6 million deaths on a worldwide basis.[153] It is reported that current smokers lose about 10 years of life compared to never smokers[153] Moreover, never smokers are show to have better diet quality compared to current daily smokers.[154] As a result, the biomonitoring of cyanide exposure from tobacco smoking and intake of specific goitrogenic foods is needed to prevent elevated mortality and chronic disease burden in vulnerable populations given large variations in smoking habits, tobacco products, food intake, and cultural norms. *Chapter III* validates the utility of urinary thiocyanate as a biomarker of cyanide exposure from active tobacco smoking. Urine samples were collected and analyzed from 1000 participants spanning 14 countries with varied smoking and income statuses from the PURE Study. Our findings reveal current smokers from high-income countries had the highest extent of cyanide exposure compared to middle- and low-income countries after adjusted for relevant confounding factors, including daily cigarette dosage.

1.4.3 Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-limited Sweat Specimens

Cystic fibrosis (CF) is a fatal and rare multisystem genetic disease of infants and young children, which primarily affects the lungs and digestive sytem.[155,156] The universal newborn screening (NBS) program allows for early detection and initiation of therapeutic interventions during infant stages for improved and prolonged quality of life.[157] Most CF newborns are screened using the two-tiered process based on increased immunoreactive trypsinogen (IRT) along with a CFTR mutation panel analysis.[158] Nevertheless, confirmatory CF diagnosis relies on the

gold standard sweat chloride test via pilocarpine iontophoresis.[59] The cystic fibrosis transmembrane conductance regulator (CFTR) channel is responsible for the transportation of chloride among other ions, including bicarbonate and sodium, in the apical cell membrane of epithelial cells.[156] Sweat chloride levels ≥ 60 mmol/L indicated a positive CF case, <30 mmol/L indicates CF unaffected, and 30 – 59 mmol/L indicates an inconclusive case in infants ≤ 6 months of age.[159] *Chapter IV* introduces a novel and rapid method based on CE-iUV for reliable quantification of chloride in volume-restricted sweat samples, while also exploring the potential utility of sweat bicarbonate in confirmatory testing of presumptive CF infants. Our preliminary results indicated good accuracy of chloride measurements using CE-iUV compared to coulometric titration methods, which may reduce testing failure rates due to insufficient sweat volumes. Although bicarbonate did not have clinical value in newborn CF diagnosis, it may prove useful in other biological samples where CFTR-dependent bicarbonate transport is relevant to CF disease severity, such as pancreatic juice and airway lung fluid.

1.4.4 New Sweat Biomarkers for Diagnostic Testing of Screen-Positive Cystic Fibrosis Infants

CF is an autosomal recessive disease that benefits from early detection and prompt treatment during infancy with improved long-term clinical outcomes later in life.[160] Newborn screening (NBS) for CF in North America often relies on a two-tiered screening strategy for immunoreactive trypsinogen (IRT) from dried blood spots, followed by detection of a panel of disease-causing *CFTR* mutations.[4] Following the two-tiered algorithm, presumptive screen-positive CF infants require sweat chloride testing via pilocarpine-stimulated iontophoresis, which remains the gold standard for confirmatory CF diagnosis.[4,7] For infants under 3 months, sweat chloride concentrations ≥ 60 mmol/L is indicative of CF, whereas unaffected carriers have sweat chloride < 30 mmol/L. That

notwithstanding, a small number of sweat tests for screen-positive CF infants remain indeterminate having intermediate sweat concentrations ranging from 30-59 mmol/L and also referred to as cystic fibrosis-screen positive inconclusive diagnosis (CF-SPID).[8] Although NBS enables pre-symptomatic diagnosis and early treatment to improve health outcomes for CF patients,[9] repeat sweat chloride testing and follow-up monitoring for ambiguous CF-SPID cases still contribute to parental anxiety and additional workload for CF clinics.[10] Recent literature has proposed that an initial sweat chloride of ≥ 40 mmol/L was an indicator for late onset CF following a longitudinal study of CF-SPID cases.[16] However, the discovery of novel diagnostic and/or prognostic biomarkers of CF that augment sweat chloride testing is urgently needed especially to aid ambiguous cases.[15] *Chapter V* applies an untargeted characterization of the sweat metabolome from presumptive CF infants when using MSI-CE-MS. A panel of sweat metabolites, including aspartic acid, glutamine, oxoproline, and pilocarpic acid, were found to discriminate screen-positive CF affected from unaffected infants. Ratiometric comparisons, such as aspartic to pilocarpic acid and glutamine to pilocarpic acid were shown to be promising sweat biomarkers in differentiating between CF sub-groups. The potential utility of these sweat metabolites to prognosticate late-onset CF infants from inconclusive sweat chloride test results was also explored.

1.5 References

(1) Strimbu, K.; Tavel, J. A. What Are Biomarkers? *Curr. Opin. HIV AIDS* **2010**, *5* (6), 463–466.

(2) Mayeux, R. Biomarkers: Potential Uses and Limitations. *NeuroRx* **2004**, *1* (2), 182–188.

(3) Atkinson, A. J.; Colburn, W. A.; DeGruttola, V. G.; DeMets, D. L.; Downing, G. J.; Hoth, D. F.; Oates, J. A.; Peck, C. C.; Schooley, R. T.; Spilker, B. A.; et al. Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework. *Clin. Pharmacol. Ther.* **2001**, *69* (3), 89–95.

(4) Gordis, L. Epidemiology and Public Policy. *Epidemiology* **1996**, 247–256.

(5) Perera, F. P.; Weinstein, I. B. Molecular Epidemiology: Recent Advances and Future Directions. *Carcinogenesis* **2000**, *21* (3), 517–524.

(6) Wild, J.; Shanmuganathan, M.; Hayashi, M.; Potter, M.; Britz-Mckibbin, P. Metabolomics for Improved Treatment Monitoring of Phenylketonuria: Urinary Biomarkers for Non-Invasive Assessment of Dietary Adherence and Nutritional Deficiencies. *Analyst* **2019**, *144* (22), 6595–6608.

(7) Macedo, A. N.; Mathiaparanam, S.; Brick, L.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P. The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Cent. Sci.* **2017**, *3* (8), 904–913.

(8) Rohner, F., Zimmerman, M., Jooste, P., Pandav, C., Caldwell, K., Raghavan, R., Raiten, D. Biomarkers of Nutrition for Development-Iodine Review. *J. Nutr.* **2014**, *144*, 1322S-1342S.

(9) Rafiq, T.; Azab, S. M.; Teo, K. K.; Thabane, L.; Anand, S. S.; Morrison, K. M.; De Souza, R. J.; Britz-Mckibbin, P. Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review. *Adv. Nutr.* **2021**, *12* (6), 2333–2357.

(10) Wellington, N.; Shanmuganathan, M.; De Souza, R. J.; Zulyniak, M. A.; Azab, S.; Bloomfield, J.; Mell, A.; Ly, R.; Desai, D.; Anand, S. S.; et al. Metabolic Trajectories Following Contrasting Prudent and Western Diets from Food Provisions: Identifying Robust Biomarkers of Short-Term Changes in Habitual Diet. *Nutrients* **2019**, *11* (10) 2407.

(11) Gill, B.; Mell, A.; Shanmuganathan, M.; Jobst, K.; Zhang, X.; Kinniburgh, D.; Cherry, N.; Britz-McKibbin, P. Urinary Hydroxypyrene Determination for Biomonitoring of Firefighters Deployed at the Fort McMurray Wildfire: An Inter-Laboratory Method Comparison. *Anal. Bioanal. Chem.* **2019**, *411* (7), 1397–1407.

(12) Perrine CG, Herrick K, Serdula MK, et al. Some Subgroups of Reproductive Age Women in the United States May Be at Risk for Iodine Deficiency. *J. Nutr.* **2010**, *140* (8), 1489–1494.

(13) Andersson M, Karumbunathan V, Z. M. Global Iodine Status in 2011 and Trends over the Past Decade. *J. Nutr.* **2012**, *142*, 744–750.

(14) WHO. *Vitamin and Mineral Requirements in Human Nutrition*, 2nd ed.; World Health Organization and Food and Agriculture Organization of the United Nations, 2004.

(15) Cavalieri, R. R. Iodine Metabolism and Thyroid Physiology: Current Concepts. **1997**, *7* (2), 177–181.

(16) de Benoist, B.; Andersson, M.; Egli, I.; Takkouche, B.; Allen, H. Iodine Status Worldwide. *WHO Glob. Database Iodine Defic. Geneca World Heal. Organ.* **2004**, 1–12.

(17) Hetzel, B. Iodine Deficiency Disorders (IDD) and Their Eradication. *Lancet* **1983**, *2*, 1126–1129.

(18) Laurberg, P., Nohr, K.M., Pedersen, A.B., Hreidarsson, S., Andersen, S., Bulow Pedersen, I., Knudsen, N., Perrild, H., Jorgensen, T., Ovesen, L. Thyroid Disorders in Mild Iodine Deficiency. *Thyroid* **2000**, *10*, 951–963.

(19) Stanbury, J.B., Ermans, A.E., Bourdoux, P., Todd, C., Oken, R., Tonglet, G., Vidor. G., Braverman, L.E., Medeiros-Neto, G. Iodine-Induced Hyperthyroidism: Occurrence and Epidemiology. *Thyroid* **1998**, *8* (1), 83–100.

(20) Dunn, J. T. Editorial : What's Happening to Our Iodine? *J. Clin. Endocrinol. Metab.* **1998**, *83* (10), 3398–3400.

(21) Als, C.; Lauber, K.; Brander, L.; Al., E. The Instability of Dietary Iodine Supply over Time in an Affluent Society. *Experientia* **1995**, *51* (6), 623–633.

(22) Laurberg, P.; Knudsen, N.; Ovesen, L.; Andersen, S. Nonmalignant Thyroid Disease. *J. Clin. Thyroidol.* **2001**, *11* (5), 457–469.

(23) Zimmermann, M. B. Symposium on 'Geographical and Geological Influences on Nutrition' Iodine Deficiency in Industrialised Countries: Conference on 'Over- and Undernutrition: Challenges and Approaches.' *Proc. Nutr. Soc.* **2010**, *69* (1), 133–143.

(24) Andersson, M.; Benoist, B.; Rogers, L. Epidemiology of Iodine Deficiency: Salt Iodisation and Iodine Status. *Best Pract. Res. Clin. Endocrinol. Metab.* **2010**, *24* (1), 1–11.

(25) Zimmermann, M. B. Iodine Requirements and the Risks and Benefits of Correcting Iodine Deficiency in Populations. *J. Trace Elem. Med. Biol.* **2008**, *22* (2), 81–92.

(26) WHO. *Assessment of Iodine Deficiency Disorders and Monitoring Their Elimination: A Guide for Programme Mangers*, 3rd ed.; World Health Organization: Geneva, 2007.

(27) American Thyroid Association. Iodine deficiency http://www.thyroid.org/iodine-deficiency/ (accessed Nov 4, 2018).

(28) Bones, A. M.; Rossiter, J. T. The Enzymic and Chemically Induced Decomposition of Glucosinolates. *Phytochemistry* **2006**, *67* (11), 1053–1067.

(29) Prue, D. M.; Martin, J. E.; Hume, A. S. A Critical Evaluation of Thiocyanate as a Biochemical Index of Smoking Exposure. *Behav. Ther.* **1980**, *11* (3), 368–379.

(30) Tsuge, Kouichiro; Mieko, Kataoka; Seto, Y. Cyanide and Thiocyanate Levels in Blood and Saliva of Healthy Adult Volunteers. *J. Heal. Sci.* **2000**, *46* (5), 343–349.

(31) Gracia, R.; Shepherd, G. Cyanide Poisoning and Its Treatment. *Pharmacotherapy* **2004**, *24* (10 II), 1358–1365.

(32) Wikström, M.; Verkhovsky, M. I.; Hummer, G. Water-Gated Mechanism of Proton Translocation by Cytochrome c Oxidase. *Biochim. Biophys. Acta - Bioenerg.* **2003**, *1604* (2), 61–65.

(33) Kammerman Sher, P. Cyanide-Induced Chronic Partial Asphyxia in Vitro: Neurochemical Abnormalities and Reversal by Magnesium. *Pediatr. Neurol.* **1987**, *3* (4), 197–202.

(34) Narkowicz, S.; Polkowska, Z.; Kiełbratowska, B.; Namieśnik, J. Environmental Tobacco Smoke: Exposure, Health Effects, and Analysis. *Crit. Rev. Environ. Sci. Technol.* **2013**, *43* (2), 121–161.

(35) Borgerding, M.; Klus, H. Analysis of Complex Mixtures - Cigarette Smoke. *Exp. Toxicol. Pathol.* **2005**, *57* (SUPPL. 1), 43–73.

(36) Jain, R. B. Trends and Variability in the Levels of Urinary Thiocyanate, Perchlorate, and Nitrate by Age, Gender, Race/Ethnicity, Smoking Status, and Exposure to Environmental Tobacco Smoke over 2005-2012. *Sci. Total Environ.* **2016**, *557–558*, 221–230.

(37) Ramsey, B. W. Disease in Patients with Cystic Fibrosis. *N. Engl. J. Med.* **1996**, *335* (3), 179–188.

(38) Talamo, R. C., Rosenstein B. J. Berninger R. W., Stanbury, J. B. *The Metabolic Basis of Inherited Disease*, 5th ed.; New York, 1983.

(39) Dupuis, A.; Hamilton, D.; Cole, D. E. C.; Corey, M. Cystic Fibrosis Birth Rates in Canada: A Decreasing Trend since the Onset of Genetic Testing. *J. Pediatr.* **2005**, *147* (3), 312–315.

(40) Zielenski, J.; Rozmahel, R.; Bozon, D.; Kerem, B.; Grzelczak, Z.; Riordan, J. R.; Rommens, J.; Tsui, L. C. Genomic DNA Sequence of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Gene. *Genomics* **1991**, *10* (1), 214–228.

(41) Rosenstein, B. J.; Cutting, G. R. The Diagnosis of Cystic Fibrosis: A Consensus Statement. *J. Pediatr.* **1998**, *132* (4), 589–595.

(42) Rommens, J. M., Iannuzzi, M. C., Kerem, B-S., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J. R., Tsui, L-C., Collins, F. S. Identification of the Cystic Fibrosis Gene : Chromosome Walking and Jumping. *Amiercan Assoc. Adv. Sci.* **1989**, *245* (4922), 1059–1065.

(43) Poulsen, J. H.; Fischer, H.; Illek, B.; Machen, T. E. Bicarbonate Conductance and PH Regulatory Capability of Cystic Fibrosis Transmembrane Conductance Regulator. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (12), 5340–5344.

(44) Reddy, M. M.; Light, M. J.; Quinton, P. M. Activation of the Epithelial Na+ Channel (ENaC) Requires CFTR Cl- Channel Function. *Nature* **1999**, *402* (6759), 301–304.

(45) Di Sant' Agnese, P. A., Darling, R. C., Perera, G. A., Shea, E. ABNORMAL ELECTROLYTE CYSTIC Significance. *Pediatrics* **1953**, *12* (5), 549–563.

(46) Cohen-Cymberknoh, M.; Shoseyov, D.; Kerem, E. Managing Cystic Fibrosis: Strategies That Increase Life Expectancy and Improve Quality of Life. *Am. J. Respir. Crit. Care Med.* **2011**, *183* (11), 1463–1471.

(47) Ross, L. F. Newborn Screening for Cystic Fibrosis: A Lesson in Public Health Disparities. *J. Pediatr.* **2008**, *153* (3), 308–313.

(48) Wilcken, B. Newborn Screening for Cystic Fibrosis: Its Evolution and a Review of the Current Situation. *Screening* **1993**, *2*, 43–62.

(49) Sahai, I.; Marsden, D. Newborn Screening. *Crit. Rev. Clin. Lab. Sci.* **2009**, *46* (2), 55–82.

(50) Munck, A.; Mayell, S. J.; Winters, V.; Shawcross, A.; Derichs, N.; Parad, R.; Barben, J.; Southern, K. W. Cystic Fibrosis Screen Positive, Inconclusive Diagnosis (CFSPID): A New Designation and Management Recommendations for Infants with an Inconclusive Diagnosis Following Newborn Screening. *J. Cyst. Fibros.* **2015**, *14* (6), 706–713.

(51) Ramsey, B.W., Welsh, M. J. Progress along the Pathway of Discovery Leading to Treatment and Cure of Cystic Fibrosis. **2017**, *195* (9), 1473–1475.

(52) Cooke, R. E. A Test for Concentration of Electrolytes in Sweat in Cycstic Fibrosis of the Pancrea Utilizing Pilocarpine by Iontophoresis. *Pediatrics* **1959**, No. 23, 545–549.

(53) LeGrys, V. A. Sweat Testing for the Diagnosis of Cystic Fibrosis: Practical Considerations. *J. Pediatr.* **1996**, *129* (6), 892–897.

(54) Ooi, C. Y.; Castellani, C.; Keenan, K.; Avolio, J.; Volpi, S.; Boland, M.; Kovesi, T.; Bjornson, C.; Chilvers, M. A.; Morgan, L.; et al. Inconclusive Diagnosis of Cystic Fibrosis After Newborn Screening. *Pediatrics* **2015**, *135* (6), e1377–e1385.

(55) Talas, U.; Dunlop, J.; Khalaf, S.; Leigh, I. M.; Kelsell, D. P. Human Elastase 1: Evidence for Expression in the Skin and the Identification of a Frequent Frameshift Polymorphism. *J. Invest. Dermatol.* **2000**, *114* (1), 165–170.

(56) Gonska, T.; Keenan, K.; Au, J.; Dupuis, A.; Chilvers, M. A.; Burgess, C.; Bjornson, C.; Fairservice, L.; Brusky, J.; Kherani, T.; et al. Outcomes of Cystic Fibrosis Screening-Positive Infants With Inconclusive Diagnosis at School Age. *Pediatrics* **2021**, *148* (6), 1–9.

(57) Monton, M. R. N.; Soga, T. Metabolome Analysis by Capillary Electrophoresis-Mass Spectrometry. *J. Chromatogr. A* **2007**, *1168* (1–2), 237–246.

(58) Ullsten, S.; Danielsson, R.; Bäckström, D.; Sjöberg, P.; Bergquist, J. Urine Profiling Using Capillary Electrophoresis-Mass Spectrometry and Multivariate Data Analysis. *J. Chromatogr. A* **2006**, *1117* (1), 87–93.

(59) Choi, D. H.; Thaxton, A.; Jeong, I. cheol; Kim, K.; Sosnay, P. R.; Cutting, G. R.; Searson, P. C. Sweat Test for Cystic Fibrosis: Wearable Sweat Sensor vs. Standard Laboratory Test. *J. Cyst. Fibros.* **2018**, *17* (4), e35–e38.

(60) Legrys, V. A.; McColley, S. A.; Li, Z.; Farrell, P. M. The Need for Quality Improvement in Sweat Testing Infants after Newborn Screening for Cystic Fibrosis. *J. Pediatr.* **2010**, *157* (6), 1035–1037.

(61) Collins, M. N.; Brawley, C. B.; McCracken, C. E.; Shankar, P. R. V.; Schechter, M. S.; Rogers, B. B. Risk Factors for Quantity Not Sufficient Sweat Collection in Infants 3 Months or Younger. *Am. J. Clin. Pathol.* **2014**, *142* (1), 72–75.

(62) Benavente, F.; Andón, B.; Giménez, E.; Barbosa, J.; Sanz-Nebot, V. Modeling the Migration Behavior of Rabbit Liver Apothioneins in Capillary Electrophoresis. *Electrophoresis* **2008**, *29* (13), 2790–2800.

(63) Quirino, J. P.; Haddad, P. R. Online Sample Preconcentration in Capillary Electrophoresis Using Analyte Focusing by Micelle Collapse. *Anal. Chem.* **2008**, *80* (17), 6824–6829.

(64) de Macedo, A. N.; Macri, J.; Hudecki, P. L.; Saoi, M.; McQueen, M. J.; Britz-McKibbin, P. Validation of a Capillary Electrophoresis Assay for Monitoring Iodine Nutrition in Populations for Prevention of Iodine Deficiency: An Interlaboratory Method Comparison. *J. Appl. Lab. Med.* **2017**, *1* (6), 649–660.

(65) Malá, Z.; Křivánková, L.; Gebauer, P.; Boček, P. Contemporary Sample Stacking in CE: A Sophisticated Tool Based on Simple Principles. *Electrophoresis* **2007**, *28* (1–2), 243–253.

(66) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. Integrative Metabolomics for Characterizing Unknown Low-Abundance Metabolites by Capillary Electrophoresis-Mass Spectrometry with Computer Simulations. *Anal. Chem.* **2007**, *79* (2), 403–415.

(67) Nevedomskaya, E.; Derks, R.; Deelder, A. M.; Mayboroda, O. A.; Palmblad, M. Alignment of Capillary Electrophoresis-Mass Spectrometry Datasets Using Accurate Mass Information. *Anal. Bioanal. Chem.* **2009**, *395* (8), 2527–2533.

(68) Piñero, M. Y.; Bauza, R.; Arce, L. Thirty Years of Capillary Electrophoresis in Food Analysis Laboratories: Potential Applications. *Electrophoresis* **2011**, *32* (11), 1379–1393.

(69) Wätzig, H.; Degenhardt, M.; Kunkel, a. Strategies for Capillary Electrophoresis: Method Development and Validation for Pharmaceutical and Biological Applications. *Electrophoresis* **1998**, *19* (16–17), 2695–2752.

(70) Frazier, R. A., Ames, J. M., Nursten, H. E. *Background Theory and Principles of Capillary Electrophoresis*; Royal Society of Chemistry, 2000.

(71) Melanson, J. E.; Baryla, N. E.; Lucy, C. A. Dynamic Capillary Coatings for Electroosmotic Flow Control in Capillary Electrophoresis. *TrAC - Trends Anal. Chem.* **2001**, *20* (6–7), 365–374.

(72) Anderson, J.L., Berthod, A., Estévez, V.P., Stalcup, A. M. *Analytical Separation Science*; Wiley-VCH, 2016.

(73) Nowak, P. M.; Woźniakiewicz, M.; Gładysz, M.; Janus, M.; Kościelniak, P. Improving Repeatability of Capillary Electrophoresis—a Critical Comparison of Ten Different Capillary Inner Surfaces and Three Criteria of Peak Identification. *Anal. Bioanal. Chem.* **2017**, *409* (18), 4383–4393.

(74) Xu, Y. Tutorial: Capillary Electrophoresis. *Chem. Educ.* **1996**, *1* (2), 1–14.

(75) Nevídalová, H.; Michalcová, L.; Glatz, Z. Capillary Electrophoresis-Based Approaches for the Study of Affinity Interactions Combined with Various Sensitive and Nontraditional Detection Techniques. *Electrophoresis* **2019**, *40* (5), 625–642.

(76) Beckers, J. L.; Boček, P. Calibrationless Quantitative Analysis by Indirect UV Absorbance Detection in Capillary Zone Electrophoresis: The Concept of the Conversion Factor. *Electrophoresis* **2004**, *25* (2), 338–343.

(77) Kitagawa, F.; Otsuka, K. Recent Applications of On-Line Sample Preconcentration Techniques in Capillary Electrophoresis. *J. Chromatogr. A* **2014**, *1335*, 43–60.

(78) Macka, M.; Johns, C.; Doble, P.; Haddad, P. R. Indirect Photometric Detection in CE Using Buffered Electrolytes - Part I, Principles. *LCGC North Am.* **2001**, *19* (1), 38–47.

(79) Chen, T.; Yao, Q.; Nasaruddin, R. R.; Xie, J. Electrospray Ionization Mass Spectrometry: A Powerful Platform for Noble-Metal Nanocluster Analysis. *Angew. Chemie - Int. Ed.* **2019**, *58* (35), 11967–11977.

(80) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246* (4926), 64–71.

(81) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. Electrospray Ionization-Principles and Practice. **1990**, 37–70.

(82) Sauer, F.; Sydow, C.; Trapp, O. A Robust Sheath-Flow CE-MS Interface for Hyphenation with Orbitrap MS. *Electrophoresis* **2020**, *41* (15), 1280–1286.

(83) Wu, H.; Tang, K. Highly Sensitive and Robust Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Interfaces, Preconcentration Techniques and Applications. *Rev. Anal. Chem.* **2020**, *39* (1), 45–55.

(84) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. Virtual Quantifcation of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Ef Ciency Without Chemical Standards. *Anal. Chem.* **2009**, *81* (7), 2506–2515.

(85) Ramautar, R.; Demirci, A.; Jong, G. J. d. Capillary Electrophoresis in Metabolomics. *TrAC - Trends Anal. Chem.* **2006**, *25* (5), 455–466.

(86) Kuehnbaum, N. L.; Kormendi, A.; Britz-Mckibbin, P. Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Anal. Chem.* **2013**, *85* (22), 10664–10669.

(87) Ioannidis, J. P. A. Why Most Published Research Findings Are False. *Get. to Good Res. Integr. Biomed. Sci.* **2018**, *2* (8), 2–8.

(88) Tolstikov, V.; James Moser, A.; Sarangarajan, R.; Narain, N. R.; Kiebish, M. A. Current Status of Metabolomic Biomarker Discovery: Impact of Study Design and Demographic Characteristics. *Metabolites* **2020**, *10* (6), 1–12.

(89) Dibattista, A.; McIntosh, N.; Lamoureux, M.; Al-Dirbashi, O. Y.; Chakraborty, P.; Britz-McKibbin, P. Temporal Signal Pattern Recognition in Mass Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Anal. Chem.* **2017**, *89* (15), 8112–8121.

(90) Smith, L.; Villaret-Cazadamont, J.; Claus, S. P.; Canlet, C.; Guillou, H.; Cabaton, N. J.; Ellero-Simatos, S. Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites* **2020**, *10* (3), 1–17.

(91) Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern Analytical Techniques in Metabolomics Analysis. *Analyst* **2012**, *137* (2), 293–300.

(92) Ribbenstedt, A.; Ziarrusta, H.; Benskin, J. P. Development, Characterization and Comparisons of Targeted and Non-Targeted Metabolomics Methods. *PLoS One* **2018**, *13* (11), 1–18.

(93) Dettmer, Katja, Hammock D., B. Metabolomics: A New Exciting Field within the "Omics" Sciences. *Environ. Health Perspect.* **2004**, *112* (7), A396–A397.

(94) Fiehn, O. Metabolomics - the Link between Genotypes and Phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171.

(95) MacEdo, A. N.; Mathiaparanam, S.; Brick, L.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P. The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Cent. Sci.* **2017**, *3* (8), 904–913.

(96) Shao, Y.; Le, W. Recent Advances and Perspectives of Metabolomics Based Investigations in Parkinson's Disease. *Mol. Neurodegener.* **2019**, *14* (1), 1–12.

(97) Varma, V. R.; Oommen, A. M.; Varma, S.; Casanova, R.; An, Y.; Andrews, R. M.; O'Brien, R.; Pletnikova, O.; Troncoso, J. C.; Toledo, J.; et al. *Brain and Blood Metabolite Signatures of Pathology and Progression in Alzheimer Disease: A Targeted Metabolomics Study*; 2018; Vol. 15.

(98) Playdon, M. C.; Moore, S. C.; Derkach, A.; Reedy, J.; Subar, A. F.; Sampson, J. N.; Albanes, D.; Gu, F.; Kontto, J.; Lassale, C.; et al. Identifying Biomarkers of Dietary Patterns by Using Metabolomics. *Am. J. Clin. Nutr.* **2017**, *105* (2), 450–465.

(99) Guasch-Ferre, M.; Bhupathiraju, S. N.; Hu, F. B. Use of Metabolomics in Improving Assessment of Dietary Intake. *Clin. Chem.* **2018**, *64* (1), 82–98.

(100) Johnson, C. H.; Patterson, A. D.; Idle, J. R.; Gonzalez, F. J. Xenobiotic Metabolomics: Major Impact on the Metabolome. *Annu. Rev. Pharmacol. Toxicol.* **2012**, *52*, 37–56.

(101) Lin, C. Y.; Viant, M. R.; Tjeederma, R. S. Metabolomics Methodologies and Applications in the Environmental Sciences. *J. Pestic. Sci.* **2015**, *31* (3), 245–251.

(102) Wishart, D. S. Emerging Applications of Metabolomics in Drug Discovery and Precision Medicine. *Nat. Rev. Drug Discov.* **2016**, *15* (7), 473–484.

(103) Long, N. P.; Nghi, T. D.; Kang, Y. P.; Anh, N. H.; Kim, H. M.; Park, S. K.; Kwon, S. W. Toward a Standardized Strategy of Clinical Metabolomics for the Advancement of Precision Medicine. *Metabolites* **2020**, *10* (2), 1–28.

(104) Wu, J.; An, Y.; Yao, J.; Wang, Y.; Tang, H. An Optimised Sample Preparation Method for NMR-Based Faecal Metabonomic Analysis. *Analyst* **2010**, *135* (5), 1023–1030.

(105) Malet-Martino, M.; Holzgrabe, U. NMR Techniques in Biomedical and Pharmaceutical Analysis. *J. Pharm. Biomed. Anal.* **2011**, *55* (1), 1–15.

(106) Emwas, A. H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Nagana Gowda, G. A.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; et al. Nmr Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9* (7), 1–39.

(107) Villas-Bôas, S. G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. Mass Spectrometry in Metabolome Analysis. *Mass Spectrom. Rev.* **2005**, *24* (5), 613–646.

(108) Tsugawa, H.; Bamba, T.; Shinohara, M.; Nishiumi, S.; Yoshida, M.; Fukusaki, E. Practical Non-Targeted Gas Chromatography/Mass Spectrometry-Based Metabolomics Platform for Metabolic Phenotype Analysis. *J. Biosci. Bioeng.* **2011**, *112* (3), 292–298.

(109) Barbas, C.; Moraes, E. P.; Villaseñor, A. Capillary Electrophoresis as a Metabolomics Tool for Non-Targeted Fingerprinting of Biological Samples. *J. Pharm. Biomed. Anal.* **2011**, *55* (4), 823–831.

(110) Fernández-Peralbo, M. A.; Luque de Castro, M. D. Preparation of Urine Samples Prior to Targeted or Untargeted Metabolomics Mass-Spectrometry Analysis. *TrAC - Trends Anal. Chem.* **2012**, *41*, 75–85.

(111) Tang, H.; Xiao, C.; Wang, Y. Important Roles of the Hyphenated HPLC-DAD-MS-SPE-NMR Technique in Metabonomics. *Magn. Reson. Chem.* **2009**, *47* (SUPPL. 1).

(112) Van Der Greef, J.; Stroobant, P.; Van Der Heijden, R. The Role of Analytical Sciences in Medical Systems Biology. *Curr. Opin. Chem. Biol.* **2004**, *8* (5), 559–565.

(113) Kell, D. B. Metabolomics and Systems Biology: Making Sense of the Soup. *Curr. Opin. Microbiol.* **2004**, *7* (3), 296–307.

(114) Adams, K. J.; Pratt, B.; Bose, N.; Dubois, L. G.; St. John-Williams, L.; Perrott, K. M.; Ky, K.; Kapahi, P.; Sharma, V.; Maccoss, M. J.; et al. Skyline for Small Molecules: A Unifying Software Package for Quantitative Metabolomics. *J. Proteome Res.* **2020**, *19* (4), 1447–1458.

(115) Katajamaa, M.; Orešič, M. Data Processing for Mass Spectrometry-Based Metabolomics. *J. Chromatogr. A* **2007**, *1158* (1–2), 318–328.

(116) Beger, R. D.; Dunn, W. B.; Bandukwala, A.; Bethan, B.; Broadhurst, D.; Clish, C. B.; Dasari, S.; Derr, L.; Evans, A.; Fischer, S.; et al. Towards Quality Assurance and Quality Control in Untargeted Metabolomics Studies. *Metabolomics* **2019**, *15* (1), 1–5.

(117) Begou, O.; Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. *Quality Control and Validation Issues in LC-MS Metabolomics*; Metabolic Profiling: Methods and Protocols, Methods in Molecular Biology, 2018; Vol. 1738.

(118) Vollmar, A. K. R.; Rattray, N. J. W.; Cai, Y.; Santos-Neto, Á. J.; Deziel, N. C.; Jukic, A. M. Z.; Johnson, C. H. Normalizing Untargeted Periconceptional Urinary Metabolomics Data: A Comparison of Approaches. *Metabolites* **2019**, *9* (10), 1–15.

(119) Gagnebin, Y.; Tonoli, D.; Lescuyer, P.; Ponte, B.; de Seigneux, S.; Martin, P. Y.; Schappler, J.; Boccard, J.; Rudaz, S. Metabolomic Analysis of Urine Samples by UHPLC-QTOF-MS: Impact of Normalization Strategies. *Anal. Chim. Acta* **2017**, *955*, 27–35.

(120) Edmands, W. M. B.; Ferrari, P.; Scalbert, A. Normalization to Specific Gravity Prior to Analysis Improves Information Recovery from High Resolution Mass Spectrometry Metabolomic Profiles of Human Urine. *Anal. Chem.* **2014**, *86* (21), 10925–10931.

(121) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. Use of a Pre-Analysis Osmolality Normalisation Method to Correct for Variable Urine Concentrations and for Improved Metabolomic Analyses. *J. Chromatogr. A* **2016**, *1431*, 103–110.

(122) Gardlo, A.; Smilde, A. K.; Hron, K.; Hrdá, M.; Karlíková, R.; Friedecký, D.; Adam, T. Normalization Techniques for PARAFAC Modeling of Urine Metabolomic Data. *Metabolomics* **2016**, *12* (7), 1–13.

(123) Barr, D. B.; Wilder, L. C.; Caudill, S. P.; Gonzalez, A. J.; Needham, L. L.; Pirkle, J. L. Urinary Creatinine Concentrations in the U.S. Population: Implications for Urinary Biologic Monitoring Measurements. *Environ. Health Perspect.* **2005**, *113* (2), 192–200.

(124) Waikar, S. S.; Sabbisetti, V. S.; Bonventre, J. V. Normalization of Urinary Biomarkers to Creatinine during Changes in Glomerular Filtration Rate. *Kidney Int.* **2010**, *78* (5), 486–494.

(125) MacPherson, S.; Arbuckle, T. E.; Fisher, M. Adjusting Urinary Chemical Biomarkers for Hydration Status during Pregnancy. *J. Expo. Sci. Environ. Epidemiol.* **2018**, *28* (5), 481–493.

(126) Miller, R. C.; Brindle, E.; Holman, D. J.; Shofer, J.; Klein, N. A.; Soules, M. R.; O'Connor, K. A. Comparison of Specific Gravity and Creatinine for Normalizing Urinary Reproductive Hormone Concentrations. *Clin. Chem.* **2004**, *50* (5), 924–932.

(127) Carrieri, M.; Trevisan, A.; Bartolucci, G. B. Adjustment to Concentration-Dilution of Spot Urine Samples: Correlation between Specific Gravity and Creatinine. *Int. Arch. Occup. Environ. Health* **2000**, *74* (1), 63–67.

(128) Hays, S. M.; Aylward, L. L.; Blount, B. C. Variation in Urinary Flow Rates According to Demographic Characteristics and Body Mass Index in NHANES: Potential Confounding of Associations between Health Outcomes and Urinary Biomarker Concentrations. *Environ. Health Perspect.* **2015**, *123* (4), 293–300.

(129) Boeniger, Mark Frederick; Lowry, Larry K.; Rosenberg, J. Interpretation of Urine Results Used to Assess Chemical Exposure with Emphasis on Creatinine Adjustments: A Review. *Am. Ind. Hyg. Assoc. J.* **1993**, *54* (10), 615–627.

(130) Davison, J. M.; Noble, M. C. B. Serial Changes in 24 Hour Creatinine Clearance During Normal Menstrual Cycles and the First Trimester of Pregnancy. *BJOG An Int. J. Obstet. Gynaecol.* **1981**, *88* (1), 10–17.

(131) Khamis, Mona M.; Adamko, Darryl J.; El-Aneed, A. Mass Spectrometric Based Approaches in Urine Metabolomics and Biomarker Discovery. *Mass Spectrom. Rev.* **2014**, *9999*, 1–20.

(132) Ryan, D.; Robards, K.; Prenzler, P. D.; Kendall, M. Recent and Potential Developments in the Analysis of Urine: A Review. *Anal. Chim. Acta* **2011**, *684* (1–2), 17–29.

(133) van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J. Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data. *BMC Genomics* **2006**, *7*, 1–15.

(134) Noble, W. S. How Does Multiple Testing Correction Work? *Nat. Biotechnol.* **2009**, *27* (12), 1135–1137.

(135) Andrade, C. Multiple Testing and Protection Against a Type 1 (False Positive) Error Using the Bonferroni and Hochberg Corrections. *Indian J. Psychol. Med.* **2019**, *41* (2), 99–100.

(136) Glickman, M. E.; Rao, S. R.; Schultz, M. R. False Discovery Rate Control Is a Recommended Alternative to Bonferroni-Type Adjustments in Health Studies. *J. Clin. Epidemiol.* **2014**, *67* (8), 850–857.

(137) Jobard, E.; Dossus, L.; Baglietto, L.; Fornili, M.; Lécuyer, L.; Mancini, F. R.; Gunter, M. J.; Trédan, O.; Boutron-Ruault, M. C.; Elena-Herrmann, B.; et al. Investigation of Circulating Metabolites Associated with Breast Cancer Risk by Untargeted Metabolomics: A Case–Control Study Nested within the French E3N Cohort. *Br. J. Cancer* **2021**, *124* (10), 1734–1743.

(138) Mahadevan, S.; Shah, S. L.; Marrie, T. J.; Slupsky, C. M. Analysis of Metabolomic Data Using Support Vector Machines. *Anal. Chem.* **2008**, *80* (19), 7562–7570.

(139) Chen, T.; Cao, Y.; Zhang, Y.; Liu, J.; Bao, Y.; Wang, C.; Jia, W.; Zhao, A. Random Forest in Clinical Metabolomics for Phenotypic Discrimination and Biomarker Selection. *Evidence-based Complement. Altern. Med.* **2013**, *2013*, 1–11.

(140) Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial. *Metabolomics* **2013**, *9* (2), 280–299.

(141) Vinayavekhin, N.; Saghatelian, A. Untargeted Metabolomics. *Curr. Protoc. Mol. Biol.* **2010**, No. SUPPL. 90, 1–24.

(142) Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M. Interferences and Contaminants Encountered in Modern Mass Spectrometry. *Anal. Chim. Acta* **2008**, *627* (1), 71–81.

(143) Domingo-Almenara, X.; Montenegro-Burke, J. R.; Guijas, C.; Majumder, E. L. W.; Benton, H. P.; Siuzdak, G. Autonomous METLIN-Guided In-Source Fragment Annotation for Untargeted Metabolomics. *Anal. Chem.* **2019**, *91* (5), 3246–3253.

(144) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012**, *2* (4), 775–795.

(145) Mahieu, N. G.; Spalding, J. L.; Gelman, S. J.; Patti, G. J. Defining and Detecting Complex Peak Relationships in Mass Spectral Data: The Mz.Unity Algorithm. *Anal. Chem.* **2016**, *88* (18), 9037–9046.

(146) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; et al. Proposed Minimum Reporting Standards for Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 211–221.

(147) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098.

(148) Creek, D. J.; Dunn, W. B.; Fiehn, O.; Griffin, J. L.; Hall, R. D.; Lei, Z.; Mistrik, R.; Neumann, S.; Schymanski, E. L.; Sumner, L. W.; et al. Metabolite Identification: Are You Sure? And How Do Your Peers Gauge Your Confidence? *Metabolomics* **2014**, *10* (3), 350–353.

(149) Kodra, D.; Pousinis, P.; Vorkas, P. A.; Kademoglou, K.; Liapikos, T.; Pechlivanis, A.; Virgiliou, C.; Wilson, I. D.; Gika, H.; Theodoridis, G. Is Current Practice Adhering to Guidelines Proposed for Metabolite Identification in LC-MS Untargeted Metabolomics? A Meta-Analysis of the Literature. *J. Proteome Res.* **2022**, *21* (3), 590-598.

(150) Katz, P, M.; Leung, A,M.; Bravermann, L, E.; Pearce, E, N.; Tomlinson, G.; He, X.; Vertes, J.; Okun, N.; Walfish, P.; Feig, D.The Status of Iodine Nutrition during Pregnancy in Toronto, Canada. *Thyroid* **2011**, *21* (2), A10.

(151) Pharoah, P.; Buttfield, I. H.; Hetzel, B. S. Neurological Damage to the Fetus Resulting from Severe Iodine Deficiency during Pregnancy. *Int. J. Epidemiol.* **2012**, *41* (3), 589–592.

(152) de Macedo, A. N.; Macri, J.; Hudecki, P. L.; Saoi, M.; McQueen, M. J.; Britz-McKibbin, P. Validation of a Capillary Electrophoresis Assay for Monitoring Iodine Nutrition in Populations for Prevention of Iodine Deficiency: An Interlaboratory Method Comparison. *J. Appl. Lab. Med. An AACC Publ.* **2017**, *1* (6), 649–660.

(153) Kondo, T.; Nakano, Y.; Adachi, S.; Murohara, T. Effects of Tobacco Smoking on Cardiovascular Disease. *Circ. J.* **2019**, *83* (10), 1980–1985.

(154) Ross MacLean, R.; Cowan, A.; Vernarelli, J. A. More to Gain: Dietary Energy Density Is Related to Smoking Status in US Adults. *BMC Public Health* **2018**, *18* (1), 1–7.

(155) Bell, S. C.; Mall, M. A.; Gutierrez, H.; Macek, M.; Madge, S.; Davies, J. C.; Burgel, P. R.; Tullis, E.; Castaños, C.; Castellani, C.; et al. The Future of Cystic Fibrosis Care: A Global Perspective. *Lancet Respir. Med.* **2020**, *8* (1), 65–124.

(156) Fonseca, C.; Bicker, J.; Alves, G.; Falcão, A.; Fortuna, A. Cystic Fibrosis: Physiopathology and the Latest Pharmacological Treatments. *Pharmacol. Res.* **2020**, *162*, 1–14.

(157) Mak, D. Y. F.; Sykes, J.; Stephenson, A. L.; Lands, L. C. The Benefits of Newborn Screening for Cystic Fibrosis: The Canadian Experience. *J. Cyst. Fibros.* **2016**, *15* (3), 302–308.

(158) Rosenfeld, M.; Sontag, M. K.; Ren, C. L. Cystic Fibrosis Diagnosis and Newborn Screening. *Pediatr. Clin. North Am.* **2016**, *63* (4), 599–615.

(159) Salinas, D. B.; Peng, Y. H.; Horwich, B.; Wee, C. P.; Frisbee, E.; Maarek, J. M. Image-Based β-Adrenergic Sweat Rate Assay Captures Minimal Cystic Fibrosis Transmembrane Conductance Regulator Function. *Pediatr. Res.* **2020**, *87* (1), 137–145.

(160) Mathiaparanam, S.; Nori, A.; Macedo, D.; Gill, B.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-Mckibbin, P. Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-Limited Sweat Specimens. *J. Cyst. Fibros.* **2022**, No. xxxx. https://doi.org/10.1016/j.jcf.2022.05.004.

Chapter II

The Prevalence and Risk Factors Associated with Iodine Deficiency in Canadian Adults

Stellena Mathiaparanam, Adriana Nori de Macedo, Andrew Mente, Paul Poirier, Scott A. Lear, Andreas Wielgosz, Koon Teo, Salim Yusuf, and Philip Britz-McKibbin

Nutrients, **2022**, *14* (13), 1-17

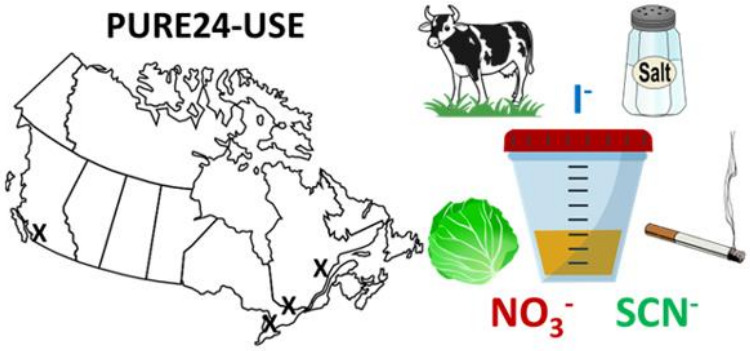
S.M. and A.N.M. performed all the experiments including sample preparation and data acquisition using CE-UV. S.M. performed the data processing, interpretation, statistical analysis, and wrote the initial draft for publication. P.B.M. conceived and supervised the study. S.Y., A.M., K.T., P.P., S.L. and A.W. were involved in participant recruitment, collection of dietary records, and coordination of 24 h urine collection. S.M. and P.B.M. were involved in further data interpretation and edited the manuscript, and with all other co-authors contributing to the final version of the manuscript.

Chapter II: The Prevalence and Risk Factors Associated with Iodine Deficiency in Canadian Adults

2.1 Abstract

Iodine is a trace micronutrient that is critical for normal thyroid function and human health. Inadequate dietary intake is associated with cognitive impairment, infertility, growth retardation and iodine deficiency disorders in affected populations. Herein, we examined the prevalence of iodine deficiency in adults (median age of 61 years) based on the analysis of 24 h urine samples collected from 800 participants in four clinical sites across Canada in the Prospective Urban and Rural Epidemiological (PURE) study. Urinary iodide together with thiocyanate and nitrate were measured using a validated capillary electrophoresis assay. Protective/risk factors associated with iodine deficiency were identified using a binary logistic regression model, whereas daily urinary iodine concentration (24 h UIC, µg/L) and urinary iodine excretion (24 h UIE, µg/day) were compared using complementary statistical methods with covariate adjustments. Overall, our Canadian adult cohort had adequate iodine status as reflected by a median UIC of 111 µg/L with 11.9% of the population <50 µg/L categorized as having moderate to severe iodine deficiency. Iodine adequacy was also evident with a median 24 h UIE of 226 µg/day as a more robust metric of iodine status with an estimated average requirement (EAR) of 7.1% (< 95 µg/day) and a tolerable upper level (UL) of 1.8% (≥1100 µg/day) based on Canadian dietary reference intake values. Participants taking iodine supplements (OR = 0.18; $p = 6.35\times10^{-5}$), had greater 24 h urine volume (OR = 0.69; $p = 4.07\times10^{-4}$), excreted higher daily urinary sodium (OR = 0.71; $p = 3.03\times10^{-5}$), and/or were prescribed thyroxine (OR = 0.33; $p = 1.20\times10^{-2}$) had lower risk for iodine deficiency. Self-reported intake of dairy products was most strongly associated with iodine status ($r = 0.24$; $p = 2.38\times10^{-9}$) after excluding for iodine supplementation and T4 use. Participants

residing in Quebec City (OR = 2.58; $p = 1.74\times10^{-4}$) and Vancouver (OR = 2.54; $p = 3.57\times10^{-4}$) were more susceptible to iodine deficiency than Hamilton or Ottawa. Also, greater exposure to abundant iodine uptake inhibitors from tobacco smoking and intake of specific goitrogenic foods corresponded to elevated urinary thiocyanate and nitrate, which were found for residents from Quebec City as compared to other clinical sites. Recent public health policies that advocate for salt restriction and lower dairy intake may inadvertently reduce iodine nutrition of Canadians, and further exacerbate regional variations in iodine deficiency risk.



2.2 Introduction

Iodine is an essential trace micronutrient in human health used in the biosynthesis of thyroid hormones, which regulate cellular metabolism, growth, and development throughout the lifespan [1,2]. Iodine deficiency remains a global public health concern since it increases neonatal mortality and is a preventable cause of cognitive impairment and developmental delays in children [3,4]. Although reproductive age women and school-aged children represent vulnerable groups [5], thyroid disorders attributed to mild to moderate iodine deficiency are also implicated in chronic disease burden in adults and older persons [6,7], which include immunomodulatory effects on immune function [8]. Nonetheless, remarkable progress has been made in expanding universal salt iodization programs to reduce the prevalence of iodine deficiency disorders worldwide [9,10]. However, several developing and developed countries still suffer from persistent iodine deficiency (e.g., Cambodia, Russia, Israel) or excessive iodine intake (e.g., South Korea, Cameroon, Columbia) that may also contribute to deleterious health outcomes [11,12]. Changing dietary patterns in contemporary societies have also altered the efficacy of iodine prophylaxis [13] through restriction of sodium intake to reduce blood pressure [14] along with increased consumption of processed foods using non-iodized salt [15]. As a result, continuous surveillance is critical to ensure optimal iodine status and to evaluate the impact of recent public health interventions, such as iodine fortification of staple foods (e.g., bread) adopted in Australia [16].

Food frequency questionnaires for estimating iodine intake in populations are limited in capturing the variable content of iodine in similar foods, which may also not be reported in national food composition databases [17]. For instance, cow's milk represents a major source of dietary iodine that varies widely in retail products based on feed composition, iodine supplementation and teat-dipping sanitation practices using iodophors [18].

As most iodine consumed is excreted as iodide, the median urinary iodine concentration (UIC) offers an objective indicator of iodine intake while also serving as a surrogate measure of the prevalence of goiter and thyroid nodules in a population [19]. However, up to ten repeat urine spot samples or 24 h urine collection is needed to reliably estimate individual iodine status [20]. According to the World Health Organization (WHO), adequate iodine intake for adults is indicated by a median UIC within 100 to 199 µg/L along with a low rate (<20%) of moderate to severe iodine deficiency (<50 µg/L) [9,21]. Higher thresholds for iodine deficiency are designated for children and pregnant/lactating women [22]. Daily iodine intake (µg/day) can be estimated from 24 h urinary iodine excretion (UIE) from spot urine iodine concentrations after correction for age, sex, ethnicity and/or anthropometric dependent creatinine adjustments. These adjustments correct for between-subject variations in urine fluid volume and muscle mass [23–25]. As urinary concentrations of iodide reflect recent dietary intake of iodine-containing foods or beverages, other factors can also affect iodine status determination when using spot urine samples, such as skipping breakfast prior to morning urine sampling [26]. Although less convenient to collect, 24 h urine samples offer greater reproducibility and accuracy than spot urine samples and are the preferred method for assessing iodine status in epidemiological studies [27].

Although populations in North America are considered to have adequate iodine status [11], differences remain across certain demographic groups and regions [28]. For instance, a median UIC of 134 µg/L from spot urine samples was reported in Canadian households surveyed from 2009 to 2011, with about 22% and 7% of Canadians at risk for mild and moderate iodine deficiency, respectively [29]. Furthermore, iodine intake is frequently inadequate for women of child-bearing age with recommendations for use of a daily multivitamin-mineral supplement containing iodine during pregnancy and breastfeeding [30]. The risk for iodine deficiency is also dependent on

environmental exposure to perchlorate, thiocyanate and nitrate that competitively inhibit active iodine uptake via the sodium-iodine symporter expressed in the intestine and thyroid gland [31]. To date, few epidemiological studies have examined the iodine status and environmental exposure to iodine uptake inhibitors in diverse populations across different regions [32]. In this work, we examined the iodine status in 24 h urine specimens collected from participants (n = 800) in the Prospective Urban and Rural Epidemiological (PURE) study [33]. A validated method based on capillary electrophoresis (CE) was used for simultaneous analysis of urinary iodide, nitrate and thiocyanate after a simple dilution step [34,35]. Our study aimed to identify risk/protective factors associated with iodine deficiency for participants residing in four communities across Canada reflecting differences in dietary habits, smoking status, and other environmental exposures.

2.3 Experimental

2.3.1 Study Design, Participant Eligibility, and 24 h Urine Sampling

Our cross-sectional study included a subset of participants from the PURE-24USE (PURE-24 Hour Urinary Sodium Excretion Survey) study [34] who were recruited from January 2012 to December 2013. This cohort included adults aged from 36 to 83 years (median of 61 years) living in four sites across Canada, Hamilton (n = 217), Vancouver (n = 200), Quebec City (n = 200), and Ottawa (n = 183). Ethical approval was provided by local research ethics boards at the four clinical sites, and all participants from each site provided signed informed consent [33]. Participants were excluded if they were diagnosed with a debilitating disease, required food restrictions due to chronic illness, as well as pregnant or breastfeeding women who are contraindicated from ingestion of *para*-aminobenzoic acid (PABA). A standardized procedure, necessary supplies, and detailed instructions for collecting 24 h urine specimens were provided to all PURE participants as

described elsewhere [33]. Participants aged under 65 years also ingested a PABA tablet (80 mg) at each of the three meals to verify adherence to 24 h urine sampling [36]. Urine samples were considered authentic for subsequent analysis when recovery was >85% for ingested PABA, and urinary 24 h creatinine excretion was within reference intervals for men (995 to 2489 mg) and women (509 to 1810 mg) [34]. However, older participants (>65 years) were exempted from PABA screening due to their delayed renal clearance [36]. On completion of collection, study staff measured and recorded the 24 h urine volume, thoroughly mixed the collection, and retained 2 mL aliquots that were stored frozen at -70˚C within the Clinical Trials and Clinical Research Laboratory in Hamilton General Hospital [37].

2.3.2 PURE Participants and Self-reported Dietary Intake

At each participant's study visit, a standardized questionnaire was used to assess an-thropogenic parameters, personal medical history, smoking status, alcohol intake, physical activity, and use of prescription medications (e.g., Levothyroxine or Synthroid, T4) and vitamin-mineral supplements [33]. Participants also completed a short questionnaire of salt exposures from foods consumed over the 24 h period when their urine sample was collected. Self-reported food intake was recorded by participants according to the type of product consumed with emphasis on assessment of sodium and/or potassium content in the PURE-24USE study [33]. In this work, foods associated with iodine nutrition and goitrogen intake were evaluated [38,39] including salty foods, dairy, eggs, fish, breads and cereals, meats (e.g., red, white, processed), processed foods, fruits, and various vegetables (e.g., green leafy, cruciferous, dark yellow). The amount of specific food products consumed daily were estimated as g/day unless otherwise indicated.

2.3.3 Iodide, Thiocyanate, Nitrate and Sodium Determination in 24 h Urine Samples

Urine samples were analyzed for iodide that is the predominant species of iodine excreted in urine, as well as nitrate and thiocyanate using a recently validated assay based on capillary electrophoresis with UV absorbance detection [34,35]. Urinary sodium concentrations were measured by indirect potentiometry using a Beckman Coulter UniCel DxC600 Synchron Clinical System [33]. All chemical reagents were purchased from Sigma-Aldrich Inc. (Oakville, ON, Canada) unless otherwise stated. Briefly, frozen aliquots of urine were thawed slowly to room temperature and then diluted two-fold in deionized water containing an internal standard, 1,5-naphthalene disulfonic acid (NDS, 40 µM). In some cases, hydrated urine samples with low ionic strength were diluted two-fold in a simulated urine matrix solution comprised of 100 mM sodium chloride and 10 mM sodium sulfate for matrix matching purposes. All diluted urine samples were vortexed and centrifuged prior to analysis. Quality control (QC) samples based on a pooled urine sample from all PURE-24USE participants were used for assessment of technical precision. All CE separations were performed using a P/ACE™ MDQ system with UV absorbance detection (SCIEX, Framingham, MA, USA). Unmodified fused-silica capillaries were purchased from Polymicro Technologies Inc. (Phoenix, AZ, USA) with an internal diameter of 75 µm and total and effective (to detector window) capillary length of 60 cm and 50 cm, respectively. The background electrolyte (BGE) was composed of 180 mM lithium hydroxide, 180 mM phosphoric acid, 46 mM α-cyclodextrin with a pH of 3.0 [35]. New capillaries were initially conditioned by flushing with methanol, deionized water, 1.0 M lithium hydroxide (5 min each), and then background electrolyte (BGE) for 20 min using a rinse pressure of 20 psi (138 kPa). The CE separations were performed at 25 ˚C under reversed polarity with an applied voltage of -18 kV, and UV absorbance was monitored at 226 nm (for iodide,

nitrate and thiocyanate) and 288 nm (for NDS). Prior to each analysis, the capillary was flushed with the BGE for 3 min at 20 psi (138 kPa) followed by a long sample injection via hydrodynamic pressure for 80 s at 0.5 psi (3.4 kPa). At the start of each day, a blank sample, a calibrant mixture, and a QC were analyzed by CE prior to a randomized analysis of a batch of individual urine samples with a QC sample repeatedly analyzed after every batch of ten runs. At the end of each day, the capillary was flushed with deionized water for 5 min and the inlet and outlet ends of the capillary were stored in vials containing deionized water overnight. Calibration curves were performed for iodide, nitrate and thiocyanate by CE, where their integrated peak areas were normalized to NDS as the internal standard. Urinary iodide, nitrate and thiocyanate concentrations from PURE-24USE participants were reported in terms of their absolute concentrations (µg/L or mg/L) or daily excretion amounts based on total volume of 24 h urine collected (µg/day or mg/day) that forgoes the need for creatinine adjustment. Dietary iodine intake estimates for PURE were derived from measured daily excretion amounts and adjusted by an iodine bioavailability of 92% [39]. Missing data following analysis of all 24 h urine samples were 0%, 2%, and 11% for nitrate, iodide and thiocyanate, respectively, if below method detection limits (S/N = 3; 0.020 µmol/L or 2.5 µg/L for iodide) or as a result of matrix interferences. For iodide non-detects, a missing value replacement was used corresponding to the lowest concentration measured in the cohort divided by 3.

2.3.4 Statistical Analyses

Descriptive statistics, box-whisker plots, and pair-wise Spearman rank correlations for urinary iodide, nitrate, thiocyanate and sodium concentrations and their equivalent daily excretion amounts were performed using MedCalc 12.5 statistical software (MedCalc Software Ltd, Ostend, Belgium). Routine data processing for population stratification (quintiles)

and least-squares linear regression for calibration curves of urinary anions measured by CE were performed using Microsoft Excel (Redmond, WA, USA). Representative electropherograms and control charts were plotted using Igor Pro 5.04B (Wavemetrics Inc., Lake Oswego, OR, USA). Protective and risk factors associated with iodine deficiency (<100 µg/L or <150 µg/day) were evaluated using a binary multivariate logistic regression with and without adjustment for covariates. An odds ratio (OR) < 1.0 corresponds to a positive association with the outcome (i.e., protective factor against iodine deficiency), whereas an OR > 1.0 indicates a negative association to the outcome (i.e., risk for iodine deficiency). Also, analysis of covariance (ANCOVA) was performed on *log*-transformed data using SPSS 23.0 statistical software (IBM SPSS, Chicago, IL, USA). Statistical tests comparing iodine status among PURE participants as a function of regional site, iodine supplementation, T4 prescription, and dietary intake of specific foods were performed unadjusted and adjusted for covariates, including age, sex, body mass index (BMI), total caloric intake, current smoking, alcohol use, education, and diet quality (Alternative Healthy Eating Index, AHEI score) unless otherwise noted. Statistical significance was set at $p < 0.05$ with a Bonferroni correction used for multiple comparisons.

2.4 Results

2.4.1 PURE Cohort Characteristics and CE Method Performance

Table S2.1 summarizes the cohort characteristics of this cross-sectional study comprising 800 PURE-24USE participants from four regional sites across Canada who completed 24 h urine sampling and a short questionnaire during their clinical visit. Overall, a sex-balanced cohort (females, $n = 412$; males, $n = 388$) of overweight Canadian adults with a mean age of (60± 9; 36 to 83) years and BMI of (28.1± 5.7; 16.3 to 59.4) kg/m^2 were recruited primarily from urban (~ 86%) regions of Vancouver, Hamilton, Ottawa and Quebec

City. Also, about 25% of PURE-24USE participants had high blood pressure based on a resting systolic blood pressure >140 mm Hg and/or diastolic blood pressure >90 mm Hg [33]. We searched the composition of all vitamin/mineral supplements reported to be taken by PURE participants and verified that iodine-containing supplements were consumed by 12.9% (n = 103). Additionally, 7.6% were prescribed thyroxine (T4, n = 61), and 7.4% (n = 59) were current smokers. However, participants with debilitating chronic diseases, restrictive diet requirements and pregnant women were excluded. All urine samples from participants were diluted minimally and then directly analyzed by CE with UV absorbance detection, which allowed for determination of iodide, as well as nitrate and thiocyanate within 10 min (**Figure S2.1**). Good technical precision was achieved following intermittent analysis of a pooled QC urine sample (n = 93) with a mean coefficient of variance (CV) < 8% as depicted in control charts. **Table S2.2** lists the figures of merit of the CE method used for reliable quantification of iodide, nitrate and thiocyanate in 24 h urine samples from the PURE-24USE study.

2.4.2 Iodine Nutritional Status of Canadian Adults from PURE-24USE

Data distribution for urinary iodide, nitrate and thiocyanate were highly skewed (Shapiro–Wilk, p > 0.05) with concentrations varying up to 1500-fold between participants. The median 24 h urinary concentrations for iodide, thiocyanate and nitrate were 111 µg/L (0.87 µmol/L, n = 800), 680 µg/L (11.7 µmol/L, n = 713), and 73.9 mg/L (587 µmol/L, n = 800), respectively. As creatinine normalization was not required with 24 h urine collection due to recording of total urine volume from each participant, the corresponding median daily amounts excreted for iodide, thiocyanate and nitrate were determined as 226 µg/day, 1.39 mg/day and 150 mg/day, respectively. **Figure 2.1A,B** summarizes the iodine status of PURE-24USE participants classified according to WHO guidelines [21] based on 24 h UIC (µg/L) and 24 h UIE (µg/day) metrics. Overall, the median UIC of 111 µg/L was within

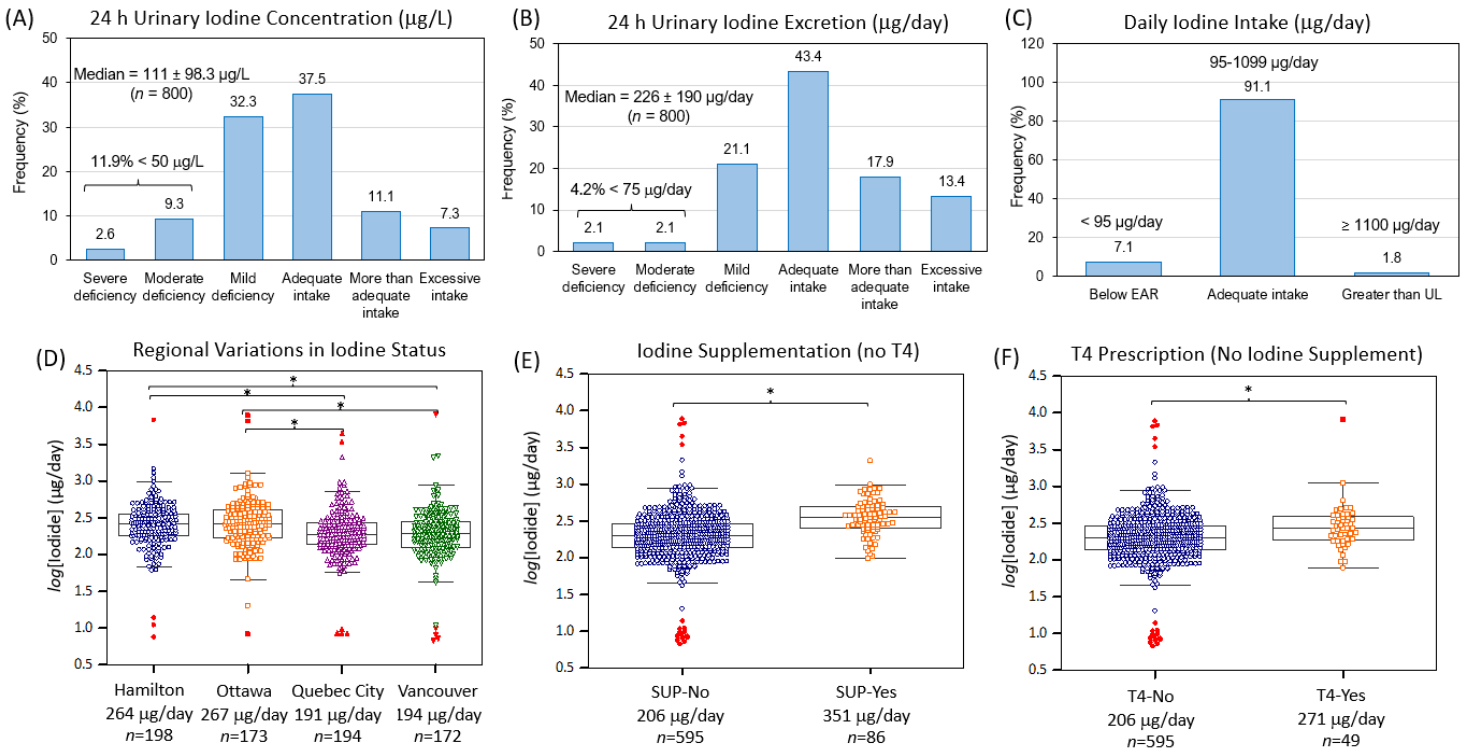


Figure 2.1. The iodine status of participants from the PURE-24USE study (n=800) based on **(A)** 24 h UIC (µg/L) and **(B)** 24 h UIE (µg/day). Median iodine concentrations (111 µg/L or 226 µg/day) indicate iodine adequacy with a low prevalence of moderate to severe iodine deficiency in the population (< 20%). **(C)** Daily iodine intake confirms only a small fraction of adult Canadians in our cohort were below EAR (7.1%, < 95 µg/day) or greater than tolerable UL (1.8%, ≥ 1100 µg/day). **(D)** ANCOVA with covariate adjustments (* $p < 0.002$ with Bonferroni correction) confirms regional variation with residents from Quebec City and Vancouver having lower iodine status than Hamilton or Ottawa ($F = 8.80$, $p = 9.82\times10^{-6}$, $n = 737$). Participants who **(E)** reported use of multivitamin supplements containing iodine, but no T4 ($F = 42.3$, $p = 1.52\times10^{-10}$, $n = 681$), or **(F)** were prescribed T4, but not taking iodine supplements ($F = 9.71$, $p = 1.91\times10^{-3}$, $n = 644$) had greater iodine status than controls.

adequacy requirements (100–199 µg/L) with 11.9% of the population (<20%) having moderate (20–49 µg/L, 9.3%) or severe (<20 µg/L, 2.6%) iodine deficiency. Similar outcomes of adequate iodine nutrition in the population were evident when using more robust 24 h median UIE of 226 µg/day that was above the recommended dietary allowance (RDA of 150 µg/day) with a much lower fraction of the population categorized with moderate (2.1%) or severe (2.1%) iodine deficiency (<75 µg/day). In contrast, more participants had excessive iodine intake based on 24 h UIE (13.4%; 450 µg/day) as compared to 24 h UIC (7.3%; 300 µg/L) metrics. **Figure 2.1C** confirms that most participants (~ 91%) had a daily iodine intake (assuming 92% bioavailability) within an acceptable interval (95–1099 µg/day) with only 7.1% below EAR (<95 µg/day) and 1.8% greater than tolerable UL (≥1100 µg/day).

2.4.3 Factors Contributing to Iodine Deficiency in Canada

Table 2.1 summarizes the major variables associated with iodine deficiency (<150 µg/day or <100 µg/L) when using a binary logistic regression model after adjustments for age, sex, BMI, total caloric intake, and diet quality (AHEI score). Overall, variables that were consistently protective against iodine deficiency (OR < 1.0, $p < 0.05$) using either UIE and UIC included use of iodine supplements, T4 prescription, site location, and dairy intake. Also, urinary sodium excretion was inversely associated with the risk of iodine deficiency based on UIE reflecting greater intake of iodized salt in foods. All other self-reports of salt intake (e.g., salty foods, table salt use at table and cooking, processed foods etc.) were not associated with iodine status. Age, as well as greater bread and cereal intake, were marginally protective against iodine deficiency based on UIC, whereas alcohol consumption increased risk for iodine deficiency. Interestingly, 24 h urine volume showed opposing trends likely reflecting a dilution effect when using UIC as a metric for iodine status resulting in an apparent risk for iodide deficiency. However,

Table 2.1. Protective and risk factors for iodine deficiency (< 150 µg/day or < 100 µg/L) among PURE-24USE participants (n = 800) using a binary linear logistic regression. Significant variables (p < 0.05) are bolded after adjustments for age, sex, BMI, total caloric intake and AHEI score.

Variable	24 h UIE (µg/day)		24 h UIC (µg/L)	
	OR (95% CI)	***p*-value**	**OR (95% CI)**	***p*-value**
Age	0.99 (0.99-1.01)	0.257	**0.98 (0.97-1.00)**	**0.0516**
Male sex	0.80 (0.56-1.14)	0.215	0.76 (0.56-1.03)	0.0750
BMI (> 27 kg/m^2)	**0.66 (0.47-0.92)**	**0.0153**	0.99 (0.96-1.02)	0.515
24 h Urine volume (L)	**0.69 (0.56-0.85)**	**4.07×10^{-4}**	**2.31 (1.91-2.81)**	**5.87×10^{-17}**
Current smoker	0.85 (0.43-1.67)	0.635	1.13 (0.64-1.99)	0.671
Current alcohol consumer	1.59 (0.98-2.47)	0.059	**1.48 (1.00-2.20)**	**0.0505**
Study site:				
Hamilton	1.00 (ref.)	--	1.00 (ref.)	--
Vancouver	**2.54 (1.52-4.23)**	**3.57×10^{-4}**	**1.83 (1.20-2.81)**	**5.31×10^{-3}**
Quebec City	**2.58 (1.57-4.22)**	**1.74×10^{-4}**	**1.89 (1.25-2.84)**	**2.41×10^{-3}**
Ottawa	1.19 (0.69-2.05)	0.531	1.28 (0.85-1.93)	0.839
Rural location	0.99 (0.69-1.60)	0.953	1.28 (0.85-1.93)	0.246
Iodine supplementation	**0.18 (0.08-0.41)**	**6.30×10^{-5}**	**0.31 (0.19-0.52)**	**8.77×10^{-6}**
T4 prescription	**0.33 (0.14-0.78)**	**1.20×10^{-2}**	**0.43 (0.23-0.79)**	**6.76×10^{-3}**
Dairy intake (g/day)	**0.999 (0.998-0.999)**	**3.94×10^{-4}**	**0.999 (0.998-1.00)**	**4.41×10^{-4}**
Starch intake (g/day)	0.999 (0.997-1.00)	0.105	**0.999 (0.997-1.00)**	**0.0470**
Sodium excretion (g/day)	**0.71 (0.61-0.84)**	**3.03×10^{-5}**	0.93 (0.83-1.05)	0.265
Salty food intake (g/day)	1.00 (1.00-1.01)	0.108	1.00 (0.97-1.00)	0.856

CI = confidence interval, OR = odds ratios, ref. = reference. Hosmer–Lemeshow goodness-of-fit logistic regression, UIE, urinary iodine excretion; UIC, urinary iodine concentration. Statistically significant variables (p < 0.05) are bolded.

correction for differences in hydration based on UIE likely accurately reflects a true protective effect due to iodine uptake from greater daily drinking water/fluid consumption.

Other factors associated with iodine deficiency (OR > 1.0, p < 0.05) were site location, where residents from Vancouver and Quebec City had about a 2.5-fold greater relative risk as compared to Hamilton or Ottawa. **Figure 2.1D** shows box plots confirming that the median 24 h UIE for residents of Hamilton (264 µg/day, n = 198) and Ottawa (267 µg/day, n = 173)

were higher than Vancouver (194 µg/day, n = 172) and Quebec City (191 µg/day, n = 194) based on ANCOVA after a Bonferroni correction and adjustments for age, sex, BMI, total caloric intake, AHEI score, education, smoking status and alcohol use (F = 8.80, p = 9.82×10^{-6}, n = 737). In contrast, iodine supplement use (F = 42.3, p = 1.52×10^{-10}, n = 681) without a T4 prescription had the greatest effect on iodine status when compared to T4 alone without iodine supplement use (F = 9.71, p = 1.91×10^{-3}, n = 644) as illustrated in **Figure 2.1E,F. Table S2.3** confirms that lower dairy and bread and cereal consumption were dietary patterns associated with iodine inadequacy when participants were categorized based on their iodine status as quintiles (Q1 vs. Q2–5) with a larger fraction from Vancouver and Quebec City, with fewer taking iodine supplements or T4 hormone therapy. **Table S2.4** summarizes a Spearman rank correlation analysis of self-reported dietary intake of specific foods as a function of UIE (µg/day) after excluding participants using iodine supplements and/or T4. Overall, dairy intake had the strongest positive correlation (r = 0.24, p = 2.38×10^{-9}, n = 611) with iodine status that was most evident for residents in Hamilton and Ottawa. In contrast, intake of bread and cereal, as well as processed food or a combination of red and processed meat were important sources related to iodine status for residents of Quebec City. Other potential iodine containing foods surveyed in this study, such as fish, eggs, and various vegetables, were not significant sources of dietary iodine.

2.4.4 Risk Assessment of Iodine Deficiency from Environmental Iodine Uptake Inhibitor Exposures

Figure 2.2A confirms that dairy consumption was an important dietary pattern associated with iodine deficiency in the PURE-24USE study (F = 18.7, p = 1.75×10^{-5}, n = 725) as compared to 24 h sodium excretion (F = 16.9, p = 4.47×10^{-5}, n = 737), and bread and cereal consumption (F = 1.44, p = 0.230,

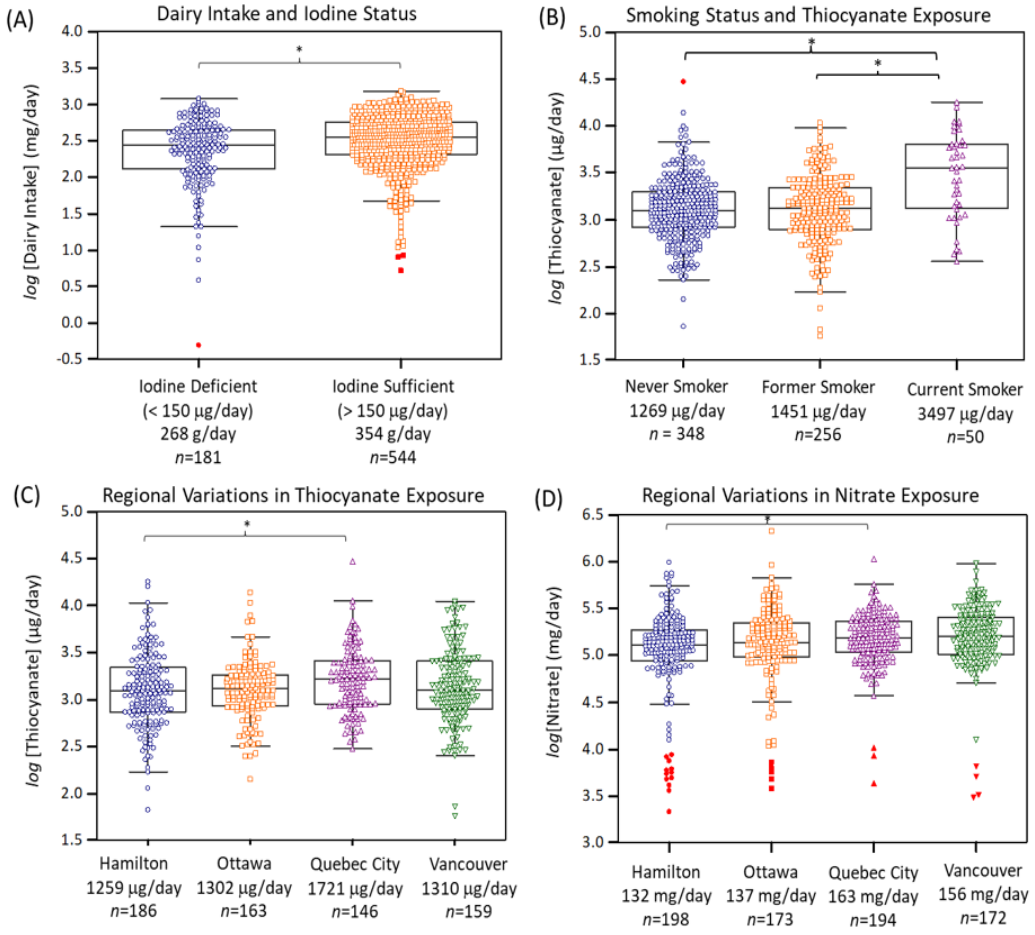


Figure 2.2. (A) Dairy intake was the most significant dietary exposure associated with iodine deficiency (F = 18.7, p = 1.75×10^{-5}, n = 725), whereas **(B)** current smoking was a lifestyle factor assocated with elevated thiocyanate exposure as compared to former or never smokers (F = 19.5, p = 5.82×10^{-9}, n = 654). Modest regional variations to exposure to environmental iodine uptake inhibitors, **(C)** thiocyanate and **(D)** nitrate were found. Overall, residents from Quebec City were exposed to both higher thiocyanate (F = 3.32, p = 0.0194, n = 654) and nitrate (F = 3.61, p = 0.0130, n = 737) relative to Hamilton when using ANCOVA after adjustment for covariates with a Bonferroni correction (* p < 0.05).

n = 728); the latter food source of iodine was only marginally significant when comparing iodine deficient vs. iodine sufficient participants in an unadjusted student's t-test (p = 0.0551). Other environmental exposures may also modulate

iodine deficiency risk despite adequate iodine nutrition. In this case, urinary thiocyanate and nitrate concentrations and their daily excretion amounts were also measured in this study. Overall, median urinary thiocyanate and nitrate concentrations were 6.1 and 670-fold higher than corresponding iodide levels. **Figure 2.2B** highlights that urinary thiocyanate was strongly dependent on smoking status (F = 19.5, p = 5.82×10^{-9}, n = 654) with median 24 h thiocyanate excretion of about 1270 µg/day (n = 348), 1450 µg/day (n = 256), and 3500 µg/day (n = 50) corresponding to never smokers, former smokers, and current smokers, respectively. Otherwise, **Table S2.5** highlights that dietary sources of thiocyanate (after excluding current smokers) were only weakly associated with intake of processed meat, cruciferous vegetables, and eggs (r ~ 0.10, p ~ 0.020, n = 620).

Table S2.6 highlights that urinary nitrate had a moderate association to the intake of (total) vegetables (r = 0.17, p = 2.46×10^{-6}, n = 759), as well as other specific vegetables (e.g., green leafy, cruciferous), and fruit. Indeed, there was a modest difference in 24 h urinary nitrate excretion when comparing low vs. high consumers of vegetables (F = 4.46, p = 0.0351, n = 728). However, there was no association of higher urinary nitrate excretion and lower blood pressure, nor was it related to hypertension prevalence [40]. There were regional variations in thiocyanate and nitrate exposures across Canada, although not as striking as for iodine status (**Figure 2.1D**). Overall, **Figure 2.2C,D** demontrate that residents from Quebec City had greater exposure to both thiocyanate (F = 3.32, p = 0.0194, n = 654) and nitrate (F = 3.61, p = 0.0130, n = 737) relative to Hamilton after covariate adjustments and Bonferroni correction. Also, residents from Vancouver had modestly elevated exposure to nitrate as compared to Hamilton in an unadjusted ANOVA model.

2.5 Discussion

2.5.1 Iodine Nutritional Status of Canadian Adults

Canada first introduced mandatory iodized salt for table or household use in 1949 as a prophylaxis to prevent iodine deficiency disorders due to the prevalence of iodine-deficient soils [41]. However, continuous monitoring of national programs is needed to optimize iodized salt content (~25 mg/kg) to ensure adequate nutrition in diverse populations with changing eating patterns [42]. A household iodine nutrition survey by the Canadian Health Measures Survey from 2007 to 2009 [29], as well as a recent study in children and young women in Canada [30], concluded adequate iodine nutrition in the population, including in high-risk demographic groups. The PURE-24USE study recently reported that daily sodium intake for Canadians was similar to other Western countries with about 47% of participants consuming <3 g/day [33] unlike other regions prone to excessive salt intake [11]. Thus, population data does not support sodium restriction as a public health policy for blood pressure reduction in Canada [43]. As a result, there is growing concern of the impact on iodine nutrition when promoting 'heart-healthy' salt-restricted diets, which may also include processed foods lacking iodized salt [44]. Additionally, vegans and vegetarians have greater risk for iodine deficiency with about half being below an EAR of 100 µg/day without iodine supplementation [45]. In fact, low iodine and selenium intake among vegans and vegetarian women represents a nutritional vulnerability [46]. To the best of our knowledge, our work is the first epidemiological study to examine iodine nutrition and exposure to iodine uptake inhibitors in Canadian adults (2012 to 2013) using a robust 24 h urine collection procedure for direct assessment of UIE [33]. In contrast, determination of iodine status from random spot urine samples is prone to significant day-to-day [47] and diurnal variations with peak concentrations excreted 4–5 h after main meals [48]. The CE assay used in this study offers a

simple and low-cost microseparation platform compared to ion chromatography-tandem mass spectrometry [49] to differentiate urinary iodide from other iodine species (e.g., iodine, iodate etc.) while also allowing for the analysis of nitrate and thiocyanate [34,35] unlike inductive coupled plasma-mass spectrometry.

We confirmed adequate iodine nutrition based on a median 24 h UIE of 226 µg/day and 24 h UIC of 111 µg/L with 4.2% and 11.9% categorized as moderately and severely deficient respectively based on WHO guidelines [21] in our cross-sectional study of community-dwelling participants living in four cities across Canada. However, dietary thresholds for estimating iodine status have assumed a mean 24 h urine volume of 1.5 L in adults. In this case, the impact of hydration status can overestimate the prevalence of iodide deficiency when relying on UIC as compared to UIE, which represents an age-old problem in urinalysis [50]. In our study, 7.1% of participants have sub-optimal iodine intake < EAR of 95 µg/day with few (1.8% > UL) prone to the deleterious effects from excessive iodine consumption of over-iodized salt, seaweeds, iodine supplements, medications, or a combination of these sources [13]. These two latter extreme conditions may be associated with adverse thyroid related health effects based on recommended Canadian dietary reference intakes [51]. However, thyroid function and other biomarkers of hyper-/hypothyroidism (e.g., thyroid-stimulating hormone) were not evaluated in the PURE-24USE study, which focused on the impact of dietary salt intake on blood pressure and hypertension [33].

Overall, use of iodine containing multivitamins was the single most important factor contributing to iodine adequacy in our study as compared to T4 use. For instance, the median UIE for individuals taking both iodine supplements and T4, iodine supplements alone (no T4), T4 hormone therapy alone (no iodine supplements) as compared to controls was 575 µg/day (n = 8),

360 µg/day (n = 95), 271 µg/day (n =53) and 206 µg/day (n = 644), respectively. The use of iodine supplements has been shown to be a strong predictor of iodine status as compared to other dietary sources [46], including in pregnant women [52]. Indeed, a large fraction of popular adult multivitamin products now contain iodine primarily as potassium iodide [53]. However, an analysis of prenatal multivitamins in the US market revealed that some products may contain more than three-times the recommended daily intake of iodine especially if derived from kelp [54]. T4 is a prescribed thyroid hormone often used for treatment of hypothyroidism due to a thyroid dysfunction impairing normal iodine uptake, such as Hashimoto's thyroiditis. Yet, certain patients may be prescribed T4 for non-thyroid indications, such as treatment for fatigue or obesity [55]. Deiodination of T4 following ingestion and metabolism (average dosage ~ 125 µg) likely results in its preferential renal excretion as iodide [56] that increases UIE to a greater extent than typical food sources of iodine in the Canadian diet. However, simultaneous intake of T4 and iodine supplements is not recommended given concerns of excessive iodine intake with potential risks for hyperthyroidism. Otherwise, only two participants in this study were reported to be using another iodine containing prescribed medication, namely amiodarone. However, recent use of iodine containing contrast agents for diagnostic imaging was not included in the questionnaire.

2.5.2 Major Dietary Sources of Iodine Nutrition in Canada

Dairy intake was the most significant food source to differentiate iodine status (p = 1.75×10^{-5}) in our cohort of Canadian adults after adjustment for covariates (**Figure 2.2A**) as compared to total salt intake (i.e., 24 h urinary sodium excretion), or bread and cereal, and processed food consumption. Milk and dairy products are important sources of iodine that contribute about 40% of total iodine nutrition in non-pregnant adults relative to about 11% for fish and seafood in the United Kingdom [57]. However, dairy may constitute a greater

fraction of total dietary iodine in other western countries given its more frequent consumption than fish [52]. In fact, postmenopausal women with reduced milk intake are at greater risk for iodine deficiency as compared to daily milk consumers, despite regular iodized salt use [58]. A survey of Canadian dairy farms reported a variable iodine content in milk in different provinces, which was dependent on feeding and sanitation practices, such as spraying or dipping teats with iodophors before milking [59]. Thus, human iodine intake from milk and dairy products arises from cattle fodder and feed fortification, as well as indirectly via transdermal uptake and/or incidental contamination of iodine containing disinfectants during milking [60]. This scenario has been characterized as an accidental public health triumph for eliminating endemic goiter in Britain [61]. The median dairy intake for PURE participants was 333 g/day (ranging from 0 to 1520 g/day), which corresponds to a theoretical iodine intake of 102 µg/day or 68% of the recommended non-pregnant adult daily requirement assuming a mean iodine concentration of 304 µg/kg in Canadian milk [59]. Data from the USA reported that iodine content of retail milk products was variable with an average 85 µg/serving (240 mL) that did not depend on milk fat content while supplying about 57% of daily iodine intake [18]. Recent studies have also demonstrated that iodide is the predominant iodine species in cow milk that has high bioavailability, which is recommended for children and pregnant women given their higher iodine nutritional requirements [62]. However, there is growing risk for iodine deficiency due to increased consumption of unfortified milk alternative drinks derived from soya, oat, hemp, rice and various nuts (e.g., almond, coconut) that contain only 1.7% of cow milk's iodine content [63]. Although most of these milk substitutes are fortified with calcium, few products are fortified with iodine.

Self-reported measures of salt use from questionnaires were not identified as significant source of dietary iodine in the PURE-24USE study with only

processed food consumption being weakly correlated to UIE. On the other hand, daily urinary sodium excretion was found to differentiate iodine status after adjustment for covariates ($p = 4.47\times10^{-5}$). For instance, iodine deficient (<150 µg/day) as compared to sufficient (≥150 µg/day) participants with completed diet records ($n = 737$) had a median urinary sodium excretion of 2.74 g/day and 3.15 g/day, respectively. Overall, only 10% of participants had excessive sodium intake of ≥5 g/day with a median sodium intake of 3.08 g/day, which highlights that public health policies to restrict sodium intake in the population [44,45] may negatively impact universal iodized salt programs unless other iodine-rich sources of food are regularly consumed [13]. This policy conflict is reflected by recent changes to Canada's Food Guide in 2020 that discourages animal protein consumption (i.e., milk and dairy) in favor of plant-based protein sources while also recommending meals to be prepared with little to no added salt [64]. Although fish consumption was not associated with iodine status in this study, bread and cereal intake was weakly protective against iodine deficiency albeit much less significant than dairy products, as well as iodine supplement or T4 use. This may reflect the declining use of iodate as a conditioner in bread and baking products in North America [12] in contrast to public health initiatives to fortify breads with iodized salt, such as in Australia [16] and Denmark [65].

2.5.3 Regional Variations in Iodine Deficiency Across Canada Modulated by Exposure to Iodine Uptake Inhibitors

An unexpected result from our multi-center cross-sectional study was the variation in iodine status for PURE participants across the four Canadian study sites, as well as their differential exposure to environmental iodine uptake inhibitors. Similar regional variations in iodine status were reported in three cities in Turkey whose population is mildly iodine deficient [32]. Also, regional variations in iodine status have been reported across four cities in

China with evidence of adequate iodine status [66]. Although there is iodine adequacy on a population level in Canada, residents from Quebec City and Vancouver (~ 191 µg/day or 92.1 µg/L, *n* = 400) were at a 2.5-fold greater relative risk for iodine deficiency as compared to Hamilton and Ottawa (~269 µg/day or 124 µg/L, *n* = 400) after covariate adjustments and Bonferroni correction. These regional variations in iodine status persisted after excluding for differential iodine supplement and T4 use. Yet, residents from Quebec City had the highest 24 h urinary sodium excretion (median of 3.70 g/day) and dairy intake (median of 379 g/day) relative to other sites, whereas Vancouver residents had average sodium excretion (median of 2.99 g/day) and a lower mean dairy consumption (median of 285 g/day). The variable iodine content of Canadian milk producers differs regionally [59] likely explains these discordant trends when relying on a standardized questionnaire for diary intake for participants from three different provinces, including Ontario, British Columbia, and Quebec.

Indeed, there was a poor correlation between self-reported dairy intake and iodine status for residents from Quebec City and Vancouver, unlike Hamilton and Ottawa (both in Ontario) implying regional variations in milk iodine content. In contrast, the iodine status of residents from Quebec City had a weak correlation with bread and cereal intake, as well as processed food and processed meat consumption reflecting distinctive eating patterns despite no overall difference in diet quality (AHEI score). However, alcohol consumption weakly increased risk for iodine deficiency in the PURE-24USE study, which is not consistent with the association of moderate alcohol use and lower rates of goiter, single thyroid nodules and autoimmune hypothyroidism [67,68]. The analysis of iodine content in household drinking water from urban and rural sites in Canada may represent an important yet unexplained dietary iodine source not examined in this study. For instance, there is growing recognition of the importance of iodine-rich spring and

ground water sources which may lead to excessive iodine intake in certain regions [69]. Indeed, there was over a ten-fold range in hydration status recorded among PURE participants with a median 24 h urine volume of 2.13 L (n = 800) ranging from 0.63 to 6.77 L that is corrected with UIE determination indicative of a protective effect against iodine deficiency due to greater daily fluid intake.

Although urinary perchlorate was not detected by the CE method, two abundant iodine uptake inhibitors were measured together with iodide, namely thiocyanate and nitrate. For instance, a median urinary perchlorate concentration of 3.2 µg/L was reported in pregnant women from Toronto, Canada that was about 100-fold lower than urinary thiocyanate [70]. Although thiocyanate is a weaker antagonist of the sodium-iodide symporter, the relative potency of perchlorate to inhibit iodine uptake is about 15- and 240-times that of thiocyanate and nitrate on a molar concentration basis [71]. Thus, the much higher concentrations of thiocyanate and nitrate render these ubiquitous thyroid antagonists essential when estimating their combined inhibitory effect based on a perchlorate equivalent concentration [72]. Although thiocyanate is biosynthesized in-vivo, exogenous sources are derived by smoke inhalation of hydrogen cyanide following combustion of nitrogen-containing tobacco alkaloids, and the digestion of certain goitrogenic foods, such as vegetables containing cyanogenic glycosides [73]. As expected, urinary thiocyanate excretion was strongly determined by smoking status [74] with smaller background dietary contributions from the intake of cruciferous vegetables and processed meat. Overall, there was a 2.6-fold (p = 5.82×10^{-9}) greater thiocyanate excretion in current smokers as compared to former/never smokers. Women who are heavy smokers have been reported to be at greater risk for hypothyroxinemia from excessive thiocyanate exposures [75]. The median urinary thiocyanate excretion of 1395 µg/day (or 680 µg/L) was 67-fold greater than iodide with higher thiocyanate exposures measured for

residents in Quebec City as compared to Hamilton, Ottawa, and Vancouver after covariate adjustment.

Similarly, residents of Quebec City also had modestly higher exposure to nitrate relative to Hamilton. Overall, there was a 664-fold higher urinary nitrate concentration than iodide with a median 24 h excretion of 150 mg/day (or 7.41 mg/L, n = 800). Unlike thiocyanate, urinary nitrate was unrelated to tobacco smoking, and it had a stronger correlation to dietary intake of vegetables, notably green leafy vegetables. However, residents from Quebec City also had nitrate exposures from processed meat and fruit intake unlike other study sites (**Table S2.6**). Similar to iodide, drinking water is likely an unaccounted source of nitrate exposures [76] in the PURE-24USE study, which may be acute in ground water contaminated by agricultural fertilizer run-off. Although greater nitrate exposure may have putative health benefits to reduce hypertension and cardiovascular disease risk reflecting higher nitric oxide levels [41], there was no significant correlation to blood pressure (systolic or diastolic) or hypertension incidence in this study. Nonetheless, nitrate and thiocyanate are anticipated to have a 96-fold and 14-fold greater inhibitory effect on iodine uptake as compared to reported urinary perchlorate levels [71] given their much higher exposure levels consistent with previous risk assessment calculations [77]. Consequently, residents from Vancouver and especially Quebec City may be at greater relative risk for iodine deficiency due to their suboptimal iodine nutritional status and greater combined exposures to thiocyanate and/or nitrate.

2.5.4 Study Strengths, Limitations and Future Perspectives

Major strengths of this study include the use of a robust 24 h urine collection procedure together with a coordinated 24 h dietary questionnaire timed during specimen drop-off at four different clinical sites across Canada. A validated CE method was also used for quantitative iodide, thiocyanate and nitrate determination in urine directly after a simple dilution step. Moreover,

complementary statistical methods were adjusted for potential confounders with key outcomes remaining robust relative to unadjusted models. Limitations include that this cross-sectional study was not representative of Canadians given selection criteria for participant selection was focused on sodium and potassium intake primarily in older persons, which excluded children and pregnant women [33]. Even though total urine volume was used to correct for variable hydration status when reporting iodine status as UIE, the recording of daily volume of water or beverages ingested, source(s) of water (e.g., tap, bottle), as well as iodide content analyses of local drinking water and commercial milk products in different sites were also study limitations. A more detailed food frequency questionnaire that included specific iodine-rich foods, such as seaweed/marine algae, and ocean fish/seafood, as well as adherence to increasingly popular diets (e.g., vegetarian, vegan, lactose-intolerant, ketogenic, low-salt, paleolithic) is recommended in future studies given their likely impact on iodine status [78]. Although nitrate exposure was not related to hypertension in this study, iodine deficiency with low urinary iodide levels is associated with hyperlipidemia and greater cardiovascular disease risk highlighting the broader public health benefits of optimal iodine prophylaxis [79,80]. Future studies involving iodine nutrition in older persons would benefit from biochemical measures of thyroid function, blood lipid panels, and inflammatory biomarkers during clinical visits. Direct analysis of circulating levels of iodide in serum may also provide deeper insights into iodine status and thyroid function than urine biomonitoring [66]. Lastly, national food guidelines recommending salt-restricted foods and plant-based protein substitutes warrant further scrutiny to the potential negative impacts on iodine deficiency in susceptible populations without iodine supplementation or mandatory iodine fortification of staple foods (e.g., bread) and commercial products (e.g., milk substitutes).

2.6 Conclusions

In summary, the iodine status of Canadian adults (n = 800) surveyed from 2012 to 2013 was determined to be adequate on a population level with a low prevalence of moderate to mild deficiency or excessive iodine intake. Overall, 24 h UIE provided a more robust indicator of iodine status than UIC that allowed for direct assessment of dietary reference intervals without creatinine adjustments. Iodine supplement use, T4 prescription, dairy intake, 24 h sodium excretion, and 24 h urine volume were key protective factors against iodine deficiency in this study. On the other hand, residents from Quebec City and Vancouver were at greater risk for iodine deficiency than Hamilton and Ottawa. These regional differences in iodine status may be further exacerbated by greater exposure to thiocyanate and nitrate as ubiquitous iodine uptake inhibitors. Continued iodine surveillance is warranted given greater consumption of processed foods, increased popularity of salt-restricted and other specialized diets, as well as emerging environmental exposures that may increase iodine deficiency risk in Canada. This work highlights that national level iodine adequacy may obscure regional differences in iodine status in local populations. Greater public awareness of the importance of optimal iodine nutrition in a healthy diet is strongly recommended along with public health guidelines that better align the optimal dietary intakes of sodium and iodine in the population.

2.7 Acknowledgements

The authors acknowledge support from the Public Health Agency of Canada and the Population Health Research Institute, as well as PURE-24USE participants who were involved in this study.

2.8 References

1. Chung H.R. Iodine and thyroid function. *Ann Pediatr Endocrinol Metab.* **2014** *19,* 8–12.

2. Larsen P.R.; Zavacki A.M. The role of the iodothyronine deiodinases in the physiology and pathophysiology of thyroid hormone action. *Eur Thyroid J.* **2012** *1,* 232–242.

3. Zimmermann M.B.; Andersson M. Assessment of iodine nutrition in populations: Past, present, and future. *Nutr Rev.* **2012** *70,* 553–570.

4. Bath S.C.; Steer C.D.; Golding J.; Emmett P.; Rayman M.P. Effect of inadequate iodine status in UK pregnant women on cognitive outcomes in their children: Results from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Lancet* **2013** *382,* 331–337.

5. Nazeri P.; Mirmiran P.; Shiva N.; Mehrabi Y.; Mojarrad M.; Azizi F. Iodine nutrition status in lactating mothers residing in countries with mandatory and voluntary iodine fortification programs: An updated systematic review. *Thyroid* **2015** *25,* 611–620.

6. Laurberg P.; Nøhr S.B.; Pedersen K.M.; Hreidarsson A.B..; Andersen S.; Bülow Pedersen I.; Knudsen N.; Perrild H.; Jørgensen T.; Ovesen L. Thyroid disorders in mild iodine deficiency. *Thyroid* **2000** 10, 951–963.

7. Zimmermann M. B.; Boelaert K. Iodine deficiency and thyroid disorders. *Lancet Diabetes Endocrinol.* **2015** *3,* 286–295.

8. Bilal M. Y.; Dambaeva S.; Kwak-Kim J.; Gilman-Sachs A.; Beaman K.D. A Role for iodide and thyroglobulin in modulating the function of human immune cells. *Front. Immunol.* **2017** *8,* 1573.

9. Pearce E. N.; Andersson M.; Zimmermann M. B. Global iodine nutrition: Where do we stand in 2013? *Thyroid* **2013** *23,* 523–528.

10. Li Y.; Teng D; Ba J.; Chen B.; Du J.; He L.; Lai X.; Teng X.; Shi X.; Li Y.; *et al.* Efficacy and safety of long-term universal salt iodization on thyroid disorders: Epidemiological evidence from 31 provinces of mainland China. *Thyroid* **2020** *30,* 568–579.

11. Zimmermann M. B.; Andersson M. GLOBAL ENDOCRINOLOGY: Global perspectives in endocrinology: Coverage of iodized salt programs and iodine status in 2020. *Eur. J. Endocrinol.* **2021** *185,* R13–R21.

12. Farebrother J.; Zimmermann M. B.; Andersson M. Excess iodine intake: sources, assessment, and effects on thyroid function. *Ann. N. Y. Acad. Sci.* **2019** 1446, 44–65.

13. Krela-Kazmierczak I.; Czarnywojtek A.; Skoracka K.; Rychter A.M.; Ratajczak A.E.; Szymczak-Tomczak A.; Ruchała M.; Dobrowolska, A. Is there an ideal diet to protect against iodine deficiency? *Nutrients* **2021**, *13*, 513.

14. Nista F.; Bagnasco M.; Gatto F.; Albertelli M.; Vera L.; Boschetti M.; Musso N.; Ferone D. The effect of sodium restriction on iodine prophylaxis: A review. *J. Endocrinol. Invest.* **2022** doi: 10.1007/s40618-022-01749-y.

15. Knowles J.; van der Haar F.; Shehata M.; Gerasimov G.; Bimo B.; Cavenagh B.; Maramag C. C.; Otico E.; Izwardy D.; Spohrer R.; Garrett G. S. Iodine intake through processed food: Case studies from Egypt, Indonesia, the Philippines, the Russian Federation and Ukraine, 2010-2015. *Nutrients* **2017** *9*, 797.

16. Charlton K.; Probst Y.; Kiene G. Dietary iodine intake of the Australian population after introduction of a mandatory iodine fortification programme. *Nutrients* **2016**, 8, 701.

17. Pehrsson P. R.; Patterson K. Y.; Spungen J. H.; Wirtz M.S.; Andrews K. W.; Dwyer J. T.; Swanson C. A. Iodine in food- and dietary supplement composition databases. *Am. J. Clin. Nutr.* **2016** *104*, 868S–876S.

18. Roseland J. M.; Phillips K. M.; Patterson K.Y.; Pehrsson P. R., Bahadur R.; Ershow A.G.; Somanchi M. Large variability of iodine content in retail cow's milk in the U.S. *Nutrients* **2020** *12*, 1246.

19. Zhao W.; Han C.; Shi X.; Xiong C.; Sun J.; Shan Z.; Teng W. Prevalence of goiter and thyroid nodules before and after implementation of the universal salt iodization program in mainland China from 1985 to 2014: A systematic review and meta-analysis. *PLoS One* **2014** *9*, e109549.

20. König F.; Andersson M.; Hotz K.; Aeberli I.; Zimmermann M. B. Ten repeat collections for urinary iodine from spot samples or 24-hour samples are needed to reliably estimate individual iodine status in women. *J. Nutr.* **2011** *141*, 2049–2054.

21. WHO. Vitamin and mineral requirements in human nutrition. 2nd ed. World Health Organization and Food and Agriculture Organization of the United Nations; **2004**. 303–314 p.

22. Li M.; Eastman C.J. The changing epidemiology of iodine deficiency. *Nat. Rev. Endocrinol.* **2012** *8,* 434–440.

23. Vejbjerg P.; Knudsen N.; Perrild H.; Laurberg P.; Andersen S.; Rasmussen L.B.; Ovesen L.; Jørgensen T. Estimation of iodine intake from various urinary iodine measurements in population studies. *Thyroid* **2009** *19,* 1281–1286.

24. Perrine C. G.; Cogswell M. E.; Swanson C. A.; Sullivan K. M.; Chen T.C.; Carriquiry A. L.; Dodd K. E.; Caldwell K. L.; Wang C. Comparison of population iodine estimates from 24-hour urine and timed-spot urine samples. *Thyroid* **2014** *24,* 748–757.

25. Montenegro-Bethancourt G.; Johner S. A.; Stehle P.; Neubert A.; Remer T. Iodine status assessment in children: Spot urine iodine concentration reasonably reflects true twenty-four-hour iodine excretion only when scaled to creatinine. *Thyroid* **2015** *25,* 688–697.

26. Doggui R.; El Ati-Hellal M.; Traissac P.; El Ati J. Pre-analytical factors influence accuracy of urine spot iodine assessment in epidemiological surveys. *Biol. Trace Elem. Res.* **2018** *186,* 337-345

27. Chen W.;Wu Y.; Lin L.; Tan L.; Shen J.; Pearce E.N.; Guo X.; Wang W.; Bian J.; Jiang W.; Zhang W. 24-hour urine samples are more reproducible than spot urine samples for evaluation of iodine status in school-age children. *J. Nutr.* **2016** *146,* 142–146.

28. Niwattisaiwong S.; Burman K. D.; Li-Ng M. Iodine deficiency: Clinical implications. *Clev. Clin. J. Med.* **2017** *84,* 236–244.

29. Government of Canada. Iodine status of Canadians, 2009 to 2011. Statistics Canada. https://www150.statcan.gc.ca/n1/pub/82-625-x/2012001/article/11733-eng.htm (Accessed 2022 May 1).

30. Bertinato J.; Qiao C.; L'Abbé M. R. Iodine status of Canadian children, adolescents, and women of childbearing age. *J. Nutr.* **2021** *151,* 3710–3717.

31. Eisenbrand G.; Gelbke H. P. Assessing the potential impact on the thyroid axis of environmentally relevant food constituents/contaminants in humans. *Arch. Toxicol.* **2016** *90,* 1841–1857.

32. Ozpinar A.; Kelestimur F.; Songur Y.; Can O.; Valentin L.; Caldwell K.; Arikan E.; Unsal I.; Serteser M.; Inal T.; Erdemgil Y.; Coskun A.; Bakirci N.; Sezgin O., Blount B. Iodine status in Turkish populations and exposure to iodide uptake inhibitors. *PLoS One* **2014** *9,* e88206.

33. Mente A.; Dagenais G.; Wielgosz A.; Lear S.A.; McQueen M. J.; Zeidler J.; Fu L.; DeJesus J.; Rangarajan S.; Bourlaud A.S.; De Bluts A.L.; Corber E.; de Jong V.; Boomgaardt J.; Shane A.; Jiang Y.; de Groh M.; O'Donnell M. J.; Yusuf S.; Teo K. Assessment of dietary sodium and potassium in Canadians using 24-hour urinary collection. *Can J Cardiol.* **2016** *32*, 319–326.

34. Nori de Macedo A.; Teo K.; Mente A.; McQueen M.J.; Zeidler J.; Poirier P.; Lear S. A.; Wielgosz A.; Britz-McKibbin P. A robust method for iodine status determination in epidemiological studies by capillary electrophoresis. *Anal. Chem.* **2014** *86*, 10010–10015.

35. Nori de Macedo A.; Macri J.; Hudecki P. L.; Saoi M.; McQueen M. J.; Britz-McKibbin P. Validation of a capillary electrophoresis assay for monitoring iodine nutrition in populations for prevention of iodine deficiency: An interlaboratory method comparison. *J. Appl. Lab. Med.* **2017** *1*, 649–660.

36. Jakobsen J.; Ovesen L.; Fagt S.; Pedersen A.N. Para-aminobenzoic acid used as a marker for completeness of 24 hour urine: assessment of control limits for a specific HPLC method. *Eur. J. Clin. Nutr.* **1997** *51*, 514–519.

37. Teo K.; Chow C. K.; Vaz M.; Rangarajan S.; Yusuf S.; PURE Investigators-Writing Group. The Prospective Urban Rural Epidemiology (PURE) study: Examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am. Heart J.* **2009** *158*, 1–7.

38. Lee K. W.; Shin D.; Cho M. S.; Song W.O. Food group intakes as determinants of iodine status among US adult population. *Nutrients* **2016** *8*, 325.

39. Bajaj J.K.; Salwan P.; Salwan S. Various possible toxicants involved in thyroid dysfunction: A review. *J. Clin. Diagn. Res.* **2016** *10*, FE01–3.

40. Mendy A. Association of urinary nitrate with lower prevalence of hypertension and stroke and with reduced risk of cardiovascular mortality. *Circulation* **2018** *137*, 2295-2297.

41. Health Canada. Dietary Sources. https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-health-measures-survey/iodine-nutrition-biomarkers-cycle-1-canadian-health-measures-survey-food-nutrition-surveillance-health-canada.html (Accessed May 1, 2022).

42. Dold S., Zimmermann M. B., Jukic T., Kusic Z., Jia Q., Sang Z., Quirino A., San Luis T. O. L., Fingerhut R., Kupka R., Timmer A., Garrett G. S.,

Andersson M. Universal salt iodization provides sufficient dietary iodine to achieve adequate iodine nutrition during the first 1000 Days: A cross-sectional multicenter study. *J. Nutr.* **2018** *148,* 587–598.

43. Mente A.; O'Donnell M.; Rangarajan S.; McQueen M.; Dagenais G.; Wielgosz A.; Lear S.; Ah S.T.L.; Wei L.; Diaz R.; et al. Urinary sodium excretion, blood pressure, cardiovascular disease, and mortality: a community-level prospective epidemiological cohort study. *Lancet* **2018** *392,* 496–506.

44. Graudal N. Population data on blood pressure and dietary sodium and potassium do not support public health strategy to reduce salt intake in Canadians. *Can. J. Cardiol.* **2016** *32,* 283–285.

45. Groufh-Jacobsen S.; Hess S.Y.; Aakre I.; Folven Gjengedal E.L.; Blandhoel Pettersen K.; Henjum S. Vegans vegetarians and pescatarians are at risk of iodine deficiency in Norway. *Nutrients* **2020** *12,* 3555.

46. Fallon, N.; Dillon, S.A. Low intakes of iodine and selenium amongst vegan and vegetarian women highlight a potential nutritional vulnerability. *Front. Nutr.* **2020**, *7*, 72.

47. Rasmussen L.B.; Ovesen L.; Christiansen E. Day-to-day and within-day variation in urinary iodine excretion. *Eur. J. Clin. Nutr.* **1999** *53,* 401–407.

48. Als C.; Helbling A.; Peter K.; Haldimann M.; Zimmerli B.; Gerber H. Urinary iodine concentration follows a circadian rhythm: A study with 3023 spot urine samples in adults and children. *J. Clin. Endocrinol. Metab.* **2000** *85,* 1367–1369.

49. Valentín-Blasini L.; Blount B. C.; Delinsky A. Quantification of iodide and sodium-iodide symporter inhibitors in human urine using ion chromatography tandem mass spectrometry. *J. Chromatogr. A* **2007** *1155,* 40–46.

50. Cheuvront S. N. Urinalysis for hydration assessment: an age-old problem. *Am. J. Clin. Nutr.* **2016** *104,* 3–4.

51. Health Canada. Dietary Reference Intakes, Updated November 2010. https://www.canada.ca/en/health-canada/services/food-nutrition/healthy-eating/dietary-reference-intakes/tables.html (Accessed 2022 May 1).

52. Dahl L.; Wik Markhus M.; Sanchez P. V. R.; Moe V.; Smith L.; Meltzer H. M.; Kjellevold M. Iodine deficiency in a study population of Norwegian pregnant women-Results from the Little in Norway Study (LiN). *Nutrients* **2018** *10,* 513.

53. Patel A.; Lee S.Y.; Stagnaro-Green A.; MacKay D.; Wong A.W.; Pearce E.N. Iodine content of the best-selling United States adult and prenatal multivitamin preparations. *Thyroid* **2019** *29*, 124–127.

54. Leung, A.M.; Pearce E.N.; Braverman L.E. Iodine content of prenatal multivitamins in the United States. *N. Engl. J. Med.* **2009** *360*, 939–940.

55. Parle J.V.; Franklyn J.A.; Cross K.W.; Jones S.R.; Sheppard M.C. Thyroxine prescription in the community: Serum thyroid stimulating hormone level assays as an indicator of undertreatment or overtreatment. *Br. J. Gen. Pract.* **1993** *43*, 107-109.

56. Shimoda S. I.; Kasai K.; Kikuchi T.; Ieiri T. Preferential renal excretion of iodide derived from thyroxine and triiodothyronine deiodination in man. *J. Clin. Endocrinol. Metab.* **1977** *44*, 137–141.

57. Bouga M., Lean M. E. J., Combet E. Contemporary challenges to iodine status and nutrition: The role of foods, dietary recommendations, fortification and supplementation. *Proc. Nutr. Soc.* **2018** *77*, 302–313.

58. Watutantrige-Fernando S., Barollo S., Bertazza L., Sensi F., Cavedon E., Censi S., Veronese N., Ceccato F., Vianello F., Boscaro M., Nacamulli D., Camozzi V., Mian C. Iodine status in the elderly: Association with milk intake and other dietary habits. *J. Nutr. Health Food Sci.* **2017** *5*, 1–5.

59. Borucki Castro S. I.; Berthiaume R.; Laffey P.; Fouquet A.; Beraldin F.; Robichaud A.; Lacasse P. Iodine concentration in milk sampled from Canadian farms. *J. Food Prot.* **2010** 73, 1658–1663.

60. Flachowsky G.; Franke K.; Meyer U.; Leiterer M,; Schöne F. Influencing factors on iodine content of cow milk. *Eur. J. Nutr.* **2014** *53*, 351–365.

61. Phillips D. I. Iodine, milk, and the elimination of endemic goitre in Britain: The story of an accidental public health triumph. *J. Epidemiol. Community Health* **1997** *51*, 391–393.

62. van der Reijden O. L., Galetti V., Bürki, S. Zeder C., Krzystek A., Haldimann M., Berard M. B., Zimmermann J., Herter-Aeberli I. Iodine bioavailability from cow milk: A randomized, crossover balance study in healthy iodine-replete adults. *Am. J. Clin. Nutr.* **2019** *110*, 102–110.

63. Bath S. C.; Hill S.; Infante H. G.; Elghul S.; Nezianya C. J.; Rayman M. P. Iodine concentration of milk-alternative drinks available in the UK in comparison with cows' milk. *Br. J. Nutr.* **2017** *118*, 525–532.

64. Government of Canada. https://food-guide.canada.ca/en/healthy-eating-recommendations/ (Accessed on May 1, 2022).

65. Rasmussen L. B.; Jørgensen T.; Perrild H.; Knudsen N.; Krejbjerg A.; Laurberg P.; Pedersen I. B.; Bjergved L.; Ovesen L. Mandatory iodine fortification of bread and salt increases iodine excretion in adults in Denmark - a 11-year follow-up study. *Clin Nutr.* **2014** *33,* 1033–1040.

66. Yu S.; Wang D.; Cheng X.; Zhang Q.; Wang M.; Guo H.; Yu B.; Zhang X.; Xia L.; Sun D. *et al.* Establishing reference intervals for urine and serum iodine levels: A nationwide multicenter study of a euthyroid Chinese population. *Clin. Chim. Acta* **2020** *502,* 34–40.

67. Knudsen N.; Bülow I.; Laurberg P.; Perrild H.; Ovesen L.; Jørgensen T. Alcohol consumption is associated with reduced prevalence of goitre and solitary thyroid nodules. *Clin. Endocrinol.* **2001** *55,* 41–46.

68. Carlé A.; Pedersen I. B.; Knudsen N.; Perrild H.; Ovesen L.; Rasmussen L. B.; Jørgensen T.; Laurberg P. Moderate alcohol consumption may protect against overt autoimmune hypothyroidism: a population-based case-control study. *Eur. J. Endocrinol.* **2012** *167,* 483–490.

69. Liu H.L.; Lam L.T.; Zeng Q.; Han S. Q.; Fu G.; Hou C. C. Effects of drinking water with high iodine concentration on the intelligence of children in Tianjin, China. *J. Public Health* **2009** *31,* 32–38.

70. Leung A. M.; Katz P. M.; He X.; Feig D. S.; Pearce E. N.; Braverman L. E. Urinary perchlorate and thiocyanate concentrations in pregnant women from Toronto, Canada. *Thyroid* **2014** *24,* 175-176.

71. Tonacchera M.; Pinchera A.; Dimida A.; Ferrarini E.; Agretti P.; Vitti P.; Santini F.; Crump K,; Gibbs J. Relative potencies and additivity of perchlorate, thiocyanate, nitrate, and iodide on the inhibition of radioactive iodide uptake by the human sodium iodide symporter. *Thyroid* **2004** 14, 1012–1019.

72. Mortensen M. E.; Birch R.; Wong L. Y.; Valentin-Blasini L.; Boyle E. B.; Caldwell K. L.; Merrill L. S.; Moye J. Jr.; Blount B. C. Thyroid antagonists and thyroid indicators in U.S. pregnant women in the Vanguard Study of the National Children's Study. *Environ Res.* **2016** *149,* 179–188.

73. Buratti M.; Xaiz D.; Caravelliand G.; Colombi A. Validation of urinary thiocyanate as a biomarker of tobacco smoking. *Biomarkers* **1997** *2,* 81–85.

74. Jain R. Trends and variability in the levels of urinary thiocyanate, perchlorate, and nitrate by age, gender, race/ethnicity, smoking status, and exposure to environmental tobacco smoke over 2005–2012. *Sci. Total Environ.* **2016** *557-558*, 221–230.

75. Vanderver G. B.; Engel A.; Lamm S. Cigarette smoking and iodine as hypothyroxinemic stressors in U.S. women of childbearing age: A NHANES III analysis. *Thyroid* **2007** *17*, 741–746.

76. Blount B. C.; Alwis K. U.; Jain R. B.; Solomon B. L.; Morrow J. C.; Jackson W. A. Perchlorate, nitrate, and iodide intake through tap water. *Environ Sci Technol.* **2010** *44*, 9564–9570.

77. De Groef B., Decallonne B. R., Van der Geyten S., Darras V. M., Bouillon R. Perchlorate versus other environmental sodium/iodide symporter inhibitors: Potential thyroid-related health effects. *Eur. J. Endocrinol.* **2006** *155*, 17–25.

78. Manousou S.; Stål, M.; Larsson C.; Mellberg C.; Lindahl B.; Eggertsen R.; Hulthén L.; Olsson T.; Ryberg M.; Sandberg S.; Nyström H. F. A Paleolithic-type diet results in iodine deficiency: a 2-year randomized trial in postmenopausal obese women. *Eur. J. Clin. Nutr.* **2018** *72*, 124–129.

79. Tran H. V.; Erskine N. A.; Kiefe C. I.; Barton B. A.; Lapane K. L.; Do V. T. H.; Goldberg R. J. Is low iodine a risk factor for cardiovascular disease in Americans without thyroid dysfunction? Findings from NHANES. *Nutr. Metab. Cardiovasc. Dis.* **2017** *27*, 651–656.

80. Herter-Aeberli I.; Cherkaoui M.; El Ansari N.; Rohner R.; Stinca S.; Chabaa L., von Eckardstein A.; Aboussad, A. Zimmermann M. B., Iodine supplementation decreases hypercholesterolemia in iodine-deficient, overweight women: A randomized controlled trial, *J. Nutr.* **2015** *145*, 2067–2075.

2.9 Supporting Information

Table S2.1. Characteristics of PURE-24USE participants categorized by study regions, including Hamilton, Vancouver, Quebec City and Ottawa.

Predictor Variable	Study Site				
	Hamilton (n=217)	**Vancouver (n=200)**	**Quebec City (n=200)**	**Ottawa (n=183)**	**Total (n=800)**
24 h iodine concentration (µg/L) – median ± IQR	122 ± 115	91 ± 84	95 ± 87	128 ± 106	111 ± 98
24 h daily iodine excretion (µg/day) – median ± IQR	272 ± 190	194 ±170	191 ±142	267 ± 242	226 ± 191
24 h thiocyanate concentration (µg/L) – median ± IQR[a]	658 ± 779	649 ± 656	846 ± 841	668 ± 904	680 ± 811
24 h daily thiocyanate excretion (mg/day) – median ± IQR[a]	1.3 ± 1.5	1.3 ± 1.0	1.7 ± 1.8	1.4 ± 1.9	1.4 ± 1.5
24 h nitrate concentration (mg/L) – median ± IQR	67 ± 57	79 ± 72	82 ± 58	67 ± 71	74 ± 64
24 h daily nitrate excretion (mg/day) – median ± IQR	133 ± 104	167 ± 154	164 ± 123	137 ± 132	150 ± 126
24 h sodium concentration (g/L) – median ± IQR	1.5 ± 1.1	1.3 ± 1.1	1.8 ± 1.2	1.3 ± 1.0	1.4 ± 1.1
24 h daily sodium excretion (g/day) – median ± IQR	3.0 ± 1.6	3.0 ± 1.7	3.7 ± 2.0	2.7 ± 1.4	3.1 ± 1.7
Age – years, – median ± IQR	62 ± 13	59 ± 12	59 ± 14	64 ± 13	61 ± 12
BMI (kg/m^2) – median ± IQR[b]	28 ± 7.0	26 ± 6.9	27 ± 6.4	27 ± 6.2	27 ± 6.7
Sex – female:male	113:104	103:97	96:104	100:83	412:388
24 h urine volume (L)	2.1 ± 1.2	2.2 ± 1.1	2.1 ± 1.2	2.3 ± 1.4	2.1 ± 1.2
Iodine supplement – no./total no. (%)	41/217 (19)	31/200 (16)	3/200 (2)	28/183 (15)	103/800 (13)
Thyroxine (T4) intake – no./total no. (%)	11/217 (5)	11/200 (6)	21/200 (11)	18/183 (10)	61/800 (8)
Hypertension/high blood pressure diagnosis – no./total no. (%)	52/217 (24)	35/200 (18)	36/200 (18)	44/183 (24)	167/800 (21)
Dairy intake (g/day) – median ± IQR[c]	345 ± 359	285 ± 330	379 ± 397	320 ± 321	329 ± 359
Starch intake (g/day) – median ± IQR[c]	299 ± 210	290 ± 193	299 ± 192	314 ± 237	299 ± 208
Sodium intake from ffq (mg) – median ± IQR[d]	2.8 ± 1.4	2.5 ± 1.3	3.0 ± 1.4	2.7 ± 1.2	2.7 ± 1.3
Current alcohol consumer – no./total no. (%)	176/217 (81)	152/200 (76)	170/200 (85)	155/183 (85)	653/800 (82)
Current smoker – no./total no. (%)	19/216 (9)	12/200 (6)	16/200 (8)	12/179 (7)	59/800 (7)
AHEI Score[e]	38 ± 11	42 ± 14	37 ± 12	41 ± 14	39 ± 13
Location – urban:rural	211:6	186:14	106:94	182:1	685:115

[a] *Hamilton: n=204, Vancouver: n=184, Quebec City: n=152, Ottawa: n=173, and total: n=713.*
[b] *Hamilton: n=204, Vancouver: n=198, Quebec City: n=195, Ottawa: n=179, and total: n=776.*
[c] *Hamilton: n=207, Vancouver: n=173, Quebec City: n=197, Ottawa: n=182, and total: n=759.*
[d] *Hamilton: n=213, Vancouver: n=174, Quebec City: n=199, Ottawa: n=181, and total: n=767.*
[e] *Hamilton: n=212, Vancouver: n=174, Quebec City: n=199, Ottawa: n=181, and total: n=766.*

Table S2.2. Summary of figures of merit of CE assay for the determination of urinary iodide, nitrate and thiocyanate from PURE-24USE study participants (n=800).

Figures of Merit	Iodide[a]	Nitrate[b]	Thiocyanate[c]
LOD (S/N = 3)	0.020 µmol/L	0.64 µmol/L	0.12 µmol/L
LOQ (S/N = 10)	0.070 µmol/L	2.12 µmol/L	0.40 µmol/L
Linearity (R^2)	0.996	0.999	0.997
Sensitivity (µmol/L) $^{-1}$	0.045	0.006	0.014
Reproducibility (CV)[d]	5.7%	7.7%	5.3%
Missing data/Non-detects[e]	2.0%	0%	11%

[a] An external calibration curve was normalized to NDS (20 µmol/L) over a 17-fold (6 calibrants) linear dynamic range, respectively. Calibrant concentrations were 0.15, 0.20, 0.50, 1.00, 1.80, 2.60 µmol/L.
[b] An external calibration curve was normalized to NDS (20 µmol/L) over a 400-fold (6 calibrants) linear dynamic range, respectively. Calibrant concentrations were 0, 50, 100, 200, 300, 400 µmol/L.
[c] An external calibration curve was normalized to NDS (20 µmol/L) over a 24-fold (6 calibrants) linear dynamic range, respectively. Calibrant concentrations were 0.5, 1.0, 2.0, 4.0, 8.0, 12.0 µmol/L.
[d] Reproducibility was assessed based on repeated analysis of a pooled urine samples from PURE as QC every batch of 10 runs.
[e] Missing data due to concentrations below method detection limit or matrix spectral interferences from PURE cohort (n=800).

Table S2.3. Characteristics of PURE-24USE participants categorized by quintiles using 24 h UIE (μg/day).

Predictor Variable	Quintiles	
	Q1 (<133.5 μg/day; n=160)	**Q2-5 (≥133.5 μg/day; n=640)**
24 h daily iodine excretion (μg/day) – median ± IQR	97 ± 40	263 ± 184
24 h daily thiocyanate excretion (mg/day) – median ± IQR[a]	1.1 ± 1.0	1.5 ± 1.6
24 h daily nitrate excretion (mg/day) – median ± IQR	148 ± 139	151 ± 122
24 h daily sodium excretion (g/day) – median ± IQR	2.7 ± 1.5	3.2 ± 1.7
Age – years, – median ± IQR	60 ± 13	61 ± 13
BMI (kg/m^2) – median ± IQR[b]	26 ± 7	27 ± 6
Sex – female:male	88:72	324:316
24 h urine volume (L)	1.9 ± 1.3	2.2 ± 1.1
Iodine supplement – no./total no. (%)	**6/160 (4)**	**97/640 (15)**
Thyroxine (T4) intake – no./total no. (%)	**4/160 (3)**	**57/640 (9)**
Dairy intake (g/day) – median ± IQR[c]	**246 ± 303**	**352 ± 368**
Starch intake (g/day) – median ± IQR[c]	**286 ± 164**	**311 ± 217**
Study City:		
Hamilton – no./total no. (%)	**29/160 (18)**	**188/640 (29)**
Vancouver – no./total no. (%)	**59/160 (37)**	**141/640 (22)**
Quebec City – no./total no. (%)	**46/160 (29)**	**154/640 (24)**
Ottawa – no./total no. (%)	**26/160 (16)**	**157/640 (25)**
Sodium intake from FFQ (g)[d] – median ± IQR[d]	2.6 ± 1.2	2.8 ± 1.3
Current alcohol consumer – no./total no. (%)	136/160 (85)	517/640 (81)
Current smoker – no./total no. (%)	11/159 (7)	48/636 (8)
AHEI Score[e]	40 ± 14	39 ± 13
Location – urban:rural	138:22	547:93

[a] *Q1: n=135 and Q2-5: n=578.*
[b] *Q1: n=154 and Q2-5: n=622.*
[c] *Q1: n=150 and Q2-5: n=609.*
[d] *Q1: n=149 and Q2-5: n=618.*
[e] *Q1: n=149 and Q2-5: n=617.*

Table S2.4. Spearman rank correlation analysis of dietary variables associated with daily iodine excretion (µg/day) of PURE-24USE participants from the four different sites in Canada after excluding for iodine supplementation and/orT4 use.

Dietary Variable	Study Site								Total (*n*=611; median=224 µg/day)	
	Hamilton (*n*=162; median=257 µg/day)		Vancouver (*n*=137; median=189 µg/day)		Quebec City (*n*=173; median=191 µg/day)		Ottawa (*n*=139; median=260 µg/day)			
	r	***p*-value**	**r**	***p*-value**	**r**	***p*-value**	**r**	***p*-value**	**r**	***p*-value**
Dairy (g/day)	**0.39****	**2.15×10^{-7}**	0.15	7.30×10^{-2}	0.09	0.266	**0.34****	**4.33×10^{-5}**	**0.24****	**2.38×10^{-9}**
Bread and cereal (g/day)	**0.18***	**2.02×10^{-2}**	0.002	0.984	**0.15**	**5.31×10^{-2}**	0.05	0.586	**0.10***	**1.63×10^{-2}**
Processed food (g/day)	0.05	0.565	0.01	0.953	**0.22****	**4.22×10^{-3}**	0.04	0.669	**0.11****	**7.93×10^{-3}**
Red and proc. meat (g/day)	0.10	0.208	0.07	0.447	**0.16***	**3.97×10^{-2}**	0.01	0.902	0.06	0.151
Red meat only (g/day)	0.13	9.66×10^{-2}	0.08	0.338	0.14	6.05×10^{-2}	-0.01	0.946	0.06	0.171
Processed meat (g/day)	-0.07	0.378	-0.04	0.613	0.12	0.118	0.06	0.508	0.02	0.679
White meat only (g/day)	0.03	0.679	0.03	0.701	0.04	0.569	-0.03	0.746	0.02	0.607
Vegetable (g/day)	**0.16***	**4.88×10^{-2}**	-0.07	0.440	-0.01	0.910	0.04	0.686	0.01	0.802
Green leafy veg. (g/day)	0.12	0.125	-0.16	5.92×10^{-2}	-0.08	0.318	0.05	0.602	-0.004	0.923
Cruciferous veg. (g/day)	0.11	0.156	0.02	0.779	0.08	0.316	0.07	0.437	0.03	0.467
Dark yellow veg. (g/day)	0.07	0.376	-0.07	0.415	0.02	0.783	-0.02	0.853	-0.01	0.800
Egg (g/day)	0.05	0.568	-0.02	0.842	0.12	0.129	-0.08	0.345	0.05	0.239
Fruit (g/day)	0.04	0.605	-0.03	0.729	0.004	0.961	-0.09	0.270	-0.02	0.676
Fish (g/day)	-0.03	0.686	-0.06	0.454	0.02	0.812	0.02	0.796	-0.02	0.709
Salty food (g/day)	-0.01	0.914	-0.14	0.107	0.10	0.183	-0.02	0.844	0.004	0.925

Participants taking iodine containing supplements and T4 were excluded from the analysis, where r denotes the Spearman rank correlation coefficient.

***Correlation is significant at the 0.01 level (2-tailed);*

** Correlation is significant at the 0.05 level (2-tailed).*

Significant p-values at the 0.01 and 0.05 levels are bolded.

Table S2.5. Spearman rank correlation analysis of dietary variables associated with daily thiocyanate excretion (µg/day) of participants across four different sites in the PURE-24USE study after excluding for current smokers, including Hamilton, Vancouver, Quebec City and Ottawa.

Dietary Variable	Study Site								Total (median=1354 µg/day; *n*=620)	
	Hamilton (median=1251 µg/day; *n*=176)		Vancouver (median=1272 µg/day; *n*=152)		Quebec City (median=1718 µg/day; *n*=136)		Ottawa (median=1281 µg/day; *n*=156)			
	r	***p*-value**	***r***	***p*-value**	***r***	***p*-value**	***r***	***p*-value**	***r***	***p*-value**
Processed meat (g/day)	0.04	0.645	0.04	0.642	0.16	6.53×10^{-2}	0.12	0.136	**0.10***	$\mathbf{1.85\times10^{-2}}$
Red and proc. meat (g/day)	0.04	0.613	0.09	0.280	-0.01	0.889	0.13	0.108	**0.09***	$\mathbf{2.03\times10^{-2}}$
Cruciferous veg. intake (g/day)	0.11	0.147	0.13	0.103	0.01	0.917	0.05	0.541	**0.09***	$\mathbf{2.08\times10^{-2}}$
Egg (g/day)	0.13	8.18×10^{-2}	0.08	0.352	**0.18***	$\mathbf{3.86\times10^{-2}}$	0.05	0.512	**0.09***	$\mathbf{3.30\times10^{-2}}$
Bread and cereal (g/day)	0.03	0.743	0.05	0.582	-0.09	0.305	**0.16***	$\mathbf{4.47\times10^{-2}}$	0.050	0.232

Participants who self-reported as currently smoking were excluded from the analysis, where r denotes the Spearman rank correlation coefficient.
***Correlation is significant at the 0.01 level (2-tailed).*
** Correlation is significant at the 0.05 level (2-tailed).*
Significant p-values at the 0.01 and 0.05 levels are bolded.

Table S2.6. Spearman rank correlation analysis of dietary variables associated with daily nitrate excretion (mg/day) of participants across from four different sites in the PURE-24USE study, including Hamilton, Vancouver, Quebec City and Ottawa.

Dietary Variable	Study Site								Total (median=147 mg/day; *n*=759)	
	Hamilton (median=133 mg/day; *n*=207)		Vancouver (median=157 mg/day; *n*=173)		Quebec City (median=162 mg/day; *n*=197)		Ottawa (median=137 mg/day; *n*=182)			
	r	***p*-value**	***r***	***p*-value**	***r***	***p*-value**	***r***	***p*-value**	***r***	***p*-value**
Vegetable (g/day)	**0.27****	**1.14×10^{-4}**	0.09	0.217	0.14	5.90×10^{-2}	0.10	0.163	**0.17****	**2.46×10^{-6}**
Green leafy veg. (g/day)	**0.25****	**3.42×10^{-4}**	0.10	0.173	**0.14***	**4.86×10^{-2}**	0.11	0.126	**0.15****	**2.90×10^{-5}**
Other veg. (g/day)	**0.16***	**2.25×10^{-2}**	0.07	0.359	0.12	9.69×10^{-2}	0.03	0.675	**0.11****	**2.77×10^{-3}**
Cruciferous veg. intake (g/day)	**0.19****	**6.68×10^{-3}**	-0.06	0.399	0.01	0.866	0.06	0.393	**0.07***	**4.99×10^{-2}**
Fruit (g/day)	0.13	6.49×10^{-2}	-0.04	0.646	**0.19****	**7.50×10^{-3}**	0.01	0.891	**0.07***	**4.20×10^{-2}**
Processed meat (g/day)	**-0.18****	**9.37×10^{-3}**	-0.09	0.228	**0.21****	**3.70×10^{-3}**	-0.003	0.963	-0.02	0.600
Processed food (g/day)	**-0.17***	**1.76×10^{-2}**	-0.02	0.812	0.04	0.540	-0.07	0.355	**-0.08***	**3.97×10^{-2}**
Dark yellow veg. (g/day)	**0.16***	**1.83×10^{-2}**	0.03	0.734	0.02	0.811	-0.06	0.417	0.06	0.105
Bread and cereal (g/day)	0.08	0.266	-0.10	0.191	0.04	0.603	**0.19****	**8.94×10^{-3}**	0.05	0.148

r denotes the Spearman rank correlation coefficient.
*** Correlation is significant at the 0.01 level (2-tailed).*
** Correlation is significant at the 0.05 level (2-tailed).*

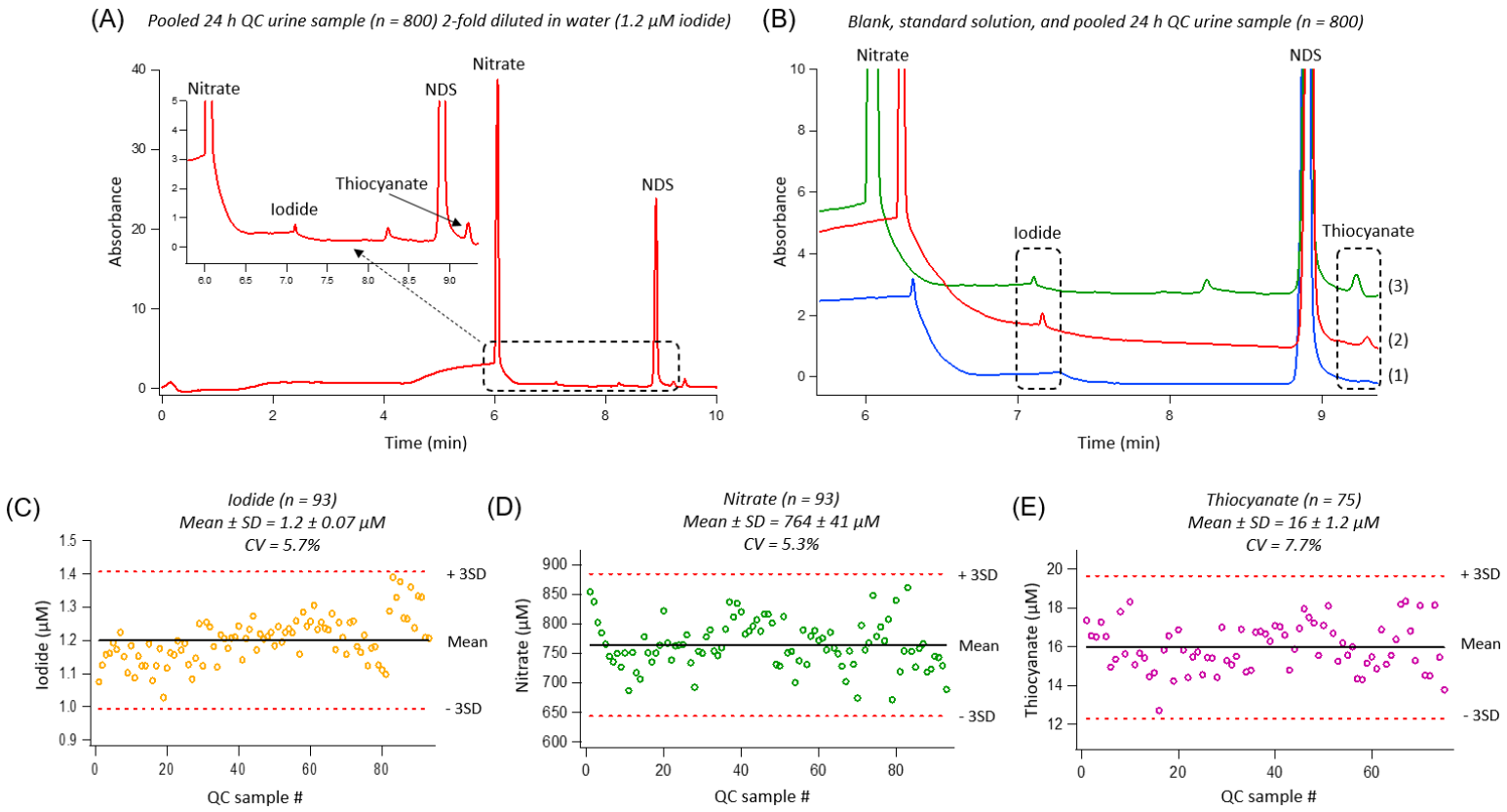


Figure S2.1. Representative electropherogram overlay of **(A)** a pooled quality control (*n*=800) and a comparison between a **(B)** quality control sample, calibrant solution, and blank solution. Also, control charts highlight reliable quantification of **(C)** iodide, **(D)** nitrate and **(E)** thiocyanate in 24 h urine samples with acceptable reproducibility (CV < 8%).

Chapter III

Validation of Urinary Thiocyanate as a Robust Biomarker of Active Tobacco Smoking in the Prospective Urban and Rural Epidemiological Study

Stellena Mathiaparanam, Biban Gill, Sathish Thirunavukkarasu, Guillaume Paré, Teo Koon, Salim Yusuf, and Philip Britz-McKibbin

Nicotine & Tobacco Research, **2022** (Under Review)

S.M. performed all the experiments including sample preparation, data acquisition using CE-UV, data processing, interpretation, statistical analysis, and wrote the initial draft for publication. B.G. performed complimentary experiments in support of method validation. S.Y., G.P., and S.T. were involved in participant recruitment, collection of dietary records, and coordination of urine collection. P.B.M. supervised the study, and with all other co-authors, contributed to the final version of the manuscript.

Chapter III: Validation of Urinary Thiocyanate as a Robust Biomarker of Active Tobacco Smoking in the Prospective Urban and Rural Epidemiological Study

3.1 Abstract

Tobacco smoking is a leading preventable cause of premature death globally. Urinary thiocyanate is a biomarker of cyanide exposure from tobacco smoke, however few studies have evaluated its utility in diverse populations of smokers. We examined the associations between urinary thiocyanate and self-reported never and current smokers among 1000 participants from 14 countries in the Prospective Urban and Rural Epidemiological study. We analyzed urinary thiocyanate in light and heavy smokers as compared to never smokers from high-, middle- and low-income countries using a validated capillary electrophoresis method in conjunction with standardized questionnaires. The median urinary thiocyanate concentration was 31 μM, which ranged from 8.6 μM to 52 μM for never smokers (n = 335) and current smokers (n = 660), respectively. Urinary thiocyanate was correlated with daily cigarette consumption (r = 0.621) and total nicotine equivalents (r = 0.514) while displaying a better dose response than urinary cotinine. A moderate association of urinary thiocyanate was found in biochemically verified never smokers (r ~ 0.38) due to intake of vegetables, fruits and dairy. Receiver-operating characteristic curves established cut-off values for urinary thiocyanate to differentiate current from never smokers with an optimal threshold of 23.9 μM (AUC = 0.861), which varied in countries with different income status. Elevated thiocyanate was evident in current smokers from high-income countries likely reflecting differences in smoking topography and greater toxicant burden. Background urinary thiocyanate in never smokers was associated with goitrogenic food intake that may obscure detection of second-hand smoke exposure.

3.2 Introduction

The tobacco epidemic remains a leading preventable cause of disability and premature deaths from cancer, cardiovascular diseases and respiratory illnesses, killing more than 8 million people every year.[1] Causal evidence of the hazards of tobacco smoking on mortality and lung cancer incidence was first established in the British Doctors Study, including the benefits of smoking cessation.[2] Globally, there are 1.14 billion tobacco users with > 80% of current smokers residing in developing countries whose absolute numbers are projected to increase due to population growth.[3,4] The main variables that impact the risks of tobacco is smoking intensity (*e.g.*, cigarettes per day or CPD), the age of initiation, and the duration of the habit.[5] Smoking topography, such as the frequency and volume of puffs, also impacts the hazards of smoking with greater toxicant exposure when using ventilated and low yield cigarette brands.[6,7] Nicotine dependence is also related to genetic factors that regulate nicotine metabolism,[8] which can guide personalized smoking cessation strategies.[9]

Self-reporting of smoking status, smoking history, and cigarette type consumed is typically evaluated by standardized questionnaires in large-scale epidemiological studies, such as the Global Adult Tobacco Survey.[10] Also, assessment of smoking dependence in clinical trials can be achieved using the Fagerström Test for Cigarette Dependence or the Cigarette Dependence Scale.[11] However, these surveys may be prone to bias and misreporting when comparing different populations, including cultural and gender differences in smoking behavior that impact tobacco smoke exposures.[12,13] Alternatively, biochemical markers (*i.e.,* biomarkers) of tobacco smoke exposure can be accurately quantified in non-invasive biofluids, such as urine.[14] For instance, tobacco-specific smoke exposure biomarkers, including urinary cotinine-verified smoking status, total nicotine equivalents and/or nicotine metabolic ratio, enable more reliable assessment recent tobacco smoke exposure, nicotine dependence, and/or

smoking behavior compared to self-reports.[15-17] Indeed, there are at least 93 other harmful and potentially harmful chemical constituents in tobacco smoke recognized by the US Food and Drug Administration,[18] including the respiratory and cardiovascular toxicant, hydrogen cyanide.

Hydrogen cyanide is readily absorbed via inhalation of tobacco smoke, as well as ingestion of certain foods or dermal contact of industrial solvents from occupational exposure. Acute exposure to hydrogen cyanide can be fatal due to inhibition of cytochrome c oxidase resulting in cellular hypoxia and anoxia.[19] The major pathway (~ 80%) for cyanide detoxification is mediated by the mitochondrial enzyme rhodanese to generate thiocyanate that is subsequently excreted in urine.[20] Given the short half-life of cyanide in circulation (< 1 h), thiocyanate is primarily used as an indirect biomarker of tobacco smoke exposure in saliva, blood and urine samples.[21-24] However, few studies to date have examined the utility of urinary thiocyanate as a viable biomarker of tobacco smoke exposure in large populations of current and never smokers from various countries.[25-27]

Herein, we analyzed thiocyanate using a validated capillary electrophoresis assay[28] from urine samples collected in an international cohort of current and never smokers (n = 1000) from the Prospective Urban and Rural Epidemiological (PURE) study.[29] For the first time, we demonstrate that urinary thiocyanate can differentiate current smokers from never smokers in a dose responsive manner in populations from low-, middle- and high-income countries despite regional variations in smoking behaviors, cigarette products consumed, and dietary patterns.

3.3 Methods

3.3.1 Study Design and Urine Collection from PURE Cohort

The PURE study is an ongoing international prospective cohort study of > 180,000 adults from 21 countries at different economic levels, with substantial heterogeneity in social and environmental circumstances.[30] A subset of 1000 participants was selected based on a nested case-cohort study having complete tobacco smoking history,[29] including never smokers (n = 335), light (< 10 CPD; n = 324) and heavy (≥ 10 CPD; n = 341) current smokers from 14 countries that differ with respect to socioeconomic status as summarized in **Table 3.1**. Questions about age, sex, education, location, smoking status, alcohol use, diet, physical activity, and use of medications were collected with standardized questionnaires as described elsewhere.[30] Detailed data on the habitual diets of participants were also collected using validated food frequency questionnaires in PURE.[31] Morning void urine samples were collected between 2005 and 2016 from consenting adults with research ethics approval obtained at local clinical sites worldwide. Urine samples were promptly stored frozen (–20 to –70 °C) on site, and then shipped frozen using liquid nitrogen vapor shippers (at –160 °C) to the Clinical Research Laboratory and Biobank at Hamilton Health Sciences (Hamilton, ON, Canada). Further description of the chemicals and reagents, urine sample preparation, as well as capillary electrophoresis assay[28,32] used for urinary thiocyanate determination are included in section **3.9 Supplemental Information**. Urinary TNE-7 was measured independently in urine samples from PURE participants using multisegment injection-capillary electrophoresis-mass spectrometry with full-scan data acquisition under positive ion mode detection as described elsewhere.[29]

Table 3.1. Summary of participant characteristics of the PURE study for urinary thiocyanate.

Parameter	Never smokers (*n* = 335)	Light smokers (*n* = 324)	Heavy smokers (*n* = 341)	Total (*n* = 1000)
Thiocyanate (µM): median ± IQR[a]	8.6 ± 18	36 ± 61	69 ± 96	31 ± 67
TNE-7 (µM): median ± IQR[b]	11 ± 36	29 ± 42	52 ± 54	39 ± 49
Age: years, median ± IQR	52 ± 17	53 ± 16	52 ± 15	52 ± 15
Sex: no./total no. (%)				
Female	219/335 (65)	159/324 (49)	102/341 (30)	480/1000 (48)
Male	116/335 (35)	165/324 (51)	239/341 (70)	520/1000 (52)
BMI: kg/m^2, median ± IQR[d]	26 ± 6.6	26 ± 6.9	25 ± 6.8	26 ± 6.9
BMI: no./total no. (%)				
<18.5 kg/m^2	18/326 (5.5)	21/317 (6.6)	27/339 (8.0)	66/982 (6.7)
18.5 – 24.9 kg/m^2	110/326 (34)	124/317 (39)	143/339 (42)	377/982 (38)
≥25 kg/m^2	198/326 (61)	172/317 (54)	169/339 (50)	539/982 (55)
Education level: no./total no. (%)				
None, Primary, or Unknown	159/334 (48)	152/323 (47)	156/341 (46)	467/998 (47)
Secondary/High School	80/334 (24)	91/323 (28)	103/341 (30)	274/998 (27)
Trade or College/University	95/334 (28)	80/323 (25)	82/341 (24)	257/998 (26)
Tobacco Type: no./total no. (%)				
Smoking Only	0/335 (0)	290/324 (90)	317/341 (93)	607/1000 (61)
Smokeless Tobacco Only	0/335 (0)	30/324 (9.2)	4/341 (1.2)	34/1000 (3.4)
Both Smoking and Smokeless	0/335 (0)	4/324 (1.2)	20/341 (5.9)	24/1000 (2.4)
Never User	335/335 (100)	0/324 (0)	0/341 (0)	335/1000 (34)
Tobacco User: no./total no. (%)				
Current User (includes quit ≤ 1 year)	0/335 (0)	324/324 (100)	341/341 (100)	665/1000 (67)
Never User	335/335 (100)	0/324 (0)	0/341 (0)	335/1000 (34)
Duration of Smoking: no./total no. (%)				
≥ 20 years	0/335 (0)	267/324 (82)	311/341 (91)	578/1000 (58)
< 20 years	0/335 (0)	57/324 (18)	30/341 (9)	87/1000 (87)
Never user	335/335 (100)	0/324 (0)	0/341 (0)	335/1000 (34)
Alcohol User – no./total no. (%)				
Current User	103/333 (31)	177/324 (55)	171/339 (50)	451/996 (45)
Former User	14/333 (4.2)	20/324 (6.2)	36/339 (11)	70/996 (7.0)
Never User	216/333 (65)	127/324 (39)	132/339 (39)	475/996 (48)
AHEI Score: median ± IQR[b]	36 ± 11	32 ± 11	30 ± 14	33 ± 13
Dietary Intake: median ± IQR[e]				
Coffee (mL/day)[f]	0 ± 281	247 ± 532	107 ± 600	125 ± 500
Vegetables (g/day)	284 ± 279	235 ± 244	227 ± 262	249 ± 260
Dairy (g/day)	170 ± 332	190 ± 334	209 ± 354	189 ± 330
Fruits (g/day)	216 ± 282	177 ± 276	178 ± 251	187 ± 270
Red and processed meats (g/day)	54 ± 112	77 ± 104	81 ± 122	71 ± 115
Diabetes: no./total no. (%)	54/335 (16)	35/324 (11)	42/341 (12)	131/1000 (13)
Hypertension: no./total no. (%)[g]	175/334 (52)	149/321 (46)	148/341 (43)	472/996 (47)
Location – no./total no. (%)				
Rural	137/335 (41)	142/324 (44)	135/341 (40)	414/1000 (41)
Urban	198/335 (59)	182/324 (56)	206/341 (60)	586/1000 (59)
Air quality ($PM_{2.5}$): median ± IQR[h]	24 ± 62	16 ± 23	16 ± 39	17 ± 37
Country Income Level: no./total no. (%)				
High-income (HIC)	111/335 (33)	126/324 (39)	139/341 (41)	376/1000 (38)
Middle-income (MIC)	113/335 (34)	142/324 (44)	126/341 (37)	381/1000 (38)
Low-income (LIC)	111/335 (33)	56/324 (17)	76/341 (22)	243/1000 (24)
Country: no./total no. (%)				
Argentina	35/335 (10)	55/324 (17)	65/341 (19)	155/1000 (16)
Bangladesh	98/335 (29)	44/324 (14)	73/341 (21)	215/1000 (22)
Brazil	8/335 (2)	11/324 (3)	22/341 (6)	41/1000 (4)
Canada	87/335 (26)	75/324 (23)	74/341 (22)	236/1000 (24)
Chile	4/335 (1)	7/324 (2)	2/341 (1)	13/1000 (1.3)
Colombia	30/335 (9)	23/324 (7)	8/341 (2)	61/1000 (6.1)
Iran	18/335 (5)	8/324 (2)	22/341 (6)	48/1000 (4.8)

Pakistan	3/335 (1)	8/324 (2)	1/341 (0.3)	12/1000 (1.2)
Philippines	9/335 (3)	2/324 (1)	4/341 (1)	15/1000 (1.5)
South Africa	9/335 (3)	37/324 (11)	3/341 (1)	49/1000 (4.9)
Sweden	11/335 (3)	48/324 (15)	59/341 (17)	118/1000 (12)
Tanzania	10/335 (3)	3/324 (1)	1/341 (0.3)	14/1000 (1.4)
United Arab Emirates	13/335 (4)	3/324 (1)	6/341 (2)	22/1000 (2.2)
Zimbabwe	0/335 (0)	1/324 (0.3)	1/341 (0.3)	2/1000 (0.20)

[a] *Light smokers, n=321; heavy smokers, n=339; total, n=995, where light smokers <10 tobacco products/day and heavy smokers ≥10 tobacco products/day.*
[b] *Never smokers, n=89; light smokers, n=275; heavy smokers, n=322; total, n=686.*
[c] *Never smokers, n=325; light smokers, n=319; heavy smokers, n=334; total, n=978.*
[d] *Never smokers, n=326; light smokers, n=317; heavy smokers, n=339; total, n=982.*
[e] *Never smokers, n=325; light smokers, n=319; heavy smokers, n=334; total, n=978.*
[f] *Never smokers, n=316; light smokers, n=308; heavy smokers, n=331; total, n=955.*
[g] *Never smokers, n=334; light smokers, n=321; total, n=996.*
[h] *Average from the last 3 years. Never smokers, n=334; heavy smokers, n=340; total, n=998.*

3.3.2 Statistical Analysis

Scatter plots, boxplots, and receiver-operating characteristic (ROC) curves were plotted using MedCalc 12.5 statistical software (Ostend, Belgium). Representative control and electropherogram overlay graphs were plotted using Igor Pro 5.04B (Wavemetrics Inc, Lake Oswego, OR, USA). All statistical analyses, including univariate general linear models and Spearman rank order correlation, were performed on SPSS 23.0 statistical software (IBM, Chicago, IL, USA). Analysis of covariance (ANCOVA) was adjusted for covariates when analyzing *log*-transformed urinary thiocyanate as a function of self-reported smoking intensity (CPD), including age, sex, body mass index (BMI), total caloric intake, alternative healthy eating index (AHEI) score, education level, location, alcohol use, history of hypertension or taking hypertensive drugs, history of diabetes or taking diabetic drugs or fasting glucose, tobacco type (smoking, smokeless, or both), and average fine particulate matter ($PM_{2.5}$) air pollution index in the last 3 years,[33] together with Bonferroni *post-hoc* testing. Similar covariate adjustments (also smoking intensity or CPD) and multiple hypothesis correction were applied when comparing urinary thiocyanate across countries at different socioeconomic development levels, including from high- (HICs), middle- (MICs) and low-income countries (LICs).[29,30] Urine samples from self-reported never smokers with measurable urinary TNE-7, and current smokers without detectable urinary

TNE-7 were excluded from subsequent statistical analyses to correct for probable misreporting of smoking behavior, such as never smokers with high levels of TNE-7 unlikely caused by passive or second-hand smoke (SHS) exposure.[29] SHS exposure was recorded at five different time intervals for PURE participants as either not exposed or exposed to passive tobacco smoke at different frequencies (e.g., 1-2 times per week to 2-3 times daily). In our study, never smokers with self-reported recent SHS exposure corresponding temporally with urine collection were compared to TNE-7 verified never smokers without self-reported SHS exposure.

3.4 RESULTS

3.4.1 PURE Cohort Characteristics and Analytical Figures of Merit for Urinary Thiocyanate

Table 3.1 summarizes the characteristics of a nested case-control cohort of PURE participants from 14 countries at different socioeconomic development (HICs, MICs, and LICs), including self-reported never smokers (n = 335), as well as light (n = 324, CPD < 10) and heavy (n = 341, CPD ≥ 10) current smokers. Morning void urine samples were collected using standardized operating protocols at various clinical sites worldwide for storage and transport. Overall, the sex-balanced cohort had a median age of 52 years and mean BMI of 26 kg/m^2 with a high prevalence of hypertension (47%) and type 2 diabetes (13%). However, there were more males (70%) classified as heavy smokers (≥ 10 CPD) and a larger fraction of females were never smokers (65%). Most current smokers reported using cigarettes alone (~ 90%) as compared to smokeless tobacco products (*e.g.*, snuff) or both cigarettes and smokeless tobacco. Otherwise, most current smokers (> 82%) had a long history (> 20 years) of smoking use primarily from urban locations (~ 60%), but there was no difference in their educational attainment relative to never smokers. However, differences in dietary habits were evident with a better diet quality (*i.e.*, higher AHEI score) for never smokers. For instance, current smokers had a greater daily intake of coffee, alcohol, as well as

red and processed meats, but a lower average consumption pattern for fruits and vegetables relative to never smokers (**Table 3.1**).

Figure S3.1 illustrates that urinary thiocyanate was reliably quantified by CE with UV absorbance detection, which provided acceptable technical precision (mean CV = 5.1%, n = 192) in a control chart based on repeated analysis of pooled QC samples. **Table S3.1** summarizes the analytical figures of merit of the CE assay for urinary thiocyanate determination that demonstrated good accuracy with a mean recovery of (94 ± 10)% when spiked at three different concentration levels. Concentration sensitivity was sufficient with a detection limit of 1.4 μM (S/N ~ 3), which resulted in no missing values. Method selectivity was also excellent with unresolved matrix interferences to thiocyanate evident in only five samples analyzed (out of 1000 urine specimens). Method validation included chemical stability studies (**Figure S3.1**) that confirmed no significant changes in urinary thiocyanate concentrations (p > 0.05) following three repeat freeze-thaw cycles of a QC sample. Urinary thiocyanate also tolerated extended delays (up to 24 h at room temperature) to storage at -80 °C after analyzing a freshly collected urine specimen.

3.4.2 Dose-response of Urinary Thiocyanate with Smoking Intensity in PURE

Figure 3.1A depicts a scatter plot with a Spearman rank correlation analysis between urinary thiocyanate concentrations (μM) and daily cigarette consumption (i.e., CPD) for light (CPD < 10) and heavy (CPD 10-20 and CPD ≥ 20) smokers, as well as never smokers from the PURE study. There was a moderate correlation (r = 0.621, p = 2.71×10^{-107}, n = 995) with urinary thiocyanate that increased progressively with greater tobacco smoking intensity. Also, urinary thiocyanate was also correlated with TNE-7 (r = 0.514, p = 3.62×10^{-47}, n=681) that is a robust biomarker of recent tobacco smoke exposure.[29] **Figure 3.1C** shows a box plot that depicts the dose response of urinary thiocyanate when comparing never smokers

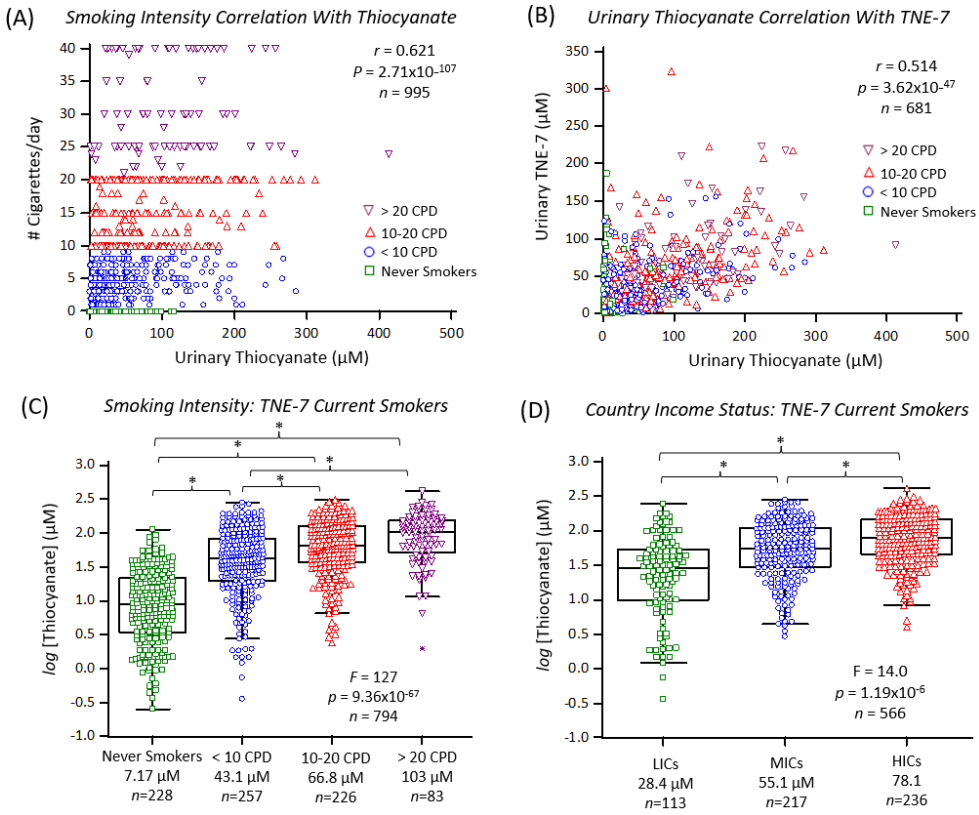


Figure 3.1. Correlation plots showing relationship **(A)** between self-reported smoking intensity (CPD) versus urinary thiocyanate concentration (r = 0.621, p = 2.71×10^{-107}, n = 995), as well as **(B)** urinary thiocyanate versus TNE-7 concentrations (r = 0.514, p = 3.62×10^{-47}, n = 681). Box plots displaying TNE-7 verified **(C)** dose response of urinary thiocyanate as a function of smoking intensity (F = 127, p = 9.36×10^{-67}, n = 794), and **(D)** urinary thiocyanate differing in current smokers across different country income groups (F = 14.0, p = 1.19×10^{-6}, n = 566) after adjustment for age, sex, BMI, alcohol use, total caloric intake, diet quality, education level, location, air quality, comorbidities, tobacco type, and CPD (**D** only). Median concentrations for urinary thiocyanate are included at bottom of graphs.

(7.2 µM), light smokers (CPD < 10, 43 µM), and heavy smokers (CPD = 10-20, 67 µM; CPD > 20, 103 µM) when using ANCOVA with Bonferroni correction and covariate adjustments (F = 127, p = 9.36×10^{-67}, n = 794), including age, sex, BMI, total caloric intake, tobacco type, alcohol intake, diet quality, education, location, air quality and disease comorbidities. In this case, urinary TNE-7 was

used to biochemically verify smoking status for never smokers, as well as current smokers. Overall, 86% and 94% of current light and heavy smokers, respectively had TNE-7 above detection limits consistent with habitual tobacco use, whereas 27% of never smokers had unexpectedly high urinary TNE-7 concentrations.[29] In fact, urinary thiocyanate displayed a greater tobacco cigarette dose response (F = 34, $p = 1.77\times10^{-14}$, n = 566) than cotinine (F = 23, $p = 3.71\times10^{-10}$, n = 518) among TNE-7 verified current smokers as shown in **Figure S3.2**. **Figure 3.1D** highlights that urinary thiocyanate in current smokers were also dependent on country income status following ANCOVA with covariate adjustments (F = 14.0, $p = 1.19\times10^{-6}$, n = 566). There was a striking increase in median urinary thiocyanate from 28 μM, 55 μM to 78 μM among TNE-7 verified tobacco smokers from LICs, MICs and HICs, respectively after adjusting for smoking intensity, tobacco type and ten other covariates.

3.4.3 Optimal Cut-off Limits for Differentiating Current Smokers Versus Never Smokers

Receiver-operating characteristic (ROC) curves were next used to determine optimal cut-off values for urinary thiocyanate to differentiate TNE-7 verified current from never smokers in diverse populations. In this case, never smokers and current smokers with and without detectable TNE-7, respectively were excluded from data analysis. **Figure 3.2A** highlights a ROC curve for urinary thiocyanate from all countries combined in the PURE study, which had an optimal cut-off of 23.9 μM (AUC = 0.861, $p = 4.30\times10^{-61}$, n = 838). The adjacent boxplot shows a 6.2-fold higher median thiocyanate concentration excreted in urine in never smokers relative to current smokers (F = 374, $p = 1.64\times 10^{-68}$, n = 807) after covariate adjustments. Factors associated with elevated urinary thiocyanate (≥ 23.9 μM) in current smokers included self-reported smoking intensity (highest for heavy smokers), country income region (highest for HICs), as well as male sex, higher education level, and urban location as summarized in

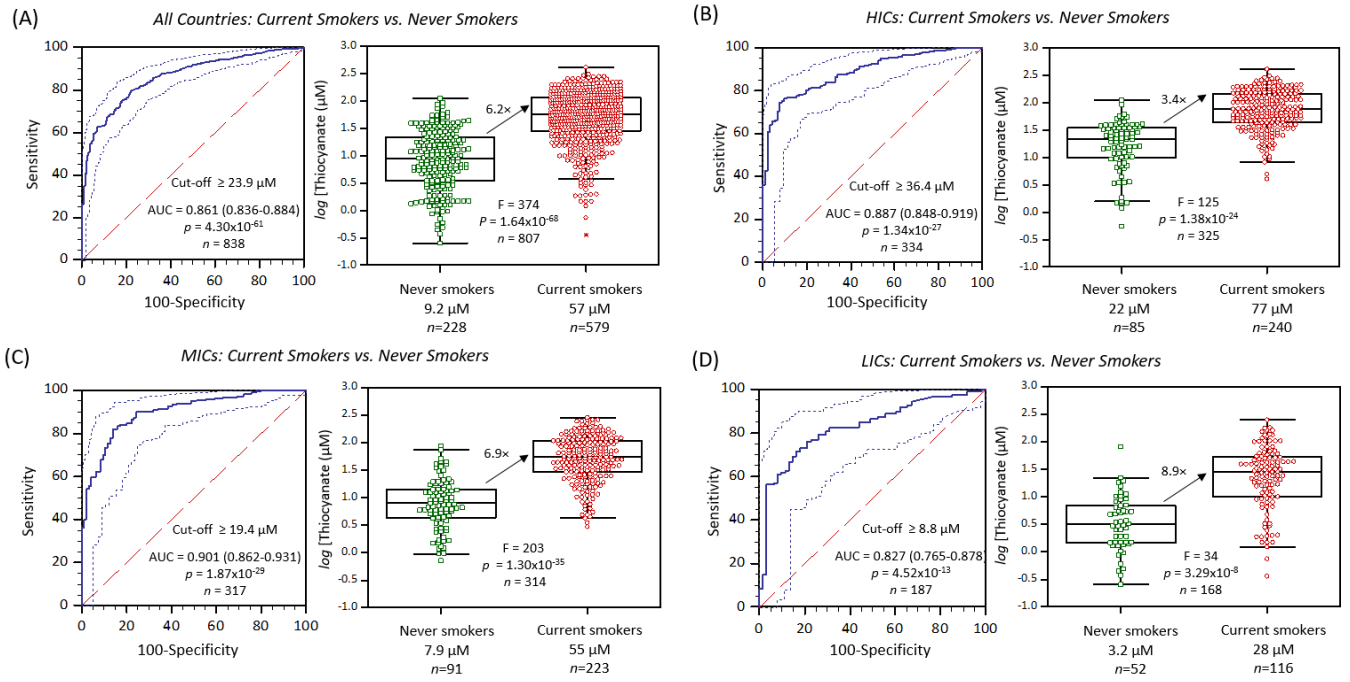


Figure 3.2. Global and regional ROC curves with boxplots used to differentiate TNE-7 verified current from never smokers in the PURE study. **(A)** The cut-off value for urinary thiocyanate for current and never smokers from all 14 countries is ≥ 23.9 µM with excellent discrimination (AUC = 0.861) as also shown in boxplots with a 6.2-fold increase in thiocyanate ($p = 1.64\times10^{-68}$) in current smokers **(B)** Current smokers from HICs had the highest urinary thiocyanate cut-off of ≥ 36.4 µM (AUC = 0.887) with a 3.4-fold increase in thiocyanate in current as compared to never smokers ($p = 1.38\times10^{-24}$). **(C)** Current smokers from HICs had a lower urinary thiocyanate cut-off of ≥ 19.4 µM (AUC = 0.901), and current smokers had a 6.9-fold increase in thiocyanate as compared to never smokers ($p = 1.30\times10^{-35)}$. **(D)** Current smokers from LICs had the lowest thiocyanate cut-off of ≥ 8.8 µM (AUC = 0.827) with an 8.9-fold increase in thiocyanate as compared to never smokers ($p = 3.29\times10^{-8}$).

Table S3.2. Given the strong dependence of thiocyanate concentrations on country income region, ROC curves and optimal cut-off values for urinary thiocyanate were also evaluated for HICs (**Figure 3.2B**), MICs (**Figure 3.2C**) and LICs (**Figure 3.2D**), which varied from 36.4 μM, 19.4 μM and 8.8 μM, respectively. In all cases, the accuracy of ROCs for smoking status classification was excellent with an AUC ranging from 0.827 to 0.901. The varying threshold values for urinary thiocyanate reflected the contribution of habitual food intake and tobacco smoke exposures in never smoker and current smokers residing in 14 countries at different socioeconomic development. However, urinary thiocyanate was found not to differentiate passive smoking ($p > 0.05$) when comparing TNE-7 verified never smokers with and without recent SHS exposures ($n = 227$) as highlighted in **Figure S3.3**.

3.4.4 Dietary Contributions to Background Urinary Thiocyanate in Never Smokers

Dietary intake patterns (g/day) related to urinary thiocyanate was also explored separately in current and never smokers using a Spearman rank correlation analysis as summarized in **Table S3.3**. There was a modest positive correlation between AHEI scores and urinary thiocyanate concentrations among TNE-7 verified never smokers ($r = 0.15$, $p = 1.95\times10^{-2}$, $n = 237$), whereas current smokers had an inverse correlation ($r = -0.20$, $p = 4.53\times10^{-7}$, $n = 611$). This suggests that never smokers in general may have healthier eating patterns as compared to heavy tobacco smokers consuming a poor diet quality. For instance, urinary thiocyanate in current smokers from HICs ($n = 240$) had the strongest negative correlation with AHEI scores ($r = -0.24$, $p = 1.59\times10^{-4}$) as compared to MICs or LICs indicating poor dietary habits for current smokers primarily from Canada and Sweden. Urinary thiocyanate concentrations among never smokers ($n = 237$) were moderately associated with the average daily intake of specific foods, including cooked and/or raw vegetables ($r = 0.45$, $p = 3.95\times10^{-10}$), dairy products

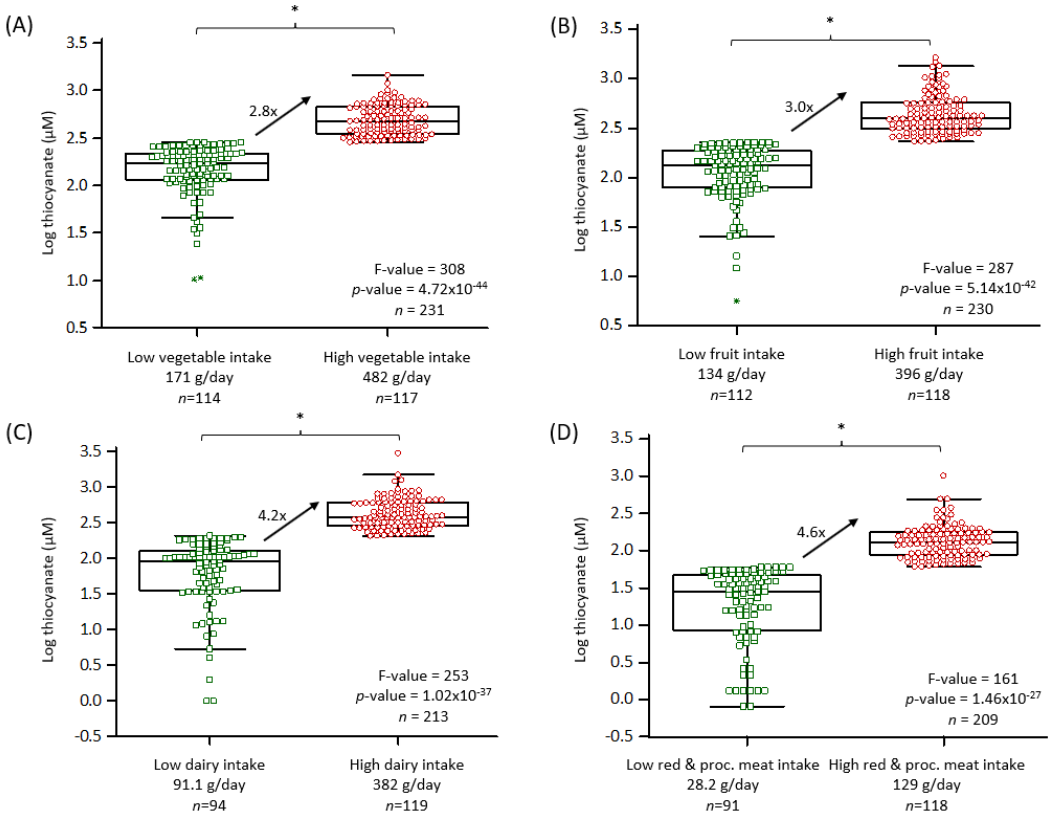


Figure 3.3. Boxplots highlighting differences in urinary thiocyanate in TNE-7 verified never smokers from the PURE study when using ANCOVA with covariate adjustments for sex, age, BMI and total caloric intake. Differences in background thiocyanate status were stratified based on low versus high consumers of distinct food sources having cyanide and/or thiocyanate content, namely **(A)** vegetables, **(B)** fruits, **(C)** dairy products and **(D)** red and processed meat intake.

(r = 0.39, p = 5.64×10^{-10}), fruit (r = 0.36, p = 1.94×10^{-8}), as well as red and processed meat (r = 0.22, p = 7.02×10^{-4}). Overall, total vegetable intake was most consistently associated with urinary thiocyanate excretion in confirmed never smokers from HICs, MICs and LICs, whereas fruit and dairy was only significant for PURE participants from MICs and LICs, respectively (**Table S3.3**). As expected, consumption patterns were regionally dependent with most never smokers from HICs and MICs eating raw vegetables as compared to cooked vegetables in LICs, such as Bangladesh. In contrast, current smokers (n = 592) from different country income regions had much weaker, non-significant or

negative correlations to eating these classes of foods. **Figure 3.3** depicts box plots for urinary thiocyanate as a promising dietary biomarker for differentiating low versus high consumers of specific goitrogenic foods in TNE-7 verified never smokers from the PURE study, including vegetables, fruits, dairy products, and well as red and/or processed meat.

3.5 Discussion

Biochemical verification of tobacco use and abstinence may better estimate the relative risks of tobacco smoking as compared to population-based surveys given the variability of cigarette product designs, their chemical composition, and individual smoking topography (*i.e.*, puffing pattern).[10,34] Also, changing societal acceptability of smoking over time has increased underreporting of smoking when compared to national sales trends for tobacco.[35] This is most pronounced in vulnerable groups (*e.g.*, youth, pregnant women) and specific cultural/religious settings since smoking for women may still be considered taboo.[36] However, successful implementation of tobacco smoke biomarkers[14] depend on the choice of specimen type, feasibility of sample collection, costs for analysis, as well as their toxicokinetic profiles and distribution in the population.[37] To date, the analysis of tobacco specific nicotine metabolites (*e.g.*, cotinine) and carcinogenic nitrosamines are widely used biomarkers of recent exposure to tobacco smoke.[14-16] Despite thiocyanate being one of the earliest reported biomarkers of tobacco smoking dose,[23,38] there have been sparse reports of urinary thiocyanate in large populations globally.[25-27] Challenges for establishing universal reference intervals for urinary thiocyanate[39] include dietary contributions from the intake of goitrogenic vegetables containing cyanogenic glycosides and glucosinolates that vary regionally.[40] Minor pathways also exist for cyanide detoxification that are not captured by thiocyanate excreted in urine, including formation of 2-amino-2-thiazoline-4-carboxylic acid[25] and protein-bound thiocyanate adducts.[41]

In this study, a validated CE assay[28,32] was used to separate and quantify thiocyanate in urine samples following a simple dilution step which provided adequate precision, accuracy, sensitivity, and selectivity. We also showed that urinary thiocyanate is a robust biomarker of tobacco smoke exposure since it is chemically stable and can tolerate pre-analytical variations related to sample collection, storage and handling. To date, various analytical methods have been applied to measure thiocyanate in human biofluids, including kinetic spectrophotometry, gas chromatography-mass spectrometry, ion-exchange chromatography with UV detection, and liquid chromatography with fluorometric detection or tandem mass spectrometry. However, these methods have several limitations when applied to large-scale epidemiological studies, including long analysis times, poor selectivity, large sample volume requirements, complex sample workup procedures, and/or high operating costs.[23] In our study, we demonstrated that urinary thiocyanate exhibited a strong dose response to self-reported daily cigarette consumption (i.e., CPD) after covariate adjustment and Bonferroni correction in a large multiethnic cohort of participants recruited from 14 countries. Moreover, urinary thiocyanate was moderately correlated to urinary TNE-7 that is a robust biomarker of recent tobacco smoke exposure independent of metabolic rate unlike the analysis of single nicotine species.[24] In fact, urinary thiocyanate has a longer half-life of tobacco smoke/cyanide exposure (~ 10-14 days) than urinary nicotine metabolites,[25] which may explain some discordant results when relying on morning spot urine sampling especially for light or casual smokers. We used urinary TNE-7 to independently verify the smoking status of self-reported never smokers since our recent report revealed a significant extent of misclassification notably among women from LICs.[29] Overall, 89 of 312 self-reported never smokers from PURE (~ 27%) were excluded from the analysis in comparison to a smaller fraction of light (49 of 324 or 15%) and heavy (19 of 341 or 5.6%) smokers who had non-detectable TNE-7. In the PURE study, the median (8.6 μM versus 52 μM), range (0.26 to 116 μM versus 1.6 to 311 μM) and

variance (CV = 45% versus 120%) for urinary thiocyanate measured in TNE-7 verified never smokers versus current smokers was consistent with previous reports comprising > 100 participants.[22,23,25,39] Creatinine normalization was not used when reporting urinary thiocyanate concentrations[22] in this work since it contributed to greater overall variance when comparing diverse populations given its dependence on muscle mass, protein intake, age, sex and kidney function.

Excellent performance was achieved when using an optimal cut-off of 23.9 μM for urinary thiocyanate to discriminate TNE-7 verified smokers from never smokers in the PURE study. However, we also determined region-specific cut-off values for urinary thiocyanate for participants from HICs, MICs and LICs since cyanide exposure from tobacco smoke and background dietary contributions to thiocyanate were dependent on country income status. For instance, the ROC for HICs, which comprised participants Canada, Sweden and the United Arab Emirates, had the highest cut-off limit of 36.4 μM with excellent discrimination (AUC = 0.887, n = 334) between current and never smokers. However, this threshold lowered progressively for participants from MICs (19.4 μM) and LICs (8.8 μM). Our results for HICs are comparable to Buratti *et al*.[25] who reported an optimal cut-off value of 38 μM for urinary thiocyanate between non-smokers and smokers from the US, which can be further stratified based on age, sex and ethnicity.[39] Other studies have reported smoker versus never smoker discrimination using thiocyanate measured in plasma or saliva with sensitivity and specificity ranging from 70 to 90% in Japan.[23] However, there is a lack of urinary thiocyanate data reported in smokers and never smokers from developing countries. The greater levels of urinary thiocyanate from current smokers in HICs as compared to MICs and LICs likely reflects greater use of ventilated and “light” cigarette brands that may paradoxically increase exposure to hydrogen cyanide due to compensatory smoking behavior with more intense and frequent puffs.[42] In contrast, hand-rolled cigarettes without filters (*i.e.,* bidis) are more frequently

consumed by the poor in Bangaldesh,[43] which may reduce toxicant exposure in tobacco smoke despite having greater nicotine and tar content than commercial cigarette brands.[44] For instance, there was more heterogenous tobacco use patterns for current smokers in LICs as 69/126 (55%) were only cigarette smokers, 22/126 (17%) were only bidi users, 13/126 (10%) were only chewing tobacco users, and 10/126 (8%) were both cigarette and bidi users. In contrast, most current smokers from HICs were primarily cigarette only smokers (195/243, 80%), with a small fraction of 5/243 (2%) cigar users, but no bidi users. The remaining fraction (18%) of current smokers from HICs consisted of a mixture of tobacco product users, including cigarettes, cigars, pipes, shisha, and/or chewing tobacco. Current smokers from MICs were predominately comprised of cigarette smokers (209/228, 92%).

Previous comparisons of various tobacco smoke biomarkers have concluded urinary cotinine as being more effective in discriminating between smokers and never smokers than urinary thiocyanate.[22] This has been largely attributed to the near zero background levels of exogenous nicotine metabolites in the urine of authentic never smokers without SHS exposure unlike several food sources, as well as endogenous contributions to urinary thiocyanate. For instance, intra-cellular thiocyanate is a substrate for lactoperoxidase to generate the antimicrobial agent hypothiocyanite, which is used by the innate immune system to combat respiratory infections.[45] In contrast to results from Muranaka *et al.*,[23] we demonstrated that the tobacco smoke dose response for urinary thiocyanate was modestly superior to urinary cotinine from TNE-7 verified current smokers in our international cohort. Additionally, urinary thiocyanate was measured in more current smokers with fewer missing data while also exhibiting a greater fold-change in concentration from light (43 μM; CPD < 10) to heavy (103 μM; CPD > 20) smokers than urinary cotinine. However, urinary thiocyanate did not detect recent self-reported SHS exposures among biochemically verified never smokers. This result may reflect reporting bias due to transient SHS exposures that were

not well matched to the longer half-life of urinary thiocyanate, as well as SHS exposures that were likely too low in hydrogen cyanide content to exceed background dietary contributions.

Although urinary thiocyanate has been examined as a biomarker of cyanide exposure from tobacco smoke, few studies have examined its potential utility as a dietary biomarker of food intake. In our study, TNE-7 confirmed never smokers had a better overall diet quality with urinary thiocyanate that was positively correlated with AHEI score.[46] This was mainly attributed to a greater intake of vegetables, fruits and dairy products unlike current smokers. In contrast, current smokers from HICs had urinary thiocyanate that was negatively correlated to AHEI score indicative a poor diet quality that may exacerbate the harms from tobacco smoking. Overall, stratifying PURE participants based on their TNE-7 confirmed smoking status revealed more consistent dietary trends in the PURE cohort (notably for vegetable intake) due to less confounding from tobacco smoke exposures. In this case, certain goitrogenic vegetables (*e.g.*, cabbage, radish, cauliflower, cassava)[40] and dairy products[47] represent major sources of urinary thiocyanate, which may incidentally also increase the risk for iodine deficiency since thiocyanate also acts as an iodine uptake inhibitor.[28] Thus, excessive intake of goitrogenic foods may counteract their salutary benefits derived from bioactive phytochemicals, essential minerals, vitamins and fiber. Nevertheless, a poor diet low in vegetables, fruit and fiber is a major preventable cause of mortality and morbidity that is as important as a risk factor in chronic diseases than tobacco smoke.[48] Thus, urinary thiocyanate may paradoxically function as a biomarker associated with a healthy eating pattern when consumed moderately, whereas it is potentially harmful with excessive intake of goitrogenic foods and/or intense compensatory smoking habits.

Our study has several strengths and limitations. First, our work represents the most extensive analysis of urinary thiocyanate in participants having large

differences in smoking habits, dietary patterns and environmental exposures. Urine specimens were collected from 14 countries using a standardized procedure and analyzed directly using a validated CE assay for robust thiocyanate determination. Also, statistical analyses with covariate adjustments were performed to reduce for potential confounding when comparing distinct populations in both urban and rural settings globally. However, socioeconomic factors were not included in our analysis, which are relevant to explain the differences in tobacco smoking related harms in vulnerable groups despite reductions in smoking rates across developing countries.[49] For instance, smoking participation and smoking intensity have been associated with income status for young urban smokers in Bangladesh.[50] Moreover, we identified dietary exposures that contributed to urinary thiocyanate status in never smokers while also establishing region-specific cut-off intervals to optimally differentiate current smokers from never smokers. Study limitations included that the lack of sensitivity of urinary thiocyanate to respond to low levels of transient SHS exposures, the need for urinary TNE-7 to independently confirm smoking status and reveal potential misreporting, as well as few participants recruited from LICs besides Bangladesh. Also, repeat urine samples collected prospectively from participants in PURE may provide more reliable indicators of tobacco smoke exposures and their potential harm than single spot urine samples.

3.6 Conclusions

This work validated the utility of urinary thiocyanate as a robust biomarker of active tobacco smoking from cyanide exposure that is applicable to an international cohort. Urinary thiocyanate exhibited a smoking dose response that increased with daily cigarette consumption, which performed better than urinary cotinine. Improved discrimination of smoking status was achieved when relying on region-specific threshold values for urinary thiocyanate for participants from HICs, MICs and LICs. Current smokers from HICs relative to MICs and LICs had

evidence of greater cyanide exposure that corresponded to higher thiocyanate concentrations excreted in urine. These results imply that current smokers from HICs may be at greater relative risk for tobacco related disease outcomes due to compensatory smoking behavior when using commercial low-yield cigarette brands. Our study also identified specific dietary exposures contributing to urinary thiocyanate in confirmed never smokers, which reflected a better diet quality due to greater intake of vegetables, fruits and dairy products. Future work will conduct an untargeted analysis of urinary metabolites that may serve as novel biomarkers of tobacco smoke exposure and its potential harm in the PURE study, including their association with clinical events and all-cause mortality. The role that healthy eating patterns have on reducing the hazards of tobacco smoking will be explored as it may mitigate the harms in former smokers even when cessation is successful.

3.7 Acknowledgements

P.B.M. acknowledges funding from the Natural Sciences and Engineering Research Council of Canada, Genome Canada, and Canada Foundation for Innovation. S.T. received funding from the Hamilton Health Sciences New Investigator Fund (NIF-20504) from Hamilton, ON, Canada. The authors wish to thank the participants and staff from the Population Health Research Institute (PHRI), as well as PURE participants who were involved in this study.

3.8 References

1. He H, Pan Z, Wu J, Hu C, Bai L, Lyu J. Health effects of tobacco at the global, regional, and national levels: Results from the 2019 Global Burden of Disease Study. *Nicotine Tob Res*. 2022;24:864–870.
2. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004;328:1519-28.
3. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990-2015: A systematic

analysis from the Global Burden of Disease Study 2015. *Lancet* 2017;389:1885–1906.

4. GBD 2019 Tobacco Collaborators. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990-2019: A systematic analysis from the Global Burden of Disease Study 2019. *Lancet* 2021;397: 2337–2360.

5. Centers for Disease Control and Prevention. Global Adult Tobacco Survey Collaborative Group. *Tobacco Questions for Surveys: A Subset of Key Questions from the Global Adult Tobacco Survey* (GATS), 2011, 2nd Edition, Atlanta, USA.

6. Matsumoto M, Inaba Y, Yamaguchi I, Endo O, Hammond D, Uchiyama S, Suzuki G. Smoking topography and biomarkers of exposure among Japanese smokers: Associations with cigarette emissions obtained using machine smoking protocols. *Environ Health Prev Med*. 2013;18:95–103.

7. Pauwels CGGM, Boots AW, Visser WF, Pennings JLA, Talhout R, Schooten FV, Opperhuizen A. Characteristic human individual puffing profiles can generate more TNCO than ISO and Health Canada regimes on smoking machine when the same brand is smoked. *Int J Environ Res Public Health* 2020;17:3225.

8. Dempsey D, Tutka P, Jacob P 3rd, Allen F, Schoedel K, Tyndale RF, Benowitz NL. Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. *Clin Pharmacol Ther*. 2004;76:64–72.

9. Wells QS, Freiberg MS, Greevy RA Jr, Tyndale RF, Kundu S, Duncan MS, King S, Abney L, Scoville E, Beaulieu DB, Gatskie V, Tindle HA. Nicotine metabolism-informed care for smoking cessation: A pilot precision RCT. *Nicotine Tob Res*. 2018;20:1489–1496.

10. Jena PK, Kishore J, Sarkar BK. Global Adult Tobacco Survey (GATS): A case for change in definition, analysis and interpretation of "cigarettes" and "cigarettes per day" in completed and future surveys. *Asian Pac J Cancer Prev*. 2013;14:3299-3304.

11. Fu M, Carnicer-Pont D, Castellano Y, Ballbè M, Sureda X, Raich A, Martínez-Sánchez JM, Martínez C, Baena A, Riccobene A, Enríquez M, Fernández E. Measuring cigarette dependence: A comparison of two scales in two different groups of smokers. *Tob Prev Cessation* 2022;8:15.

12. Melikian AA, Djordjevic MV, Hosey J, Zhang J, Chen S, Zang E, Muscat J, Stellman SD. Gender differences relative to smoking behavior and emissions of toxins from mainstream cigarette smoke. *Nicotine Tob Res*. 2007;9:377 387.

13. Amos A, Greaves L, Nichter M, Bloch M. Women and tobacco: A call for including gender in tobacco control research, policy and practice. *Tobacco Control* 2012;21:236–243.

14. Benowitz NL, Bernert JT, Foulds J, Hecht SS, Jacob P, Jarvis MJ, Joseph A, Oncken C, Piper ME. Biochemical verification of tobacco use and abstinence: 2019 Update. *Nicotine Tob Res*. 2020;22:1086-1097.

15. Hong JW, Noh JH, Kim DJ. The prevalence of and factors associated with urinary cotinine-verified smoking in Korean adults: The 2008-2011 Korea National Health and Nutrition Examination Survey. *PLoS One* 2018;13:e0198814.

16. Van Overmeire IP, De Smedt T, Dendale P, Nackaerts K, Vanacker H, Vanoeteren JF, Van Laethem DM, Van Loco J, De Cremer KA. Nicotine dependence and urinary nicotine, cotinine and hydroxycotinine levels in daily smokers. *Nicotine Tob Res* 2016;18:1813-1819.

17. Benowitz NL, St. Helen G, Nardone N, Sanderson Cox L, Jacob 3rd P. Urine metabolites for estimating daily intake of nicotine from cigarette smoking. *Nicotine Tob Res* 2020;22:288–292.

18. U.S. Food and Drug Administration. Harmful and Potentially Harmful Constituents in Tobacco Products and Tobacco Smoke: Established List. April 2012. https://www.fda.gov/tobacco-products/rules-regulations-and guidance/harmful-and potentially-harmful-constituents-tobacco-products-and tobacco-smoke-established list (accessed 1 May 2022).

19. Logue BA Hinkens DM, Baskin SI, Rockwood GA. The analysis of cyanide and its breakdown products in biological samples. *Crit Rev Anal Chem*. 2010;40:122–147.

20. Bhandari RK, Oda RP, Petrikovics I, Thompson DE, Brenner M, Mahon SB, Bebarta VS, Rockwood GA, Logue BA. Cyanide toxicokinetics: the behavior of cyanide, thiocyanate and 2-amino-2-thiazoline-4-carboxylic acid in multiple animal models. *J Anal Toxicol*. 2014; 38:218-225.

21. Prue DM, Martin JE, Krapfl JE. Brand fading: The effects of gradual changes to low tar and nicotine cigarettes on smoking rate, carbon monoxide, and thiocyanate levels. *Behav Ther*. 1981;12:400–416.

22. Hill P, Haley NJ, Wynder EL. Cigarette smoking: carboxyhemoglobin, plasma nicotine, cotinine and thiocyanate vs self-reported smoking data and cardiovascular disease. *J Chronic Dis*. 1983;36:439-449.

23. Muranaka H, Higashi E, Itani S, Shimizu Y. Evaluation of nicotine, cotinine, thiocyanate, carboxyhemoglobin, and expired carbon monoxide as biochemical tobacco smoke uptake parameters. *Int Arch Occup Environ Health* 1988;60:37-41.

24. Scherer, G. Carboxyhemoglobin and thiocyanate as biomarkers of exposure to carbon monoxide and hydrogen cyanide in tobacco smoke. *Exp Toxicol Pathol*. 2006;58:101–24.

25. Buratti M, Xaiz D, Caravelliand G, Colombi A. Validation of urinary thiocyanate as a biomarker of tobacco smoking. *Biomarkers* 1997;2:81–85.

26. Vinnakota CV, Peetha NS, Perrizo MG, Ferris DG, Oda RP, Rockwood GA, Logue BA. Comparison of cyanide exposure markers in the biofluids of smokers and non-smokers. *Biomarkers* 2012;17:625–633.

27. Jain RB. Use of urinary thiocyanate as a biomarker of tobacco smoke. *Epidemiology*: *Open Access* 2016;6:268.

28. Nori de Macedo A, Teo KK, Mente A, McQueen MJ, Zeidler J, Poirier P, Lear SA, Wielgosz A, Britz-McKibbin P. A robust method for iodine status determination in epidemiological studies by capillary electrophoresis. *Anal Chem* 2014;86:10010–10015.

29. Sathish T, Teo KK, Britz-McKibbin P, Gill B, Islam S, Paré G, Rangarajan S, Duong M, Lanas F, Lopez-Jaramillo P, et al. Variations in risks from smoking between high-income, middle-income, and low-income countries: an analysis of data from 179 000 participants from 63 countries. *Lancet Global Health* 2022;10:e216–e226.

30. Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, Brauer M, Kutty VR, Gupta R, Wielgosz A, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle income, and low-income countries (PURE): A prospective cohort study. *Lancet* 2020;395:795–808.

31. Miller V, Mente A, Dehghan M, Rangarajan S, Zhang X, Swaminathan S, Dagenais G, Gupta R, Mohan V, Lear S, et al. Prospective Urban Rural Epidemiology (PURE) study investigators. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study. *Lancet* 2017;390:2037–2049.

32. Nori de Macedo A, Macri J, Hudecki PL, Saoi M, McQueen MJ, Britz McKibbin, P. Validation of a capillary electrophoresis assay for monitoring iodine nutrition in populations for prevention of iodine deficiency: An interlaboratory method comparison. *J App. Lab Med* 2017; 1:649–660.

33. Arku RE, Brauer M, Ahmed SH, AlHabib KF, Avezum Á, Bo J, Choudhury T, Dans AM, Gupta R, Iqbal R, et al. Long-term exposure to outdoor and household air pollution and blood pressure in the Prospective Urban and Rural Epidemiological (PURE) study. *Environ Pollut*. 2020;262:114197.

34. Hammond D, Fong GT, Cummings KM, Hyland A. Smoking topography, brand switching, and nicotine delivery: Results from an in vivo study. *Cancer Epidemiol Biomarkers Prev*. 2005;14:1370–1375.

35. Gallus S, Tramacere I, Boffetta P, Fernandez E, Rossi S, Zuccaro P, Colombo P, La Vecchia 39. Temporal changes of under-reporting of cigarette consumption in population-based studies. *Tob Control*. 2011;20:34-9.

36. Bush J, White M, Kai J, Rankin J, Bhopal R. Understanding influences on smoking in Bangladeshi and Pakistani adults: Community based, qualitative study *BMJ* 2003;326 :962.

37. Benowitz NL, Bernert JT, Foulds J, Hecht SS, Jacob III P, Jarvis MJ, Joseph A, Oncken C, Piper ME. Biochemical verification of tobacco use and abstinence: 2019 update. *Nicotine Tob Res*. 2020;22:1086–1097.

38. Mendel LB, Schneider EC. On the sulphocyanide-content of human saliva. *Am J Physiol* 1900;4:vii–viii.

39. Jain RB. Trends and variability in the levels of urinary thiocyanate, perchlorate, and nitrate by age, gender, race/ethnicity, smoking status, and exposure to environmental tobacco smoke over 2005-2012. *Sci Total Environ.* 2016;557-558:221–230.

40. Malik T. Goitrogenic content of common vegetables in sub-Himalayan Tarai region of eastern Uttar Pradesh. *Int J Med Health Sci*. 2012,1:32–38.

41. Youso SL, Rockwood GA, Logue BA. The analysis of protein-bound thiocyanate in plasma of smokers and non-smokers as a marker of cyanide exposure. *J Anal Toxicol* 2012;36:265–269.

42. Gendreau PL, Vitaro F. The unbearable lightness of "light" cigarettes: A comparison of smoke yields in six varieties of Canadian "light" cigarettes. *Can J Public Health* 2005;96:167–172.

43. Roy A. Tobacco consumption and the poor: An ethnographic analysis of hand-rolled cigarette (Bidi) use in Bangladesh. *Ethnography* 2012, 13:162 188.

44. Watson CH, Polzin GM, Calafat AM, Ashley DL. Determination of tar, nicotine, and carbon monoxide yields in the smoke of bidi cigarettes. *Nicotine Tob Res* 2003;5:747–753.

45. Day BJ, Bratcher PE, Chandler JD, Kilgore MB, Min E, LiPuma JJ, Hondal RJ, Nichols DP. The thiocyanate analog selenocyanate is a more potent antimicrobial pro-drug that also is selectively detoxified by the host. *Free Radic Biol Med*. 2020;146:324–332.

46. Schwingshackl L, Bogensberger B, Hoffmann G. Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: An updated systematic review and meta-analysis of cohort studies. *J Acad Nutr Diet* 2018 118:74–100.

47. Laurberg P, Andersen S, Knudsen N, Ovesen L, Nøhr SB, Bülow Pedersen I. Thiocyanate in food and iodine in milk: From domestic animal feeding to improved understanding of cretinism. *Thyroid* 2002;12:897–902.

48. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019;393:1958–1972.

49. Corsi DJ, Boyle MH, Lear SA, Chow CK, Teo KK, Subramanian SV. Trends in smoking in Canada from 1950 to 2011: Progression of the tobacco epidemic according to socioeconomic status and geography. *Cancer Causes Control*. 2014;25:45–57.

50. Ahmed MU, Pulok MH, Hashmi R, Hajizadeh M, Nargis N. Price and income elasticities of cigarette smoking demand in Bangladesh: Evidence from urban adolescents and young adults. *Nicotine Tob Res*. 2022;24:826–833.

3.9 Supporting Information

3.9.1 Chemicals, Reagents, and Analysis of Urine Samples from the PURE Cohort

Most chemicals were purchased from Sigma-Aldrich Inc. (Oakville, ON, Canada) for preparation of buffer and calibrant solutions. The chemicals for the preparation of the simulated urine matrix (100 mM sodium chloride, 10 mM sodium sulfate) and calibrant solutions were purchased from Alfa Aesar Inc. (Tewksbury, MA, USA). All frozen urine specimens were thawed slowly at room temperature and then diluted 2-fold in deionized water containing 2.0 μM of 1,5-naphthalene disulfonic acid (NDS) as internal standard unless otherwise noted. A pooled quality control (QC) urine sample was also prepared by mixing equal volumes from all 1000 urine specimens to assess technical precision in this PURE study, as well as repeat freeze-thaw and extended room temperature stability studies.

3.9.2 Capillary Electrophoresis Assay for Urinary Thiocyanate Determination

All separations were performed using an Agilent 7100 capillary electrophoresis (CE) instrument coupled to a photodiode array detector (Agilent Technologies Inc., Santa Clara, CA, USA). Data was acquired and processed using vendor software, including MassHunter Workstation LC/MS Data Acquisition (B.06.01) and extracted using MassHunter Qualitative Analysis (B.06.00) programs. Uncoated fused-silica capillaries having 75 μm internal diameter were purchased from Polymicro Technologies Inc. (Phoenix, AZ, USA). The total capillary length was 80 cm, whereas the effective capillary length to detector was 71.5 cm. The background electrolyte (BGE) for anion separation consisted of 180 mM lithium hydroxide, 230 mM phosphoric acid, 56 mM alpha-cyclodextrin (α-CD), with a final pH of 3.0 ± 0.1.[1] All electrophoretic separations were performed under reversed polarity with an applied voltage of -21 kV and the UV absorbance was monitored at 200 nm (thiocyanate) and 288 nm (NDS). All CE separations were

analyzed under an ambient temperature of 25 °C and the data was collected at a rate of 2.5 Hz with a total analysis time under 15 min. New capillaries were initially conditioned with methanol, deionized water, 1.0 M lithium hydroxide for 5 min each, and then BGE for 20 min. Prior to sample injection, the capillary was rinsed with the BGE for 3 min, followed by sample introduction into capillary using a hydrodynamic pressure for 80 s at 100 mbar (10 kPa). On-line sample preconcentration via sample self-stacking was performed by CE in order to improve concentration sensitivity for the analysis of thiocyanate, as well as sub-micromolar levels of urinary iodide.[1,2] At the beginning of each day, the capillary was flushed at 950 mbar (95 kPa) with BGE for 15 min, then a blank sample, a calibrant mixture, and a QC were analyzed by CE prior to a randomized analysis of a batch of random urine samples from PURE with a QC repeated after every 10 runs. The capillary was rinsed with deionized water for 5 min at the end of each day and the inlet and outlet ends of the capillary were stored in vials containing deionized water overnight. Urinary thiocyanate was quantified by peak integration in CE with normalization to NDS as the internal standard (1.0 μM). Spike and recovery studies were also performed on pooled urine specimens from PURE participants by adding thiocyanate at 3 different concentration levels (1.5-fold or 14 μM; 3-fold or 56 μM; 5-fold or 112 μM) in triplicate. Urinary thiocyanate, an iodine uptake inhibitor, was recently analyzed by CE to be elevated in current smokers as compared to former and never smokers when using 24 h urine specimens collected from four different clinical sites across Canada.[3] Urinary TNE-7 was measured independently in urine samples from PURE participants using multisegment injection-capillary electrophoresis-mass spectrometry with full-scan data acquisition under positive ion mode detection as described elsewhere.[4]

3.10 Supporting References

1. Nori de Macedo A, Macri J, Hudecki PL, Saoi M, McQueen MJ, Britz McKibbin, P. Validation of a capillary electrophoresis assay for monitoring iodine nutrition in populations for prevention of iodine deficiency: An interlaboratory method comparison. *J App. Lab Med.* 2017; 1:649–660.

2. Nori de Macedo A, Teo KK, Mente A, McQueen MJ, Zeidler J, Poirier P, Lear SA, Wielgosz A, Britz-McKibbin P. A robust method for iodine status determination in epidemiological studies by capillary electrophoresis. *Anal Chem*. 2014;86:10010–10015.

3. Mathiaparanam S, Nori de Macedo A, Mente A, Poirier P, Lear SA, Wielgosz A, Teo KK, Yusuf S, Britz-Mckibbin P. The prevalence and risk factors associated with iodine deficiency in Canadian adults. *Nutrients* 2022;14:1-17.

4. Sathish T, Teo KK, Britz-McKibbin P, Gill B, Islam S, Paré G, Rangarajan S, Duong M, Lanas F, Lopez-Jaramillo P, et al. Variations in risks from smoking between high-income, middle-income, and low-income countries: an analysis of data from 179 000 participants from 63 countries. *Lancet Global Health* 2022;10:e216–e226.

Table S3.1. Summary of figures of merit of CE assay for the determination of urinary thiocyanate in the PURE cohort with varying tobacco smoke exposure and country income status.

Figures of Merit	**Thiocyanate (μM)**
Inter-day precision (n=192)[a]	5.1%
Linearity (R^2)[b]	0.612
Sensitivity (mmol/L)$^{-1}$ [c]	0.999
Mean recovery in urine (%)[d]	94 ± 10
LOQ (S/N = 10)	4.5
LOD (S/N = 3)	1.4

[a] *Inter-day precision of a pooled urine sample as quality control over 44 days (including re-runs), which was analyzed intermittently between batches of 10 random urine samples in the study.*
[b] *An external calibration curve was normalized to NDS (1.0 μM) over a 15-fold (6 calibrants) linear dynamic range, respectively. Units of the slope is reported in μM.*
[c] *Method sensitivity for thiocyanate based on the slope of the external calibration curve following least-squares linear regression.*
[d] *Recovery studies were performed in triplicated by spiking thiocyanate into pooled urine samples from the PURE study participants at three different concentration levels (14 μM for 1.5×, 56 μM for 3×, and 112 μM for 5×).*

Table S3.2. Potential risk factors determined using odds ratio calculated by binary linear logistic regression with cut-off value of ≥ 23.9 µM contributing to elevated thiocyanate levels in current smokers *vs.* never smokers. Significant variables ($p < 0.05$) are bolded after adjustment for age, sex, BMI, total caloric intake, and AHEI score.

Predictor Variables	**Urinary Thiocyanate Concentration (µM)**	
	OR (95% CI)	***p*-value**
Age	1.01 (0.994-1.03)	0.243
Male sex	**2.02 (1.50-2.72)**	**3.34×10^{-6}**
Body mass index, BMI (>27 kg/m^2)	0.989 (0.964-1.02)	0.419
Education:		
None, primary, or unknown	1.00 (ref)	
Secondary/High/Higher secondary	**1.60 (1.12-2.27)**	**9.84×10^{-3}**
Trade or College/University	**2.44 (1.66-3.57)**	**5.42×10^{-6}**
Urban location	**1.51 (1.11-2.04)**	**8.26×10^{-3}**
Number of tobacco products used per day	**1.14 (1.11-1.17)**	**2.78×10^{-21}**
Country income level:		
LICs	1.00 (ref)	
MICs	**5.54 (3.32-9.23)**	**5.06×10^{-11}**
HICs)	**14.4 (8.35-24.8)**	**7.17×10^{-22}**
Smoking status:		
Never smokers	1.00 (ref)	
Light smokers (<10 CPD)	**7.07 (4.71-10.6)**	**3.00×10^{-21}**
Heavy smokers (≥10 CPD)	**15.7 (10.0-24.6)**	**3.70×10^{-33}**
Number of alcoholic drinks per day	1.16 (0.98-1.37)	0.0832
Coffee (mL/day)	**1.002 (1.001-1.002)**	**1.25×10^{-10}**

CI = confidence interval, OR = odds ratios, ref. = reference. Hosmer-Lemeshow goodness-of-fit logistic regression.

Table S3.3. Correlations between self-reported dietary intake and urinary thiocyanate concentrations from PURE participants as a function of smoking status as confirmed by independent TNE-7 measurements.

Dietary Variable[a]	Total				HICs				MICs				LICs			
	Never smokers[c] (*n*=237)		Current smokers[d] (*n*=592)		Never smokers (*n*=86)		Current smokers (*n*=240)		Never smokers (*n*=94)		Current smokers[e] (*n*=227)		Never smokers (*n*=57)		Current smokers (*n*=127)	
	r	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value
Vegetables[b]	**0.45****	**3.95×10^{-10}**	**0.13****	**5.65×10^{-3}**	**0.36****	**7.99×10^{-4}**	**-0.14***	**3.15×10^{-2}**	**0.32****	**1.46×10^{-3}**	0.08	0.219	**0.30***	**2.54×10^{-2}**	0.05	0.558
Dairy	**0.39****	**5.64×10^{-10}**	**0.19****	**1.87×10^{-6}**	0.16	0.145	-0.03	0.699	0.19	0.062	0.07	0.273	**0.29***	**3.01×10^{-2}**	-0.01	0.907
Fruits	**0.36****	**1.94×10^{-8}**	-0.02	0.633	-0.07	0.528	**-0.29****	**5.14×10^{-6}**	**0.49****	**5.48×10^{-7}**	-0.06	0.404	0.07	0.583	-0.01	0.921
Red & proc. meats	**0.22****	**7.02×10^{-4}**	**0.23****	**3.12×10^{-8}**	-0.05	0.624	0.002	0.972	-0.19	0.062	-0.02	0.773	**0.27***	**4.53×10^{-2}**	0.06	0.525
AHEI score	**0.15***	**1.95×10^{-2}**	**-0.20****	**4.53×10^{-7}**	-0.05	0.626	**-0.24****	**1.59×10^{-4}**	**0.41****	**4.59×10^{-5}**	0.04	0.569	0.21	0.118	0.08	0.346

Spearman rank order correlation analysis was performed on the non-parametric dataset, where r = Spearman rank correlation coefficient and AHEI = Alternative healthy eating index.
Current smokers are classified as ≥1 tobacco products/day with detected urinary TNE-7 levels and biochemically verified never smokers had no measurable TNE-7 levels.
*** Correlation was significant at the 0.01 level (2-tailed), whereas *correlation was significant at the 0.05 level (2-tailed).*
[a] *Intake reported as grams/day.*
[b] *Data comprised of raw (for total, HICs, and MICs) and cooked vegetables (for LICs).*
[c] *Raw vegetables (n = 178).*
[d] *AHEI score (n = 611) and raw vegetables (n = 476).*
[e] *Raw vegetables (n = 223).*

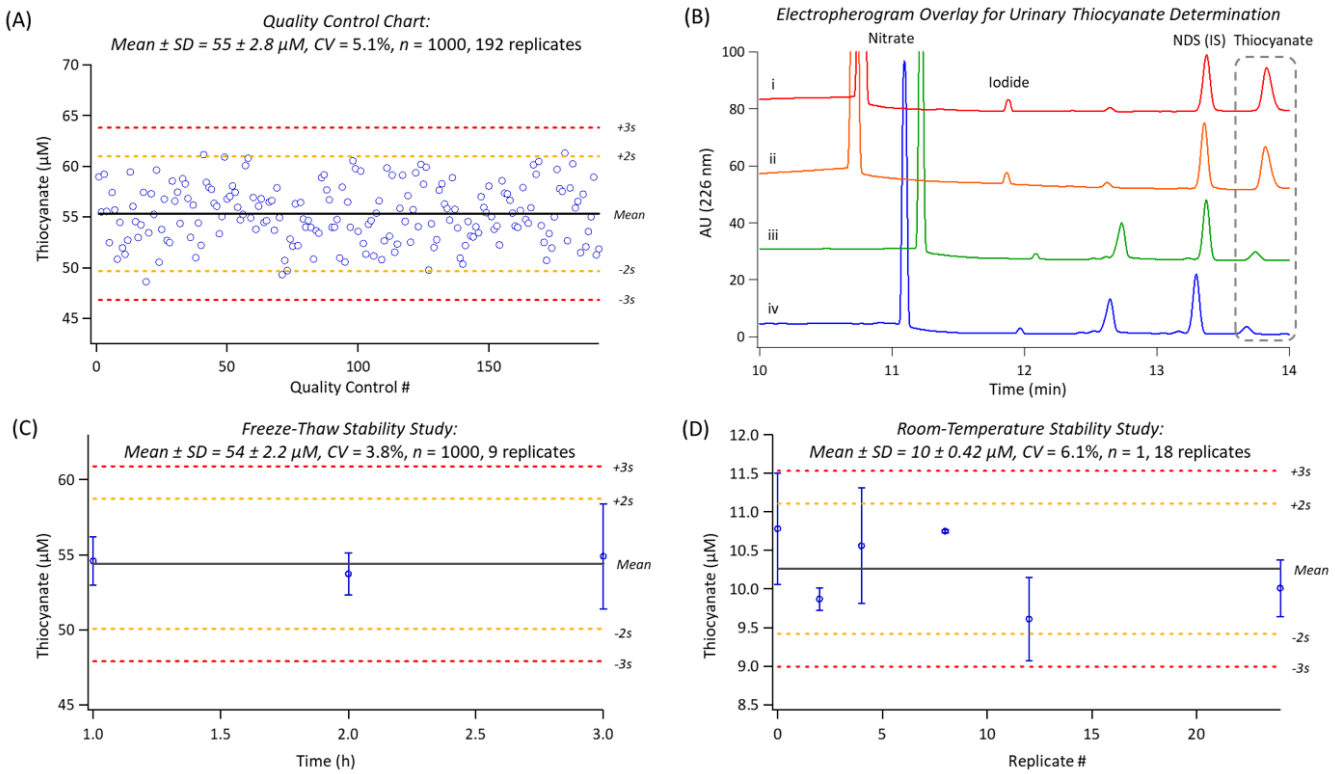


Figure S3.1. The determination of urinary thiocyanate and stability studies to demonstrate accurate and precise analytical measurements by capillary electrophoresis. **(A)** Quality control (QC) chart outlining adequate precision (mean concentration = 55 ± 2.8 µM, mean CV = 5.1%, *n* = 1000, 192 replicates). **(B)** An overlay of electropherograms representing i) QC sample (*n* = 1000), ii) after a third repeated freeze-thaw cycle, iii) freshly collected urine sample analyzed instantly, and iv) after 24 h at room temperature. **(C)** A scatter plot determining the stability of thiocyanate in pooled urine from the PURE study undergoing three repeated freeze-thaw cycles (mean concentration = 54 ± 2.2 µM, mean CV = 3.8%, *n* = 9). **(D)** A scatter plot determining the stability of thiocyanate in freshly collected urine over a 24 h period, including time points 0, 2, 4, 8, 12, and 24 h (mean concentration = 10 ± 0.42 µM, mean CV = 6.1%, *n* = 18).

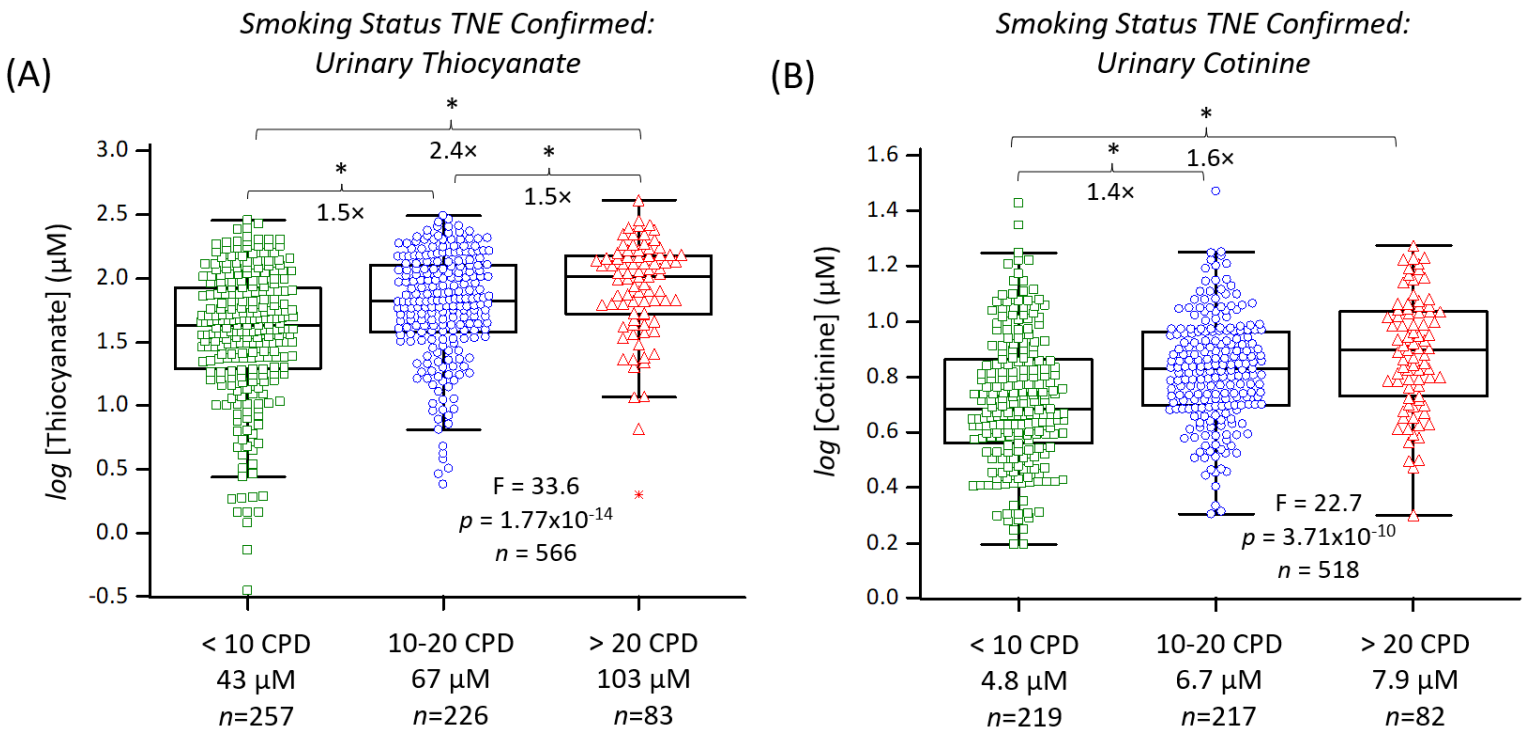


Figure S3.2. Box whisker plots comparing the tobacco smoke dose response of urinary **(A)** thiocyanate relative to **(B)** cotinine measured by CE and MSI-CE-MS methods, respectively in TNE-7 verified current smokers from the PURE cohort. Overall, urinary thiocyanate had a more significant response changes as a function of self-reported smoking intensity as compared to cotinine when using ANCOVA with covariate adjustment for age, sex, BMI, tobacco type, total caloric intake, AHEI score, education, location, alcohol use, history of diabetes, history of hypertension, and $PM_{2.5}$. Also, there were fewer non-detected urine samples in the case of thiocyanate than cotinine, which is only one of up to 7 or more major nicotine metabolites excreted in urine (* $p < 0.05$ with Bonferroni correction).

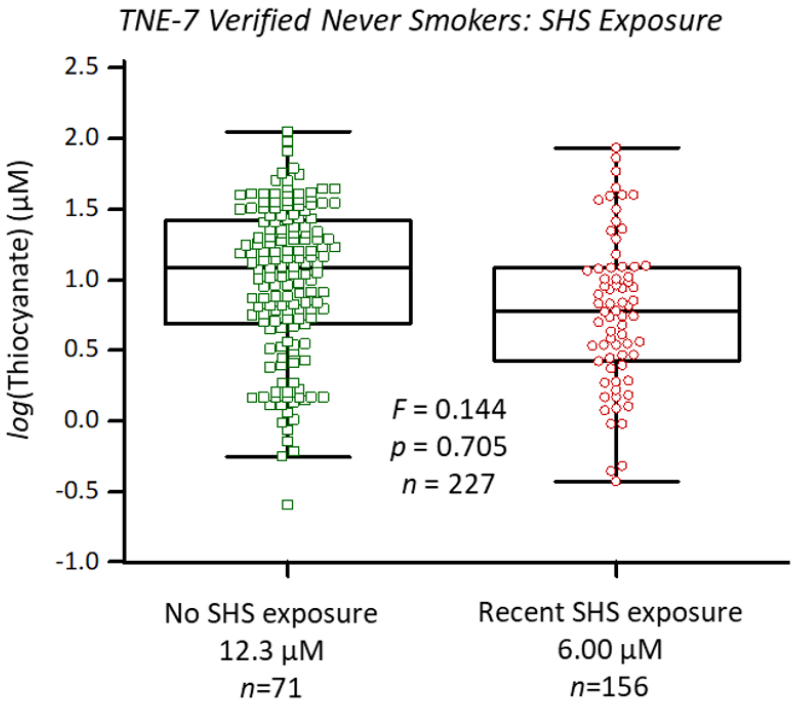


Figure S3.3. Box whisker plot highlights that urinary thiocyanate was not significantly different for participants in the PURE study when comparing TNE-7 verified never smokers (*n*=227) who self-reported SHS versus without SHS exposure when using ANCOVA with covariate adjustment.

Chapter IV

Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-Limited Sweat Specimens

Stellena Mathiaparanam, Adriana Nori de Macedo, Biban Gill, Katherine Keenan, Tanja Gonska, Linda Pedder, Stephen Hill, and Philip Britz-McKibbin

J. Cyst. Fibros, **2022** (https://doi.org/10.1016/j.jcf.2022.05.004)

S.M. performed all the experiments including sample preparation, data acquisition using CE-iUV, data processing, interpretation, statistical analysis, and wrote the initial draft for publication. A.N.M and B.G. performed complimentary experiments. S.H., L.P., T.G., and K.K. were involved in participant recruitment, collection of patient metadata, and coordination of sweat samples. P.B.M. supervised the study, and with all other co-authors, contributed to the final version of the manuscript.

Chapter IV: Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-Limited Sweat Specimens

4.1 Abstract

Cystic fibrosis (CF) is a debilitating genetic disorder that benefits from early detection. CF diagnosis relies on measuring elevated sweat chloride that is difficult in neonates with low sweat rates. We introduce a new method for sweat chloride determination from volume-limited specimens, and explore the potential utility of sweat bicarbonate in neonatal CF screening. A rapid assay (< 5 min) was developed to analyze chloride and bicarbonate using capillary electrophoresis with indirect UV detection (CE-iUV). Pilocarpine-stimulated sweat samples from screen-positive CF infants were collected at two hospital sites, including confirmed CF (n=12), CF screen-positive inconclusive diagnosis (n=4), and unaffected non-CF cases (n=37). All sweat chloride samples were analyzed by a coulometric titrator and CE-iUV, and the viability to measure acid-labile bicarbonate was also evaluated. Stability studies revealed that bicarbonate can be reliably assessed in sweat if acidification and heating were avoided. Method validation demonstrated that sweat chloride and bicarbonate were quantified with acceptable accuracy (recovery of 102%), precision (CV = 3.7%) and detection limits (~ 0.1 mM). An inter-laboratory comparison confirmed a mean bias of 6.5% (n=53) for sweat chloride determination by CE-iUV relative to a commercial chloridometer. However, sweat bicarbonate did not discriminate between CF and non-CF infants (AUC = 0.623, p = 0.215) unlike chloride (AUC = 1.00, $p = 3.00 \times 10^{-7}$). CE-iUV offers a robust method for sweat chloride testing from presumptive CF infants that may reduce testing failure rates. However, sweat bicarbonate does not have clinical value in newborn CF diagnosis.

4.2 Introduction

Cystic fibrosis (CF) is a life-shortening autosomal recessive disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) channel, which leads to chronic malnutrition, progressive pulmonary injury, and other health complications if left untreated [1]. The advent of universal newborn screening (NBS) for CF enables initiation of therapeutic interventions during infancy for improved long-term health outcomes as compared to symptomatic diagnosis [2]. Most CF neonates are screened using a multi-tiered algorithm based on elevated immunoreactive trypsinogen (IRT) along with a panel of one or two disease-causing *CFTR* mutations from dried blood spots, [3] which may also include next generation sequencing to improve overall positive predictive value [4]. Nevertheless, confirmatory diagnosis of presumptive CF cases relies on measuring elevated sweat chloride (≥ 60 mM) using a quantitative pilocarpine-stimulated iontophoresis test [5]. This is required to differentiate CF cases from a large fraction of screen-positive infants who are unaffected carriers or false-positives (< 30 mM) [6]. However, diagnostic dilemmas in CF remain [7], such as ambiguous test results for infants denoted as cystic fibrosis screen-positive inconclusive diagnosis (CF-SPID) with intermediate sweat chloride concentrations (30-59 mM), a subset who may later be diagnosed with CF [8].

Standardized sweat testing for CF is thus critical for clinical decision making that is hampered by variations in sweat collection methods implemented between laboratories when using coulometric titrators or conductivity analyzers for sweat chloride determination [9,10]. This is important since sweat chloride concentrations depend on the rate of secretion from the human eccrine gland, and its active reabsorption by CFTR in the dermal duct that is defective in CF [11]. Most clinical laboratories use proprietary coiled microbore collection tubes following sweat stimulation via pilocarpine iontophoresis performed at the inner forearm [12]. Although the rate of quantity not sufficient sweat samples collected via the Macroduct® collection system generally satisfies recommendations by the

Cystic Fibrosis Foundation (≤ 10% in infants under 3 months), delays in diagnosis are more frequent among low birthweight, premature, and non-Caucasian infants who generate inadequate sweat volumes for testing due to their lower sweat rates [13]. Alternatively, *in-situ* conductivity measurements using a Nanoduct® sweat test system requires less sweat volume, but suffers from similar failure rates and lower specificity for newborn CF diagnosis as compared to conventional methods [14]. As a result, rapid and accurate methods that can analyze smaller volumes of sweat (< 10 μL) are needed for reliable confirmatory testing of CF infants to reduce the potential harms from delayed treatment, premature discharge and prolonged uncertainty contributing to parental anxiety [15].

Active salt and water transport across epithelial cells are largely mediated by chloride conductance via the CFTR channel, which also regulates the secretion of several other anions with different permeability selectivity, such as thiocyanate [16]. Recent sweat metabolome analyses also revealed differential secretion of exogenous anionic compounds between CF and non-CF infants, including a drug metabolite of pilocarpine, namely pilocarpic acid [17]. Thus, the complex pathophysiology of CF is likely dependent on the impaired transport of other anions since CFTR functions not only as a chloride channel [18]. For instance, lower epithelial bicarbonate efflux results in increased acidification and reduced buffer capacity of airway surface liquid impacting mucus production in CF (*i.e.*, mucoviscidosis), which is needed to prevent recurrent bacterial infections and chronic inflammation [19,20]. However, the clinical potential of sweat bicarbonate as compared to chloride to diagnose presumptive CF infants has not been explored to date [21]. Herein, we present a robust assay to simultaneously measure chloride and bicarbonate in volume-limited sweat samples from screen-positive CF neonates when using capillary electrophoresis with indirect UV absorbance detection (CE-iUV).

4.3 Methods

4.3.1 Study Design and Sweat Collection from Screen-Positive CF Infants

This study was approved by the Hamilton Integrated Research Ethics Board and The Hospital for Sick Children's Research Ethics Board in accordance with the Declaration of Helsinki principles. All sweat specimens were analyzed from screen-positive CF infants following standardized pilocarpine-stimulated sweat chloride collection at two different pediatric hospital sites. Sweat samples were de-identified and relabeled with a research ID from McMaster Children's Hospital (Hamilton, ON) and The Hospital for Sick Children (Toronto, ON) to protect patient privacy. Sweat collection was performed using a Model 3700 Macroduct Advanced® Sweat Capillary Collection System (ELITechGroup Wescore Biomedical Systems, Logan, UT, USA). Electrodes and agar gel discs containing 0.5% w/w pilocarpine nitrate were placed on the infant's forearm for sweat stimulation after carefully cleaning the region with ethanol and distilled water. A low current (1.5 mA) was applied for 5 min to deliver pilocarpine to the sweat gland via iontophoresis. Sweat was collected for 30 min using a coiled microbore tube containing a blue dye to facilitate visualization of the sample volume. Samples were transferred to a 0.5 mL centrifuge tube, and then stored in a fridge (+4 °C) after taking an aliquot for analysis via a coulometric titrator using a FDA/Health Canada-approved Model 3400 ChloroChek® Chloridometer® (ELITechGroup Wescore Biomedical Systems). In general, the commercial chloridometer required a minimum of 10 μL of sweat and analyses were performed in duplicate. Residual sweat samples were then stored frozen at -80 °C prior to thawing and subsequent CE-iUV analysis. Overall, 53 presumptive/screen-positive CF infants (< 3 months old) identified via a two-tiered algorithm based on elevated immunoreactive trypsinogen (IRT) and genetic mutation panel at NSO were recruited in this study as summarized in **Table 4.1**.

Table 4.1. Clinical characteristics of screen-positive CF infants recruited in this study.

Characteristic	Non-CF (*n*=37)	CF-SPID (*n*=4)	CF (*n*=12)
Female:Male	18:19	0:4	5:8
Age (days)[*]	22 ± 6	34 ± 7	19 ± 14
Birth weight (g)	3550 ± 487	3750 ± 61	3514 ± 412
Gestational age (weeks)[b]	39 ± 1	--	40 ± 1
Chloride (mM)	13 ± 6	34 ± 8	93 ± 21
IRT (μg/L)	64 ± 33	81 ± 1	149 ± 49
CFTR genotype:			
0 mutations	2	--	--
1 mutation: delF508/null	11	--	--
1 mutation: non-delF508/null [c]	3	--	--
2 mutations: delF508/delF508	--	--	2
2 mutations: delF508/non-delF508[d]	12	3	7
2 mutations: non-delF508[e]	9	1	3
Collection site:			
McMaster	22	--	11
Sick Kids	15	4	2
Pancreatic status:[f]			
Pancreatic sufficient	--	4	3
Moderate pancreatic insufficient			2
Pancreatic insufficient	--	--	7

[a] *All data is reported as (mean ± 1s)* except age, chloride, and IRT measurements (median ± IQR).*

[b] *For non-CF cases, 15 samples were excluded due to insufficient information.*

[c] *For non-CF cases, variants of unknown significance: n=3, non-delF508/null.*

[d] *For non-CF cases, variants of unknown significance: n=12, delF508/-. For CF-SPID cases, CFTR variants of unknown significance: n=3, delF508/non-delF508. For CF cases, CFTR disease-causing variants: n=2, delF508/G551D; n=1, delF508/G1244E; n=1, delF508/G85E; n=1, delF508/3199del6. Variants under evaluation: n=1, delF508/2789+2insA. Variants of unknown significance: n=1, delF508/non-delF508.*

[e] *For non-CF cases, CFTR disease-causing when combined with another pathogenic variant: n=1, 621+1G>T/-; n=1, N1303K; n=1, Y1092X/-. Variants of varying clinical consequences: n=2, R117H/-. Variants of unknown significance: n=4, -/-. For CF-SPID cases, CFTR variants of unknown significance: n=1, non-delF508/non-delF508. For CF cases, CFTR disease-causing variants: n=1, Y569D/Y569D. Variants of varying clinical consequence: n=1, R117H/G551D. Variants of unknown significance: n=1, non-delF508/non-delF508.*

[f] *Pancreatic status classified for screen-positive CF infants by fecal elastase measurements, where pancreatic sufficient (> 200 μg/g), moderate pancreatic insufficient (100-200 μg/g) and pancreatic insufficient (< 100 μg g/). Pancreatic status was not tested for non-CF infants.*

This included 37 unaffected infants with normal sweat chloride (< 30 mmol/L), 12 CF infants with elevated sweat chloride (≥ 60 mmol/L), as well as 4 CF-SPID cases with intermediate sweat chloride results (30-59 mmol/L). Pancreatic status

was measured for confirmed CF and CF-SPID cases based on fecal elastase measurements and classified as pancreatic sufficient (> 200 μg/g), moderate pancreatic insufficient (100-200 μg/g) or pancreatic insufficient (< 100 μg g/g).

4.3.2 Chemicals, Reagents, and Analysis of Infant Sweat Samples from Presumptive CF Infants

All the chemicals were purchased from Sigma-Aldrich Inc. (St. Louis, MO, USA) for preparation of buffer and calibrant solutions. All frozen sweat specimens were thawed and then diluted 4-fold in double deionized water containing 3 mmol/L of sodium formate as internal standard, followed by sample mixing using a vortex for 30 s prior to CE-iUV analysis. A pooled quality control (QC) sample was prepared by mixing sweat samples collected from 8 screen-positive infants, including 4 non-CF and 4 CF infants. Overall, residual sweat sample volumes available for analysis by CE-iUV ranged from 0.5 to 10 μL, which would be considered insufficient volume for testing by the commercial coulometric titrator.

4.3.3 Instrumentation and CE-iUV Operating Conditions

All CE separations were conducted using a P/ACE^{TM} MDQ system with UV absorbance detection (SCIEX, Framingham, MA, USA). Unmodified fused-silica capillaries having outer polyimide coating were purchased from Polymicro Technologies Inc. (Phoenix, AZ, USA). The internal diameter of the fused-silica capillaries was 50 μm internal diameter with a total and effective (to detector window) capillary length of 60 cm with 50 cm, respectively. New capillaries were initially conditioned by flushing with methanol, deionized water, 1.0 M lithium hydroxide (5 min each), and then background electrolyte (BGE) for 20 min using a rinse pressure of 20 psi (138 kPa) as programmed by the 32 Karat^{TM} software for the CE system. The BGE consisted of 5.0 mmol/L sodium chromate, 5.0 mmol/L naphthalene trisulfonic acid, 0.5 mmol/L tetradecyl trimethyl ammonium bromide (TTAB), that was adjusted with 10 mM diethylamine with a pH of 11.30.

The CE separations were performed under reversed polarity with a maximum applied voltage of -30 kV, and UV absorbance was monitored at 280 nm. For clarity purposes, electropherograms were inverted as positive peaks for UV-transparent electrolytes. All CE separations were analyzed at 25 °C and the data was collected at a rate of 4 Hz with a total run time of 5 min. Prior to each analysis, the capillary was rinsed with the BGE for 3 min followed by sample injection via hydrodynamic pressure for 15 s at 0.5 psi (3.4 kPa). At the start of each day, a blank sample, a calibrant mixture, and a QC were analyzed by CE-iUV prior to a randomized analysis of a batch of individual sweat samples with a QC sample repeatedly analyzed after every 9 runs. At the end of each day, the capillary was flushed with deionized water for 5 min and the inlet and outlet ends of the capillary were stored in vials containing deionized water overnight. Calibration curves were performed for chloride and bicarbonate by CE-iUV, where their integrated peak areas were normalized to formate as an internal standard. Recovery studies were performed on pooled sweat samples from CF infants by spiking in chloride and bicarbonate at 3 different concentration levels (1.5× or 7 mM and 2 mM; 3× or 28 mM and 8 mM; 5× or 56 mM and 16 mM, respectively) in triplicate. As a comparison, two other validated CE assays were also used to measure sweat thiocyanate and sodium from presumptive CF infants as described previously [22,23].

4.3.4 Bicarbonate Stability Study and Acid-Heat Treatment

Stability studies for sweat bicarbonate and chloride determination by CE-iUV were performed by collecting sweat via pilocarpine iontophoresis from a healthy adult volunteer at McMaster Children's Hospital. This sweat sample was aliquoted, kept on ice and then promptly analyzed by CE-iUV under different conditions, including repeat analysis over a 10 h period at room temperature, as well as after three repeat freeze-thaw cycles. Also, an acidified heat treatment was performed on pooled sweat QC samples after addition of 8 mmol/L acetic acid

with heating at a temperature of 60 °C for 25 min in a heat block to assess the stability of bicarbonate.

4.3.5 Statistical Analyses

Electropherograms and calibration curves were plotted on Igor Pro 5.04B (Wavemetrics Inc, Lake Oswego, OR, USA). Percent difference Bland-Altman plots, boxplots with scatter plot overlays, and receiver operating characteristic (ROC) curves were plotted using MedCalc 12.5 statistical software (Ostend, Belgium). Univariate statistical analysis, such as student's t-test with unequal variance and ROC curves were also performed on SPSS 18.0 statistical software (IBM SPSS, Chicago, IL, USA).

4.4 Results

4.4.1 Rapid and Robust Analysis of Sweat Chloride and Bicarbonate by CE-iUV

Figure 4.1A depicts a series of electropherograms for simultaneous analysis of sweat chloride and bicarbonate within 5 min by CE-iUV. Given the acid lability of bicarbonate (pH < 6.0), CE separations were performed under strongly alkaline conditions, where bicarbonate migrates predominately as carbonate with a larger negative electrophoretic mobility than chloride. Pilocarpine-stimulated sweat specimens were found to be weakly alkaline (pH ~ 8) despite human sweat being considered a weakly acidic (pH ~ 6.3) biofluid as compared to blood [24]. Stability studies demonstrated that delays to analysis of recently collected sweat samples over a 10 h period when left at room temperature or following three repeat freeze-thaw cycles, did not impact the response of either chloride or bicarbonate (p > 0.05) as shown in **Figure 4.1A**. In fact, control charts for repeated analysis of sweat chloride (**Figure 4.1B**) and bicarbonate (**Figure 4.1C**) demonstrated excellent precision with a coefficient of variance (CV) < 5% with no outliers exceeding warning limits (± 2 s). As expected, heating (60 °C) of

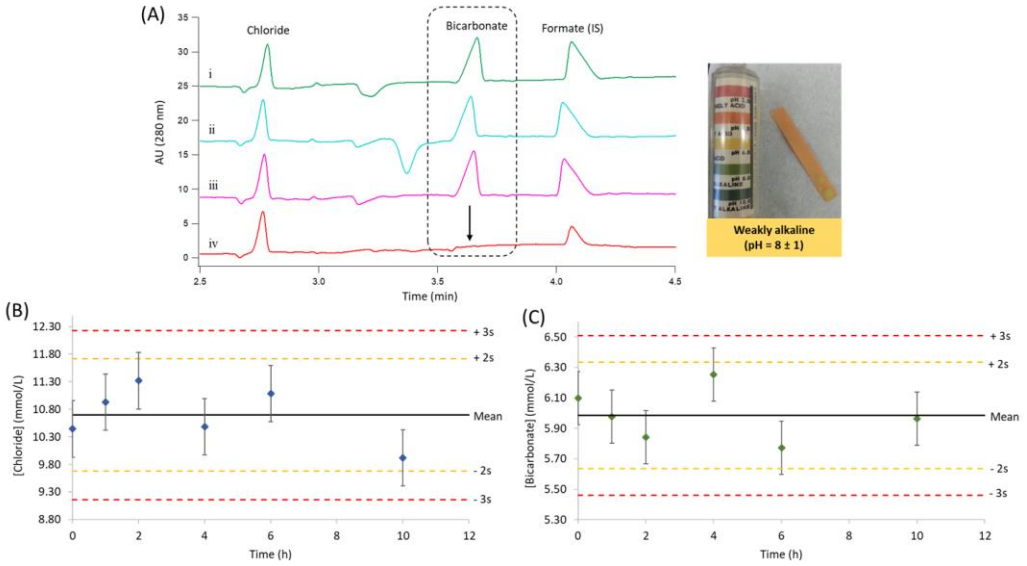


Figure 4.1. Rapid and selective determination of sweat chloride and bicarbonate by CE-iUV and stability studies to demonstrate robust analytical performance. **(A)** Electropherogram overlay displaying a i) freshly collected pilocarpine-stimulated sweat sample analyzed promptly after collection, ii) after 10 h at room temperature, iii) after a third repeated freeze-thaw cycle and iv) after acidification and heat treatment. Control charts for repeated analysis of sweat **(B)** chloride (mean CV = 4.8%, *n*=18) and **(C)** bicarbonate (mean CV = 2.9%, *n*=18) during stability studies confirm good precision over a 10 h time period with no outliers beyond warning limits (± 2 s). Recently collected pilocarpine sweat specimens were confirmed to be weakly alkaline (pH ~ 8) that allows for quantitative measurement of bicarbonate from screen-positive CF infants provided that excessive acidification and heating are avoided.

acidified sweat samples resulted in bicarbonate decomposition and its loss as carbon dioxide gas unlike chloride. The latter experiment also demonstrated good method selectivity with no evidence of other involatile anions co-migrating with bicarbonate in diluted sweat samples as highlighted in **Figure 4.1A**.

4.4.2 Method Validation and Inter-Method Comparison of Sweat Chloride

Table 4.2 summarizes the figures of merit of the CE-iUV assay for quantitative sweat chloride and bicarbonate determination. Inter-day precision was performed

Table 4.2. Summary of figures of merit of CE-iUV assay for the determination of sweat chloride and bicarbonate in screen-positive CF infants.

Figures of Merit	**Chloride (mmol/L)[a]**	**Bicarbonate (mmol/L)[a]**
Mean ± 1s (CF, *n*=12)	93 ± 21	5.1 ± 4.1
Mean ± 1s (non-CF, *n*=37)	12.5 ± 5.2	6.7 ± 4.1
Linearity (R^2)[b]	0.999	0.999
Sensitivity (mmol/L^{-1})[c]	0.211	0.591
LOQ (*S/N* = 10)	0.47	0.17
LOD (*S/N* = 3)	0.14	0.05
Inter-day precision (*n*=9)[d]	1.5%	5.9%
Mean recovery (%)[e]	103 ± 5	100 ± 8

[a] There was a mean fold-change in the ratio of CF to non-CF of 7.4 and 0.75 for sweat chloride and bicarbonate with a significance of $p = 2.59 \times 10^{-8}$ and $p = 0.237$ based on student's t test with unequal variance, respectively. For bicarbonate measurements in the non-CF category (n=35), two samples were exclude as they were < LOQ. *[b] Calibration curves were normalized to formate (3.0 mmol/L) over a 70-fold (chloride, 7 calibrants) and 250-fold (bicarbonate, 8 calibrants) linear dynamic range, respectively.*

[c] Method sensitivity for each anion based on slope of calibration curve following least-squares linear regression.

[d] Interday precision of a pooled sweat sample as quality control over 3 days, which was analyzed intermittently between batches of 9 random infant sweat samples in study.

[e] Recovery studies were performed in triplicate by spiking chloride and bicarbonate into pooled sweat samples from screen-positive CF infants at three different concentration levels.

following an intermittent analysis of a pooled sweat sample as a QC over a 3 day period, which generated a mean CV of 1.5% and 5.9% (*n*=9) for sweat chloride and bicarbonate, respectively similar to stability study outcomes. The linearity of calibration curves (R^2 > 0.999) was excellent over a wide dynamic range (70 to 250-fold) with nearly a 3-fold greater sensitivity for bicarbonate than chloride reflecting its lower detection (*S/N* ~ 3; 0.05 vs. 0.14 mmol/L) and quantification limits (*S/N* ~ 10; 0.17 vs. 0.47 mmol/L). Also, spike-recovery studies confirmed excellent method accuracy as reflected by a mean recovery of (103 ± 5)% and (100 ± 8)% performed at three different concentration levels for chloride and bicarbonate in pooled infant sweat samples, respectively. Furthermore, an inter-laboratory comparison for sweat chloride determination from presumptive CF infants using CE-iUV and coulometric titration are depicted in **Figure 4.2**. A

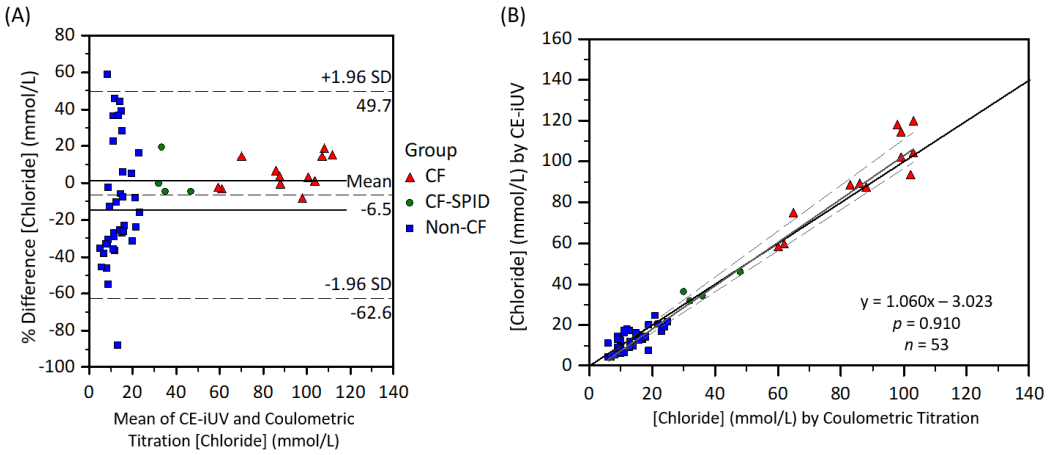


Figure 4.2. Good mutual agreement in sweat chloride determination was demonstrated for confirmatory testing of presumptive CF infants (n=53) based on the CE-iUV method as compared to coulometric titration as highlighted in a **(A)** % difference Bland-Altman plot and **(B)** Passing-Bablok linear regression analysis. Overall, there was a mean bias of 6.5 % with two outliers exceeding agreement limits (p = 0.0377) along with a slope of 1.06 that was not significantly different from the line of unity (p = 0.910). CE-iUV enabled the analysis of residual sweat samples as low as 0.5 µL that may reduce testing failure rates due to inadequate sweat volumes collected from infants.

Bland-Altman plot confirmed a modest bias of 6.5% with few outliers beyond agreement limits (p = 0.0374). Similarly, a linear correlation between CE-iUV and chloridometer measurements for sweat chloride are shown in a Passing-Bablok regression plot with a slope of 1.06 that did not differ from the line of unity (p = 0.910). Overall, the mean concentrations of sweat chloride as measured by CE-iUV in CF, CF-SPID and non-CF infants were (93 ± 21 mmol/L, n=12), (37.1 ± 6.1 mmol/L, n=4), and (12.5 ± 5.2 mmol/L, n=37), respectively. Importantly, residual sweat volumes from infants as low as 0.5 µL were reliably analyzed by this method.

4.4.3 Diagnostic Performance for Differentiating Presumptive CF Infants

Table 1 summarizes the clinical characteristics of a cohort of screen-positive CF

infants (n=53) with similar birth weight and gestational age (p > 0.05). Confirmatory CF diagnosis by sweat chloride testing was performed at two different pediatric hospital sites with a median age of 22 days (ranging from 11 to 95 days). As expected, CF as compared to non-CF infants had higher IRT and notably elevated sweat chloride (≥ 60 mmol/L) with two confirmed disease-causing *CFTR* mutations, including 7 delF508 heterozygotes, 2 delF508 homozygotes, and 3 non-delF508 mutations. In contrast, unaffected carriers with low sweat chloride (< 30 mmol/L) primarily had only one delF508 *CFTR* mutation detected unlike CF-SPID cases. Most confirmed CF infants had severe (< 100 μg/g) or moderate (100-199 μg/g) pancreatic insufficiency, whereas all CF-SPID infants having an initial inconclusive CF diagnosis were pancreatic sufficient (> 200 μg/g) based on fecal elastase measurements. **Figure 4.3** summarizes box-whisker plots and their corresponding receiver operating characteristic (ROC) curves for sweat chloride and bicarbonate from screen-positive CF infants as measured by CE-iUV. There was an average 7.4-fold elevated sweat chloride in CF as compared to non-CF infants with excellent diagnostic accuracy with an area under the curve (AUC = 1.00, $p = 3.00 \times 10^{-7}$). Also, no CF infants were misclassified by CE-iUV as compared to confirmatory testing by coulometric titration. In contrast, there was no difference in sweat bicarbonate concentrations between CF and non-CF infants as shown in the ROC curve (AUC = 0.623, p = 0.215). Also, a Pearson correlation analysis confirmed no association between sweat chloride and bicarbonate (r = 0.193) from screen-positive CF infants. For instance, the mean sweat bicarbonate concentrations were not different between the three screen-positive CF sub-groups (p > 0.05) tested with CF, CF-SPID and non-CF infants having (5.1 ± 4.1 mmol/L, n=12), (5.0 ± 3.3 mmol/L, n=4) and (6.7 ± 4.1 mmol/L, n=35), respectively. We also independently measured sweat thiocyanate and sodium using two other validated CE assays [22,23] that confirmed excellent discrimination (AUC > 0.935)

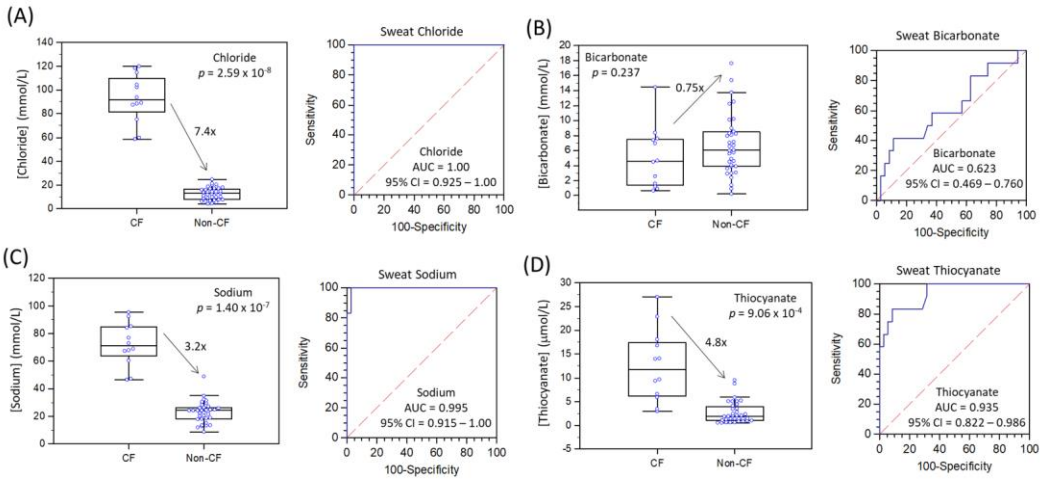


Figure 4.3. Boxplots and ROC curves used to differentiate CF from unaffected non-CF infants identified from newborn screening when using CE-iUV. **(A)** Overall, sweat chloride provided the strongest discrimination between presumptive CF infants based on an area under the curve, AUC = 1.00 and $p = 3.00 \times 10^{-7}$ with a mean fold-change (FC) of 7.4. **(B)** In contrast, sweat bicarbonate did not differentiate between screen-positive CF sub-groups as reflected by AUC = 0.623 and $p = 0.215$ with a mean FC of 0.75. Two other CE assays were also applied to measure sweat electrolytes whose transport is impaired by disease-causing *CFTR* mutations in CF as a comparison, including **(C)** sodium having an AUC = 0.995 and $p = 3.88 \times 10^{-7}$ with a mean FC = 3.2, as well as **(D)** thiocyanate having an AUC = 0.935 and $p = 8.49 \times 10^{-6}$ with a mean FC = 4.8.

between CF and non-CF infants as shown in **Figure 4.3**. Yet unlike bicarbonate, both sodium ($r = 0.898$) and thiocyanate ($r = 0.796$) had a much stronger linear correlation with sweat chloride concentrations in screen-positive CF infants (n=53).

4.5 Discussion

We have demonstrated that the weakly alkaline conditions in pilocarpine-stimulated sweat specimens from infants enable reliable analysis of acid-labile bicarbonate together with chloride by CE-iUV that also tolerates delays in storage and multiple freeze-thaw cycles. High concentrations of pilocarpine were previously measured in neonatal CF sweat specimens [17] that likely modulate

pH and buffer capacity following iontophoretic drug delivery for sweat stimulation. Earlier studies reported that local stimulation of sweating from intradermal injection of pilocarpine or acetylcholine generated sweat with an alkaline pH in contrast to heat-stimulated acidic sweat collected without cholinergic drug stimulation [25]. Overall, an inter-laboratory comparison for sweat chloride by CE-iUV and coulometric titration demonstrated good mutual agreement with no differences in CF diagnostic classification. However, the modest positive bias and greater variance for chloride measurements notably among non-CF cases (< 10 mmol/L) are likely attributed to the lower selectivity of silver-based titration methods with amperometric detection from potential matrix interferences in sweat specimens that also form insoluble silver salts, including minor halides (e.g., iodide) and other anions (e.g., thiocyanate). A major advantage of CE-iUV as compared to coulometric titrators, conductivity analyzers and ion selective electrodes is that extremely low volumes of sweat (~ 0.5 μL) can be analyzed from low birth weight and premature infants, who are prone to insufficient sweat production requiring repeat testing with delayed CF diagnoses [13,26]. This microseparation platform also complements CE with contactless conductivity detection reported for assessment of sweat sodium, potassium, and chloride [27,28], as well as inductively coupled plasma-mass spectrometry that suffers from more expensive infrastructure and high operating costs [29].

Despite the key role of impaired bicarbonate transport by CFTR in CF pathophysiology [19,20], pilocarpine-stimulated sweat bicarbonate was found not to differentiate between CF and non-CF infants while also lacking any correlation with sweat chloride concentrations. These results are consistent with sweat bicarbonate measurements performed by X-ray analysis following barium precipitation without pilocarpine stimulation, which reported no differences in adult CF patients relative to healthy control irrespective of sweat rate [21]. Overall, sweat bicarbonate concentrations (~ 5 mmol/L) were considerably lower

than reported in pancreatic juice that can exceed 140 mmol/L as mediated by a CFTR-mediated chloride/bicarbonate exchange in pancreatic duct cells, which are reduced to about 20 mmol/L in CF patients with pancreatic insufficiency [30]. Analysis of sweat sodium and thiocyanate confirmed that impaired reuptake of these electrolytes allowed for excellent differentiation of CF from non-CF infants unlike bicarbonate. As we have shown that potential bias from pre-analytical variations in sweat collection and storage do not impact bicarbonate stability in well-matched CF infant sub-groups, reabsorption of bicarbonate is likely mediated by other mechanisms other than CFTR conduction within human sweat gland ducts that require further exploration.

In summary, CE-iUV offers a rapid, selective, precise and accurate approach for simultaneous analysis of sweat chloride and bicarbonate from volume-limited infant sweat samples. This latter feature may significantly reduce test failure rates as compared to coulometric titration, which needs to be evaluated within a clinical setting. Although sweat bicarbonate lacked diagnostic value in confirmatory sweat testing of presumptive CF infants, this method may prove useful when characterizing other epithelial cells where CFTR mutations have been shown to impair bicarbonate secretion and contribute to disease burden in CF, such as airway lung fluid and pancreatic juice.

4.6 Acknowledgements

The authors wish to thank the families of presumptive CF infants, and staff from CF clinics at McMaster Children's Hospital and The Hospital for Sick Children.

4.7 References

[1] Goetz D, Ren CL, Review of cystic fibrosis. *Pediatr Ann.* 2019; **48**: e154–e161.

[2] Mak DYF, Sykes J, Stephenson AL, Lands LC, The benefits of newborn screening for cystic fibrosis: The Canadian experience, *J Cystic Fibrosis.* 2016; **15**: 302–308.

[3] Rosenfeld M, Sontag MK, Ren CL, Cystic fibrosis diagnosis and newborn screening. *Pediatr Clin North Am.* 2016; **63**: 599–615.

[4] Baker M, Atkins A, Cordovado S, et al., Improving newborn screening for cystic fibrosis using next-generation sequencing technology: A technical feasibility study. *Genet Med* 2016; **18**: 231–238.

[5] Quinton PM, Chloride impermeability in cystic fibrosis. Nature 1983; **301**: 421–422.

[6] Collie JTB, Massie RJ, Jones OAH, et al., Sixty-five years since the New York heat wave: Advances in sweat testing for cystic fibrosis. *Pediat Pulmonol.* 2014; **49**: 106–117.

[7] Simmonds NJ, Is it cystic fibrosis? The challenges of diagnosing cystic fibrosis. *Paediatr Respir Rev.* 2019; **31**: 6–8.

[8] Gonska T, Keenan K, Au J, et al. Outcomes of cystic fibrosis screening-positive infants with inconclusive diagnosis at school age. *Pediatrics* 2021; **148**: e2021051740.

[9] Servidoni MF, Gomez CCS, Marson FAL, et al., Sweat test and cystic fibrosis: Overview of test performance at public and private centers in the state of São Paulo, Brazil. *J Bras Pneumol.* 2017; **43**: 121–128.

[10] Gokdemir Y, Vatansever P, Karadag B, et al. Performance evaluation of a new coulometric endpoint method in sweat testing and its comparison with classic Gibson&Cooke and chloridometer methods in cystic fibrosis. *Front Pediatr.* 2018; **6**: 133.

[11] Sonner Z, Wilder E, Heikenfeld J, et al., The microfluidics of the eccrine sweat gland, including biomarker partitioning, transport, and biosensing implications. *Biomicrofluidics* 2015; **9**: 031301.

[12] Laguna TA, Lin N, Wang Q, et al., Comparison of quantitative sweat chloride methods after positive newborn screen for cystic fibrosis. *Pediat Pulmonol.* 2012; **47**: 736–742.

[13] McColley SA, Elbert A, Wu R, et al., Quantity not sufficient rates and delays in sweat testing in US infants with cystic fibrosis. *Pediatr Pulmonol.* 2020; **55**: 3053–3056.

[14] Rueegg CS, Kuehni CE, Gallati S, et al., Comparison of two sweat test systems for the diagnosis of cystic fibrosis in newborns. *Pediatr Pulmonol.* 2019; **54**: 264–272.

[15] Chakraborty P, Potter BK, Hayeems RZ, Maximizing benefits and minimizing harms: Diagnostic uncertainity arising from newborn

screening. *Pediatrics* 2021; **148**: e2021052822.

[16] Linsdell P. Relationship between anion binding and anion permeability revealed by mutagenesis within the cystic fibrosis transmembrane conductance regulator chloride channel pore. *J Physiol.* 2001; **531**: 51–66.

[17] Nori de Macedo A, Mathiaparanam S, Brick L, et al., The sweat metabolome of screen-positive cystic fibrosis infants: Revealing mechanisms beyond impaired chloride transport. *ACS Cent Sci.* 2017; 23: 904–913.

[18] Hanssens LS, Duchateau J, Casimir GJ, CFTR protein: Not just a chloride channel?. *Cells* 2021; **10**: 2844.

[19] Quinton PM. The neglected ion: HCO3− *Nat Med.* 2001; **7**: 92–293.

[20] Kunzelmann K, Schreiber R, Hadorn HB, Bicarbonate in cystic fibrosis. *J Cyst Fibros.* 2017; **16**: 653–662.

[21] Bijman J, Quinton PM, Lactate and bicarbonate uptake in the sweat duct of cystic fibrosis and normal subjects. *Pediatr Res.* 1987; **21**: 79–82.

[22] Nori de Macedo A, Teo K, Mente A, et al., A robust method for iodine status determination in epidemiological studies by capillary electrophoresis. *Anal Chem.* 2014; **286**: 10010–10015.

[23] Saoi M, Percival M, Nemr C, et al., Characterization of the human skeletal muscle metabolome for elucidating the mechanisms of bicarbonate ingestion on strenuous interval exercise. *Anal Chem.* 2019; **91**: 4709–4718.

[24] Jadoon S, Karim S, Akram MR, et al., Recent developments in sweat analysis and its applications. *Int J Anal Chem.* 2015; **2015**: 164974.

[25] Herrmann F, Mandol L, Studies of pH of sweat produced by different forms of stimulation. *J Invest Dermatol.* 1955; **24**: 225–246.

[26] Kleyn M, Korzeniewski S, Grigorescu V, et al., Predictors of insufficient sweat production during confirmatory testing for cystic fibrosis. *Pediatr Pulmonol.* 2011; **46**: 23–30.

[27] Kubán P, Greguš M, Pokojová E, et al., Double opposite end injection capillary electrophoresis with contactless conductometric detection for simultaneous determination of chloride, sodium and potassium in cystic fibrosis diagnosis. *J Chromatogr A.* 2014; **1358**: 293–298.

[28] Ďurč P, Foret F, Pokojová E, et al., New approach for cystic fibrosis diagnosis based on chloride/potassium ratio analyzed in non-invasively obtained skin-wipe sweat samples by capillary electrophoresis with contactless conductometric detection. *Anal Bioanal Chem.* 2017 **409**:

3507–3514.

[28] Marvelli A, Campi B, Mergni G, et al., Sweat chloride assay by inductively coupled plasma mass spectrometry: A confirmation test for cystic fibrosis diagnosis. *Anal Bioanal Chem.* 2020; **412**: 6909–6916.

[30] Park HW, Nam JH, Kim JY, et al., Dynamic regulation of CFTR bicarbonate permeability by [Cl-]i and its role in pancreatic bicarbonate secretion. *Gastroenterology* 2010; **139**: 620–631.

Chapter V

New Sweat Biomarkers for Diagnostic Testing of Screen-Positive Cystic Fibrosis Infants

Stellena Mathiaparanam, Adriana Nori de Macedo, Katherine Keenan, Tanja Gonska, Linda Pedder, Stephen Hill, and Philip Britz-McKibbin

J. Cyst. Fibros, **2022** (Under Review)

S.M. performed all the experiments including sample preparation, data acquisition using CE-iUV, data processing, interpretation, statistical analysis, and wrote the initial draft for publication. A.N.M. performed preliminary characterization. S.H., L.P., T.G., and K.K. were involved in participant recruitment, collection of patient metadata, and coordination of sweat samples. P.B.M. supervised the study, and with all other co-authors, contributed to the final version of the manuscript.

Chapter V: New Sweat Biomarkers for Diagnostic Testing of Screen-Positive Cystic Fibrosis Infants

5.1 Abstract

Cystic fibrosis (CF) is an autosomal recessive disease that benefits from early detection with newborn screening. However, ambiguous diagnoses can occur with confirmatory sweat chloride testing, such as cystic fibrosis-screen positive inconclusive diagnosis (CF-SPID) cases. We characterized the sweat metabolome from screen-positive CF infants to independently validate putative biomarkers that may complement sweat chloride. Multisegment injection-capillary electrophoresis-electrospray ionization-mass spectrometry (MSI-CE-ESI-MS) method was used for the analysis of polar/ionic sweat metabolites in screen-positive CF (n = 25), non-CF (n = 53), and ambiguous CF-SPID (n = 17) infants that included three late-onset CF diagnoses with follow-up testing. Complementary multivariate and univariate statistical methods were used to identify phenotypic biomarkers that differentiate CF from non-CF controls, and potentially provide prognostic insight into CF-SPID trajectories. Sum normalization of sweat metabolite responses reduced between-subject variance due to different sweat rates following standardized pilocarpine-stimulated iontophoresis. Overall, 36 sweat metabolites were measured consistently (CV < 35%) with adequate frequency (> 50%) from screen-positive CF infants when using MSI-CE-MS under positive and negative ion mode conditions. Top-ranked sweat metabolites that differentiated CF from non-CF infants after covariate adjustment (sex, age at sweat collection, and birth weight) with Bonferroni correction, included aspartic acid, glutamine, pilocarpic acid, and oxoproline (effect size > 0.200; adjusted $p < 1.0\times10^{-4}$). Ratiometric sweat biomarkers provided better discriminatory performance when using ROC curves, including aspartic:pilocarpic acid (AUC = 0.896, $p = 2.11\times10^{-8}$) and glutamine:pilocarpic acid (AUC = 0.893, $p = 3.42\times10^{-8}$). The sweat chloride levels for all screen-positive CF cases at baseline collection was also compared with aspartic

acid:pilocarpic acid. A panel of endogenous sweat metabolites and exogenous drug responses to sweat stimulation can provide new insights into CF pathophysiology while augmenting CF diagnosis and potentially late-onset CF prognosis of ambiguous CF-SPID cases than sweat chloride alone.

5.2 Introduction

Cystic fibrosis (CF) is a life-shortening and pleiotropic disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) channel that regulates the transport of chloride and other anions.[1,2] Affected children with CF suffer from malnutrition, stunted growth along with recurrent bacterial infections and progressive pulmonary inflammation if left untreated.[3] Newborn screening (NBS) for CF in North America often relies on a two-tiered screening strategy for immunoreactive trypsinogen (IRT) from dried blood spots, followed by detection of a panel of disease-causing *CFTR* mutations, such as delF508.[4] There is growing interest in incorporating a third-tier screen based on next generation sequencing for more comprehensive screening of more than 2000 *CFTR gene* mutations.[5,6] Irrespective of the exact screening algorithm, all presumptive screen-positive CF infants require sweat chloride testing via pilocarpine-stimulated iontophoresis, which remains the gold standard for confirmatory CF diagnosis.[4,7] For infants under 3 months, sweat chloride concentrations ≥ 60 mmol/L is indicative of CF, whereas unaffected carriers or false-positives (*i.e.*, non-CF) have sweat chloride < 30 mmol/L. However, a minor fraction of sweat tests for screen-positive CF infants remain indeterminate having intermediate sweat concentrations (30-59 mmol/L), referred to as cystic fibrosis-screen positive inconclusive diagnosis (CF-SPID).[8] Although NBS enables pre-symptomatic diagnosis and early treatment to improve health outcomes for CF patients,[9] repeat sweat chloride testing and follow-up monitoring for ambiguous CF-SPID cases contributes to parental anxiety and CF clinic workloads despite evolving guidelines for patient management.[10]

The term “CFTR-related metabolic syndrome” was proposed for CF-SPID infants; however, it has not been universally accepted since infants are asymptomatic and otherwise healthy.[11] Indeed, prospective studies have demonstrated that about 11% of CF-SPID cases are subsequently diagnosed with CF, but often do not require therapeutic interventions, such as pancreatic enzyme

and fat-soluble vitamin supplementation.[12] CF-SPID infants who are not diagnosed with CF may also develop a CFTR-related disease (CFTR-RD).[13] Gene sequencing within NBS may allow for early identification of CFTR-RD from CF-SPID neonates,[14] which is also used to diagnose infants who progress onto CF when also supported by two pathological sweat chloride results.[15] Recently, Gonska *et al.*[16] proposed that an initial sweat chloride of ≥ 40 mmol/L was an indicator for late onset CF following a longitudinal study of CF-SPID cases over an average of 7.7 years. Due to the ongoing diagnostic dilemmas in NBS for CF,[14] the discovery of novel diagnostic and/or prognostic biomarkers of CF that augment sweat chloride testing is urgently needed notably for resolving ambiguous cases.[15] This may also reduce false-positive and false-negative sweat test results,[17] since elevated sweat chloride may not always be diagnostic of CF.[18]

Herein, we describe a comprehensive analysis of the sweat metabolome from screen-positive CF infants to differentiate CF, non-CF and CF-SPID cases when using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS). CE is an ideal microseparation platform for the analysis of volume-restricted infant sweat specimens (~ 0.5 μL) that may reduce testing failure rates for sweat chloride in premature or low birth weight neonates due to insufficient volume.[19] This work represents a follow-up replication of our previous sweat metabolome study[20] that first reported a panel of metabolic signatures associated with CF infants besides chloride[21] and other electrolytes, such as sodium and thiocyanate, but not bicarbonate.[19] In this work, we confirmed a panel of robust sweat biomarkers associated with chloride that differentiate true CF from non-CF controls after covariate adjustments. Sum normalization and sweat metabolite ratios were investigated to improve discrimination by correcting for variations in sweat rate between infants. Also, the potential of certain sweat metabolites to prognose late onset CF or non-CF among a subset of CF-SPID cases from their initial ambiguous sweat collection is preliminarily explored to further augment sweat chloride testing.

5.3 Methods

5.3.1 Study Design and Sweat Collection from Screen-Positive CF Infants

All residual sweat samples were analyzed from screen-positive CF infants following pilocarpine-stimulated sweat chloride collected at two pediatric CF clinics located at McMaster Children's Hospital (Hamilton, ON) and The Hospital for Sick Children (Toronto, ON). This study was approved by the Hamilton Integrated Research Ethics Board and The Hospital for Sick Children's Research Ethics Board. To protect the privacy of the patients, sweat specimens were de-identified and relabeled with a research ID. Sweat collection was performed using a Model 3700 Macroduct Advanced® Sweat Capillary Collection System (ELITechGroup Wescore Biomedical Systems, Logan, UT, USA). After proper sanitation protocols were performed, electrodes and agar gel discs containing 0.5% w/w pilocarpine nitrate were placed on the infant's forearm for sweat stimulation. A low current of 1.5 mA was applied for 5 min to deliver pilocarpine to the sweat gland via iontophoresis. Sweat was collected for 30 min using a coiled microbore tube containing a blue dye to facilitate visualization of the sample volume. Samples were transferred to a 0.5 mL centrifuge tube, and then stored in a fridge (+4 °C) after taking an aliquot for analysis via a coulometric titrator using a FDA/Health Canada-approved Model 3400 ChloroChek® Chloridometer® (ELITechGroup Wescore Biomedical Systems). The commercial chloridometer required a minimum of 10 μL of sweat and analyses were performed in duplicate. Residual sweat samples were then stored frozen at -80 °C prior to thawing on ice and analysis using MSI-CE-MS. The study consisted of a total of 95 screen-positive CF infants from newborn screening Ontario (NSO) were recruited in this study as summarized in **Table 5.1** with the exception of a 5 year old child with intermediate sweat chloride levels who was later diagnosed with CF. The study cohort consisted of 53 unaffected non-CF (< 30 mmol/L), 25

Table 5.1. Clinical characteristics of screen-positive CF infants included in this study.

Variable	Non-CF (*n*=53)	CF (*n*=25)	CF-SPID Unaffected (*n*=14)[a]	CF-SPID Late CF (*n*=3)[b]
Female:Male	25:28	9:16	7:7	3:0
Age of collection (days)	22 ± 8	20 ± 22	18 ± 13	358 ± -
Birth weight (g)	3347 ± 1093	3042 ± 1010	3480 ± 614	3092 ± 1314
Gestational age (weeks)[c]	40 ± 1.2	39 ± 1.3	40 ± 1.2	36 ± 5.3
Chloride (mM)	13 ± 8.5	94 ± 20	42 ± 12	38 ± -
IRT (µg/L)[d]	62 ± 34	152 ± 75	82 ± 106	94 ± -
CFTR genotype:				
0 mutations	2			
1 mutation: delF508/null	14		8	
1 mutation: non-delF508/null[e]	5		2	
2 mutations: delF508/delF508		9		1
2 mutations: delF508/non-delF508[f]	19	13	1	2
2 mutations: non-delF508/non-delF508[g]	13	3	1	
Collection site:				
McMaster	34	24	12	2
Sick Kids	19	1	2	1
Pancreatic status:[h]				
Pancreatic sufficient	-	4	2	1
Moderate pancreatic insufficient		2		
Pancreatic insufficient	-	17	-	1

All data reported as mean ± 1s, except age of collection, chloride, and IRT measurements (median ± IQR).
[a] No genotype was reported for n=1 and n=1 had no detected pathogenic variants (7T/7T both alleles).
[b] IQR not determined due to sample size consisting of n=3. Alternate reports as mean ± 1s: age of collection (782 ± 1047 days), sweat chloride level (40 ± 4.9 mM), and IRT level (94 ± 14 µg/g).
[c] 21 samples for non-CF, 4 samples for CF, and 1 sample for CF-SPID late onset CF were excluded due to insufficiency.
[d] For CF-SPID late onset CF cases, 1 sample was excluded due to insufficient information.
[e] For non-CF cases, CFTR disease-causing when combined with another pathogenic variant: n=1, 621+1G>T/null (7T/7T); n=1, c.1521_1523delCTT (7T/9T). Variants of unknown significance: n=4, non-delF508/null. For CF-SPID late unaffected cases, variants under evaluation: n=1, 0/C.2657+2_2657+3insA, c.2562T>G, c.4389G>A.
[f] For non-CF cases, CFTR disease-causing when combined with another pathogenic variant: n=17, delF508/-. Variants of varying clinical consequences: n=2, delF508/R117H (7T/9T).
For CF-SPID late onset affected cases, variants of varying clinical consequences: n=2, delF508/R117H (7T/9T) and delF508/5T9T (5T on opposite alleles).
For CF-SPID unaffected cases, variants of unknown significance: n=1, delF508/non-delF508.
For CF cases, CFTR disease-causing variants: n=4, delF508/G551D; n=1, delF508/G1244E; n=1, delF508/G85E (7T/9T); n=1, delF508/G542X; n=1, delF508/Lys710 (8T-track); n=1, delF508/3199del6. Variants under evaluation: n=2, delF508/C.2657+2_2657+3insA; n=1, delF508/C.2562T>G, C.2657+2_2657+3insA, C.4389G>A. Variants of unknown significance: n=1, delF508/non-delF508.
[g] For non-CF cases, CFTR disease-causing when combined with another pathogenic variant: n=1, delI507/-; n=1, 621+1G>T/-; n=1, N1303K/-; n=1, Y1092X/-. Variants of varying clinical consequences: n=2, R117H/-. Variants under evaluation: n=1, 4 variants - C.132-329A>G, C.256T>G, C.4389G>A, C.1584G>A/-. Variants of unknown significance: n=6, -/-.
For CF-SPID unaffected cases, variants of unknown significance: n=1, non-delF508/non-delF508.
For CF cases, CFTR disease-causing variants: n=2, delI507/621+1G->T and Y569D/Y569D (7T/7T). Variants of varying clinical consequence: n=1, R117H/G551D (5T/7T).
[h] Pancreatic sufficiency (> 200 µg/g), moderate pancreatic insufficiency (100-200 µg/g), and pancreatic insufficiency (< 100 µg/g) based on fecal elastase measurements. Pancreatic status was not assessed for non-CF and non-CF late onset unaffected infants. Fecal elastase data not determined for n=1 CF-SPID late onset CF and n=2 CF affected infants.

CF (≥ 60 mmol/L), and 17 CF-SPID (30-59 mmol/L) infants, who had two or more repeated sweat testing with follow-up monitoring for more than one year. A subset of 3 CF-SPID cases later progressed to a late onset CF diagnosis. However, two of these cases included an infant (11 mo) who had an increasing sweat chloride trajectory close to the CF cut-off (54 mmol/L), as well as an older child (5 years) with an initial intermediate sweat chloride result that was later diagnosed as late onset CF (≥ 60 mM). Sweat chloride trajectories for all the CF-SPID cases are depicted in **Figure S5.1**.

5.3.2 Chemicals, Reagents, and Analysis of Sweat Samples from Screen-Positive CF Patients

All the reagents were purchased from Sigma-Aldrich Inc. (St. Louis, MO, USA) for preparation of buffer and stock solutions. All frozen sweat specimens were thawed on ice and then simply diluted either by 5- or 10-fold in ultra-pure grade water containing internal and recovery standards, with a final concentration of 10 μM for 3-chlorotyrosine (Cl-Tyr, positive ion mode) and naphthalene sulfonic acid (NMS, negative ion mode) followed by sample mixing using a vortex for 30 s prior to MSI-CE-MS analysis. Due to the small residual volumes of sweat, pipette tip inserts were adapted[22] and further modified for the analysis of as low as 0.5 μL of sweat from CF infants after dilution as depicted in **Figure S5.2.** Rigorous validation of the pipette tip inserts was assessed on standard mixtures consisting of volumes 5, 2.5, and 1 μL in repeated technical replicates (n = 10) with good reproducibility as reflected by an average coefficient of variance (CV) of 3.8%. A pooled quality control (QC) sweat sample was prepared from 15 screen-positive infants, including 5 non-CF, 5 CF, and 5 CF-SPID infants, which was used for nontargeted screening of the sweat metabolome, and assessment of technical precision throughout the study.

5.3.3 Instrumentation and MSI-CE-ESI-MS Operating Conditions

All CE separations were conducted using an Agilent 7100 capillary electrophoresis (CE) instrument coupled to a time-of-flight (TOF) mass spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA). Fused-silica capillaries having outer polyimide coating were purchased from Polymicro Technologies Inc. (Phoenix, AZ, USA) and the polyamide coating covering 2 cm of both ends of the capillary were burned off to reduce sample carry-over during repeated serial injections in MSI-CE-MS. The internal diameter of the fused-silica capillaries were 50 μm internal diameter with a total capillary length of 135 cm and a separation voltage of 30 kV was used for both ionization modes, respectively. For positive ion mode, the background electrolyte (BGE) consisted of 1 mmol/L formic acid and 15% vol acetonitrile with a final pH of 1.80 for the detection of cationic species. MSI-CE-MS used a serial injection format comprising 13 sweat samples in a single run, which each sample alternately injected hydrodynamically for 5 s followed by a 75 s electrokinetic injection with BGE as illustrated in **Figure S5.3**. [23,24] The sheath liquid composition for positive ion mode consisted of 60:40 methanol:water for optimal spray stabilization to achieve adequate sensitivity. For negative ion mode, the BGE was composed of 50 mmol/L ammonium bicarbonate, with a final pH of 8.5 (adjusted with 10% vol ammonium hydroxide) for the detection of anionic/weakly acidic metabolites. In this case, the sheath liquid composition consisted of 70:30 methanol:water without any electrolyte additive. For both ionization modes, a temporal signal pattern (1:2 1:1 2:1) was implemented on the samples in order to filter out any spurious and/or artifact signals and reduce for false discoveries when performing untargeted metabolite profiling.[25] For instance, a typical run would comprise a randomized injection of 6 pairs of sweat samples diluted in a specific temporal pattern (1:2, 1:1, 2:1) together with a QC sample as shown in **Figure S5.4**.

5.3.4 Statistical Analyses

The extracted-ion chromatograms (EIC) were collected using Agilent MassHunter Workstation Software, Qualitative Analysis B.06.00 and the raw data was extracted using Agilent MassHunter Workstation Software, Qualitative Analysis 10.0 (Santa Clara, CA, USA). The features were extracted in profile format using 10 ppm mass window and labeled according to their *m/z*, RMT, and ionization modes (p = ESI+ and n = ESI-). Peaks were integrated after smoothing using quadratic/cubic Savitzky-Golay, 7 points. Migration times, peak areas, heights, widths, full width at half maximum (FWHM), and signal-to-noise ratio (SNR) were transferred to Microsoft Excel for calculation of relative integrated peak area (RPA), relative migration time (RMT), coefficient of variation (CV), as well as conversion into a data matrix comma-separated values (CSV) format for multivariate analyses. Data cleanup and normalization, partial least squares-discriminant analysis (PLS-DA): 2D scores plot, and derived variables importance plot (VIP) scores ranking were created on Metaboanalyst 5.0.[26] Univariate statistical analyses, including student's t-test with unequal variance, receiver operating characteristic (ROC) curves, ANOVA, and ANCOVA with a general linear model to adjust for confounding variables (*e.g.*, age of sweat collection, sex, and birth weight) were performed using SPSS 23.0 statistical software (IBM SPSS, Chicago, IL, USA). ROC curve visualizations and related boxplots were plotted using MedCalc 12.5 statistical software (Ostend, Belgium). Representative electropherograms and recovery control charts were generated using Igor Pro 5.04B (Wavemetrics Inc, Lake Oswego, OR, USA).

5.4 Results

Table 5.1 highlights that there were no significant differences in gestational age (~ 40 weeks), birth weight (~ 3200 g) nor age at time of sweat collection (~ 20 days) in our cohort of screen-positive CF infants. However, sweat chloride confirmed CF infants (≥ 60 mM) had a higher incidence of severe or moderate

pancreatic deficiency (i.e., fecal elastase), and were frequently homozygous or heterozygous with F508 as compared to non-CF or CF-SPID cases. All sweat samples from screen-positive CF infants were analyzed in duplicate in randomized order when using MSI-CE-MS with full-scan data acquisition under positive and negative ion mode conditions. The technical precision of our method was assessed by the analysis of a recovery standard (F-Phe, 10 μM) added to all diluted sweat samples which had a mean CV of 10.4% under both ionization modes overall as shown the control charts in **Figure S5.5**. Moreover, pooled QC samples repeatedly analyzed in each MSI-CE-MS run also demonstrated acceptable technical precision with a median CV of 23% (n = 20), as shown in **Figure S6.6** for a 2D scores plot from PCA. Overall, a total of 36 sweat metabolites were consistently detected (> 50% of all samples) with adequate precision based on repeated analysis of pooled QC samples (mean CV < 25%). All sweat metabolites were annotated based on their accurate mass, relative migration time, ionization mode for detection (*m/z:*RMT:p or n), as well as most likely molecular formula as summarized in **Table S5.1**. Most sweat metabolites were also identified with high confidence after spiking with authentic standards (*i.e.,* co-migration). Normalization of sweat metabolome data based on the sum response of all 36 sweat metabolites by MSI-CE-MS (i.e., sum-normalization) was used to correct for hydration status and sweat rate variations when using stimulated sweat collection via pilocarpine iontophoresis. However, highly abundant endogenous metabolites and exogenous drug metabolites in sweat contributed to most of the sum response in our study, including lactic acid, pilocarpine, oxoproline, etc.

Figure 5.1A depicts a supervised multivariate analysis of the sweat metabolome based on PLS-DA, which was used to differentiate CF from non-CF infants following *log*-transformation and autoscaling of sum-normalized data. A variable importance in projection (VIP) was used to rank order sweat metabolites

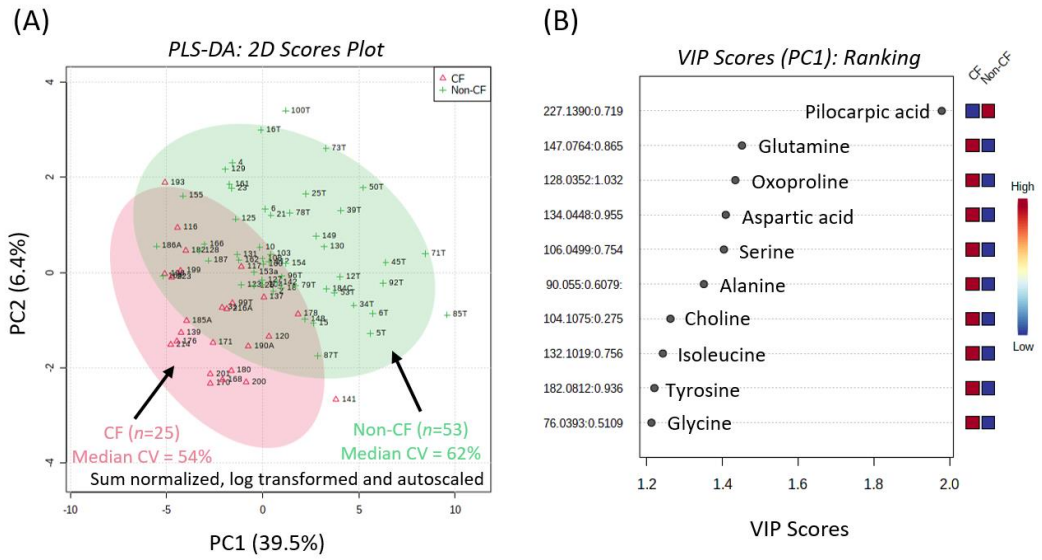


Figure 5.1. Sum-normalized multivariate analysis, showing (A) PLS-DA: 2D scores plots with overall median biological variance CV of 54 % for screen-positive CF infants (n = 25) and 62% for unaffected infants (n = 53). (B) VIP scores plot ranking pilocarpic acid and glutamine as the most significant metabolites, similar to not sum-normalized results. (C) Boxplots displaying the significant differences between screen-positive CF and non-CF infants based on pilocarpic acid and glutamine levels.

associated with CF as compared to non-CF infants as shown in **Figure 5.1B**, including pilocarpic acid, glutamine, oxoproline and aspartic acid. Sum normalization had the effect of reducing biological variance (from a mean CV of 80 to 58%) with only minor changes in the rank ordering of CF-related sweat metabolites as shown in **Figure S5.7**. **Table 5.2** summarizes 12 top-ranked sum-normalized sweat metabolites that discriminate between CF and non-CF (*i.e.*, unaffected carrier or false-positive) that also satisfied a Bonferroni correction (adjusted $p < 0.05$) with covariate adjustments for age at sweat collection, sex, and gestational weight. An unadjusted model without covariate adjustment resulted in only slightly varied outcomes for sweat metabolites as shown in **Table S5.2**. Similar to PLS-DA results, aspartic acid, glutamine and oxoproline were significantly elevated in sweat excreted from CF as compared to non-CF (effect

Table 5.2. Top-ranked sweat metabolites that differentiate CF from non-CF infants based on their *log*-transformed, sum-normalized responses after adjusting for sex, birth weight and age at sweat collection.[a]

***m/z*:RMT:mode**	**Compound ID**	***p*-value**[d]	**Effect size**[e]	**Median FC**	***r*** [f]
-	Aspartic acid/Pilocarpic acid	8.29×10^{-8}	0.364	6.12	0.533 (6.18×10^{-7})
-	Glutamine/Pilocarpic acid	2.22×10^{-7}	0.336	10.2	0.484 (1.09×10^{-5})
134.0448:0.9558:p	Aspartic acid[b]	1.24×10^{-5}	0.253	1.77	0.359 (1.36×10^{-3})
227.1390:0.7199:p	Pilocarpic acid	1.72×10^{-5}	0.243	3.44	-0.484 (7.31×10^{-6})
147.0764:0.8654:p	Glutamine[c]	2.61×10^{-5}	0.243	1.91	0.260 (2.41×10^{-2})
128.0352:1.0323:n	Oxoproline[b]	1.15×10^{-4}	0.203	2.24	0.481 (9.54×10^{-6})
104.1075:0.2750:p	Choline	3.09×10^{-4}	0.178	1.55	0.095 (0.410)
132.1019:0.7565:p	Isoleucine	4.40×10^{-4}	0.170	1.44	0.390 (4.20×10^{-4})
182.0812:0.9364:p	Tyrosine	4.24×10^{-4}	0.170	1.50	0.404 (2.47×10^{-4})
106.0499:0.7544:p	Serine	5.21×10^{-4}	0.166	1.46	0.430 (8.41×10^{-5})
90.055:0.6079:p	Alanine	5.58×10^{-4}	0.164	1.79	0.378 (6.42×10^{-4})
76.0393:0.5109:p	Glycine	1.42×10^{-3}	0.142	1.27	0.358 (1.28×10^{-3})
148.0604:0.8869:p	Glutamic acid	1.65×10^{-3}	0.138	1.44	0.254 (2.50×10^{-2})
156.0768:0.3622:p	Histidine	5.29×10^{-3}	0.108	2.20	0.286 (1.11×10^{-2})

For reference: log-transformed sweat chloride was $p = 3.02\times10^{-32}$, effect size = 0.88, and median FC = 7.23.
[a] *CF affected (n=23) and unaffected (n=49) infants.*
[b] *CF affected (n=23) and unaffected (n=48) infants.*
[c] *CF affected (n=23) and unaffected (n=46) infants.*
[d] *Univariate general linear model analysis performed on log-transformed data and adjusted for sex, birth weight, and age of collection when using a Bonferroni correction applied at the 0.05 level.*
[e] *Effect size calculated from partial Eta squared.*
[f] *r corresponds to Spearman rank order correlation coefficient with p-values with chloride as a reference (1.00).*

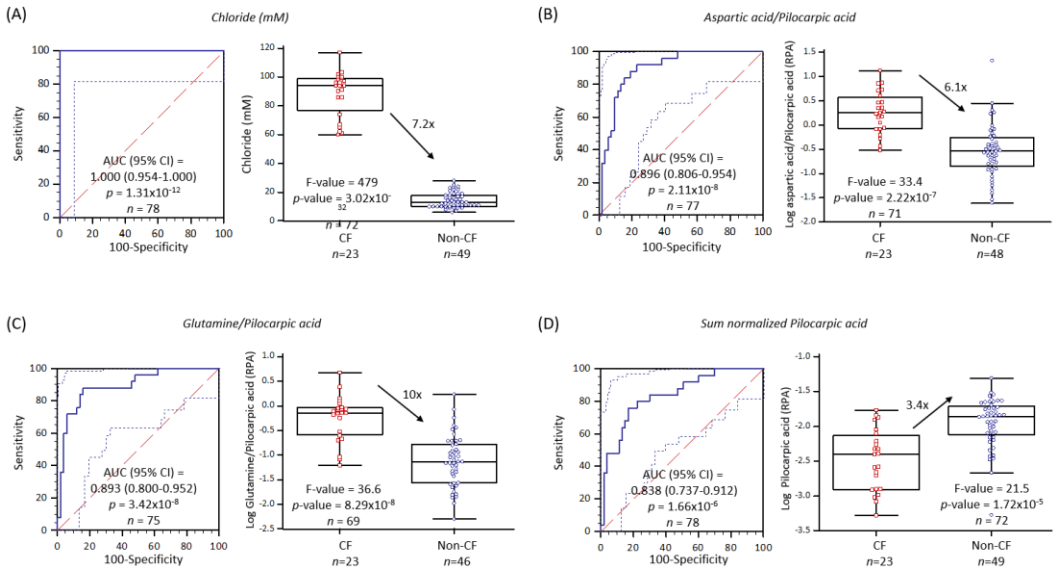


Figure 5.2. ROC curves and boxplots of top-ranked significant biomarkers that discriminate screen-positive CF affected from unaffected infants in reference to gold standard, sweat chloride levels. (A) Sweat chloride demonstrated perfect discrimination between both groups (AUC = 1.00, p = 1.31x10^{-12}, n = 78, and median FC = 7.2x). Excellent discriminators included ratiometric biomarkers: (B) aspartic acid/pilocarpic acid (AUC = 0.896, p = 2.11x10^{-8}, n = 77, and median FC = 6.1x) and glutamine/pilocarpic acid (AUC = 0.893, p = 3.42x10^{-8}, n = 75, and median FC = 10x). (C) Pilocarpic acid (exogenous metabolite) alone showed to be an excellent diagnostic biomarker that goes beyond impaired chloride transport (AUC = 0.838, p = 1.66x10^{-6}, n = 78, and median FC = 3.4x).

sizes > 0.200; $p < 1 \times 10^{-4}$), whereas pilocarpic acid had a median 3.4 fold-change lower abundance in CF than non-CF infants. In addition, most sweat metabolites associated with CF also had moderate positive correlations to chloride measured by a commercial chloridometer (r = 0.254 to 0.430), with the exception of pilocarpic acid that had a stronger inverse association (r = -0.484). We also found that the ratio of aspartic acid to pilocarpic acid, and glutamine to pilocarpic acid was able to further enhance the effect sizes (> 0.340) and strength of correlation to chloride (r > 0.480) while self-correcting for hydration status and/or sweat rate without the need for sum normalization. **Figure 5.2** depicts ROC curves for the ratio of aspartic acid to pilocarpic acid (AUC = 0.896, p = 2.11×10^{-8}, n = 77), and

glutamine to pilocarpic acid (AUC = 0.893, $p = 3.42\times10^{-8}$, $n = 75$), which provided excellent discrimination between CF and non-CF infants. In contrast, sum-normalized sweat aspartic acid and glutamine alone proved to have lower classification performance using ROC curves (AUC ~ 0.770) as shown in **Figure S5.8**.

5.4.1 Qualitative Perspective on Potential Prognostic Biomarker Discovery in CF-SPID Cases

Due to the very few CF-SPID cases who subsequently progressed with a late diagnosis with CF ($n = 3$) in this study, the potential prognostic performance of the top-ranked ratiometric biomarker for CF (*i.e.*, aspartic:pilocarpic acid) was compared to chloride based on the initial sweat collected in overall CF-SPID infants. **Figure 5.3A** outlines sweat chloride insignificantly differentiating between CF-SPID late onset CF *vs.* CF-SPID late onset non-CF or remain CF-SPID cases ($p = 0.641$, $n = 17$), where **Figure 5.3B** shows similar results using the top-ranked ratiometric biomarker aspartic acid to pilocarpic acid ($p = 0.0847$, $n = 17$).

5.5 Discussion

We have demonstrated a high-throughput method to reliably identify novel biomarkers in volume-limited sweat specimens (≥ 0.5 μL) from presumptive CF infants. We have demonstrated the clinical utility of both endogenous and exogenous metabolites serving as diagnostic biomarkers of CF as compared to unaffected carriers that may also potentially better characterize ambiguous CF-SPID cases in conjunction with sweat chloride. Overall, the ratio of aspartic acid to pilocarpic acid, and glutamine to pilocarpic acid were sweat metabolites that

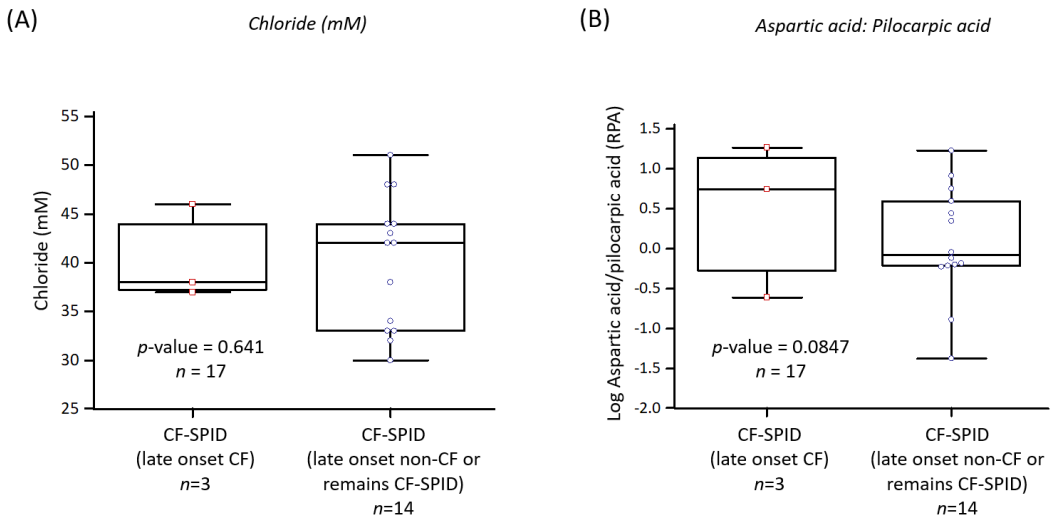


Figure 5.3. Preliminary boxplots displaying the trends of CF-SPID late onset CF *vs.* CF-SPID late onset non-CF or remain CF-SPID cases in **(A)** sweat chloride levels: (p = 0.641, n = 17) and **(B)** sweat ratiometric comparison Aspartic:pilocarpic acid: (p = 0.0847, n = 17).

optimally differentiated CF from non-CF infants as compared to sum-normalized pilocarpine, aspartic acid or glutamine alone. Our previous work also showed that sweat pilocarpic acid and glutamine were significant metabolites in differentiating infant CF sub-groups.[20] The human enzyme paraoxonase is responsible for metabolizing the sweat stimulating and muscarinic cholinergic agent, pilocarpine into pilocarpic acid as a hydrolysis byproduct of this arylesterase/lactonase enzyme involved in xenobiotic detoxification and bacterial biofilm regulation via hydrolysis of quorum sensing lipids.[20,27] Therefore, CF infants are likely deficient in this enzyme and more prone to intoxification and recurrent bacterial infections as reflected by their lower enzyme activity in the sweat gland to form pilocarpic acid following standardized pilocarpine iontophoresis. As further supporting evidence, a reported linkage was determined between the paraoxonase system and CF dating back to 1985.[28] Our work also demonstrated that the median fraction of pilocarpine to pilocarpic acid in the gel discs (*i.e.,* background) was only 0.3% proving likely *in-vivo* hydrolysis within the sweat gland of screen-positive CF

infants following drug delivery.[20] Interestingly, pilocarpic acid had the strongest negative correlation with chloride that reflects impaired CFTR re-uptake from the sweat gland in CF patients.

Our work also demonstrated that glutamine was a significant endogenous sweat metabolite to differentiate between non-CF and CF infants. Although glutamine is a conditionally essential amino acid that is the most abundant amino acid in circulation, it has been reported to become deficient with malnutrition and inflammatory stress in CF patients.[29,30] DiBattista *et al.*[31] reported that glutamine was lower in dried blood spot extracts collected shortly after birth from CF as compared to non-CF neonates. Also, neutrophil glutamine depletion has been implicated in CF children with disease-causing *CFTR* mutations that may impair innate immune responses to infections and promote respiratory inflammation.[32] However, a randomized placebo-controlled clinical trial involving glutamine supplementation (21 g/day for 8 weeks) did not reduce biomarkers of pulmonary inflammation in CF patients.[33]Aspartic acid is considered a non-essential amino acid[34] and to the best of our knowledge has not been reported in CF apart from our previous sweat metabolome study.[20] Oxoproline in sweat was also determined to differentiate CF from non-CF infants and also known as a biomarker of glutathione depletion.[35] In fact, CFTR has been reported to regulate the transport of reduced glutathione that is impaired and may play an important role in CF pathology based on mucous build-up at surface of epithelial cell lining.[36] Lastly, the use of ratiometric biomarker analyses have been shown to be greatly effective in clinical/diagnostic research, as it compensates for intra- and inter-individual variations caused by hydration status.[37] As demonstrated in our work, the ratio between aspartic acid to pilocarpic acid and glutamine to pilocarpic acid were promising ratiometric biomarkers in differentiating true CF cases from unaffected screen-positive CF infants after covariate adjustments. Indeed, the trends between overall CF-SPID cases based on sweat chloride and aspartic:pilocarpic acid ratiometric levels were similar. However, both biomarkers were not able to

discriminate disease progression from baseline sweat measurement. As a result, due to the modest sample size in overall CF-SPID cases, these preliminary results may provide incentive to perform prospective studies. Nevertheless, its clinical utility to improve CF-SPID prognosis to further augment sweat chloride testing warrants further investigation given the limited cohort size in our study.

Limitations in this study includes the modest sample size comprising CF-SPID cases with only 3 out of 17 infants/children progressing to be later diagnosed with CF. As a result, prospective studies are needed to validate the clinical utility of prognostic biomarkers of late onset CF among CF-SPID cases that can be replicated in multiple centres across Canada. This study also did not include CFTR-RD cases that further contribute to diagnostic dilemmas in CF, which could further be explored based on their characteristic sweat metabolic phenotype. Lastly, there is a need for reference intervals for novel sweat metabolites reported in concentrations, such as pilocarpic acid, which can then establish optimal cut-off intervals independent from sweat chloride diagnostic criteria for CF and non-CF. Overall, this work demonstrated that metabolic signatures of CF can provide new insights into disease pathophysiology in CF, as well as novel therapeutic targets for treatment, such as small molecules that can function as allosteric activators of paraoxanase.

5.6 Acknowledgements

The authors wish to thank the families of screen-positive CF children, and staff from CF clinics at McMaster Children's Hospital and The Hospital for Sick Children.

5.7 References

(1) Linsdell, P.; Evagelidis, A.; Hanrahan, J. W. Molecular Determinants of Anion Selectivity in the Cystic Fibrosis Transmembrane Conductance Regulator Chloride Channel Pore. *Biophys. J.* **2000**, *78* (6), 2973–2982.

(2) Linsdell, P. Anion Conductance Selectivity Mechanism of the CFTR Chloride Channel. *Biochim. Biophys. Acta - Biomembr.* **2016**, *1858* (4), 740–747.

(3) Culhane, S.; George, C.; Pearo, B.; Spoede, E. Malnutrition in Cystic Fibrosis: A Review. *Nutr. Clin. Pract.* **2013**, *28* (6), 676–683.

(4) Mak, D. Y. F.; Sykes, J.; Stephenson, A. L.; Lands, L. C. The Benefits of Newborn Screening for Cystic Fibrosis: The Canadian Experience. *J. Cyst. Fibros.* **2016**, *15* (3), 302–308.

(5) Baker, M. W.; Atkins, A. E.; Cordovado, S. K.; Hendrix, M.; Earley, M. C.; Farrell, P. M. Improving Newborn Screening for Cystic Fibrosis Using Next-Generation Sequencing Technology: A Technical Feasibility Study. *Genet. Med.* **2016**, *18* (3), 231–238.

(6) Sicko, R. J.; Stevens, C. F.; Hughes, E. E.; Leisner, M.; Ling, H.; Saavedra-Matiz, C. A.; Caggana, M.; Kay, D. M. Validation of a Custom Next-Generation Sequencing Assay for Cystic Fibrosis Newborn Screening. *Int. J. Neonatal Screen.* **2021**, *7* (4), 1–14.

(7) Pagaduan, J. V.; Ali, M.; Dowlin, M.; Suo, L.; Ward, T.; Ruiz, F.; Devaraj, S. Revisiting Sweat Chloride Test Results Based on Recent Guidelines for Diagnosis of Cystic Fibrosis. *Pract. Lab. Med.* **2018**, *10* (October 2017), 34–37.

(8) Vermeulen, F.; Lebecque, P.; De Boeck, K.; Leal, T. Biological Variability of the Sweat Chloride in Diagnostic Sweat Tests: A Retrospective Analysis. *J. Cyst. Fibros.* **2017**, *16* (1), 30–35.

(9) McKay, K.; Wilcken, B. Newborn Screening for Cystic Fibrosis Offers an Advantage over Symptomatic Diagnosis for the Long Term Benefit of Patients: The Motion For. *Paediatr. Respir. Rev.* **2008**, *9* (4), 290–294.

(10) Barben, J.; Castellani, C.; Munck, A.; Davies, J. C.; de Winter–de Groot, K. M.; Gartner, S.; Kashirskaya, N.; Linnane, B.; Mayell, S. J.; McColley, S.; et al. Updated Guidance on the Management of Children with Cystic Fibrosis Transmembrane Conductance Regulator-Related Metabolic Syndrome/Cystic Fibrosis Screen Positive, Inconclusive Diagnosis (CRMS/CFSPID). *J. Cyst. Fibros.* **2021**, *20* (5), 810–819.

(11) Borowitz, D.; Parad, R. B.; Sharp, J. K.; Sabadosa, K. A.; Robinson, K. A.; Rock, M. J.; Farrell, P. M.; Sontag, M. K.; Rosenfeld, M.; Davis, S. D.; et al. Cystic Fibrosis Foundation Practice Guidelines for the Management of Infants with Cystic Fibrosis Transmembrane Conductance Regulator-

Related Metabolic Syndrome during the First Two Years of Life and Beyond. *J. Pediatr.* **2009**, *155* (6 SUPPL.), S106–S116.

(12) Ooi, C. Y.; Castellani, C.; Keenan, K.; Avolio, J.; Volpi, S.; Boland, M.; Kovesi, T.; Bjornson, C.; Chilvers, M. A.; Morgan, L.; et al. Inconclusive Diagnosis of Cystic Fibrosis After Newborn Screening. *Pediatrics* **2015**, *135* (6), e1377–e1385.

(13) Levy, H.; Nugent, M.; Schneck, K.; Stachiw-Hietpas, D.; Laxova, A.; Lakser, O.; Rock, M.; Dahmer, M. K.; Biller, J.; Nasr, S. Z.; et al. Refining the Continuum of CFTR-Associated Disorders in the Era of Newborn Screening. *Clin. Genet.* **2016**, *89* (5), 539–549.

(14) Castaldo, A.; Cimbalo, C.; Castaldo, R. J.; D'Antonio, M.; Scorza, M.; Salvadori, L.; Sepe, A.; Raia, V.; Tosco, A. Cystic Fibrosis-Screening Positive Inconclusive Diagnosis: Newborn Screening and Long-Term Follow-up Permits to Early Identify Patients with CFTR-Related Disorders. *Diagnostics* **2020**, *10* (8), 1–11.

(15) Terlizzi, V.; Mergni, G.; Buzzetti, R.; Centrone, C.; Zavataro, L.; Braggion, C. Cystic Fibrosis Screen Positive Inconclusive Diagnosis (CFSPID): Experience in Tuscany, Italy. *J. Cyst. Fibros.* **2019**, *18* (4), 484–490.

(16) Gonska, T.; Keenan, K.; Au, J.; Dupuis, A.; Chilvers, M. A.; Burgess, C.; Bjornson, C.; Fairservice, L.; Brusky, J.; Kherani, T.; et al. Outcomes of Cystic Fibrosis Screening-Positive Infants With Inconclusive Diagnosis at School Age. *Pediatrics* **2021**, *148* (6), 1–9.

(17) Guglani, L.; Stabel, D.; Weiner, D. J. False-Positive and False-Negative Sweat Tests: Systematic Review of the Evidence. *Pediatr. Allergy, Immunol. Pulmonol.* **2015**, *28* (4), 198–211.

(18) Cimbalo, C.; Tosco, A.; Terlizzi, V.; Sepe, A.; Castaldo, A.; Salvadori, L.; Raia, V. Elevated Sweat Chloride Test: Is It Always Cystic Fibrosis? *Ital. J. Pediatr.* **2021**, *47* (1), 1–5.

(19) Mathiaparanam, S.; Nori, A.; Macedo, D.; Gill, B.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-mckibbin, P. Rapid Chloride and Bicarbonate Determination by Capillary Electrophoresis for Confirmatory Testing of Cystic Fibrosis Infants with Volume-Limited Sweat Specimens. *J. Cyst. Fibros.* **2022**, No. xxxx. https://doi.org/10.1016/j.jcf.2022.05.004.

(20) Macedo, A. N.; Mathiaparanam, S.; Brick, L.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P. The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Cent. Sci.* **2017**, *3* (8), 904–913.

(21) Gibson, L. E.; Cooke, R. E. A Test for Concentration of Electrolytes in Sweat in Cystic Fibrosis of the Pancreas Utilizing Pilocarpine by Iontophoresis. *Pediatrics* **1959**, No. March, 545–549.

(22) Šlampová, A.; Kubáň, P. Injections from Sub-ML Sample Volumes in Commercial Capillary Electrophoresis. *J. Chromatogr. A* **2017**, *1497*, 164–171.

(23) Kuehnbaum, N. L.; Kormendi, A.; Britz-Mckibbin, P. Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Anal. Chem.* **2013**, *85* (22), 10664–10669.

(24) Saoi, M.; Li, A.; McGlory, C.; Stokes, T.; von Allmen, M. T.; Phillips, S. M.; Britz-Mckibbin, P. Metabolic Perturbations from Step Reduction in Older Persons at Risk for Sarcopenia: Plasma Biomarkers of Abrupt Changes in Physical Activity. *Metabolites* **2019**, *9* (7), 1–19.

(25) Dibattista, A.; McIntosh, N.; Lamoureux, M.; Al-Dirbashi, O. Y.; Chakraborty, P.; Britz-McKibbin, P. Temporal Signal Pattern Recognition in Mass Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Anal. Chem.* **2017**, *89* (15), 8112–8121.

(26) Pang, Z.; Chong, J.; Zhou, G.; De Lima Morais, D. A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P. É.; Li, S.; Xia, J. MetaboAnalyst 5.0: Narrowing the Gap between Raw Spectra and Functional Insights. *Nucleic Acids Res.* **2021**, *49* (W1), W388–W396.

(27) Hioki, T.; Fukami, T.; Nakajima, M.; Yokoi, T. Human Paraoxonase 1 Is the Enzyme Responsible for Pilocarpine Hydrolysis. *Drug Metab. Dispos.* **2011**, *39* (8), 1345–1352.

(28) Schmiegelow, K.; Eiberg, H.; Tsui, L. -C; Buchwald, M.; Phelan, P. D.; Williamson, R.; Warwick, W.; Niebuhr, E.; Mohr, J.; Schwartz, M.; et al. Linkage between the Loci for Cystic Fibrosis and Paraoxonase. *Clin. Genet.* **1986**, *29* (5), 374–377.

(29) Zhao, H.; Dennery, P. A.; Yao, H. Metabolic Reprogramming in the Pathogenesis of Chronic Lung Diseases, Including BPD, COPD, and Pulmonary Fibrosis. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **2018**, *314* (4), L544–L554.

(30) Morris, C. R.; Hamilton-Reeves, J.; Martindale, R. G.; Sarav, M.; Ochoa Gautier, J. B. Acquired Amino Acid Deficiencies: A Focus on Arginine and Glutamine. *Nutr. Clin. Pract.* **2017**, *32* (1_suppl), 30S-47S.

(31) Dibattista, A.; McIntosh, N.; Lamoureux, M.; Al-Dirbashi, O. Y.; Chakraborty, P.; Britz-Mckibbin, P. Metabolic Signatures of Cystic Fibrosis Identified in Dried Blood Spots for Newborn Screening Without Carrier Identification. *J. Proteome Res.* **2019**, *18* (3), 841–854.

(32) D'Eufemia, P.; Finocchiaro, R.; Celli, M.; Tote, J.; Ferrucci, V.; Zambrano, A.; Troiani, P.; Quattrucci, S. Neutrophil Glutamine Deficiency in Relation to Genotype in Children with Cystic Fibrosis. *Pediatr. Res.* **2006**, *59* (1), 13–16.

(33) Forrester, D. L.; Knox, A. J.; Smyth, A. R.; Barr, H. L.; Simms, R.; Pacey, S. J.; Pavord, I. D.; Honeybourne, D.; Dewar, J.; Clayton, A.; et al. Glutamine Supplementation in Cystic Fibrosis: A Randomized Placebo-Controlled Trial. *Pediatr. Pulmonol.* **2016**, *51* (3), 253–257.

(34) Appleton, H.; Rosentrater, K. A. Sweet Dreams (Are Made of This): A Review and Perspectives on Aspartic Acid Production. *Fermentation* **2021**, *7* (49), 1–12.

(35) Shanmuganathan, M.; Sarfaraz, M. O.; Kroezen, Z.; Philbrick, H.; Poon, R.; Don-Wauchope, A.; Puglia, M.; Wishart, D.; Britz-McKibbin, P. A Cross-Platform Metabolomics Comparison Identifies Serum Metabolite Signatures of Liver Fibrosis Progression in Chronic Hepatitis C Patients. *Front. Mol. Biosci.* **2021**, *8*, 1–15.

(36) Hudson, V. M. Rethinking Cystic Fibrosis Pathology: The Critical Role of Abnormal Reduced Glutathione (GSH) Transport Caused by CFTR Mutation. *Free Radic. Biol. Med.* **2001**, *30* (12), 1440–1461.

(37) Lee, J. C.; Kim, S. J.; Hong, S.; Kim, Y. S. Diagnosis of Alzheimer's Disease Utilizing Amyloid and Tau as Fluid Biomarkers. *Exp. Mol. Med.* **2019**, *51* (5), 1–10.

5.8 Supporting Information

Table S5.1. Summary of 36 metabolites detected in sweat from screen-positive CF patients using MSI-CE-MS. Metabolites are listed according to their mass-to-charge ratio (*m/z*), RMT, ionization mode (p = ESI+, n = ESI-), molecular formula, compound ID, confidence level for identification, and chemical classification.

***m/z*:RMT:mode**	**Formula[a]**	**Compound ID**	**Confidence level**	**Classification**
76.0393:0.5109:p	$C_2H_5NO_2$	Glycine	1	Amino acid
90.0550:0.6079:p	$C_3H_7NO_2$	Alanine	1	Amino acid
104.1075:0.2750:p	$C_5H_{14}NO$	Choline	1	Quaternary ammonium salt
106.0499:0.7544:p	$C_3H_7NO_3$	Serine	1	Amino acid
114.0662:0.3612:p	$C_4H_7N_3O$	Creatinine	1	Heterocyclic compound
116.0706:0.9200:p	$C_5H_9NO_2$	Proline	1	Amino acid
118.0863:0.7339:p	$C_5H_{11}NO_2$	Valine	1	Amino acid
120.0655:0.8222:p	$C_4H_9NO_3$	Threonine	1	Amino acid
132.0655:0.5917:p	$C_5H_9NO_3$	Hydroxyproline	1	Amino acid derivative
132.0766:0.5917:p	$C_4H_9N_3O_2$	Creatine	1	Amino acid
132.1019:0.7565:p	$C_6H_{13}NO_2$	Isoleucine	1	Amino acid
132.1019:0.7748:p	$C_6H_{13}NO_2$	Leucine	1	Amino acid
133.0608:0.9120:p	$C_4H_8N_2O_3$	Asparagine	1	Amino acid
133.0972:0.645:p	$C_5H_{12}N_2O_2$	Ornithine	1	Amino acid
134.0448:0.9558:p	$C_4H_7NO_4$	Aspartic acid	1	Amino acid
147.0764:0.8654:p	$C_5H_{10}N_2O_3$	Glutamine	1	Amino acid
148.0604:0.8869:p	$C_5H_9NO_4$	Glutamic acid	1	Amino acid
156.0768:0.3622:p	$C_6H_9N_3O_2$	Histidine	1	Amino acid
162.1125:0.5573:p	$C_7H_{15}NO_3$	Carnitine	1	Hydroxy fatty acid
166.0863:0.8815:p	$C_9H_{11}NO_2$	Phenylalanine	1	Amino acid
168.0770:0.4592p	$C_7H_9N_3O_2$	Amino acid	3	Amino acid
176.1030:0.9063:p	$C_6H_{13}N_3O_3$	Citrulline	1	Amino acid
182.0812:0.9364:p	$C_9H_{11}NO_3$	Tyrosine	1	Amino acid
194.1380:0.6003:p	$C_8H_{19}NO_4$	Unknown	4	Unknown
205.0972:0.8901:p	$C_{11}H_{12}N_2O_2$	Tryptophan	1	Amino acid
209.1290:0.5842:p	$C_{11}H_{16}N_2O_2$	Pilocarpine	1	Exogenous
213.0990:0.635:p	$C_8H_{12}N_4O_3$	Glycylhistidine	1	Dipeptide
227.1390:0.2425:p	$C_{11}H_{18}N_2O_3$	Pilocarpic acid	1	Exogenous
89.0244:1.2834:n	$C_3H_6O_3$	Lactic acid	1	Hydroxy organic acid
128.0352:1.0323:n	$C_5H_7NO_3$	Oxoproline	1	Amino acid derivative
167.0211:0.9043:n	$C_5H_4N_4O_3$	Uric acid	1	Purine derivative
171.1391:0.7537n	$C_{10}H_{20}O_2$	Capric acid	1	Medium-chain fatty acid
188.0929:0.7207:n	$C_8H_{15}NO_4$	N-butyrylhomoserine	2	Amino acid derivative
257.1146:0.6696:n	$C_{11}H_{18}N_2O_5$	Unknown	4	Unknown
329.2333:0.5152:n	$C_{18}H_{34}O_5$	TriHOME	2	Hydroxy fatty acid
373.0720:0.7938:n	$C_{37}H_{37}N_2O_9S_3$	FD&C blue dye no. 1	1	Exogenous

[a] *Most probable formula is presented in the case of unknowns or tentatively identified compounds.*

Table S5.2. Top-ranked sweat metabolites when comparing between *log*-transformed, sum-normalized responses between CF and non-CF infants without covariate adjustments.[a]

***m/z*:RMT:mode**	**Compound ID**	***p*-value**[b]	**Effect size**[c]	**Median FC**	***r***[d]
227.1390:0.7199:p	Pilocarpic acid	3.41×10^{-6}	1.44	0.29	-0.484 (7.31×10^{-6})
147.0764:0.8654:p	Glutamine	5.18×10^{-5}	1.01	1.70	0.260 (2.41×10^{-2})
128.0352:1.0323:n	Oxoproline	5.46×10^{-6}	1.00	2.24	0.481 (9.54×10^{-6})
134.0448:0.9558:p	Aspartic acid	2.25×10^{-5}	0.98	1.73	0.359 (1.36×10^{-3})
106.0499:0.7544:p	Serine	8.43×10^{-5}	0.94	1.46	0.430 (8.41×10^{-5})
90.055:0.6079:p	Alanine	1.13×10^{-4}	0.91	1.58	0.378 (6.42×10^{-4})
104.1075:0.2750:p	Choline	6.27×10^{-4}	0.83	1.48	0.095 (0.410)
132.1019:0.7565:p	Isoleucine	3.37×10^{-4}	0.82	1.41	0.390 (4.20×10^{-4})
76.0393:0.5109:p	Glycine	7.08×10^{-4}	0.80	1.29	0.358 (1.28×10^{-3})
182.0812:0.9364:p	Tyrosine	4.11×10^{-4}	0.79	1.47	0.404 (2.47×10^{-4})
156.0768:0.3622:p	Histidine	9.74×10^{-4}	0.77	2.42	0.286 (1.11×10^{-2})
148.0604:0.8869:p	Glutamic acid	1.82×10^{-3}	0.72	1.46	0.254 (2.50×10^{-2})

Log-transformed chloride: p-value = 5.09×10^{-43}, effect size = 5.71, and median FC for CF/non-CF = 7.23.

[a] *CF (n=25) versus non-CF (n=53) infants as confirmed by sweat chloride confirmatory testing.*

[b] *Univariate student t-Test analysis with an unequal variance assumption at the 0.05 level.*

[c] *Hedges' G test used to determined effect sizes.*

[d] *r corresponds to Spearman correlation rank coefficient and their statistical significance (p-values) using sweat chloride as a reference (1.00).*

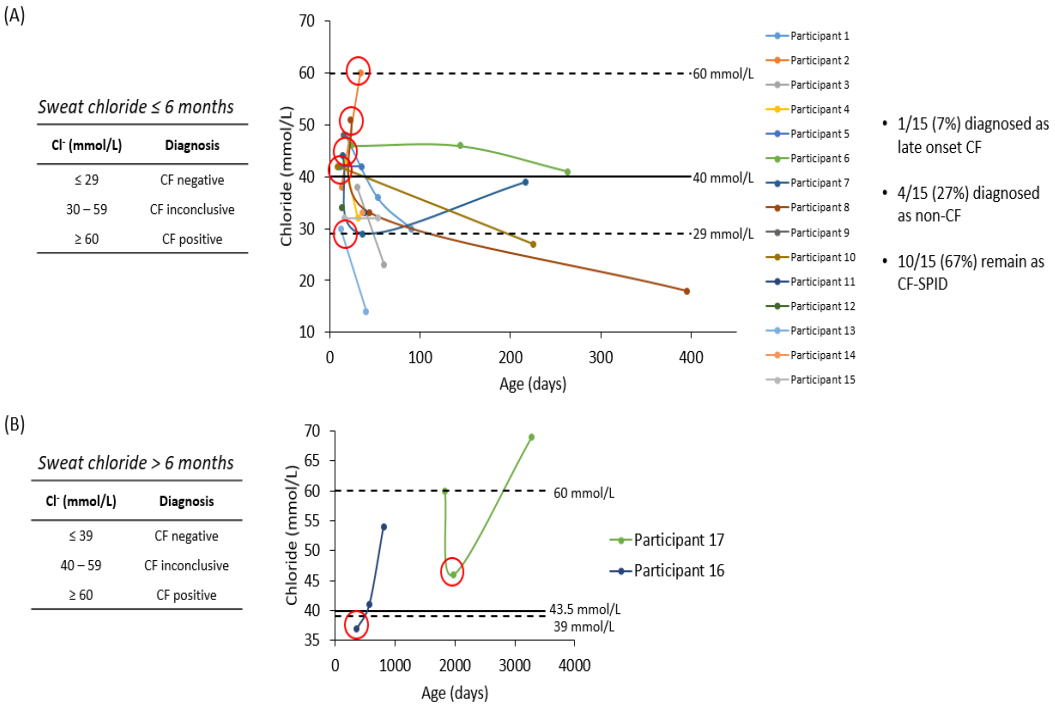


Figure S5.1. Trajectories of CF-SPID cases: **(A)** confirmed late onset non-CF infants (n = 4) and 1 late onset CF infant. **(B)** 1 assumed late onset CF infant with last sweat chloride reported at 54 mmol/L and one 5 yo child who was diagnosed with last onset CF were used in the study. The circled points indicate the sweat samples that were analyzed at baseline collection or the next collection if initial samples were not available.

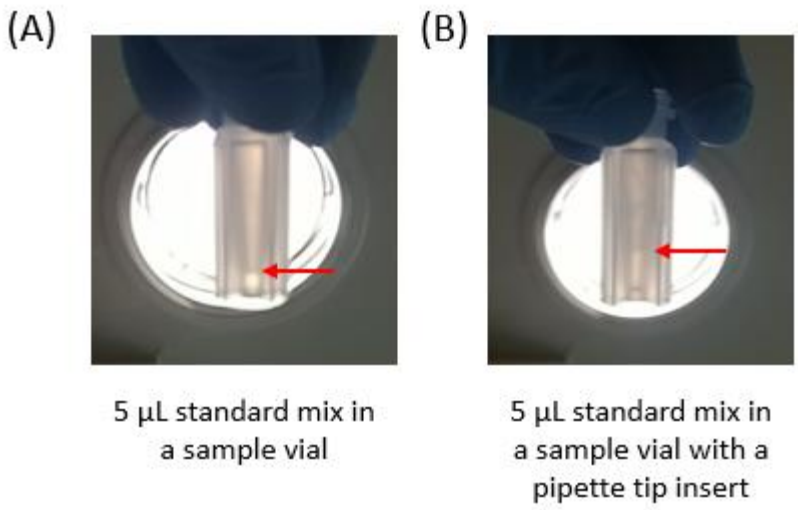


Figure S5.2. Photos illustrating a special vial sample adaptor was developed to reproducibly introduce small volumes of sweat specimens on-capillary when using MSI-CE-MS in this study. The adapted and further modified pipette tip inserts, where **(A)** corresponds to a regular sample vial holding 5 µL of standard mix solution and **(B)** corresponds to the pipette tip insert present in the sample vial present with 5 µL of the same solution but lifted higher up in the sample due to the conical property of the tip insert.

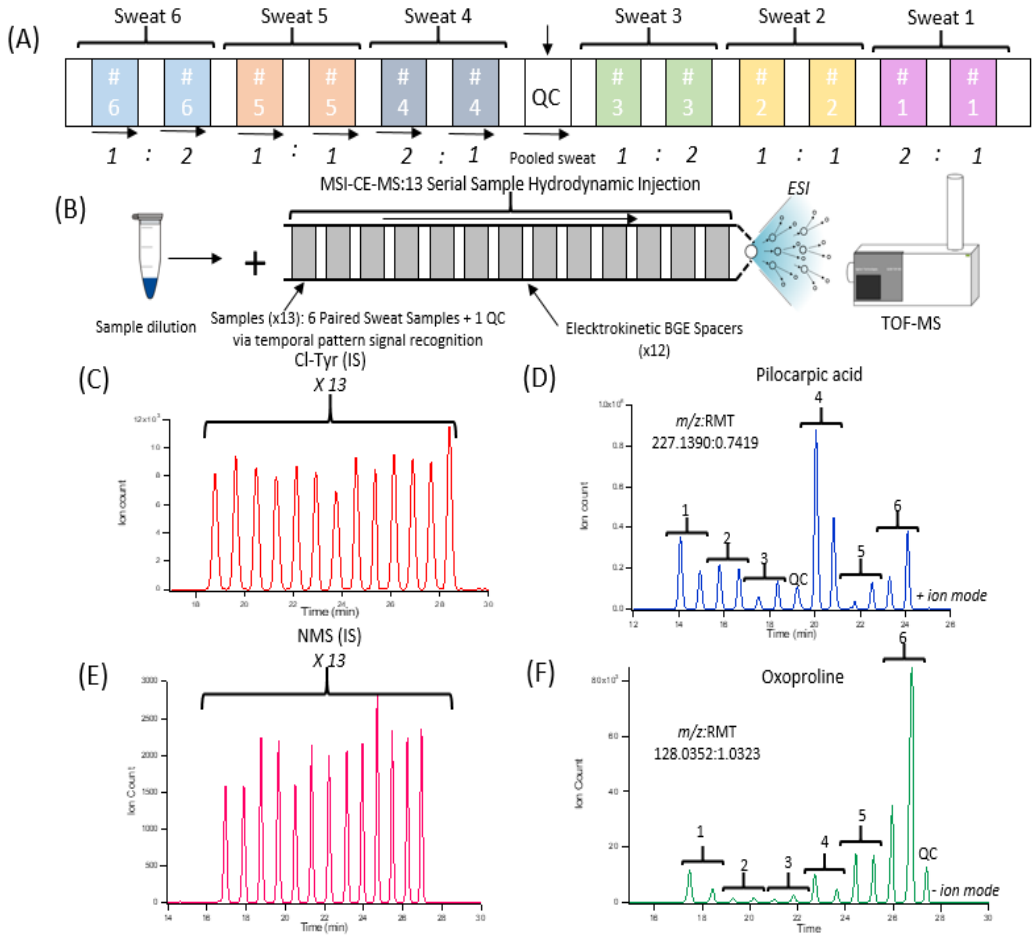
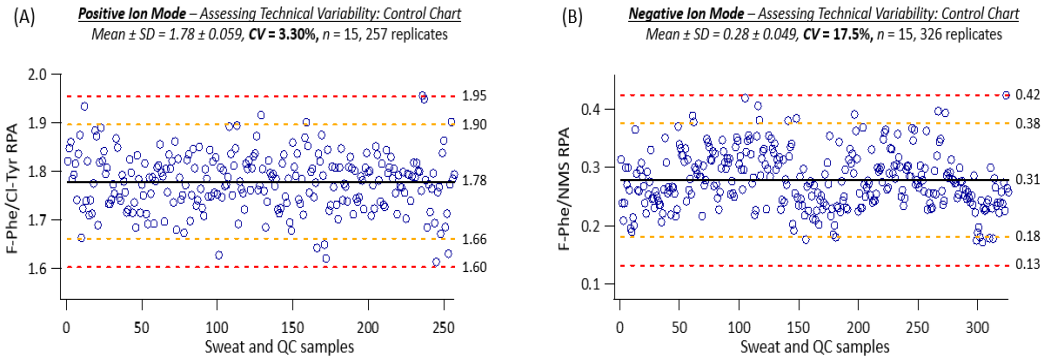


Figure S5.3. Multiplexed separations when using MSI-CE-MS that enables the analysis of 13 diluted sweat specimens within a single run, which also allows for unique study designs when performing nontargeted metabolite profiling. A schematic illustrating **(A)** a temporal signal pattern recognition was applied to the analysis of six paired samples (i.e., measured in duplicate) from screen-positive CF infants and a single pooled QC sample when using MSI-CE-MS. **(B)** Sweat samples were simply diluted either by 5× or 10× (1:2, 2:1 or 1:1 ratio) to enable sample pattern recognition with addition of 10 µM Cl-Try and F-Phe as internal and recovery standards, respectively. **(C)** An extracted ion electropherogram (EIE) displaying Cl-Tyr consistently measured in all 13 samples injections by MSI-CE-MS under acidic buffer conditions with positive ion mode detection. **(D)** An EIE of an cationic sweat metabolite, pilocarpic acid (227.1390:0.7419:p) highlighting its large biological variance between screen-positive infants. **(E)** An EIE displaying NMS consistently measured in all 13 samples injections by MSI-CE-MS under alkaline buffer conditions with negative ion mode detection. **(F)** An EIE of an anionic sweat metabolite, oxoproline (128.0352:1.0323:n) highlighting its large biological variance between screen-positive infants.

		Injection													
Experiment	Run	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	0	QC	QC	QC	QC	QC	QC	Blank	QC	QC	QC	QC	QC	QC	*New QC
2	1	148	148	31	31	QC	168	168	15	15	207A	207A	210	210	
3	2	QC	92T	92T	166	166	211A	211A	12T	12T	139	139	22T	22T	
4	3	203C	203C	QC	155	155	214	214	215B	215B	129	129	187	187	
5	4	85T	85T	162	162	170	170	QC	6T	6T	32T	32T	211B	211B	
6	5	154	154	10T	10T	141	141	184C	184C	182	182	199	199	QC	
7	6	45T	45T	224A	224A	171	171	190A	190A	QC	5T	5T	201	201	*New QC
8	7	216A	216A	96T	96T	78T	78T	QC	178	178	100T	100T	208C	208C	
9	8	220D	220D	180	180	QC	202B	202B	142	142	87T	87T	137	137	
10	9	200	200	204B	204B	185A	185A	52T	52T	160	160	QC	75T	75T	
11	10	53T	53T	161	161	50T	50T	QC	204C	204C	209	209	128	128	
12	0	QC	QC	QC	QC	QC	QC	Blank	QC	QC	QC	QC	QC	QC	
14	0	QC	QC	QC	QC	QC	QC	Blank	QC	QC	QC	QC	QC	QC	*New QC
15	11	215A	215A	QC	4	4	149	149	120	120	202E	202E	207B	207B	
16	12	126	126	204A	204A	34T	34T	103	103	202D	202D	QC	110	110	
17	13	212A	212A	111	111	223	223	131	131	QC	QC	108	108	QC	
18	14	104	104	195	195	220B	220B	123	123	QC	203A	203A	39T	39T	
19	15	QC	208B	208B	16T	16T	112	112	10	10	208A	208A	125	125	*New QC
20	16	73T	73T	QC	191	191	23	23	116	116	99T	99T	18	18	
21	17	224B	224B	QC	QC	192B	192B	QC	23T	23T	164	164	71T	71T	
22	18	205	205	21	21	25T	25T	193	193	74T	74T	QC	31T	31T	
23	19	QC	127	127	203B	203B	117	117	2	2	79T	79T	186A	186A	
24	20	176	176	82T	82T	QC	6	6	130	130	198	198	153A	153A	
25	0	QC	QC	QC	QC	QC	QC	Blank	QC	QC	QC	QC	QC	QC	*New QC

Figure S5.4. Summary of experimental plan and randomized serial sample injection sequence when using multiplexed separations in MSI-CE-MS for the sweat metabolome characterization of screen-positive CF infants and pooled QC samples under both positive and negative ionization mode detection.



Appendix S5.5. Control charts for recovery standards used to assess method reproducibility when using MSI-CE-MS for characterization of the sweat metabolome of screen-positive CF infants. **(A)** Positive ionization mode using F-Phe/Cl-Tyr of individual sweat and QC samples with a mean CV of 3.30% (*n* = 257). **(B)** Negative ionization mode using F-Phe/NMS of individual sweat and QC samples with a mean CV of 17.5% (*n* = 326).

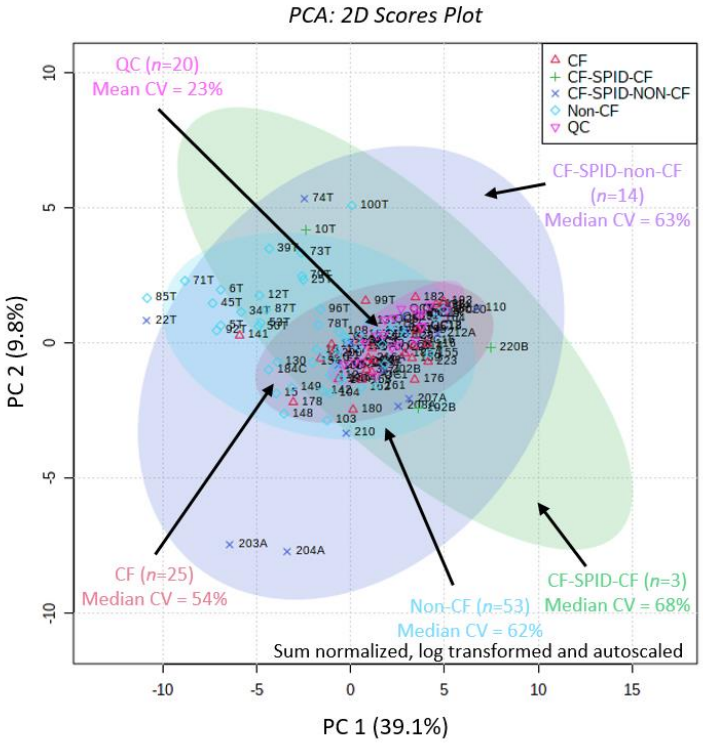


Figure S5.6. A PCA:2D scores plot outlining the median technical CV of the pooled QCs imbedded in each MSI-CE-MS run (median CV = 23%, n = 20), as well as the median biological CV of the four different screen-positive CF groups, including non-CF (median CV = 62%, n = 53), CF (median CV = 54%, n = 25), CF-SPID late onset unaffected or remains CF-SPID (median CV = 63%, n = 14), and CF-SPID late onset CF (68%, n = 3).

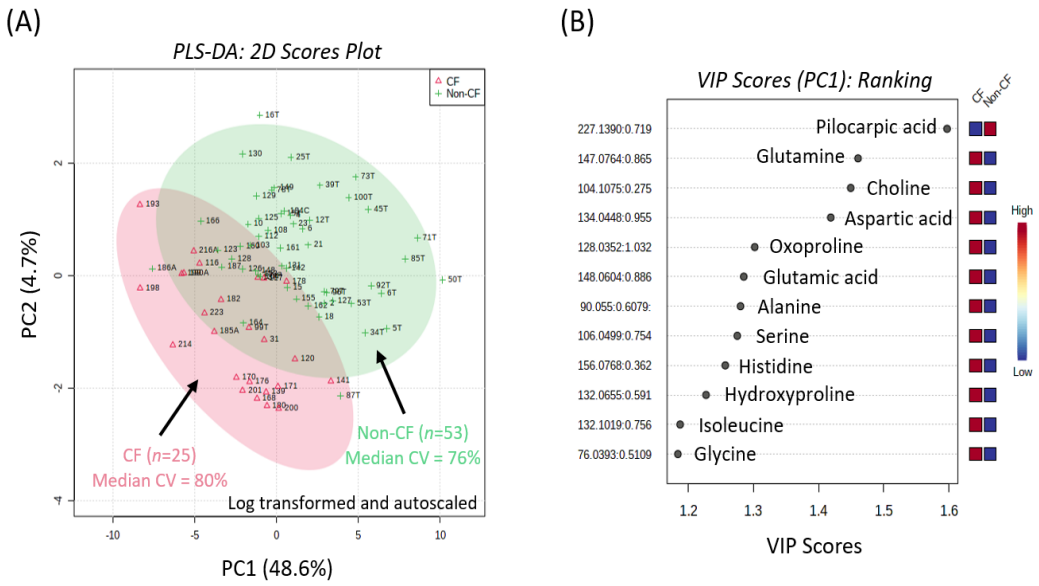


Figure 5S.7. Supervised multivariate data analysis without sum normalization of the sweat metabolome and its impact on **(A)** PLS-DA 2D scores plots and **(B)** VIP scores plot ranking, where pilocarpic acid and glutamine remain among the most significant sweat metabolites to differentiate CF cases from non-CF controls.

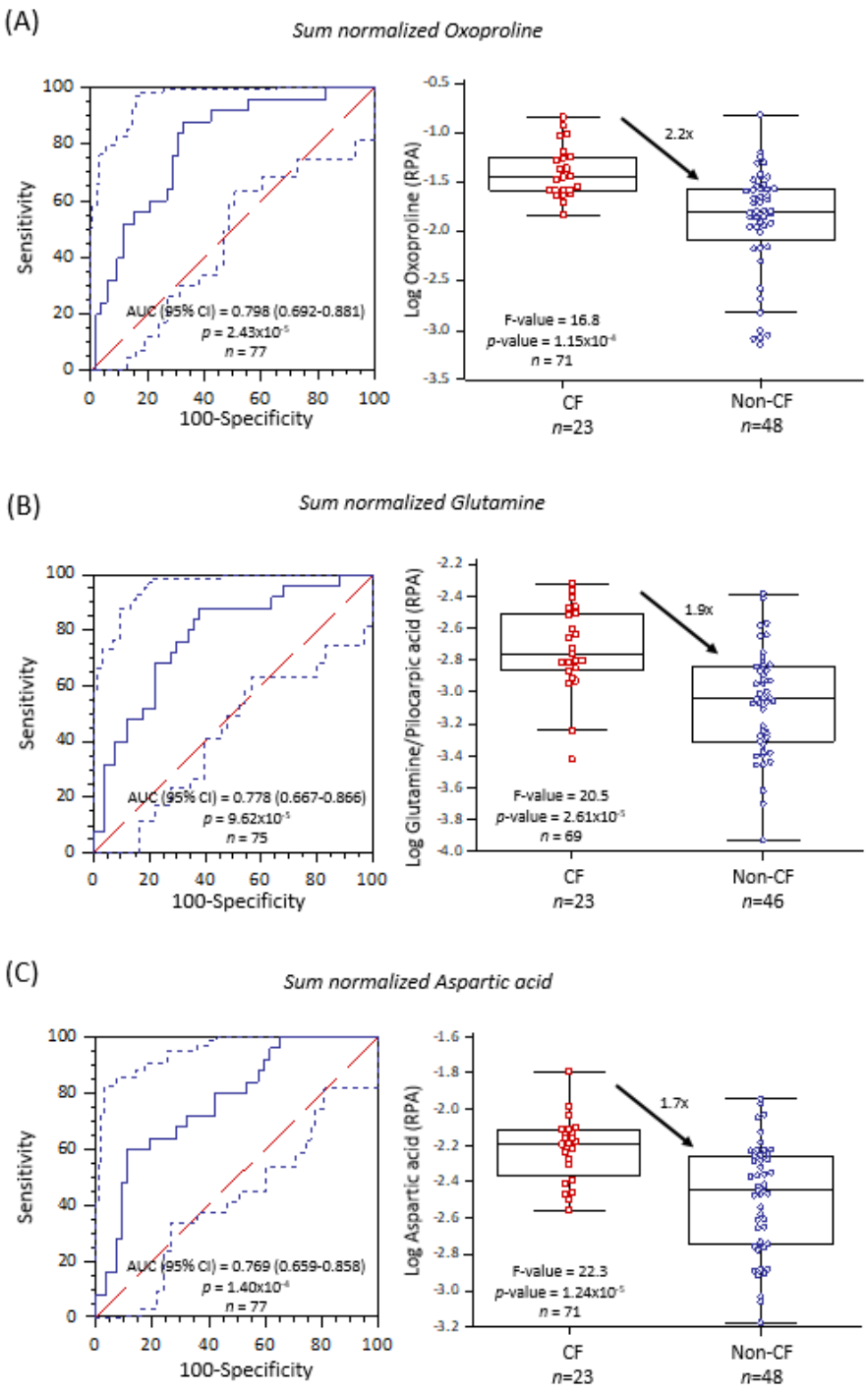


Figure S5.8. Receiver-operating characteristic (ROC) curves and boxplots for top-ranked sweat metabolites that differentiate CF from non-CF infants, **(A)** oxoproline, **(B)** glutamine, and **(C)** aspartic acid.

Chapter VI

Future Directions in New Advances in Capillary Electrophoresis for Biomonitoring in Population Health and Newborn Screening of Cystic Fibrosis

Chapter VI: Future Directions in New Advances in Capillary Electrophoresis for Biomonitoring in Population Health and Newborn Screening of Cystic Fibrosis

6.1 Overview of Major Thesis Contributions

The work in this thesis has contributed to novel insights into the measurement and validation of biomarkers in clinical or epidemiological studies when using high efficiency capillary electrophoresis separations. The thesis outlined several different applications of biomarkers that were used for more objective assessment of nutritional status, diet quality, and toxicant exposure from tobacco smoke as compared to self-reports. The thesis also disproved the clinical utility of sweat bicarbonate in CF diagnosis while providing a robust method to reliably analyze sweat chloride from volume-restricted specimens in a clinical setting not feasible by a commercial chloridometer. Furthermore, several novel sweat metabolites were identified using nontargeted screening that may augment sweat chloride testing for unambiguous CF diagnosis and potentially CF-SPID prognosis. Overall, the thesis comprised four major contributions that have relied on state-of-the-art CE separations, including: (1) rapid iodide screening based on analysis of 24 h urine samples collected from 800 adults residing in four different cities across Canada to assess for iodine deficiency, (2) validation of urinary thiocyanate as a robust biomarker of tobacco smoke exposure and diet quality in an international cohort consisting of 1000 participants from 14 countries worldwide, (3) a rapid assay to measure sweat chloride and bicarbonate simultaneously from volume-limited screen-positive CF sweat specimens, and (4) a high-throughput method to identify novel sweat metabolites to discriminate presumptive CF infants that complements sweat chloride testing.

Chapter I of this thesis provided an overview of biomarker discovery and use, relevant in clinical and epidemiological research, including the utility of biomarkers in clinical medicine, used to monitor iodine nutritional status, tobacco

smoke exposure, overall diet quality, as well as the diagnostic dilemmas related to universal newborn screening for CF. *Chapter II* introduced a previously validated CE assay that takes advantage of sample self-stacking[1] to analyze sub-micromolar levels of iodide from 24 h urine specimens from participants across Canada. This work was aimed for evaluating the prevalence and risk factors associated with iodine deficiency in the population, including exposure to two ubiquitous environmental iodine uptake inhibitors, namely nitrate and thiocyanate. Iodine deficiency remains a major global health issue that impacts approximately 2 billion people worldwide.[2] Health consequences of iodine deficiency in adults include, impaired mental function, hypothyroidism, decreased educability, apathy, and impaired social and economic development due to reduced work productivity.[3] Excessive exposure to iodine symporter inhibitors, such as perchlorate, thiocyanate, and nitrate, can impact the sensitivity of the central thyroid hormones in adults and should be closely monitored.[4] Our method is able to selectively and reliably quantify urinary iodide as well as thiocyanate and nitrate with minimal sample workup. In our work, participants residing in Vancouver and Quebec City were shown to be at greater risk for iodine deficiency compared to participants from Hamilton and Ottawa. In addition, higher levels of thiocyanate and nitrate were determined for participants notably from Quebec City. Overall, iodine supplement and thyroxine use were protective factors to prevent iodine deficiency. The increased consumption of dairy products was also protective against iodine deficiency given that it is a major yet variable source of elemental iodine. However, participants from Quebec City still had lower iodine status despite their higher intake of dairy and salt (from 24 h urine excretion) as compared to other regions. Further investigation is urgently needed to examine regional variations in iodine status and environmental iodine uptake exposure specially in vulnerable populations, such as children and pregnant women. Limitations of this study include not having iodine specific dietary questions, such

as seaweed intake. Also, biomarkers of thyroid function were not directly assessed, including potential drinking water sources of exposure.

Chapter III described a validated CE-UV assay previously developed for iodine status determination to reliably quantify urinary thiocyanate as a biomarker of tobacco smoke exposure across 14 countries that differ with respect to socioeconomic status, including self-reported smokers and never smokers. Tobacco smoking is known to be the leading preventable cause of cancer death.[5] which also contributing to overall chronic disease burden globally. Thiocyanate has been used as a biomarker for smoking dating back to the 1900s when using saliva as a biological sample and into the 1980s when using urine as a biological specimen.[6,7] However, there have been sparse studies using urinary thiocyanate as a biomarker of tobacco smoke exposure due to several reasons, including endogenous thiocyanate and thiocyanate from food exposures that may act as confounding factors.[8,9] Our results indicate that current smokers from high-income countries (HICs), such as Canada, have greater risk for chronic disease and clinical events due to their greater exposure to the toxicants, such as cyanide as compared to smokers from middle- and low-income countries after appropriate covariate adjustments. Ironically, our results also indicated better diet quality in confirmed never smokers as compared to current smokers since most background sources of thiocyanate originate from the intake of vegetables that promote health.

Chapter IV introduced a rapid (< 5 min) and selective assay to simultaneously measure chloride and bicarbonate using CE with indirect UV detection. Cystic fibrosis (CF) is a heterogeneous recessive genetic disorder that is initially screened days after birth through the newborn screening (NBS) program.[10,11] NBS for CF is currently based on a two-tiered immunoreactive trypsinogen (IRT) followed by CFTR mutational analysis screen prior to a confirmatory sweat chloride test for all presumptive screen-positive infants.[12] Authentic CF infants have chloride concentrations ≥ 60 mmol/L, whereas

unaffected non-CF have sweat chloride < 30 mmol/L with inconclusive cases having intermediate concentrations (30 – 59 mmol/L).[12] CF patients are known to have insufficient pancreatic function due to the impaired transport of bicarbonate.[13] As a result, we developed a simple and rapid assay to diagnose CF using volume-limited pilocarpine stimulated collected sweat specimens. Our assay was rigorously validated by performing stability, recovery, precision studies, and inter-method comparison of sweat chloride to a commercial coulometric titration technique. Stability studies indicated sweat bicarbonate was stable even after prolonged delays to room temperature and multiple freeze-thaw cycles. Also, an inter-method comparison revealed good mutual agreement n sweat chloride determination using CE-iUV as compared to coulometric titration. That notwithstanding, our results showed no correlation was determined between bicarbonate levels and CF status or pancreatic function that was consistent with a previous report.[14] Nevertheless, the CE assay may prove useful in a clinical setting by reducing test failure rates due to insufficient sweat volumes collected from premature or underweight infants, which avoids retesting and delayed diagnoses. Although sweat bicarbonate lacked clinical diagnostic value, this assay may be useful in biological samples showed to have a direct correlation to impaired bicarbonate transport, such as airway lung fluid and pancreatic juice.

Chapter *V* describes a high-throughput method to characterize the sweat metabolome in volume-restricted (≥ 0.5 μL) sweat samples collected from screen-positive CF affected, unaffected, and inconclusively diagnosed infants based on MSI-CE-MS. Although the initial sweat chloride test may potentially be used for risk assessment of probable late onset CF among CF-SPID cases,[15] there may exist mechanisms beyond impaired chloride transport or re-uptake in sweat that could aide CF prognosis and disease progression. Our previously study included a nontargeted analysis of sweat metabolites from screen-positive CF affected and unaffected infants. As a result, two endogenous (glutamine and asparagine) and two exogenous biomarkers, pilocarpic acid (PA) and mono(2-ethylhexyl)phthalic

acid (MEHP) were shown to be significant at discriminating both groups (< 3 months).[16] Since our previous work, we have increased our throughput by almost double with similar analysis times. This work was reproducible and determined aspartic acid, pilocarpic acid, glutamine, and oxoproline to be top biomarkers to discriminate CF from non-CF infants with minor limitations involving isobaric interferences and lower sensitivity in negative ion mode. Overall, the top-ranked ratiometric biomarkers in sweat were determined to be aspartic acid/pilocarpic acid and glutamine/pilocarpic acid, which did not require sum normalization to adjust for differences in sweat rate/hydration status. Our major goal was to investigate potential prognostic potential of these CF biomarkers to further augment chloride as applied to disease progression in CF-SPID. However, due to the limitation of sample size with only 3 CF-SPID cases being diagnosed with CF over the time frame of our study, future prospective studies are needed to better evaluate the clinical utility of a panel of sweat metabolites in CF diagnosis, prognosis, which also may be useful when monitoring treatment responses to therapy.

6.2 General Conclusion

In summary, this thesis contributes innovative CE methods for biomarker screening and discovery relevant to population health and clinical studies, which may allow for early detection that benefit from optimal nutrition or smoking cessation.

6.3 References

(1) Macedo, A. N.; Macri, J.; Hudecki, P. L.; Saoi, M.; McQueen, M. J.; Britz-McKibbin, P. Validation of a Capillary Electrophoresis Assay for Monitoring Iodine Nutrition in Populations for Prevention of Iodine Deficiency: An Interlaboratory Method Comparison. *J. Appl. Lab. Med. An AACC Publ.* **2017**, *1* (6), 649–660.

(2) Krela-Kaźmierczak, I.; Czarnywojtek, A.; Skoracka, K.; Rychter, A. M.; Ratajczak, A. E.; Szymczak-Tomczak, A.; Ruchała, M.; Dobrowolska, A. Is There an Ideal Diet to Protect against Iodine Deficiency? *Nutrients* **2021**, *13* (2), 1–15.

(3) Zimmermann, M. B.; Jooste, P. L.; Pandav, C. S. Iodine-Deficiency Disorders. *Lancet* **2008**, *372* (9645), 1251–1262.

(4) King, L.; Huang, Y.; Li, T.; Wang, Q.; Li, W.; Shan, Z.; Yin, J.; Chen, L.; Wang, P.; Dun, C.; et al. Associations of Urinary Perchlorate, Nitrate and Thiocyanate with Central Sensitivity to Thyroid Hormones: A US Population-Based Cross-Sectional Study. *Environ. Int.* **2022**, *164*, 1–14.

(5) Samet, J. M. Tobacco Smoking. The Leading Cause of Preventable Disease Worldwide. *Thorac. Surg. Clin.* **2013**, *23* (2), 103–112.

(6) Buratti, M.; Xaiz, D.; Caravelli, G.; Colombi, A. Validation of Urinary Thiocyanate as a Biomarker of Tobacco Smoking. *Biomarkers* **1997**, *2* (2), 81–85.

(7) Haley, N. J.; Axelrad, C. M.; Tilton, K. A. Validation of Self-Reported Smoking Behavior: Biochemical Analyses of Cotinine and Thiocyanate. *Am. J. Public Health* **1983**, *73* (10), 1204–1207.

(8) Heming, T. A.; Thurston, R. V.; Meyn, E. L.; Zajdel, R. K. Acute Toxicity of Thiocyanate to Trout. *Trans. Am. Fish. Soc.* **1985**, *114* (6), 895–905.

(9) Chang, W. H.; Chen, P. H.; Herianto, S.; Chen, H. L.; Lee, C. C. Aggregating Exposures and Toxicity Equivalence Approach into an Integrated Probabilistic Dietary Risk Assessment for Perchlorate, Nitrate, and Thiocyanate: Results from the National Food Monitoring Study and National Food Consumption Database. *Environ. Res.* **2022**, *211*, 1–13.

(10) Hirsch, M. S. What Is Cystic Fibrosis? *N. Engl. J. Med.* **2002**, *347* (6), 439–443.

(11) Bell, S. C.; Mall, M. A.; Gutierrez, H.; Macek, M.; Madge, S.; Davies, J. C.; Burgel, P. R.; Tullis, E.; Castaños, C.; Castellani, C.; et al. The Future of Cystic Fibrosis Care: A Global Perspective. *Lancet Respir. Med.* **2020**, *8* (1), 65–124.

(12) De Boeck, K.; Wilschanski, M.; Castellani, C.; Taylor, C.; Cuppens, H.; Dodge, J.; Sinaasappel, M. Cystic Fibrosis: Terminology and Diagnostic Algorithms. *Thorax* **2006**, *61* (7), 627–635.

(13) Kunzelmann, K.; Schreiber, R.; Hadorn, H. B. Bicarbonate in Cystic Fibrosis. *J. Cyst. Fibros.* **2017**, *16* (6), 653–662.

(14) Bijman, J.; Quinton, P. M. Lactate and Bicarbonate Uptake in the Sweat Duct of Cystic Fibrosis and Normal Subjects. *Pediatric Research*. **1987**, *21* (1), 79–82.

(15) Gonska, T.; Keenan, K.; Au, J.; Dupuis, A.; Chilvers, M. A.; Burgess, C.; Bjornson, C.; Fairservice, L.; Brusky, J.; Kherani, T.; et al. Outcomes of Cystic Fibrosis Screening-Positive Infants With Inconclusive Diagnosis at School Age. *Pediatrics* **2021**, *148* (6), 1–9.

(16) Macedo, A. N.; Mathiaparanam, S.; Brick, L.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P. The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Cent. Sci.* **2017**, *3* (8), 904–913.