IODINE NUTRITION IN EARLY LIFE AND CHILDHOOD NEURODEVELOPMENT AND GROWTH

Molla Mesele Wassie

(BSc in Public Health, MSc in Applied Human Nutrition)

The University of Adelaide

School of Agriculture, Food and Wine

Submitted in partial fulfilment of the Degree of Doctor of Philosophy

The University of Adelaide

June 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SUPPLEMENTAL TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL FIGURES

ABSTRACT

Iodine is a micronutrient required for optimal development and growth. Mandatory iodine fortification of bread was implemented in Australia in October 2009 following the re-emergence of iodine deficiency. However, there are limited data on the iodine status of South Australians and populations at risk, particularly pregnant women. There is a concern about the inconsistent classification of iodine status in the general population and pregnant women by different markers. Evidence on the association between iodine nutrition in early life, and neurodevelopment and growth in populations with mild iodine deficiency to iodine sufficiency, is inconsistent. Newborn thyroid stimulating hormone (TSH) concentration has been suggested as a marker of iodine nutrition in early life.

This thesis aims to assess: 1) the agreement between markers of iodine status, 2) iodine status (including pregnant women) in South Australia, before and after the mandatory iodine fortification using the newborn TSH concentration and urinary iodine concentration (UIC), and 3) associations between newborn TSH concentration and childhood neurodevelopment and growth.

A systematic review was conducted to assess agreement between markers of population-level iodine status (median UIC, newborn TSH concentration and goitre prevalence), while urinary iodine data, including an estimated 24-h urinary iodine excretion (UIE), UIC, iodine-to-creatinine ratio (I/Cr), and UIC-corrected for creatinine, were used to assess agreement between markers of individual-level iodine status in pregnant women. The results of the systematic review showed that at a population level, newborn TSH concentration >5 mIU/L greater than 3% had a better agreement with goitre prevalence than median UIC to define iodine deficiency in populations. At an individual level, I/Cr from spot urine samples had a better agreement with the estimated 24-h UIE compared with UIC or UIC~Cr markers in pregnant women.

Based on the newborn TSH concentration, South Australia was classified as mildly iodine deficient in both pre- and post-fortification periods in contrast to iodine sufficiency defined by median UIC post-fortification. Iodine status of pregnant women was classified as iodine deficiency pre-fortification and iodine sufficiency post-fortification by the UIC marker.

Utilising developmental outcome data from two studies, a null association was observed between newborn TSH and childhood neurodevelopment or growth at 18 months. However, poorer neurodevelopment or growth in infants with high TSH in a borderline iodine deficient setting and better neurodevelopment in infants with high TSH in iodine sufficient setting cannot be excluded.

In conclusion, monitoring of iodine status using multiple markers is required to identify populations or population groups at increased risk of iodine deficiency disorders. Re-evaluation of current TSH criteria for classifying iodine status in populations is suggested. As neurodevelopmental assessments at 18 months of age may not be stable, data-linkage studies that utilise newborn TSH data and examine neurodevelopmental outcomes at later ages are warranted in populations where newborn screening is routinely performed.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Molla Mesele Wassie

Signature__ __ Date 15/06/2020

ACKNOWLEDGEMENTS

I want to express my deep and sincere gratitude to my principal supervisor Dr Shao Jia (Jo) Zhou, for your guidance throughout each stage of my PhD. I am so grateful for your kindness, patience, advice and expertise every time we met to discuss my thesis and while supervising my work. This would have been impossible without your unconditional support and continued encouragement.

My co-supervisors, Associate Professor Lisa Gaye Smithers and Dr Lisa Nicole Yelland – you were instrumental in my PhD journey. Lisa S, you inspired my interest in recent epidemiological thinking; I am so thankful for your insightful thoughts and expertise. Thank you Lisa Y for your support in statistical analysis methods and for reviewing my work.

I want to acknowledge the staff of the Healthy Mothers, Babies and Children Theme and the then Foodplus Research Centre at the University of Adelaide, for their support and motivation in the regular science meetings and trainee forums. I am profoundly grateful to Dr Dao Huyhn for training me in the urinary iodine analysis methods. I would also like to extend my heartfelt thanks to my postgraduate coordinator, Associate Professor Matthew Denton for his kindness and encouragement.

I would like to thank Professors Maria Makrides and Claire Roberts for allowing me to access data from their projects, and for their support in reviewing my work. Thanks also go to Dr Enzo Renieri who gave me access to the newborn screen data and reviewing the work generated from the newborn screening data. I also thank the librarians at the University of Adelaide, Michael Draper and Angela Mills, in assisting me to build my search terms for my systematic reviews. I would also like to thank Dr Yvonne Miels for her editorial support in the final draft of my thesis.

My fellow previous and current students and friends (Gabi, Erandi, Chang, Khaled, Gizachew, John, Abel, Yonatal, Daniel, Anteneh, and Anwar) for all the chat and fun we have had in the last couple of years; I learned a lot from each of your experiences.

To my wife Eskedar Friew, thank you for unconditional love, understanding and continuing support to complete my thesis. Kukuye – your continuous motivation helps me to be a better version of myself in many aspects of life. To my daughter Haphseba Molla, you are a source of happiness since you born midway in my candidature. I am incredibly grateful to my mother, Zewditu Assefa and my father, Mesele Wassie, for your prayers, love and all the sacrifices you made when raising me. This work is dedicated to you all.

LIST OF PUBLICATIONS

List of publications from this thesis

1. **Wassie MM**, Middleton, P., & Zhou, S. (2019). Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review. *The American Journal of Clinical Nutrition,* 110(4), 949-958. DOI: https://doi.org/10.1093/ajcn/nqz118

2. **Wassie MM**, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. (2019). Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: A population study using newborn thyroid-stimulating hormone concentration as a marker. *Public Health Nutrition,* 22(16),3063-3072. DOI: https://doi.org/10.1017/S1368980019001915

List of other publications during PhD Candidature

1. Derso, T., Tariku, A., Biks, G., & **Wassie MM** (2017). Stunting, wasting and associated factors among children aged 6-24 months in Dabat health and demographic surveillance system site: A community based cross-sectional study in Ethiopia. *BMC Pediatrics*, *17*(1), 9 pages. DOI: https://doi.org/10.1186/s12887-017-0848-2

   Contributions: Conception, critical revision and editing of the manuscript.

2. **Wassie MM**, Abebe, Z., Trariku, A., Gebeye, E., Awoke, T., Gete, AA, . . . Zhou, S. J. (2018). Iodine status five years after the mandatory salt iodisation legislation indicates above requirement: a cross sectional study in Northwest Ethiopia. *BMC Nutrition*, 4(52). DOI: https://doi.org/10.1186/s40795-018-0261-8

   Contributions: Conception, design, statistical analysis, interpretation of data, manuscript preparation, critical revision, editing of the manuscript and a corresponding author.

3. Melaku, Y., **Wassie MM**, Gill, T., Zhou, S., Tessema, G., Amare, A., . . . Deribew, A. (2018). Burden of disease attributable to suboptimal diet, metabolic risks and low physical

activity in Ethiopia and comparison with Eastern sub-Saharan African countries, 1990-2015: findings from the Global Burden of Disease Study 2015. *BMC Public Health*, 18(1), 552-1-552-20. DOI: https://doi.org/10.1186/s12889-018-5438-1

Contributions: Conception, data provision, interpretation of data, manuscript preparation, critical revision and editing of the manuscript.

4. Engidaw, M., **Wassie MM**, & Teferra, A. (2018). Anaemia and associated factors among adolescent girls living in Aw-Barre refugee camp, Somali regional state, Southeast Ethiopia. *PLoS ONE,* 13(10), 12 pages. DOI: https://doi.org/10.1371/journal.pone.0205381

   Contributions: design, data analysis, supervision, manuscript preparation, critical revision, editing of the manuscript.

5. Geberselassie SB, Abebe SM, Melsew YA, Mutuku SM, & **Wassie MM** (2018). Prevalence of stunting and its associated factors among children 6-59 months of age in Libo-Kemekem district, Northwest Ethiopia; A community based cross sectional study. *PLoS ONE,* 13(5), 1-11. DOI: https://doi.org/10.1371/journal.pone.0195361

   Contributions: Conception, design, data analysis, supervision, manuscript preparation, critical revision, editing of the manuscript and corresponding author.

6. Tariku, A., Belew, A., Gonete, K., Hunegnaw, M., Muhammad, E., Demissie, G., . . . **Wassie MM** (2019). Stunting and Its Determinants among Adolescent Girls: Findings from the Nutrition Surveillance Project, Northwest Ethiopia. *Ecology of Food and Nutrition*, 58(5), 481-494. DOI: https://doi.org/10.1080/03670244.2019.1636793

   Contributions: Conception, design, supervision, critical revision and editing of the manuscript.

7. Dagne, S., Gelaw, Y., Abebe, Z., & **Wassie MM** (2019). Factors associated with overweight and obesity among adults in northeast Ethiopia: A cross-sectional study. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, 12, 391-399. DOI: https://doi.org/10.2147/DMSO.S179699

Contributions: Conception, design, data analysis, interpretation of findings, supervision, critical revision and editing of the manuscript.

8. Abebe Z, Tariku A, Bikes GA, **Wassie MM**, *et al.* (2019). Poor child complementary Feeding Practices in northwest Ethiopia: Finding from the Baseline Survey of Nutrition Project, 2016. *Italian Journal of Pediatrics* 45(1):154. DOI: https://doi.org/10.1186/s13052-019-0747-2

   Contributions: Conception, design, data analysis, interpretation of findings, supervision, critical revision and editing of the manuscript.

LIST OF RESEARCH CONFERENCE PRESENTATIONS

Conference presentations

1. **Wassie MM**, Middleton, P., & Zhou, S. Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review. The Nutrition Society of Australia 42nd Annual scientific meeting. Canberra, Australia. 27-30 November 2018. Oral presentation.
2. **Wassie MM**, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: A population study using newborn thyroid-stimulating hormone concentration as a marker. The Nutrition Society of Australia 42nd Annual scientific meeting. Canberra, Australia. 27-30 November 2018. Oral presentation.
3. **Wassie MM**, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. A population study examined the effect of mandatory iodine fortification of bread on the iodine status. Australian Institute of Food Science and Technology Student summer school, The University of Adelaide, Australia. 26 February 2019. Oral presentation.
4. **Wassie MM**, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: A population study

using newborn thyroid-stimulating hormone concentration as a marker. American Society of Nutrition Conference. Baltimore, Maryland. 8-11 June 2019. Oral presentation.

5. **Wassie MM**, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. Effect of mandatory iodine fortification on the iodine status in Australia: an interrupted time series study. The 5th Emerging Leaders in Nutrition Science Poster Competition. American Society of Nutrition Conference. Baltimore, Maryland. 8 June 2019. Poster presentation.

LIST OF ABBREVIATIONS/ACRONYMS

AUC, Area Under Curve

Bayley-III, Bayley Scales of Infant and Toddler Development third edition

BMI, Body Mass Index

CI, Confidence Intervals

CH, Congenital Hypothyroidism

DOMInO, Docosahexaenoic acid to Optimize Mother Infant Outcome

g, Hedge's g

GA, Gestational Age

HCZ, head circumference-for-age

ICCIDD, International Counsel for Control of Iodine Deficiency Disorders

I/Cr, Iodine-to-Creatinine ratio

IDD, Iodine Deficiency Disorders

IGN, Iodine Global Network

IQ, Intelligence Quotient

IRR, Incidence Rate Ratio

IQR, Inter Quartile Range

LAZ, length-for-age

MD, Mean Difference

OR, Odds Ratio

PINK, Pregnancy Iodine and Neurodevelopment in Kids

RCT, Randomized Controlled Trials

RDI, Recommended Dietary Intake

ROC, Receiver Operating Curve

RR, Relative Risk

SAC, School Age Children

SD, Standard Deviation

SES, Socio-Economic Status

SCOPE, screening for pregnancy end points

STOP, screening tests to identify poor outcome in pregnancy

T3, 3, 5, 3/ Tri-iodothyronine

T4, Thyroxine

Tg, Thyroglobulin

TNH, Transient Neonatal Hyperthyrotropinemia

TSH, Thyroid Stimulating Hormone

TSH~age, thyroid-stimulating hormone corrected for age at blood sampling

UIC, Urinary Iodine Concentration

UIC~Cr, urinary iodine concentration corrected for creatinine

UIE, Urinary Iodine Excretion

UNICEF, United Nations Children's Fund

USI, Universal Salt Iodisation

WAZ, weight-for-age

WHO, World Health Organization

WLZ, weight-for-length

CHAPTER 1: INTRODUCTION

Iodine is a trace mineral necessary for the synthesis of thyroid hormones which are required for optimal growth and development [1, 2]. Insufficient iodine intake may lead to goitre and hypothyroidism in all ages, as well as impaired growth, mental retardation and cretinism during critical periods of early life [3]. The critical period in child growth and development starts from conception and continues to early childhood [4]. A systematic review of studies in severe iodine deficient settings in China showed a 12.5 points lower Intelligence Quotient (IQ) points in children with severe iodine deficiency compared with children living in an iodine sufficient setting [5]. Therefore, it is crucial to ensure iodine sufficiency at all times, especially during critical periods of life to prevent the unwanted consequences of iodine deficiency.

The goal of eliminating iodine deficiency in the year 2000 was adopted by the world leaders who attended the World Summit of Children through the introduction of a universal salt iodisation program in their countries [6]. Despite the global initiative to eliminate iodine deficiency, and evidence of improvement in iodine status of populations, it remains a public health problem in several countries even two decades since it had been assumed to be virtually eliminated worldwide. Recently, iodine deficiency has been reported in several countries [7, 8] and iodine deficiency remains a public health issue in both the developed and developing worlds. A recent Iodine Global Network (IGN) report using a nationally representative data until 2018 showed that 105 countries had optimal iodine intake, 14 countries had insufficient iodine intake, and nine countries had excessive iodine intake [7]. The epidemiology of iodine deficiency has shifted from severe forms of iodine deficiency to mild and moderate forms of iodine deficiency and excessive iodine intake due to the introduction of the iodine fortification programs [9-11]. Although many countries have declared iodine sufficiency, pregnant women remained at a higher risk of iodine deficiency in some countries, including Australia. This may be due to higher iodine requirements during

pregnancy (250 µg/day) than for non-pregnant women and adults (150 µg/day) [12]. Lack of considering the higher iodine demand during pregnancy in the existing fortification programs may also contribute for the existence of iodine deficiency in pregnant women.

The damaging effect of severe iodine insufficiency on the brain is recognised, but mild to moderate iodine deficiency and iodine excess during pregnancy may also harm the neurodevelopmental and growth outcomes in children [2, 13]. These, in turn, may hurt the economic performance and productivity of the affected nations as well as their quality of life [14]. Therefore, public health efforts towards sustainable elimination of iodine deficiency should continue to target at-risk population groups such as infants and pregnant women. As confounders of neurodevelopmental and anthropometric outcomes are very complex, a well-designed follow-up studies that adjust for key confounders are required.

The ability to accurately assess iodine status is crucial in preventing both iodine deficiency and excess. The World Health Organization (WHO), United Nations Children's Fund (UNICEF) and International Council for Control of Iodine Deficiency Disorders (ICCIDD) recommend the median urinary iodine concentration (UIC), goitre prevalence, and newborn thyroid stimulating hormone (TSH) concentration >5 mIU/L as appropriate markers to define iodine status in populations [12]. Most countries apply the median UIC to monitor the iodine status of populations using a representative sample of School Age Children (SAC). However, some evidence showed inconsistent classification of the iodine status of populations between these markers of iodine status [15].

In Australia, the mandatory iodine fortification of bread was in place by October 2009 [16] to control the re-emergence of iodine deficiency [17]. The intake of an iodine supplement of 150 µg per day has also been recommended to pregnant and breastfeeding women since January 2010 [18]. Studies in Australia that used median UIC in SAC showed iodine sufficiency in the

post-fortification periods, but there is little evidence on the iodine status of pregnant women using a representative sample. The use of multiple markers of iodine status and linking the iodine exposure status with functional outcomes like neurodevelopmental scores will be of paramount importance in the prevention and control of iodine deficiency disorders.

The monitoring system that includes iodine nutrition of populations and at-risk population groups is essential to make appropriate adjustments on the existing iodine intervention programs like the mandatory iodine fortification of bread in Australia. Although different markers are recommended to assess iodine status, the agreement between these markers to consistently classify iodine status is not known. Sustained adequate iodine intake in all population groups will have a vital role in the prevention of neurodevelopmental impairment and in boosting economic productivity. However, evidence of the effect of iodine nutrition on development and growth of children in populations of mild iodine deficiency or iodine sufficiency, is inconclusive. Therefore, the aims of this thesis were:

1) To evaluate the agreement between different markers of iodine status.
2) To examine the impact of the iodine fortification program in Australia on the iodine status of populations and pregnant women.
3) To investigate the associations between early life iodine nutrition and neurodevelopment and growth of children.

CHAPTER 2: LITERATURE REVIEW

The literature review covers the role of iodine nutrition, markers of iodine status, the prevalence of iodine deficiency in the general population and pregnant women worldwide and in Australia, and the associations between iodine nutrition in early life and neurodevelopment and growth of children. A systematic review and meta-analysis was performed to assess the relationship between newborn TSH and child neurodevelopment and growth.

2.1 The role of iodine

Iodine, an essential micronutrient, is needed for the synthesis of thyroid hormones (thyroxine (T4) and 3, 5, 3′ tri-iodothyronine (T3)). The thyroid hormones have a significant role in many physiological processes in the body [19-21]. One of the key roles of thyroid hormones is to maintain a healthy brain development as they are involved in the myelination, differentiation, migration, and signalling of the neural cells of the human brain in fetal and early postnatal periods [2]. Therefore, optimal iodine nutrition is required during the fetal and early critical periods of life [22].

2.2 Consequences of iodine deficiency or excess

Both iodine deficiency and excess has adverse effects on human health. Inadequate iodine intake impairs thyroid hormone production and leads to adverse health outcomes, which are collectively termed ‘Iodine Deficiency Disorders’ (IDD) [3]. IDD include goitre, hypothyroidism, impaired growth, mental retardation, cretinism, deaf-mutism, congenital anomalies, abortion and stillbirth [3]. IDD affects any stages of the life cycle. For example, poor iodine nutrition during pregnancy and early childhood may cause impaired psycho-neuromotor and intellectual development in children [23, 24]. Iodine deficiency remains a cause of preventable neurodevelopmental delay in a growing child [25]. Similarly, excessive iodine intake impairs the thyroid system in all stages of life and may cause adverse health outcomes like iodine-induced hypothyroidism, iodine-induced hyperthyroidism, goiter, and

thyroid autoimmunity [26]. Furthermore, excessive iodine intake during pregnancy may also negatively affect the neurodevelopmental outcomes of children [10]. Therefore, adequate iodine intake is necessary to improve neurodevelopment and somatic growth of children, and reduces the risk of adverse birth outcomes [22].

2.3 Iodine requirements and food sources

Iodine needs to be obtained from the diet, and its daily requirements vary according to gender and stage of life [1]. Infants and children have high iodine requirements relative to their body weight. Furthermore, pregnant women and breastfeeding women have higher iodine requirements and are at risk of developing IDD. In Australia, the recommended dietary intake (RDI) of iodine for pregnant and breastfeeding women is 220 μg/day and 270 μg/day, respectively [27], which is significantly higher than the RDI of 150 μg/day for adults including non-pregnant women (**Table 2.1**).

Table 2. 1 Reference iodine intake by life stage in Australia [27].

Life stage	RDI (μg/day)	EAR (μg/day)	UL (μg/day)
Infants			
0-6 month	90 (AI)		
7-12 month	110 (AI)		
Children			
1-3 years	90	65	200
4-8 years	90	65	300
Adolescents			
Boys			
9-13 years	120	75	600
14-18 years	150	95	900
Girls			
9-13 years	120	75	600
14-18 years	150	95	900
Adults	150	100	1100
Pregnant women	220	160	1100[1]
Lactating women	270	190	1100[1]

AI, adequate intake; EAR, estimated average requirement; RDI, recommended dietary intake and UL, upper level of intake.

[1]The UL is 900 μg/day for pregnant and lactating women aged 14-18 years [27].

Iodine is a trace element mainly found in the shallow layers of soil and prone to washout from the soil [1]. The natural iodine content of food is influenced by the iodine content of the environment where the food is grown. Most foods are naturally low in iodine content; only a few foods are good sources of iodine such as fish and seafood, eggs and milk products [28, 29]. Therefore, the iodine requirements may be challenging to achieve in people who consume few or none of the iodine-rich food.

Iodine can be obtained from food fortified with iodine to fulfil the daily iodine requirements. Universal salt iodisation (USI) and mandatory iodine fortification of bread or other condiments are strategies recommended by the World Health Organization (WHO) to improve the iodine status of the general population. In countries that practise the USI program, all salt available on the market for human consumption, as well as feed stock for animals, need to be iodised, at a level recommended by the WHO and ICCIDD [12]. Similarly, in the mandatory iodine fortification of bread, all bread makers are mandated to use iodised salt with a few exceptions [30]. Besides salt, iodine can also be fortified with other food, beverages and condiments [31]. Iodine supplementation is also a valuable source of iodine in vulnerable groups, which include infants, pregnant women, and breastfeeding women.

As the iodine requirement is higher in pregnant women than the general population, they will significantly benefit from the USI designed for the general population or the targeted iodine supplementations during pregnancy. If population-wide iodine deficiency persists (median UIC <50 μg/L in SAC or goitre prevalence >20% in SAC) following the introduction of the USI, the WHO and UNICEF jointly recommend intake of iodine-containing supplements for all pregnant and breastfeeding women, as well as children 7-24 months of age [32]. For example, with the absence of a USI program and the decreased iodine intake in pregnant

women, the American Thyroid Association recommended intake of iodine-containing supplements for pregnant and breastfeeding women in the US and Canada [33].

2.4 Assessment of iodine status

2.4.1 Assessment of population-level iodine status

Iodine status of populations can be assessed using a physical examination of the thyroid gland and biochemical markers including UIC, as well as TSH concentration [12].

Goitre or thyroid volume by ultrasound

Enlargement of the thyroid gland reflects longer-term exposure to low iodine intake [12]. Even though palpation of the thyroid gland is a commonly used technique, the WHO also recommend ultrasonography to measure the thyroid size. Therefore, goitre or thyroid volume by ultrasound can be used as epidemiological markers to assess or monitor population iodine status.

As shown in **Table 2.2**, iodine deficiency is classified by the WHO as a mild, moderate and severe public health problem when the prevalence of goitre is 5-19%, 20-29% and ≥30% in SAC, respectively. Goitre is also defined as thyroid volume >97th percentile in SAC [34]. The percentiles of thyroid size by ultrasound were obtained from iodine sufficient SAC worldwide.

Prevalence of goitre by palpation or ultrasound is recommended to be used for assessing baseline levels of IDD and the long-term impact of iodine intervention programs [12, 35]. Assessment of goitre by palpation or ultrasound is non-invasive, simple and quick [36]. However, goitre is not the appropriate marker for monitoring iodine status shortly after iodine interventions, since it takes longer for changes in the prevalence of goitre to occur [35].

Goitre is not a good marker in neonates and young children since the gland is too small to be detected by inspection and palpation in these groups of the population. SAC are important target groups for assessment of goitre prevalence because of their vulnerability, ease of access and representativeness of the population [37]. Measuring thyroid size by ultrasound is a more precise method, but it needs an experienced assessor [38]. Therefore, unlike goitre prevalence by palpation, measuring thyroid size by ultrasound may not be feasible to assess IDD in field surveys of low-income countries.

Urinary iodine concentration

UIC reflects recent iodine intake in populations over a short period of time. UIC is not normally distributed in the population and the median value of UIC is used to classify population iodine status [12, 37]. Median UIC has been recommended by the WHO, UNICEF and ICCIDD to be used as a first choice to assess population-level iodine status [12] and becomes the most widely used marker to assess iodine status of populations [17, 29, 39, 40]. UIC is convenient and cheap marker, since only a small amount of urine (0.5-1.0 ml) is needed, and the assessment can be done with simple methods using the Sandell-Kolthoff reaction.

As depicted in **Table 2.2**, populations are classified as iodine sufficient if the median UIC is ≥100 μg/L in SAC [12]. This criterion was established in 1994 from the lowest level of 24 hours (24-h) urinary iodine excretion (UIE) associated with the absence of goitre, a functional outcome of iodine deficiency, from a study conducted in Central America [37]. The study showed that areas with 24-h UIE >100 μg/day had no goitre in adults and SAC [41]. As the average 24-h urine volume in SAC is 1 L, the UIC in SAC would be comparable with 24-h UIE in SAC. Hence, the WHO recommends using UIC instead of 24-h UIE in population surveys [37]. To obtain a representative national sample of SAC, the WHO recommends selecting at least 1200 SAC from 30 clusters and at least 40 SAC per cluster to classify

population iodine status using the median UIC. However, cluster sampling technique may not be feasible when including populations of pregnant women in household-based surveys as the number of pregnant women may be small in the clusters [12].

Table 2. 2 Criteria for assessing iodine status in populations[1]

		Iodine deficiency		
	Iodine sufficient	Mild	Moderate	Severe
Proportion of neonatal TSH >5 mIU/l (%)	<3	3-19	20-39	≥40
Median UIC in SAC (µg/L)	≥100	50-99	20-49	<20
Goitre prevalence in SAC (%)[2]	<5	5-19	20-29	≥30

[1]SAC, school age children; TSH, thyroid stimulating hormone; UIC, urinary iodine concentration.

[2]Goitre prevalence assessed by either palpation or ultrasound [12].

Thyroid stimulating hormone concentration

The TSH concentration of newborns can be affected by maternal iodine intake during pregnancy and is used to monitor the iodine status of populations [12, 37, 42]. Newborns have higher iodine turnover than adults due to lower iodine stores in their thyroid gland [25]. Hence, newborns are highly sensitive to even mild iodine deficiency, and newborn TSH concentration can indicate the iodine status of populations [37]. According to WHO, UNICEF and ICCIDD joint criteria, populations are classified as iodine deficient if more than 3% of newborns have TSH concentration >5 mIU/L in whole blood samples collected at 3-4 days after birth **(Table 2.2)** [37].

The TSH marker is practical for monitoring iodine status of countries which have a neonatal screening program to screen newborns for congenital hypothyroidism (CH) [25]. Neonatal TSH concentration may be affected by several factors including the timing of neonatal blood sample collection (early vs late) [43, 44], the type of blood sample (cord vs heel) [45], mode of delivery (vaginal vs caesarean section) [45, 46], gestational age (term vs preterm) [47], birthweight (normal vs low birthweight) [43, 47, 48], sex [48], season of birth [43, 49], use of iodine-containing antiseptics [45], organochlorine pesticides exposure in the perinatal period

[50], and the TSH assay used [44, 51]. Thus, appropriate caution is needed to interpret the findings of TSH as a marker of iodine status in populations. Due to the cost associated with screening all children for TSH levels, this marker is not widely used to report the iodine status of populations in developing countries.

Questions have been raised on whether the current criterion based on the proportion of newborns with TSH concentration >5 mIU/L is appropriate in mildly iodine deficient or iodine sufficient settings [45, 51-54]. Burns et al. [54] and Vandevijvere et al. [51] suggested using the trend over time, rather than a numerical cut-off, to see changes in the iodine status of populations. Evans et al. [53] also suggested the WHO modify the criterion, and recommended using a lower than 5 mIU/L cut-off to identify mild iodine deficiency in populations. The time of blood sample collection and other determinants of newborn TSH concentration needs to be considered in the interpretation of newborn TSH results. Despite the criticism, the proportion of the population with newborn TSH concentration >5 mIU/L has been used as an effective measure of monitoring population iodine status in a number of countries [44, 55-60].

2.4.2 Assessment of individual-level iodine status

Currently there is lack of recommended markers to classify iodine status in individuals; UIC and newborn TSH has been widely used to define individual-level iodine status in populations and at-risk population groups including pregnant women.

Urinary iodine concentration

Although UIC is not a suitable marker to determine the iodine status of individuals due to large day-to-day and within-day variations in urinary iodine excretion [12], the within-day variation in the iodine excretion minimises when assessed as 24-h UIE or as iodine to creatinine ratio (I/Cr) [12, 61]. To overcome the drawbacks of spot UIC, measuring UIC from

repeated urine samples over the 24-h period [62, 63] or adjusting for urinary creatinine concentration [64-66] has been recommended. Creatinine adjusted urinary iodine biomarkers are considered as alternatives to the actual 24-h UIE, since creatinine excretion is constant throughout the day [67]. However, urinary creatinine may vary depending on the age, sex, ethnicity and anthropometric status of individuals, and the different methods that have been used to estimate 24-h urinary creatinine [68, 69]. Although I/Cr and 24-h UIE minimise the within-day variation, these markers do not address the day-to-day variation in iodine excretion.

Despite the fact that spot UIC is not an appropriate marker of individual iodine status [12], several studies applied the WHO cut-off of median UIC <150 μg/L to divide pregnant women into groups of iodine deficiency (UIC <150 μg/L) and iodine sufficiency (UIC ≥150 μg/L) **(Table 2.3)** [37]. This classification has been used frequently to assess maternal exposure to iodine nutrition during pregnancy, and its association with child development [70] and health outcomes [71].

Table 2. 3 Criteria for assessing iodine status in pregnant women[1]

Median UIC in μg/L	Iodine status[2]
>500	Iodine excess
250-499	Above requirement
150-249	Adequate iodine nutrition
<150	Iodine deficiency

[1]UIC, urinary iodine concentration.

[2]The classification of iodine status is based on the WHO criteria [12].

Agreements between the urinary iodine markers in assessing iodine status of individuals

Only a few of the previous studies assessed agreement between spot urinary iodine markers and actual 24-h UIE in populations [66, 72-74], infants and adolescents [75] and lactating women [76]. A comparative study of population iodine status from 400 volunteers aged 18-39 years showed a better agreement between actual 24-h UIE and estimated 24-h UIE than actual 24-h UIE and I/Cr or actual 24-h UIE and UIC [72]. The estimated 24-h UIE was based on

24-h creatinine estimations using published equations by Mage et al. that consider the age, sex, ethnicity and anthropometric status from healthy adult populations [69]. A study on 123 urine samples from 31 hospital staff aged 27-71 years also showed the estimated 24-h UIE had a better agreement with actual 24-h UIE than the agreement between actual 24-h UIE and spot UIC or actual 24-h UIE and I/Cr [66]. In that study, the estimated 24-h UIE was obtained by adjusting the I/Cr by age and sex using Kesteloot and Jansson's [68] published values. The study concluded that both UIC and I/Cr underestimated the iodine values compared with the 24-h UIE.

In populations with different socio-economic status (SES) and anthropometry, I/Cr poorly predicts 24-h UIE unless the necessary adjustments have been made for confounders [73, 76]. A study of 22 healthy adults with 95 24-h urine samples found a higher correlation coefficient for UIC ($r = 0.83$) than crude I/Cr ($r = 0.69$) when both markers were compared with actual 24-h UIE [73]. The significant difference in creatinine excretion in the study population may have contributed to the decreased correlation between I/Cr and 24-h UIE. The crude I/Cr also showed poor agreement with actual 24-h UIE in a study from 16 breastfeeding women with 52 24-h urine samples [76].

There are no published studies examining the agreement between spot urinary iodine biomarkers like I/Cr and estimated 24-h UIE and 24-h UIE during pregnancy. Several studies have assessed iodine status at the first trimester of pregnancy using both the UIC and I/Cr markers [67, 71, 77-81] but none of them reported the agreement between these markers. Determining UIE from 24-h urine collection is labour intensive and needs several assessments to check the completeness of 24-h urine samples [74]. Thus, further study assessing the agreement between urinary iodine markers including UIC, I/Cr and actual or estimated 24-h UIE in pregnant women is required. If spot urine markers have a better agreement with UIE, collecting 24-h urine samples may not be required to assess individual-level iodine status.

Thyroid stimulating hormone concentration

As the newborn TSH concentration reflects the levels of thyroid hormones synthesised by the thyroid gland during pregnancy [20], and it can be used as a marker of iodine nutrition in early life [48]. The TSH concentration of newborns will be high in early neonatal life when there is insufficient iodine intake during the late pregnancy period [42, 48]. Moreover, excessive iodine intake during pregnancy may also increase the newborn TSH concentration, as infants are sensitive to inhibitory effects of excess iodine intake during pregnancy [82]. The classifications of iodine status were the same based on newborn TSH and UIC in pregnancy in few previous studies that measured both newborn TSH and median UIC in pregnant women from the same area and at the same time period [83-85]. Therefore, newborn TSH could be used as a marker of iodine nutrition during pregnancy [36]. In addition to newborn TSH's role in screening children with CH, it may also help to identify children at risk of impaired growth and neurodevelopment as there is existing evidence showing associations between borderline newborn TSH concentration below the newborn screening cut-off, and developmental delay and poor school achievement [86].

2.4.3 Other markers of iodine status

Thyroid hormones, T3 and T4 are more expensive and less sensitive methods of measuring iodine concentration and are not recommended by the WHO to assess the iodine status of populations [12, 87]. Serum thyroglobulin (Tg) is also considered to be a good indicator of iodine nutrition in children [12, 88] and adults [89]. Based on a study by Zimmerman et al. [88], populations are classified as iodine sufficient if the median Tg is <13 μg/L in SAC. Unlike the thyroid hormones that are suggested as markers of iodine status in individuals, thyroglobulin can be used as a functional marker of long-term exposure to low or excessive iodine intake both in individuals and at population-level.

In summary, multiple markers including median UIC in SAC, prevalence of goitre in SAC, and the percentage of newborns with TSH concentration >5% have been recommended to assess iodine status in populations. However, these markers have different applications and target groups. Whilst each marker has its own strength, there are situations where the markers may not be applicable, such as the assessment of recent iodine intake by goitre prevalence. Hence, it is important to apply the most suitable marker of population iodine status, when appropriate. These biomarkers are also helpful in assessing exposure to low or excessive iodine intake at individual-level, especially in vulnerable groups including pregnant women and infants.

2.5 Iodine nutrition status worldwide

Globally, IDD was a public health problem in 118 countries and approximately 1.6 billion people were living in an iodine deficient environment at the beginning of 1990s [90]. By 2004, due to the widespread implementation of the USI, the number of countries with IDD had reduced to 54 countries [91]. The role of USI to improve iodine status and prevent mental retardation in children is one of the great public health triumphs of the 21st century.

Multiple studies have demonstrated that the USI was effective in preventing and controlling IDD in general populations. A systematic review of Randomised Controlled Trials (RCTs) and observational studies showed that salt iodisation increased UIC, increased I/Cr and decreased goitre prevalence in adults and SAC [92]. Studies published after that systematic review also showed that iodine intervention programs were successful in improving the iodine status of the general population in a number of countries [93, 94]. A multicentre cross-sectional study on SAC in three countries (China, Philippines and Croatia) that implemented USI reported adequate iodine intake in China and Croatia but excessive iodine intake in the Philippines [93]. This multicentre study found a sufficient iodine status in women of reproductive age in all three countries. Doggui et al. [94] also reported iodine sufficiency

(median UIC of 220 μg/L) in SAC after two decades of USI in Tunisia. Zimmerman et al. [55] reported that the USI program in Switzerland in 1999 decreased the proportion of newborns with a TSH concentration >5 mIU/L from 2.9% to 1.7%.

The USI is designed for general populations and may not be enough for pregnant women. If the iodine level in the USI is high enough to fulfil the high requirement of pregnant women, there will be a high risk of excessive iodine intake when using for young children. Using surveys of iodine nutritional status in pregnant women, several countries documented improved iodine status in pregnant women following the introduction of the USI for the general population [29, 84, 93, 95]. For instance, the salt iodisation program in India [95] and Switzerland [55] prevented iodine deficiency in pregnant women. A cross-sectional study in India revealed adequate iodine status (median UIC of 172 μg/L) of pregnant women after the salt iodisation program [95]. Likewise, the salt iodisation program of Switzerland in 1999 also improved the iodine status of pregnant women from 115 μg/L in the pre-fortification period to 249 μg/L in the post-fortification period [55]. Also, iodine supplementation during pregnancy has improved the iodine status of pregnant women in countries with mild to moderate iodine deficiency [96, 97].

In contrast, maternal iodine deficiency during pregnancy remains a public health problem in many countries, including Denmark [98], Croatia [93], Turkey [99], Madagascar [100] and the United States [101, 102], even with USI or iodine fortification programs. This may be partly due to inadequate monitoring of iodine status in pregnant women in many countries, although even mild-to-moderate iodine deficiency during pregnancy is linked to poor developmental outcomes later in life.

Despite global initiatives to eliminate IDD and evidence of improvement in iodine status of populations, it remains a public health problem worldwide. Recently, iodine deficiency has

been reported in a number of industrialised countries [7, 8] and iodine deficiency remains a significant public health issue in both the developed and developing worlds. A recent IGN report using nationally representative data up until 2018 documented either iodine deficiency or excess in several countries [7]. Several factors may contribute to the delay in global progress towards eliminating all forms of IDD. These include a low coverage and access of iodised salt or iodine fortified food, lack of commitment from the government to take legal action against manufacturers who are not compliant, lack of integration of iodine interventions with other public health programs like salt reduction, low awareness about iodine nutrition and its consequences, lack of surveillance systems, and resistance of governments to introduce iodine intervention programs [103, 104]. Discrepancies in the iodine status of populations according to different markers is also a challenge to properly monitor and take appropriate public health measures [105]. In recent years, the epidemiology of iodine deficiency has shifted from severe forms of iodine deficiency to mild-to-moderate forms of iodine deficiency and excessive iodine intake [9-11].

Though the public health importance of the USI outweighs its adverse effects at a population level, there are few countries that has been classified with excessive iodine intake following the implementation of the USI [7]. The USI program may also contribute to the rise in cardiovascular mortality associated with increased salt intake unless the USI program is integrated with the public health measures to reduce salt intake [106]. Therefore, a strong monitoring and evaluation system should be in placed to monitor the iodine as well as the salt intake of individuals from the iodised salt.

2.6 Iodine nutrition status in Australia

Mandatory iodine fortification of bread was introduced in Australia in October 2009 [16]. The program was in place after the results of the national survey that revealed the re-emergence of iodine deficiency [17]. The fortification program aimed to prevent iodine deficiency in the

general population and mandated the use of an iodised salt instead of non-iodised salt in bread-making, with a few exception such as organic bread. Whilst the type of iodine could be potassium/sodium iodide/iodate, the recommended level of iodine is 25-65 mg of iodine per kg of salt [30]. In January 2010 the fortification program was followed by the National Health and Medical Research Council recommendation that pregnant and breastfeeding women take an iodine supplement of 150 µg/day [107]. This recommendation was partly due to concerns about mandatory fortification not being enough to satisfy the iodine requirements of pregnant women, and fear that even mild iodine deficiency during pregnancy may have adverse health outcomes [108]. Despite the fact that there is the lack of consideration and data on food patterns of pregnant women or women of child bearing age, encouraging the intake of iodine rich food may be a sustainable approach to prevent iodine deficiency in all population groups.

2.6.1 Iodine status pre-fortification in the general population

Studies that used UIC marker

The national iodine nutrition survey conducted in 2003-2004 showed that the population weighted median UIC of SAC in mainland Australia was 98 µg/L with significant geographical variation in major states [17, 109]. The change in sanitary practices in the dairy industry, from iodophors to other chemicals, and the use of non-iodised salt instead of iodised salt were suggested as causes for insufficient iodine intake in Australians [81]. Other studies in pre-fortification periods that used median UIC in SAC and adults also reported mild iodine deficiency in several states in Australia, including Victoria [110], New South Wales [81, 111-113] and Tasmania [114]. All of these studies were conducted in different states of Australia before the national iodine nutrition survey, in a sample size ranged from 94 to 607 participants.

Studies that used newborn TSH marker

Population studies using newborn TSH concentration in Victoria [115], Western Australia [52] and New South Wales [116] showed mild iodine deficiency consistent with results from median UIC in SAC and adults. Moreover, a cross-sectional study in northern Sydney (n = 84) in pre-fortification periods, showed mild iodine deficiency using the proportion of the population having newborn TSH concentration >5 mIU/L [117].

Overall, studies using either median UIC or newborn TSH showed iodine deficiency in Australians in the pre-fortification periods.

2.6.2 Iodine status post-fortification in the general population

Iodine status by UIC marker

Although Australia was classified as borderline iodine deficient in 2003-2004 [17], the iodine status of Australian children and non-pregnant women has changed to sufficient since the mandatory iodine fortification of bread in October 2009 [118]. Difference in iodine status of the population across all states and between major cities and inner regional areas has been observed in the Australian Health Survey in 2011-2012 [118]. This is the only national survey of iodine status of Australians since the fortification program took place. The samples were nationally representative and were collected using a multistage area sample design. Adult populations of Western Australia recorded the highest median UIC in the national survey post-fortification, followed by the Northern Territory and South Australia, respectively; while adults in Tasmania had the lowest median UIC [119]. However, UIC data were not collected in a representative sample of highly vulnerable population groups including infants, pregnant women and breastfeeding women in the Australian Health Survey.

Iodine status by newborn TSH marker

A study in Western Australia showed mild iodine deficiency after two years of the fortification program using TSH as a marker [52]. In the Western Australia study, newborns conceived after fortification were grouped with those born after fortification (they used June 2010 as a cut-off). This may be the reason why the post-fortification group was classified as iodine deficient. The authors also reported a difference in the proportion of the population having TSH concentration >5 mIU/L while using blood samples collected at 3 and 4 days after birth, such that the proportion was higher in day 3 samples. These studies are reported by researchers and there has been no nationally representative survey in Australia before and after the implementation of the fortification program using newborn TSH as a marker.

Monitoring of the iodine status of the population using multiple markers will contribute to early detection of IDD and enable appropriate public health measures to be put in place. However, little evidence is available as to whether overall Australia and South Australia, in particular, are iodine sufficient post-fortification by multiple markers.

2.6.3 Iodine status in pregnant women in Australia: pre- and post-fortification

As shown in **Table 2.4,** studies before the introduction of mandatory iodine fortification of bread, and the routine recommendation of taking iodine supplements, showed iodine deficiency in pregnant women in New South Wales [40, 81, 117, 119-121] and other states [122-124]. All of these studies showed mild to moderate iodine deficiency, in which the median UIC ranged from 24 μg/L to 109 μg/L. All studies were conducted using small sample sizes of less than 300 participants, except for a study in New South Wales by Travers et al. [40] using 796 participants. There had been no nationally representative studies of iodine status in pregnant women before mandatory iodine fortification of bread, and the extent of iodine deficiency among pregnant Australian women was not known.

Following the introduction of the two strategies between 2009 and 2010 to improve iodine status, there was widespread renewed interest in the iodine status of pregnant women in Australia. Discrepant results were reported with iodine sufficiency reported in studies from South Australia [29] and New South Wales [119], but a deficiency in four studies [39, 125-127], including a small study (n = 196) by Clifton et al. [39] in northern Adelaide among women of low SES, and a small number of pregnant women (n = 53) who participated in the Australian Health Survey in 2011-2012 [125]. The reason for the different levels of iodine status in pregnant women across states in Australia remains unclear. Disparities in dietary pattern and in the number of pregnant women taking iodine-containing supplements may contribute to variations in the median UIC across different regions. Pregnant women who did not take iodine-containing supplements were classified as iodine deficient with median UIC <150 μg/L post-fortification [39, 119]. The cost of iodine-containing supplements may constrain pregnant women from taking these supplements. However, women with the lowest SES had higher median UIC compared with women from the highest SES in the Australian Health Survey [125]. The success of both strategies to improve the iodine status of pregnant women remains uncertain in Australia.

Table 2. 4 Studies on iodine status of pregnant women in Australia in the pre- and post-fortification periods[1]

	Author, year	Study design	Setting, year and GA of urine sample collection, sample size	UIC Assessment methods	Median UIC (IQR) in μg/L	Limitations/gap of the study/comment
1	Hynes, 2019 [127]	Cross-sectional survey	Southern Tasmania March 2014 to November 2015 All GA eligible n = 255 Trimester 1 (0-13 weeks) n = 42 Trimester 2 (14-26 weeks) n = 144 Trimester 3 (≥27 weeks) n= 69	Spectrophotometric method	All GA: 133 (82, 233) Trimester 1: 161 (92, 233) Trimester 2: 137.5 (81.5, 243) Trimester 3: 119 (78, 185)	Small sample size at each trimester
2	Singh, 2019 [126]	Analysis of cross-sectional data from a prospective cohort	Northern Territory 2013-2015 n = 24	Spectrophotometric method	93 (62, 171)	Small sample size
3	Condo, 2017 [29]	Analysis of cross-sectional data from a prospective cohort	South Australia August 2011 to December 2012 Assessments at study entry (<20 weeks) and 28 weeks of gestation n = 783	Spectrophotometric method for UIC	Median UIC 20 weeks GA: 189 28 weeks GA: 172	
4	Australian Institute of Health and Welfare, 2016 [125]	Australian health survey	All states of Australia 2011-2012 Small number of pregnant women from the Australian health survey All GA aged 16-44 years n = 53	Not reported	116	Small sample size GA not reported Specific states not reported

5	Clifton, 2013 [39]	Multiple cross-sectional survey	Northern Adelaide, South Australia January 2009 to July 2010 GA at urine sample collection: 12, 18, 30, 36 weeks n = 196	Spectrophotometric method	12 weeks: 73 18 weeks: 68 30 weeks: 84 36 weeks: 118	In women of low socio-economic status
6	Charlton, 2013 [128]	A before-after survey	Regional area of NSW 2008 (n = 110) 2011 (n = 106) 2012 (n = 95) All GA	Spectrophotometric method for	2008: 87.5 (62, 123.5) 2011: 145.5 (91, 252) 2012: 166 (97, 237)	
7	Blumenthal, 2012 [129]	cross-sectional	North Western Sydney, NSW November 2007 to March 2009 GA: 7-11 weeks n = 367	ICPMS	81 (41, 169)	
8	Rahman, 2011 [130]	Cross-sectional analysis of a cohort study	Gippsland, Victoria January 2009 to February 2010 GA ≥28 weeks n = 86 (pre-fortification n = 24, post-fortification n= 62)	Spectrophotometric method	Pre-fortification: 96 (45, 153) Post-fortification 95.5 (60, 156)	Small sample size
9	Mackerras, 2011 [124]	Cross-sectional analysis of Aboriginal Birth Cohort Study	Darwin region, Northern Territory 1987-1990 n = 24	Spectrophotometric method	49 (40, 72)	Small sample size GA not reported
10	Charlton, 2010 [121]	Cross-sectional	Illawarra region, NSW August to September 2008 All GAs n = 110	Spectrophotometric method	87.5 (range 18-325 μg/L)	

11	Burgess, 2007 [122]	Cross-sectional	Tasmania Royal Hobart Hospital Pre-voluntary fortification October 2000 to September 2001 n = 285 Post-voluntary fortification 2006 n = 229	Spectrophotometric method	Pre-voluntary fortification 76 (43,189) Post-voluntary fortification 2006: 86 (57, 160)	Voluntary iodine fortification of bread was in place in October 2001.
12	Travers, 2006 [40]	Cross-sectional	Central Coast area, NSW March to June 2004 GA ≥28 weeks n = 796	Spectrophotometric method	85 (range 19-1510 μg/L)	
13	Hamrosi, 2005 [123]	Cross-sectional analysis of a prospective cohort study	Melbourne, Victoria 1998 - 2002 Caucasian n = 277 Indian/Sri Lankan n = 262 Vietnamese n = 263	In-house semi-automated	Caucasian: 52 (28, 80) Vietnamese: 58 (35, 91) Indian/Sri Lankan: 61 (30, 95)	Participants were from a Down Syndrome screening study
14	McElduff, 2002 [117]	Cross-sectional	Northern Sydney, NSW n = 84	ICPMS	109 (65, 168)	Small sample size GA not reported
15	Li, 2001 [120]	Cross-sectional	Sydney, NSW 1998-1999 Term pregnant women n = 101	In-house semi-automated	88 (range 20-448 μg/L)	
16	Gunton, 1999 [81]	Cross-sectional	Sydney, NSW 1998-1999 n = 81	ICPMS	104 (89, 129)	Participants were from high-risk pregnancy clinic

[1]Values are median (IQR) if not specified. GA, gestational age; ICPMS, Inductively Coupled Plasma Mass Spectrometry; IDD, iodine deficiency disorders; IQR, interquartile range; n, sample size; NA, not reported; NSW, New South Wales; UIC, urinary iodine concentration.

In summary, even though iodine intervention programs have been successful in improving the iodine status of the population in general [131], the effect of both interventions to improve the iodine status of pregnant women in Australia is inconclusive. Frequent monitoring of iodine status in pregnant women is important to prevent adverse effects of both inadequate and excessive iodine intake [132]. There is no nationally representative data on the iodine status of pregnant women in Australia, either overall or from low SES areas, using a representative sample.

2.7. Association of iodine nutrition in pregnancy (as indicated by newborn TSH) with child neurodevelopment and growth

This part of the literature review is presented in a systematic review form according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Checklist [133]. The protocol of this review is published in PROSPERO with a registration number of CRD42020152878 [134].

Abstract

Iodine nutrition during pregnancy can affect newborn TSH concentration (TSH). Studies that assessed associations of newborn TSH with the neurodevelopment and growth of children are inconsistent. The literature on associations between newborn TSH and neurodevelopment and growth was systematically reviewed. Databases including PubMed, Scopus, CINAHL, Embase, PsycINFO, WHO and Iodine Global Network were searched for eligible studies. Seventeen studies were included. The studies applied different ways of newborn TSH classification to indicate iodine deficiency during pregnancy. Neurodevelopment was assessed using different tools in children aged 1-12 years of old. Most of the included studies had an inadequate adjustment for confounding or not adjusted for confounding. The associations between newborn TSH and cognitive development were negative in studies from iodine deficient populations, while null or negative associations were found in studies from iodine

sufficient populations. A null association was reported between TSH and psychomotor development regardless of iodine status. There was no evidence of an association between newborn TSH and child anthropometry, but evidence of negative association was found between newborn TSH and birthweight. The associations between newborn TSH and neurodevelopment may differ based on the iodine status of populations. Long-term follow-up studies that utilise newborn TSH data from newborn screening and adjust for potential confounding variables are required to identify the association between newborn TSH and neurodevelopment and growth in all settings.

Key words: TSH, iodine, growth, development, child

2.7.1 Introduction

Thyroid Stimulating Hormone (TSH) is regulated by the hypothalamic-pituitary axis and modulates the synthesis of thyroid hormones thyroxine and 3, 5, 3/ tri-iodothyronine through a negative feedback mechanism [135]. Low thyroid hormone production increases TSH to stimulate the synthesis of thyroid hormones and vice versa. Thyroid dysfunction can affect thyroid hormone level, but both iodine deficiency and excess can also affect thyroid hormone synthesis [20]. Therefore, both thyroid dysfunction [136] and poor iodine nutrition during pregnancy [137] may lead to elevated TSH in pregnant women as well as their newborns.

Newborn TSH concentration is used to screen children with congenital hypothyroidism (CH), mainly caused by inadequate thyroid hormone secretion in infants due to structural abnormality in their thyroid gland [138]. It is well known that children with untreated CH are at higher risk of impaired neurodevelopment and growth as thyroid hormones are required for optimal growth and neurodevelopment in early life [139].

Evidence showed elevated TSH even below the CH cut-off, which is TSH below <15 mIU/L from whole blood [140], associated with poor child neurodevelopment [86] and anthropometry [141]. Transient increase in newborn TSH concentration may also affect neurodevelopmental outcomes [142]. Identification of children at risk of impaired neurodevelopment may help for subsequent developmental screening and treatment [143]. Hence, newborn TSH concentration may be used as a potential screening tool for identifying children at risk of impaired neurodevelopment and growth [144]. Newborn TSH data can be obtained from routine newborn screening databases in countries where newborn screening programs are practised.

To the best of our knowledge, a review by Bougma et al. [145] was the only published review that assessed associations between newborn TSH and neurodevelopment of children under

five years of age, but this review did not include the associations of newborn TSH with child growth. The review showed that children with elevated TSH or with CH had poorer mental development compared with children with normal TSH or without CH. While the review included a meta-analysis, half of the studies (4/8 included studies) were on children born with CH. Hence, the relationship between newborn TSH and neurodevelopment in children without CH is uncertain as there may be different causes for elevated TSH. That review combined studies that used infant blood TSH and cord blood TSH as well as newborn TSH and infant thyroxine in the meta-analysis. A subgroup analysis including by the degree of iodine status (iodine deficiency, sufficiency and excess) was also not performed [145].

Whether the association between newborn TSH and child development differ by iodine status of populations remain unclear [143]. Furthermore, the newborn TSH cut-off associated with poor developmental or anthropometric scores is unclear. Hence, this systematic review examined the association between newborn TSH concentration and child neurodevelopment and growth in different settings.

2.7.2 Methods

Study Selection

Studies that reported the relationship between newborn TSH concentration (from either infant or cord blood) and the development or education performance or growth of children, were included. Participants were children up to 18 years of age. Studies assessing the association between TSH concentration and development and growth in children with newborn errors of metabolism including CH were excluded. Studies with newborn TSH data obtained from either the newborn screening database or measured as part of the studies were eligible. Neurodevelopmental outcomes included cognitive, motor and behavioural outcomes. Studies were included in anthropometry outcomes if they reported any measures of weight, length,

head circumference, waist circumference and skinfolds, or converted such measures into z-scores.

Data sources and Search strategy

A systematic literature search was conducted in June 2018, then last updated in August 2019. Databases including PubMed, Embase, PsycINFO, Scopus, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and WEB OF SCIENCE were searched and the search was limited to the English language. The World Health Organization (WHO) and Iodine Global Network databases were also reviewed. The key terms applied in database search included: TSH, thyrotropin, thyroid stimulating hormone, iodine, iodides, iodate, iodine deficiency, iodine status, transient neonatal hyperthyrotropinemia and hypothyroxinaemia, child development, Bayley outcomes, cognition, intelligence, motor skills, psychomotor performance, IQ, neurological development, cognition disorders, cognitive function, neural development, intellectual disability, post-natal/postnatal development, memory, executive function, mental process, learning, child behaviour, developmental delay, brain, language, growth, child anthropometry, weight, height and head circumference. The search was conducted using the predefined search strategy in all selected databases. Details of the search strategies for each database are depicted in **Supplemental Table 2.1**.

Data extraction

All retrieved articles were exported to an Endnote Library. After removing duplicate records, eligibility of articles was screened by title and abstract (by MMW). A further review of full-text articles was performed to determine eligibility in studies with insufficient information in the abstract to judge for eligibility. Uncertainties in eligibility were resolved by discussion with a co-author (SJZ). MMW extracted the data using a data extraction form. Data on the first author, year of publication, the country where the study was conducted, study design, iodine status of the study population, sample size, exposure assessment (the type of blood for

newborn TSH assessment and age at blood sampling), outcome assessment (type, measurement or scale and age at assessment), measures of association and effect size (small and adequate), and information on adjustment of confounders were extracted. The key confounders of neurodevelopment and growth identified from the literature include maternal socio-economic status, maternal smoking, home environment, maternal IQ, maternal body mass index and dietary intake, gestational age, birthweight and infant's age at TSH assessment.

Quality of evidence

We used the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) assessment tool [146] to assess the quality of evidence. The ROBINS-I tool has seven domains to assess the risk of bias in included studies. The domains include confounding, selection bias, bias in classification of interventions, bias due to deviations from intended interventions, missing data, bias in the measurement of outcomes and in the selection of the reported result (**Supplemental Table 2.2**).

Data analysis

The analysis was performed according to the analysis plan [134]. A narrative synthesis was performed when data were not suitable for meta-analysis. Studies were classified based on newborn TSH concentration and neurodevelopmental domains. Meta-analysis was performed to assess associations between TSH and outcomes when TSH was treated as a continuous variable or classified based on different categories (dichotomised to <5 vs ≥5 mIU/L or in quartiles. Meta-analysis was performed separately for different neurodevelopmental domains.

Meta-analysis was performed using STATA. Random effects models were used to synthesise results in a meta-analysis as the included studies were heterogeneous. Heterogeneity between the studies was assessed by I^2 and classified as low (I^2 : <25%), moderate (I^2: 25-75%) and

high (I^2: >75%) [147, 148]. If the included studies reported effect measures, we used them in the meta-analysis as reported by the original studies (odds ratio (OR) or Coefficient). When effect measures were not reported, effect size as Hedges' g (g) with 95% CIs was calculated to include the studies in the meta-analysis [149, 150]. The 95% CI of effect size were calculated from p values when any measures of variance were not reported [151]. Classification of effect sizes was based on literature rather than using the small (~0.2 standard deviation (SD)), medium (~0.5 SD), and large (~0.8 SD) classification suggested by Cohen [152]. In nutrition literature, a five-point difference on average in the neurodevelopmental score between children with adequate vs deficient iodine or iron intake was considered as clinically significant [70, 153, 154]. The effect size (g) reported can be converted into points on a standardised neurodevelopmental scale by multiplying with a SD of 15 as a standardiser. The minimum sample size required to detect a five-point (or standardised neurodevelopment points = 0.33) difference (SD = 15 points, two-tailed $\alpha = 0.05$ and Power = 90%) in a standardised neurodevelopmental scale between children with normal TSH or adequate maternal iodine intake during pregnancy vs elevated TSH or deficient/excess maternal iodine intake during pregnancy was 388 participants. Therefore, studies that had a sample size of less than 388 participants were considered as small.

The risk of bias was reported as low, moderate, serious and critical at domain and study level. Studies were classified as low risk of bias if all domains had a low risk of bias and are considered as comparable to well-performed RCT; moderate risk of bias if all domains had a mixed of low or moderate risk of bias and studies cannot be considered as comparable to a RCT; serious risk of bias if at least one domain had a serious risk of bias; and critical if at least one domain had a critical risk of bias [146].

Narrative synthesis, as well as meta-analyses, were performed according to 1) the types of blood for TSH assessment (infant's blood or cord blood). 2) Level of TSH elevation

(Dichotomised to TSH <5 vs ≥5 mIU/L or normal TSH vs transient neonatal hyperthyrotropinemia (TNH)). TNH was defined as a transient elevation of TSH concentration, even a TSH value higher than the CH cut-off, after 2 days of life but returns to the normal range at re-examination without treatment [155]. Classification of newborns for TNH was based on authors' reports, and the TSH cut-off to classify newborns with TNH may differ across studies. 3) Types of neurodevelopmental scale (cognitive, psychomotor, language or behaviour). If two or more items were reported for each domain of neurodevelopmental assessment, the combined score was used for the meta-analysis. 4) Iodine status of the study population (iodine deficient vs sufficient). We used the country-level iodine status data from the WHO database [156, 157] if available, if not, we used the Iodine Global Network data [158, 159] to classify iodine status of the study populations. If country-level data were not reported in either of the two databases, we used the authors' reports.

2.7.3 Results

2.7.3.1 Characteristics of the included studies

As shown in **Figure 2.1**, a total of 3,984 records were identified in database searching and 17 were eligible for inclusion. As depicted in **Table 2.5**, a majority of the studies reported TSH from the infant's blood and used standard scales to assess neurodevelopment. The age range for the neurodevelopmental assessment was between 1 and 12 years of age, and birth to 6 years of age for the anthropometric assessment.

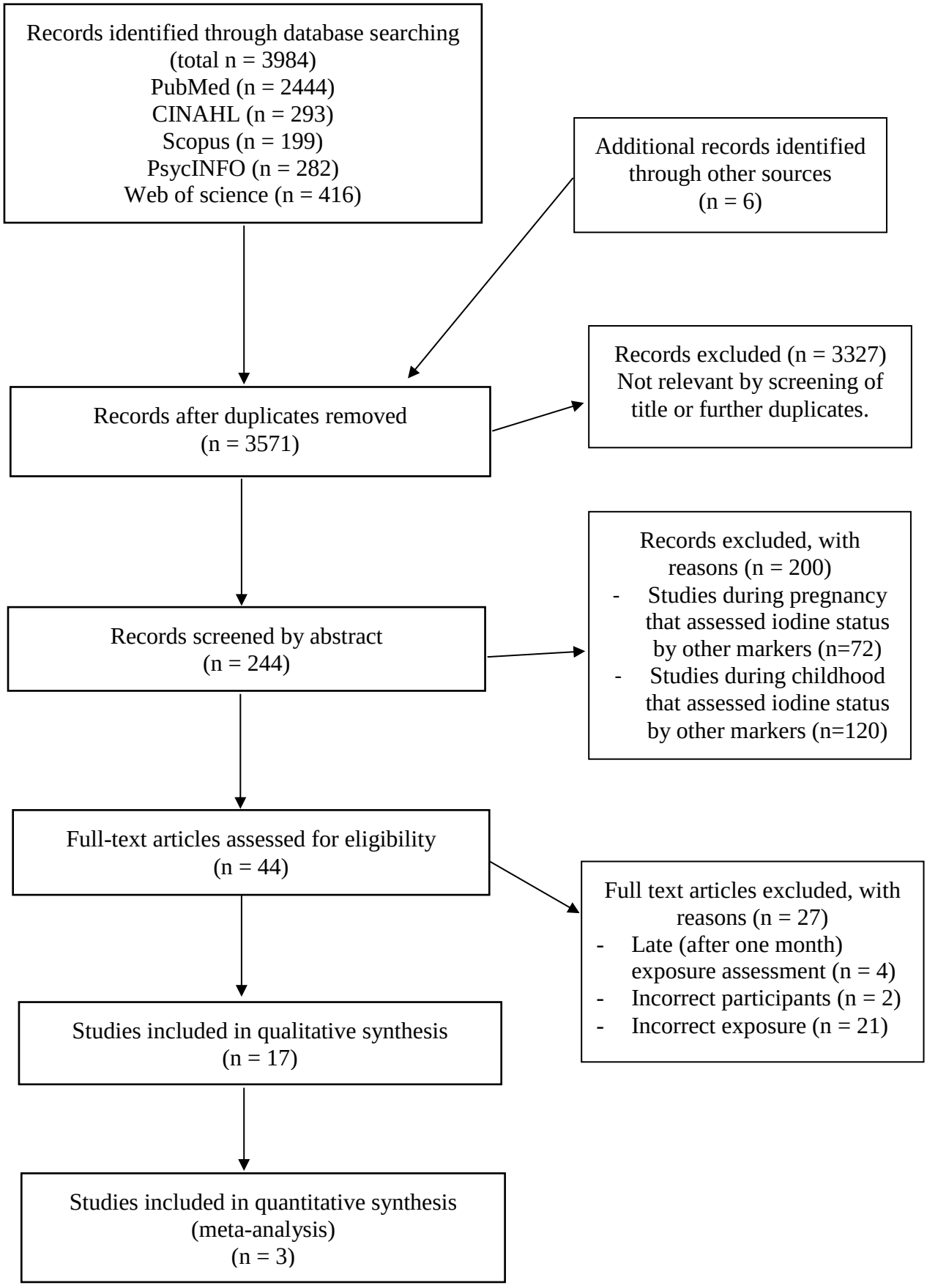


Figure 2. 1 PRISMA Flow Diagram for selection of studies.

2.7.3.2 Risk of bias assessment

Using the ROBINS-I tool, eight studies [86, 160-166] had a moderate risk of bias and nine [141, 142, 167-173] had a serious risk of bias. The main reasons for bias were: 1) did not control confounding or had an inadequate adjustment for confounding 2) used an arbitrary classification of exposure or did not describe the actual TSH cut-off points to define TNH, and 3) did not account for missing data. Full details of the risk of bias assessments are shown in **Supplemental Table 2.3**.

2.7.3.3 Associations between TSH concentration and neurodevelopment

Studies using TSH concentration from newborns blood

Cognitive outcome

As shown in **Table 2.5**, ten studies [86, 141, 142, 160, 161, 167, 169-172] reported associations between newborn TSH concentration and cognitive development. Of studies conducted in iodine deficient populations (5/10), the associations were negative in 4/5 studies [86, 167, 169, 170] and null in one study [160]. Of the studies in iodine sufficient populations (5/10), the associations were null in 3/5 studies [141, 171, 172], negative in one study [142] and both low (<5 mIU/L) and high (>10 mIU/L) TSH was associated with poorer outcome in one study [161].

Only three studies (n ranged from 61 to 691) compared with the mean cognitive scores between two TSH categories: ≥5 mIU/L vs <5 mIU/L. **Figure 2.2** displays a meta-analysis on three studies. In two studies from iodine deficient populations, children with newborn TSH concentration ≥5 mIU/L had 0.41 SD (equivalent to 6 cognitive points) lower cognitive score on average compared with children whose newborn TSH concentration was <5 mIU/L (g: -0.41; 95% CI: -0.74, -0.09; I^2 = 56.0%). However, a null association was found between the two TSH categories (g: 0.22; 95% CI: -0.02, 0.46) in a study from an iodine sufficient

population. In the pooled result across all three studies, there were no difference in mean cognitive scores between the two categories.

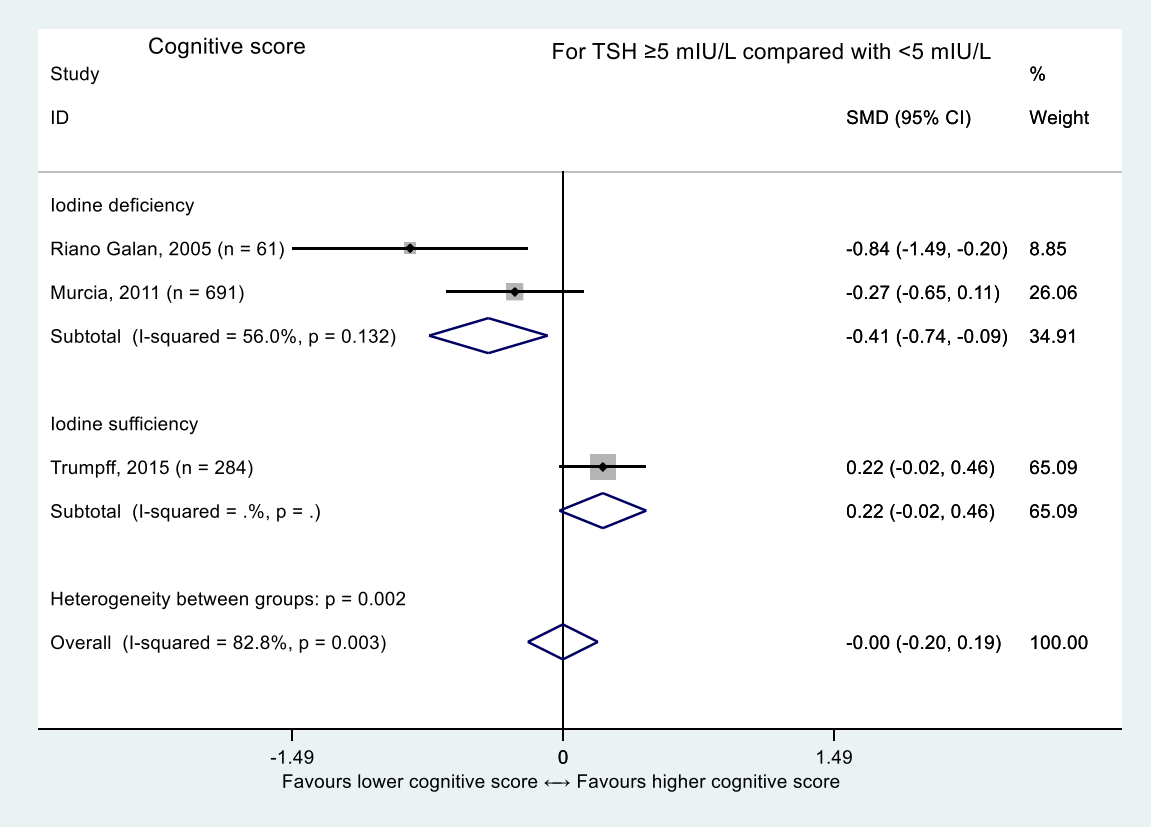


Figure 2. 2 Associations between dichotomised TSH from infant blood (≥5 mIU/L vs <5 mIU/L) and cognitive scores.

CI, confidence interval; SMD, standardise mean difference; TSH, thyroid-stimulating hormone.

Table 2. 5 Summary of key literatures for associations between newborn TSH and neurodevelopment

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
	Studies on Infant blood					
1	Riano Galan, 2005 [167], Northern Spain	Prospective cohort N = 61 Hospital setting Full-term healthy newborns Mild iodine deficiency	3 days of age G1: >5 mIU/L (n = 12) G2: <5 mIU/L (n = 49)	MSCA 3 years verbal, perceptual, quantitative, cognitive general index, memory and motor	MSCA scores: Verbal (± SD): 53.8 ± 7.4 G1 = 49.2 (n = 12), G2 = 56.9 (n = 49); P >0.05 Perceptual (± SD): 52.9 ± 8.7 G1 = 48.6 (n = 12), G2 = 54.2 (n = 49); P <0.05 Quantitative: 46.9 ± 7.2 G1 = 43.7 (n = 12), G2 = 47.3 (n = 49); P >0.05 Cognitive general index (± SD): 103.1 ± 11.6 G1 = 95.1 (n = 12), G2 = 104.9 (n = 49); P <0.05 Memory (± SD): 48.7 ± 8.1 G1 = 43.1 (n = 12), G2 = 50.1 (n = 49); P <0.05 Motor (± SD): 54.6 ± 9.9 G1 = 50.9 (n = 12), G2 = 55.1 (n = 49); P >0.05	Small sample size. Did not adjust for SES, maternal IQ, home-environment.
2	Murcia, 2011 [160], Spain	Cohort N = 691 Children from the INMA population based cohort in Valencia Healthy infants	1-7 days G1: TSH >4 mIU/L (n = 28) G2: TSH ≤4 mIU/L (n = 652)	Bayley-I 1 year Continuous score	MDI (n = 669): G1: 96.2 ± 17.0 G2: 100.2 ± 14.9 PDI (n = 669): G1: 95.8 ± 15.4 G2: 100.3 ± 15.0	Did not adjust for maternal IQ and home-environment. N for Bayley-I assessment was not reported for G1 and G2 separately.

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
		Mild iodine deficiency		Delay (<85) vs Normal MDI PDI	G1: β = -3.2 (-8.9, 2.6) G1: OR = 1.4 (0.5, 3.7)	
3	Belcari, 2011 [169], Italy	Cohort N = 55 Preterm infants admitted to the neonatal intensive care unit (GA 31-32 weeks) Mild iodine deficiency	3-4 days of life G1: TSH >4.3 mIU/L (n = 10) G2: TSH <4.3 mIU/L (n = 45)	Griffiths Scales of Mental Development 18 months corrected age Composite MD	OR (95% CI) G1 (n = 10): 14.6 (2.49, 86.02) Ref: G2 (n = 45)	Small sample size. Confounders like maternal IQ were not controlled.
4	Costeira, 2011 [170], Portugal	Cohort N = 86 Healthy term infants Hospital setting Excluded: use of iodinated antiseptics, endocrine dysfunction, heavy smoking, previous diabetes Moderate iodine deficiency	3 days of age Continuous (n = 86)	BSID-I 18 months MDI	r = -0.27 (n = 86) d= -0.56082901	Small sample size Only correlation test was conducted. Mean TSH and MDI were not reported. Outcome assessment were not blinded.
5	Korzeniewski, 2014 [141], ELGAN Multicentre study, US	Cohort N = 120 Hospital setting	14 days of age G1: newborns with TNH	BSID-II 2 years	OR, 95% CI (n = 120) Mental delay G1: 1.2 (0.8, 1.9) Motor delay	Preterm infants. Late TSH sampling outside the WHO recommended range

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
		Preterm infants (GA <28 weeks) Iodine sufficient population	G2: Normal	Cognitive and Motor developmental delay (scores <55)	G1: 1.3 (0.9, 2.0)	and age at sampling was not considered in the analysis. Did not adjust for SES, maternal IQ and home environment. The TSH value for each quartile cut-points was not reported.
6	Cuestas, 2015 [142], Argentina	Prospective cohort N = 250 Healthy infants Excluded: preterm birth, birthweight <2500 g, Down's syndrome, descendants of mothers with immune thyroid disease, congenital Malformations Iodine sufficient population	2-3 days of age G1: TNH (newborn TSH ≥10 mU/l) (n = 65) G2: normal (newborn TSH <10 mIU/L) (n = 185)	Parents' Evaluation of Developmental Status (PEDS) 6 years Developmental delay	N, % for developmental delay (95% CI) G1: 15/65 (23%, 95% CI 12, 34.1) G2: 21/185 (11.3%, 95% CI 6.5, 16.2) P = 0.03 G1: Crude OR 95% CI 2.0 (1.0, 4.2) P = 0.05 (n = 65) NB: OR was not reported (crude OR calculated manually by the reviewers).	Outcome from parents report. Did not adjust for SES, maternal IQ, and home environment.
7	Trumpff, 2015 [161], Belgium	Retrospective cohort N = 315 Healthy infants from the PsychoTSH study Exclusion criteria: LBW, preterm, CH or	3-5 days of age G1: <5 mIU/L (n = 198) G2: 5-9 mIU/L (n = 102)	WPPSI-III 4-6 years of age IQ	Mean ± SD Full Scale IQ: G1: 95.9 ± 16.9 (n = 198) G2: 99.4 ± 14.2 (n = 102) G3: 88.9 ± 26.1 (n = 11) P = 0.06	On highly educated women (73% of mothers completed university education).

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
		known neurological diseases Iodine sufficient population	G3: ≥10 mIU/L (n = 11)		Performance IQ: G1: 98.0 ± 16.3 (n = 198) G2: 99.4 ± 13.1 (n = 102) G3: 94.5 ± 21.6 (n = 11) P = 0.53 Verbal IQ: G1: 94 ± 19.8 (n = 198) G2: 100.8 ± 14.2 (n = 102) G3: 86.7 ± 22.8 (n = 11) P = 0.006	The effect measures from the regression models were not reported. N for G3 was very small
8	Trumpff, 2016 [168], Belgium	Retrospective Cohort N = 310 Healthy infants from the PsychoTSH study Exclusion criteria: TSH >15 mIU/L, low birthweight, premature Iodine sufficient population	3-5 days G1: ≥5 mIU/L (n = 113) G2: <5 mIU/L (n = 197)	CBCL 4-5 years Psychosocial development	Mean (SD) score G2 = 48.6 (10.3) (n = 197) G1 = 46.9 (9.7) (n = 113) P = 0.15 CBCL, β (SE) G1: -1.57 (1.14) (n = 113) P= 0.17	Parent reported outcome. Did not adjust for home environment and maternal IQ.
9	Trumpff, 2016 [162], Belgium	Retrospective cohort N = 284 Healthy infants from the PsychoTSH study Exclusion criteria: TSH >15 mIU/L, low birthweight, premature Iodine sufficient population	3-5 days Per unit increase in TSH	Charlop-Atwell scale of motor coordination -Average z scores 4-6 years Psychomotor development	β (95% CI) - Total motor score : 0.04 (-0.01 to 0.09) P = 0.10 - Objective motor score: 0.02 (-0.02 to 0.07) P = 0.26 - Subjective motor score: 0.03 (-0.01 to 0.08) P = 0.11	Multiple imputation was performed. Did not adjust for home environment and maternal IQ.

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
10	Lain, 2016 [86], New south wales, Australia	Record-linkage cohort Population based newborn TSH data (1994-2008) N of infants linked with AEDC = 149 569 N of infants linked with NAPLAN = 354 137 Excluded: VLBW (<1500g) Mildly iodine deficient population	Cut-off for CH was 20 mIU/L 3-5 days of age TSH categories based on percentiles: G1 (reference category): <75th (<~4 mIU/L) G2: 99.5th to 99.9th (~7-12.4 mIU/L)	NAPLAN -Poor developmental outcome if the student scores below the national minimum standard. 5-6 years reading or numeracy AEDC "Developmentally high risk" while the developmental scores were below 10% in ≥2 out of the 5 domains. 8-12 years Developmental outcome	Numeracy performance [OR 95% CI] G2: 1.57 (1.29, 1.90) (n = 1378) Reading [OR 95% CI] G2: 1.42 (1.18, 1.69) (n = 1392) Developmentally high risk [OR 95% CI] G2: 1.52 (1.20, 1.93) (n = 510) Special needs [OR 95% CI] G2: 1.68 (1.23, 2.30) (n = 557)	The mean/median score for educational outcomes was not described. National minimum value was not reported. Infants lost to follow up for outcome assessment but unlikely to introduce bias. Developmental outcomes assessed by teachers.

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
1 1	Leviton, 2016 [171], ELGAN Multicentre study, US	Cohort day 21 (N = 749) and day 28 (N = 697) Hospital setting Preterm infants (GA <28 weeks) Iodine sufficient population	21 days 28 days G1: Newborns with TNH G2: normal	BSID-II 2 years MDI (scores <55 and scores 55-69) PDI (scores <55 and scores 55-69)	OR, 95% CI MDI (scores <55) G1 (Day 21): 1.1 (0.7, 1.7) G1 (Day 28): 1.2 (0.8, 1.9) MDI (scores 55-69) G1 (Day 21): 1.1 (0.6, 1.8) G1 (Day 28): 1.0 (0.6, 1.7) PDI (scores <55) G1 (Day 21): 1.1 (0.7, 1.8) G1 (Day 28): 1.3 (0.8, 2.1) PDI (scores 55-69) G1 (Day 21): 1.0 (0.6, 1.5) G1 (Day 28): 1.0 (0.6, 1.6)	Preterm infants. Late TSH sampling outside the WHO recommended range and the late sampling was not considered in the analysis. Did not adjust for SES, maternal IQ and home environment. The TSH value for each quartile cut-points was not reported. N for G1 and G2 was not separately reported.
1 2	Kuban, 2017 [172], ELGAN Multicentre study, US	Prospective cohort N = 873 Hospital setting Preterm infants Iodine sufficient population	First 1 month G1: elevated TSH (n = 164) G2: normal TSH (n = 645)	School-Age differential scales-II (DAS-II) 10 years DASS IQ z score -2 to ≤-1 SD ≤-2 SD	OR (95% CI) DASS IQ -2 to ≤-1 z score G1: 1.0 (0.6, 1.5) (n = 164) DASS IQ ≤-2 z score G1: 1.2 (0.7, 1.9) (n = 645)	On extremely preterm infants. Late TSH sampling outside the WHO recommended age range. Did not adjust for home environment, SES, and Maternal IQ. The TSH value for TSH categories was not reported.

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
	Studies on Cord blood TSH					
1	Calaciura,1995 [163] Sicily, Italy	Retrospective cohort N = 18 Healthy infants Participants were from rural areas Severe iodine deficiency	G1: elevated TSH due to TNH (TSH >50 mIU/L), (n = 9) G2: normal TSH (TSH <10 mIU/L), (n = 9)	WISC-revised 8.5 y IQ (GIQ, VIQ, PIQ)	GIQ G1: 78.3 ± 11.1 (n = 9) G2: 90.9 ± 14.2 (n = 9) P <0.05 VIQ G1: 84.4 ± 15.4 (n = 9) G2: 96.2 ± 14.8 (n = 9) P >0.05 PIQ G1: 75.0 ± 8.5 (n = 9) G2: 89.2 ± 12.5 (n = 9) P <0.01	Small sample size Did not adjust for SES, maternal IQ and home-environment.
2	Azizi, 2001 [164], Iran	Historic cohort N = 37 Healthy infants Hospital based Iodine sufficient population	G1: newborns with TNH (TSH >18 mIU/L), (n = 18) G2: newborns without TNH (TSH <5 mIU/L), (n = 19)	Raven 9 years IQ Bender-Gestalt tests 9 years Psychomotor performance	IQ G1: 98 ± 11 (n = 18) G2: 106 ± 8 (n = 19) P <0.01 Psychomotor performance No correlation between serum TSH and psychomotor performance	Non-representative sample (no sample size calculation). Did not adjust for SES, maternal IQ and home-environment. Effect size was not reported for psychomotor performance.
3	Choudhury, 2003 [165], Northern China	Cohort N = 284 Full-term and normal birthweight infants Community setting	Continuous (all newborns, n = 275) G1 (TSH <5 mIU/L) (n = 58)	FTI-I 7 months Information processing task	Continuous exposure Information processing skills: β = -0.14, P <0.05) 95% CI (-0.21, -0.08) MDI: β = -0.34, P <0.05) 95% CI(-0.50, -0.18)	Follow-up not long enough for outcome to occur.

	Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
		Mild Iodine deficient population	G2 (TSH 10.0-19.9 mIU/L) (n = 94) G3 (TSH 20.0-29.9) (n = 82) G4 (TSH ≥30 mIU/L) (n = 41)	BSID-II 13 months MDI PDI	PDI: β = not reported, P = 0.56 Categorical exposure Information processing skills: G1 (n = 58) = 59.6 ± 3.0 G2 (n = 94) = 58.9 ± 4.3 G3 (n = 82) = 57.7 ± 5.6 G4 (n = 41) = 57.5 ± 3.1 MDI G1 (n = 58) = 102.5 ± 8.2 G2 (n = 94) = 98.2 ± 8.3 G3 (n = 82) = 98.7 ± 9.3 G4 (n = 41) = 93.5 ± 11.1 PDI G1 (n = 58) = 102.7 ± 11.3 G2 (n = 94) = 105.1 ± 12.2 G3 (n = 82) = 107.8 ± 13.9 G4 (n = 41) = 104.8 ± 12.1	
4	Freire, 2010 [166], Spain	Prospective cohort N = 178 Newborns from INMA-Granada population cohort Healthy infants Excluded: mothers with DM, HTN and thyroid disease, non-permeant residents, pregnancy complication	Mean TSH: 3.55 mU/l (range 0.24–17 mU/L) Per unit increase in TSH Categories G1: <2.05 mIU/L G2: 2.05–2.95 mIU/L	MSCA 4 years MSCA scores (general cognitive, verbal, perceptual performance, quantitative, memory, motor	Continuous exposure General cognitive (β) -3.52 (-6.81, -0.23) Memory (OR) 1.61 (0.60, 3.80) Motor skills (OR) 0.65 (0.32, 1.47) Fine motor 1.48 (0.63, 3.66) Gross motor 0.69 (0.26, 1.99) Executive function (β)	None representative sample (only boys). Did not control maternal IQ. N was not reported for each TSH category. More than one effect measure used to report results.

Author, year, Country	Design and setting	TSH assessment	Neurodevelopment assessment	Key Finding	Limitations/gap of the study/comment
	Mild iodine deficiency	G3: 2.96–4.18 mIU/L G4: 4.19–17.0 mIU/L	skills, executive function) Developmental delay vs normal Continuous scores	-3.15 (-6.66, -0.19) Categorical exposure General cognitive (B) G2: -0.36 (-6.29, 5.67) G3: -3.03 (-8.78, 2.74) G4: -5.42 (-11.30, -0.61) Memory (OR) G2: 0.83 (0.13, 3.60) G3: 0.92 (0.18, 3.74) G4: 3.51 (0.65, 15.26) Motor skills (OR) G2: 0.39 (0.11, 1.87 G3: 0.42 (0.13, 1.73) G4: 0.80 (0.24, 3.10) Executive function G2: -0.30 (-6.51, 5.96) G3: -3.37 (-9.46, 2.56) G4: -4.29 (-10.56, 1.86)	

Abbreviations: AEDC, Australian Early Development Census; BSID, Bayley Scales of Infant Development; CBCL, Child Behaviours Check List; CI, confidence internal; CH, congenital hypothyroidism; DAS, Differential Ability Scale; ELGAN, extremely low gestational age newborns; FIQ, full intelligence quotient; G, group; GA, gestational age; ID, intellectual disability; IQ, intelligence quotient; MDI, mental development index; MSCA scores, McCarthy scales of children's abilities, N, sample size; NAPLAN, National Assessment Program Literacy and Numeracy; OR, odds ratio; P, p value; PARCA, Parent Report of Children's Abilities; PDI, Psychomotor developmental index; PIQ, performance IQ; RR, risk ratio; SBIS, Stanford–Binet Intelligence Scale; TNH, transient neonatal hyperthyrotropinemia; VIQ, verbal IQ; WISC, Wechsler Intelligence Scale for children; WPPSI, Wechsler Preschool and Primary Scale of Intelligence

Five studies compared the odds of cognitive delay between newborns who had TNH and newborns without TNH at birth were pooled in a meta-analysis (**Figure 2.3**). Children with TNH had 33% higher odds of developmental delay (OR: 1.33; 95% CI: 1.10, 1.63; I^2 = 12.6%) compared with children without TNH in a pooled analysis of studies from both iodine deficient and sufficient populations. However, 3/5 studies [141, 171, 172] in the pooled meta-analysis were in preterm newborns. Besides the results of developmental delay included in the pooled meta-analysis, the population-based study in Australia [86] showed that infants in the highest percentile of TSH approaching the CH screening cut-off had higher odds of poor educational performance. The study reported that infants whose TSH concentration was between the 99.5th and 99.9th percentile (~7-12.4 mIU/L) had a 1.6 higher odds of poor numeric performance (OR: 1.57; 95% CI: 1.29, 1.90) and a 1.4 higher odds of poor reading score (OR: 1.42; 95% CI: 1.18, 1.69) than infants with <75th TSH percentile (equates TSH concentration <~4 mIU/L).

Two small studies were not suitable for meta-analysis. The first study [169] (n = 55) showed higher odds of composite mental development delay (OR: 14.6; 95% CI: 2.5, 86.0) at 18 months corrected age in children with TSH >4.3 mIU/L compared with TSH <4.3 mIU/L and the other study [170] (n = 86) showed a negative correlation between newborn TSH and mental development index at 18 months corrected age (r = -0.27).

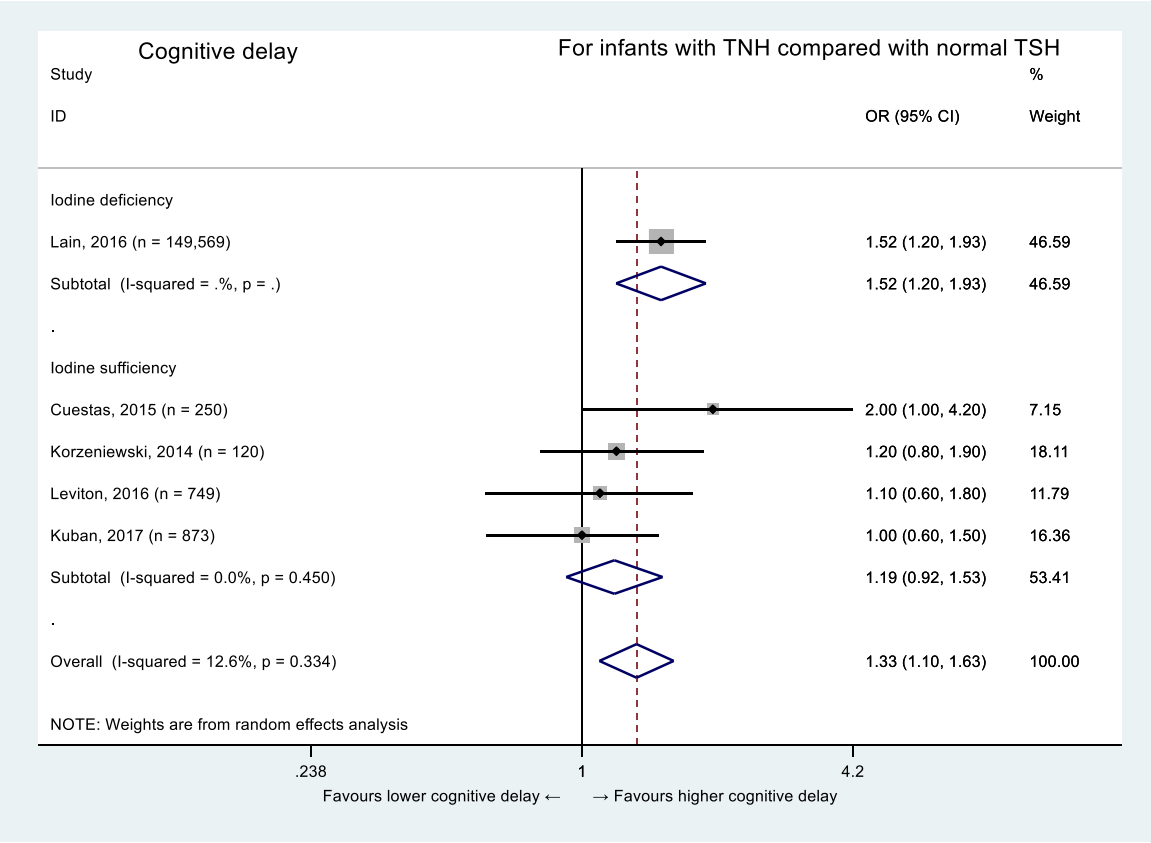


Figure 2. 3 Associations between TNH defined by elevated TSH from infant blood and cognitive delay.

CI, confidence interval; OR, odds ratio; TNH, transient neonatal hyperthyrotropinemia; TSH, thyroid-stimulating hormone.

Psychomotor outcome

All included studies that reported psychomotor outcomes (5/5) showed a null association between newborn TSH concentration and psychomotor development and 3/5 [160, 162, 167] were in full-term infants while 2/5 [141, 171] were in preterm infants. Data were not suitable for one pooled meta-analysis as the included studies assessed newborn TSH in different ways which include <5 mIU/L vs ≥5 mIU/L categories (Riano Galan et al., 2005, Murcia et al., 2011 or TNH vs without TNH [141, 171] or assessing as a continuous variable [162].

Adaptive behaviour and social emotional outcomes

The only study [168] (n = 310) that assessed psychosocial development at 4-5 years of age reported a null association between newborn TSH concentration and psychosocial development.

Studies using TSH concentration from cord blood

Cognitive outcome

All of the studies (4/4) [163-166] reported a negative association between cord blood TSH and cognitive development. Meta-analysis was performed on two studies that compared infants with TNH and without TNH and **Figure 2.4** reports children with TNH had 0.88 SD (equivalent to 13.2 cognitive points) lower cognitive score on average compared with children without TNH (g: -0.88; 95% CI: -1.44, -0.33; I^2 = 0%) irrespective of iodine status in the study populations. Based on the pooled analysis from two studies [165, 166] that analysed TSH as a continuous exposure, a 1 mIU/L increase in cord blood TSH was associated with a 0.25 point lower cognitive score (β: -0.25, 95% CI: -0.49, -0.001) (**Figure 2.5**).

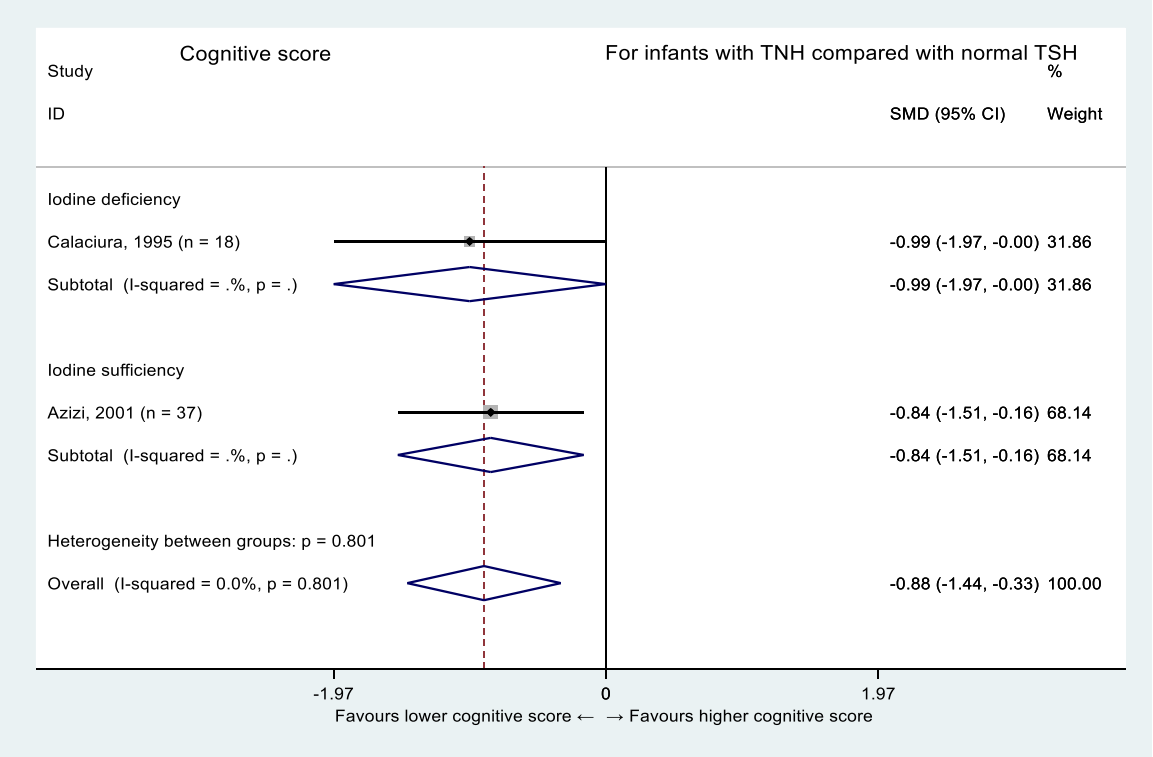


Figure 2. 4 Associations between TNH defined by elevated TSH from cord blood and cognitive scores.
CI, confidence interval; SMD, standardise mean difference; TNH, transient neonatal hyperthyrotropinemia; TSH, thyroid-stimulating hormone.

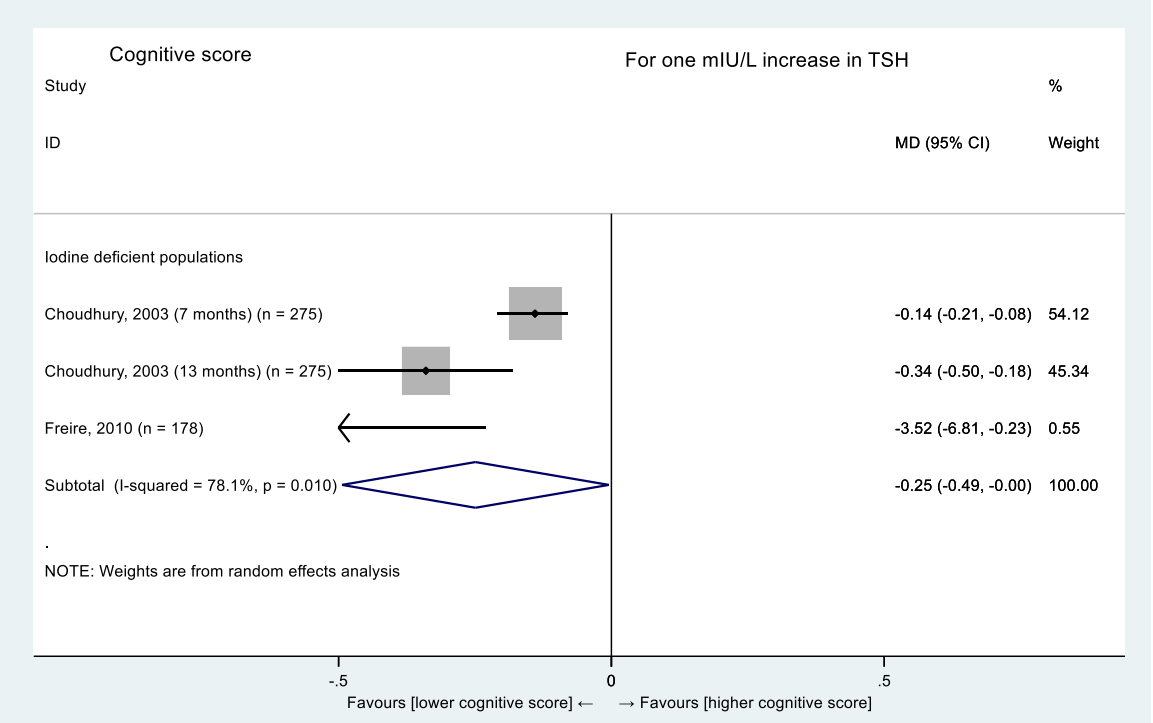


Figure 2. 5 Associations between cord blood TSH as continuous exposure and cognitive scores.
CI, confidence interval; MD, mean difference; TSH, thyroid-stimulating hormone.

Psychomotor outcome

The studies in iodine deficient populations showed negative [166] or null [165] association while the study from the iodine sufficient population reported a null association [164].

2.7.3.4 Associations between newborn TSH and child anthropometry

As depicted in **Table 2.6**, 4 studies reported anthropometric outcomes. Two studies reported associations between newborn TSH and head circumference at birth [141] and 2 years of age [171] and showed a null association, one study (n = 250) [142] assessed associations between newborn TSH and child anthropometry at 6 years of age and showed no associations, and one study (n = 616) [173] assessed the correlation between cord blood TSH and anthropometric outcomes at birth and showed no correlation. One of the studies [141] showed newborns at the highest TSH quartile had a 2.2 higher odds of a birthweight z score <-2 SD (OR: 2.2; 95% CI: 1.2, 4.1) compared with infants in the lower three TSH quartiles.

Table 2. 6 Summary of key literature for assessment of associations between TSH and growth[1]

	Author, year, Country	Design, sample size, setting, iodine status of country/region	TSH assessment(type of blood, age at assessment, classification)	Neurodevelopment assessment tool, age in years at assessment Outcome type	Key Finding	Limitations/gap of the study/comment
1	Shields, 2011 [173], UK	Cohort N = 616 Healthy infants participated in the Exeter Family Study of Childhood Health (EFSOCH) Excluded: preterm, Congenital anomalies, twins Mild iodine deficient population	Cord blood TSH Log transformed value	Birthweight Birth length Birth head circumference Skinfold thickness At birth Fetal growth	Birthweight (r = 0.04; P >0.05), Birth length (r = -0.02; P >0.05), Birth head circumference (r = 0.05, P >0.05) Sum of skinfolds (r = -0.02; P >0.05).	No adjustment for major confounders. Follow-up not long enough for the outcomes to occur.
2	Korzeniewski, 2014 [141], ELGAN Multicentre study, US	Cohort N = 120 Hospital setting Preterm infants (GA <28 weeks) Iodine sufficient population	Newborn TSH 14 days of age G1: newborns at 4th TSH quartile G2: newborns at 1st, 2nd and 3rd TSH quartiles	Birthweight ≤750 g Birthweight Z score <-2 SD Head circumference z score <-2 SD	OR (95% CI) Low birthweight G1:2.0 (1.3, 2.9) Birthweight z score <-2 SD G1:2.2 (1.2, 4.1) Head circumference z score <-2 SD G1:1.4 (0.8, 2.3)	On preterm infants. Late TSH sampling and did not considered in the analysis. Did not adjust for SES. The TSH value for each quartile cut-points was not reported.

3	Cuestas, et al 2015 [142] Argentina	Prospective cohort N = 250 Healthy infants Iodine sufficient population	Newborn TSH 2-3 days of age G1: TNH (newborn TSH ≥10 mIU/L) (n = 65) G2: normal (newborn TSH <10 mIU/L) (n = 185)	Weight Height 6 years	Height G1: 110.3 ± 16.8 G2: 109.5 ± 17.1 P = 0.74 Weight G1:19.6 ± 1.7 G2:19.3 ± 1.5 P = 0.21	Did not adjust for confounders.
4	Leviton, 2016, [171], ELGAN Multicentre study, US	Cohort day 21 (N = 749) and day 28 (N = 697) Hospital setting Preterm infants (GA <28 weeks) Iodine sufficient population	Newborn TSH 21 days 28 days G1: newborns at 4th TSH quartile G2: newborns at 1st, 2nd and 3rd TSH quartiles	head circum-ference Z-score <-2 SD 2 years	OR, 95% CI Day 21 G1: 1.1 (0.7, 1.8) Day 28) G1: 1.6 (0.9, 2.6)	Preterm infants. Measurement bias due late TSH sampling and confounding due to not considering late sampling in the analysis. Did not adjust for SES, maternal IQ, and home environment. The TSH value for each quartile cut-points was not reported.

[1]CI, confidence internal; ELGAN, extremely low gestational age newborns; GA, gestational age; N, sample size; OR, odds ratio; RR, risk ratio; TNH, transient neonatal hyperthyrotropinemia

2.7.4 Discussion

Overall, the association between newborn TSH and neurodevelopmental outcomes was inconclusive. Higher newborn TSH was associated with lower cognitive developmental scores in studies from mild-to-moderate iodine deficient populations, while a null association was observed in studies from iodine sufficient populations, however, this results should be interpreted with caution as all the studies were assessed as having a moderate or serious risk of bias. Studies using the cord blood TSH as the exposure measure showed a negative association between TSH and cognitive development regardless of the population iodine status. We found a null association between TSH and psychomotor development. Although few studies assessed associations between TSH and anthropometry, there was evidence of lower anthropometric outcomes at birth in children with elevated TSH at birth, but this may not affect long-term growth.

Among iodine deficient populations, the finding of the negative association between newborn TSH and neurodevelopment may be attributable to poor maternal iodine intake in pregnancy. There is evidence of negative consequences of severe [145, 174] and mild-to-moderate iodine deficiency [96] in pregnancy on child development. While meta-analysis was planned, only two studies [160, 167] (n = 752) with moderate and serious risk of bias in iodine deficient populations that dichotomised TSH to ≥5 mIU/L vs 5 mIU/L were suitable for the meta-analysis. The result showed that children with TSH ≥5 mIU/L had six cognitive points lower than infants with TSH <5 mIU/L. Even though there was no established TSH cut-off to classify iodine status at individual-level, children with mildly elevated newborn TSH concentration above 5 mIU/L or increased TSH concentration due to TNH showed lower cognitive performance in most studies included in the review. A new newborn screening cut-off, which considers the risk of impaired neurodevelopment due to mildly elevated newborn TSH concentration and the economic implications of lowering the newborn screening cut-off,

needs to be established [143]. Therefore, until the new newborn screening cut-off is established, infants with mildly elevated newborn TSH concentration ≥5 mIU/L in iodine deficient populations could benefit from continuous monitoring and early intervention, as needed, to prevent cognitive impairments in childhood.

The null findings in the iodine sufficient populations could be due to change in the distribution of TSH concentration as there is a transient increase in newborn TSH following the change in iodine status of populations from deficiency to sufficiency [82]. However, identifying the TSH threshold where the association between TSH and neurodevelopmental scores was lost is beyond the scope of the review. In iodine sufficient populations, maternal iodine status during pregnancy, defined as iodine-to-creatinine ratio <150 µg/g or UIC <150 µg/L, were not associated with non-verbal IQ and language development [175, 176]. An infant's thyroid gland is sensitive to both deficient and excess levels of iodine intake during pregnancy [177]. Therefore, in iodine sufficient populations, the mild TSH elevation or TNH may not be due to iodine deficiency and may be associated with higher iodine intake from iodine-containing supplements during pregnancy and may not affect childhood neurodevelopmental outcomes. Most of the included studies in the iodine sufficient populations were on preterm infants, small sample size, assessed neurodevelopmental outcomes at early ages and did not adjust for key confounding variables. Adequately powered studies and controlled for key confounders are required to identify a TSH threshold, in children without CH, associated with neurodevelopmental deficits in iodine sufficient populations.

While the level of TSH is different in cord and newborn blood [137], the relationship between TSH and neurodevelopment was consistent regardless of the type of blood. Based on the meta-analysis of two studies that used cord blood TSH, children with TNH had 13.2 cognitive

points (n = 55) lower than children without TNH. Moreover, a 1 mIU/L increase in cord blood TSH was also associated with a 0.3 point lower cognitive score in studies from iodine deficient populations. However, the evidence was from low-quality studies with a small sample size ranging from 18 to 275 participants, and did not adjust for key confounding variables.

Most of the included studies used standard scales to assess neurodevelopment. More data on the association between newborn TSH and other developmental outcomes like psychosocial and adaptive behaviour scales are needed. A null association was found between newborn TSH and psychosocial development but only one study [168] was included in the review. However, a parent-reported scale was used to assess the outcome and may introduce bias in the measurement of the outcome variable. Further study using standard psychosocial and adaptive behaviour assessment tools is warranted to confirm this finding in all settings.

Of the studies that assessed the association between newborn TSH and child anthropometry, only one study [141] showed a negative association and the other two [142, 171] showed a null association. The study [141] from an iodine sufficient population showed that infants in the fourth TSH quartile had higher odds of low birthweight and birthweight z score < -2 SD compared with the other quartiles, but participants were preterm infants with late TSH sampling outside of the WHO's recommended age range. Moreover, the study did not adjust for key confounders of child growth including maternal socio-economic status. Key confounders of child growth were not also controlled in the two studies [142, 171] that reported a null association. The only study that used cord blood TSH reported no correlation between cord blood TSH and birth anthropometry [173]. Evidence from large studies that used population-based newborn TSH data and adjust for key confounders of child

anthropometry are required to examine the association of newborn TSH concentration and child growth.

There is no established newborn TSH cut-off to screen children at risk of impaired neurodevelopment and growth other than to screen children with CH [140] and to define iodine status in populations [12]. There are a paucity of studies that use newborn TSH from newborn screening and link to neurodevelopmental or growth outcomes. More quality population-based studies on the associations between newborn TSH and neurodevelopment and growth that adequately adjust for key confounders in populations with different iodine status are needed to define a newborn TSH cut-off associated with poor outcomes in different settings.

2.7.5 Conclusion

There is evidence of a negative association between newborn TSH and cognitive development in iodine deficient populations. However, newborn TSH shows a null association with cognitive and psychomotor development in iodine sufficient populations. There is little evidence that newborn TSH concentration is associated with birth anthropometry, but there is limited evidence on associations between TSH and child anthropometry at later ages.
More quality follow-up studies that assess associations between newborn TSH and neurodevelopmental and anthropometric outcomes by adjusting all potential confounding variables are warranted.

2.8 Summary of literature review and the rationale for the study presented in this thesis

IDD remains a public health problem worldwide, and iodine deficiency in pregnancy remains a concern in several countries, including Australia. In Australia, this is partly because the fortification is aimed at the general population, acknowledging that it might not be enough for pregnant women. Therefore, it may be difficult to meet the RDI of iodine in pregnant women and breastfeeding women in Australia from the fortification program alone.

UIC, I/Cr and UIC-corrected for creatinine and predicted 24-h UIE markers have been widely used in the literature despite their specific use and applications. No study has assessed agreement between urinary iodine measurements from spot urine samples in pregnant women, including discrepancies in iodine status of populations as measured by different markers. To date, there has been no systematic examination of agreement between different markers of population iodine status. Comparison of markers to classify iodine status will help to investigate whether these markers consistently classify iodine status in populations or pregnant women and to identify appropriate markers in different settings.

In Australia, although the recent National Health Survey conducted in Australia between 2011 and 2012 showed improvement in the iodine status of the population based on UIC in SAC, assessing iodine status using multiple markers of population iodine status including newborn TSH and goitre prevalence is limited. There is only one study [52] in Western Australia that evaluated the effect of the mandatory iodine fortification on the iodine status of the population using the newborn TSH concentration as a marker, but the study was conducted in the transitional period; and there is also a geographical variation in iodine status in Australia. Assessing iodine status in SAC and at-risk populations, including pregnant women, breastfeeding women and infants, gives a better picture of population-level iodine nutrition. Pregnant women from low SES may be at risk of iodine deficiency post-fortification, as

reported in a small study in northern Adelaide in the transitional period. Therefore, evaluating the iodine status of pregnant women of low socio-economic position in pre- and post-iodine fortification periods with adequate sample size is needed to evaluate the impact of fortification on the iodine status of this at-risk population.

Newborn TSH concentration could be a useful marker to identify children at risk of suboptimal neurodevelopmental or growth outcomes in countries where newborn screening is routinely practised. There is evidence of a negative association between newborn TSH concentration and the cognitive development of children in iodine deficient populations, and no associations in iodine sufficient populations though the evidence was inconclusive for other developmental scales and child growth. However, the quality of the studies that assessed the associations between newborn TSH and child development and growth is limited, mainly due to inadequate adjustment of confounding variables and small sample size. Hence, prospective follow-up studies with an adequate sample size investigating the association between newborn TSH concentration and the neurodevelopmental and growth outcomes with adequate adjustment of confounding variables in both iodine deficient and sufficient settings are required.

2.9 Research questions and objectives of the thesis

The thesis addressed the following questions:

1. What is the agreement between the different markers used to assess the iodine status of general populations, populations of pregnant women and individual pregnant women?
2. What is the impact of the mandatory iodine fortification program in Australia on the iodine status of the general population and pregnant women from a low socio-economic position in South Australia?

3. What is the relationship between newborn TSH concentration and childhood growth and development?

Objectives of the thesis are:

1. To evaluate the agreement between different markers used to define iodine status:
    a) in the populations (Chapter 3)
    b) in pregnant women (Chapter 4)
2. To examine the impact of the iodine fortification program in Australia by comparing the iodine status before and after the mandatory iodine fortification of bread:
    a) in the general South Australian population using newborn TSH concentration as a marker (Chapter 5)
    b) in pregnant women from a low SES region in South Australia using UIC as a marker (Chapter 6)
3. To investigate the associations between newborn TSH concentration and development and growth of children at 18 months of age (Chapter 7)

CHAPTER 3 AGREEMENT BETWEEN MARKERS OF POPULATION IODINE STATUS: A SYSTEMATIC REVIEW

This chapter has been published in the *American Journal of Clinical Nutrition* [15].

3.1 Publication

Wassie MM, Middleton, P., & Zhou, S. (2019). Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review. *The American Journal of Clinical Nutrition*, 110(4), 949-958. DOI: https://doi.org/10.1093/ajcn/nqz118

To adapt to the format of the thesis, the form of the published version of the paper has been slightly modified in the following ways without any changes to the text content;

1. Re-numbered the consecutive page numbering.
2. Used Australian English instead of American English.
3. Re-numbered the consecutive reference numbering and used the numbered format of the references to be consistent with the rest of the chapters.
4. Re-numbered the tables, figures and supplemental tables.
5. Embedded the tables and figures into the text.

Statement of Authorship

Title of Paper	Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review. The American journal of clinical nutrition, 2019. doi: 10.1093/ajcn/nqz118
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wassie MM, Middleton P, Zhou SJ. Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review. Am J Clin Nutr 2019. doi: 10.1093/ajcn/nqz118

Principal Author

Name of Principal Author (Candidate)	Molla Mesele Wassie		
Contribution to the Paper	Conception, design, data collection, statistical analysis, interpretation of results, manuscript preparation and critical revision of the manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	13/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Philippa Middleton

Contribution to the Paper	Design, interpretation of results and critical revision of the manuscript		
Signature		Date	16/03/2020

Name of Co-Author	Shao Jia Zhou		
Contribution to the Paper	Conception, design, data collection, statistical analysis, interpretation of results and critical revision of the manuscript		
Signature		Date	16/03/2020

Please cut and paste additional co-author panels here as required.

Title Page

Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review

Author Names

Molla Mesele Wassie, Philippa Middleton, Shao Jia Zhou

Authors' affiliations

School of Agriculture Food and Wine, Faculty of sciences, The University of Adelaide, Australia (MMW, SJZ).

Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia (MMW).

Robinson Research Institute, School of Paediatrics and Reproductive Health, The University of Adelaide, Adelaide, Australia (PM, SJZ).

South Australian Health and Medical Research Institute, *Adelaide, Australia (PM).*

Authors' last names

Wassie, Middleton, Zhou

Disclaimers

None.

Corresponding Author

Shao Jia Zhou

School of Agriculture, Food & Wine

University of Adelaide, Waite Campus

PMB1, Glen Osmond, SA 5064

Phone: +61 8 8313 2065

Email: jo.zhou@adelaide.edu.au.

Sources of Support

MMW was supported by The University of Adelaide, Australian Government Research Training Program.

Short running head

Agreement between markers of iodine status

Abbreviations

AUC, Area Under Curve; MUIC, Median Urinary Iodine Concentration; ROC, Receiver Operating Curve; TSH, Thyroid Stimulating Hormone; WHO, World Health Organization

Registry number for systematic review

The protocol is published at PROSPERO with registration number of CRD42018091247.

Abstract

Background: Population iodine deficiency is indicated by > 3% of the population with newborn thyroid stimulating hormone (TSH) concentration >5 mIU/L, median urinary iodine concentration (MUIC) <100 µg/L or prevalence of goitre > 5% in school age children. However, the agreement between these population markers has not been systematically investigated.

Objective: To assess the agreement between TSH, MUIC and goitre as markers of population iodine status.

Design: We performed a systematic search for studies published on PubMed, Scopus, CINAHL, Embase, and PsycINFO to 29 October 2018. Studies assessing iodine status in the population using the TSH marker and either MUIC or goitre prevalence in school age children were included. The agreement between markers in classifying iodine status of the population was assessed. The sensitivity and specificity of the TSH marker was determined against the MUIC and goitre prevalence as the reference markers.

Results: Of 17,435 records identified by the search strategy, 57 eligible studies were included in the review. The agreement between markers in classifying the iodine status of populations into the same category was 65% for TSH and MUIC, and 83% for TSH and goitre prevalence. The TSH marker had a sensitivity of 0.75 and specificity of 0.53 when compared with MUIC, and 0.86 and 0.50 when compared with goitre prevalence.

Conclusion: The TSH marker has a better agreement with goitre prevalence than MUIC when classifying the iodine status of populations. Re-evaluation of the current criteria for classifying iodine status of populations using the TSH marker is warranted.

Key words: iodine deficiency, thyroid stimulating hormone concentration, urinary iodine concentration, goitre, iodine status

Introduction

Iodine is an essential micronutrient which is important for growth and development. Iodine deficiency can lead to goitre, hypothyroidism and hyperthyroidism; miscarriage, congenital anomalies, cretinism and deaf mutism in newborns and impaired physical growth and mental function in children which are collectively termed as iodine deficiency disorders [3]. Iodine fortification of staple foods and iodine supplementation of at risk population groups are the common strategies to prevent iodine deficiency [1]. The World Health Organization (WHO) recommends the use of the median urinary iodine concentration (MUIC) and prevalence of goitre in school age children as the principal markers of population iodine status [1, 12, 178]. It also recommends the use of thyroid stimulating hormone (TSH) concentration in newborns to classify iodine status of populations [12, 178].

Populations are classified as iodine sufficient if the proportion of the newborn TSH >5 mIU/L in the population is < 3%, or the MUIC of school age children is 100 - 299 μg/L, or goitre prevalence < 5% (by palpation or ultrasound examination) in school age children aged 6-12 years [1, 12, 34]. Ideally classification of iodine status of populations at the same time point should be consistent regardless of which of the markers is used. However, inconsistent classification of population iodine status has been reported when using the TSH marker versus MUIC [45, 52, 54, 179, 180] or goitre [181-184].

There is no single gold standard marker for population iodine status. The MUIC reflects recent iodine intake in populations while the percentage of newborn TSH >5 mIU/L and goitre prevalence reflect longer-term exposure (months to years) to iodine intake [12]. The MUIC requires a large representative sample from school age children to account for the diurnal variation in urinary iodine excretion [87]. Goitre prevalence is a less sensitive marker of iodine status in populations with mild-to-moderate iodine deficiency and populations with

recent iodine interventions [185]. The newborn TSH concentration can be affected by factors such as exposure to iodine-containing antiseptics during delivery [183, 186] and timing of newborn's blood sampling [52]. The current TSH cut-off has been criticized for its inability to correctly classify iodine status in populations [45]. The WHO recommends the use of newborn's blood collected between three and four days after birth for newborn TSH determination [12] but this recommendation is not always followed. Some studies used newborn blood samples collected before day three [183, 187, 188] or after day four [54, 60, 179, 181] to assess iodine status of populations using TSH as a marker.

To the best of our knowledge there has been no systematic evaluation of the consistency between the markers of population iodine status, TSH, MUIC and goitre prevalence, in classifying the iodine status of populations. The aims of this systematic review were to investigate the agreement between markers of population iodine status, and the sensitivity and specificity of the TSH compared with MUIC and goitre prevalence.

Methods

Study selection

Studies which assessed the iodine status of populations using the proportion of neonatal (or cord blood) TSH concentration >5 mIU/L and at least one of the other recommended markers of population iodine status, including MUIC or prevalence of goitre in school age children (aged 6-12 years) were included. Goitre was defined according to the WHO criteria as school age children having either grade 1 or grade 2 goitre by palpation or thyroid volume > 97th percentile by ultrasound. Studies in which MUIC or goitre was assessed in school aged children outside of the recommended 6-12 years were also be eligible for inclusion if the school children were below 18 years of age. Studies where data on the iodine status of a population classified by different markers (eg. TSH and MUIC) in the same time period were

reported in separate publications, were also eligible to be included in the review. Results from multiple surveys in different regions of a country, or different time points, or different neonatal age at sampling reported in a same publication were counted as separate studies to assess the agreement between markers.

Search strategy

A systematic literature search was conducted on August 1, 2017, then last updated on October 29, 2018. Databases included PubMed, Scopus, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, and PsycINFO were searched. Reference lists of included studies and other databases including WHO and Iodine Global Network were also reviewed. The key terms used in literature search included: Thyroid stimulating hormone, TSH, thyrotropin, thyrotrophin, thyreotropin, hypothyroidism, hypothyroxine, iodine deficiency, goitre, thyroid volume, urinary iodine concentration, iodine nutritional status, iodide status, infant, newborn, neonate, child, children and school aged children. The search was limited to human studies, but no other restrictions were applied. See **Supplemental Table 3.1** for more details of the search strategies for each database.

Data collection

All retrieved articles from the search strategy were exported to endnote X7 library and duplicates were removed. Articles were then screened by title and abstract for eligibility (by MMW), if not enough information in the abstract to determine eligibility, then review of full text was performed. Uncertainty in eligibility was resolved by discussion with a co-author (SJZ). Data were extracted (by MMW) using a data extraction form and was double-checked by the co-author (SJZ).

Quality assessment

We used the American Dietetic Association's "Evidence Analysis Library" manual [189] to evaluate the quality of the included studies. Quality of the studies was assessed using the 10 item questions **(Supplemental Table 3.2)**. Four of the 10 items were defined as important items for assessing the quality of studies, which included item 2 (study subjects), item 3 (comparability of study groups), item 6 (description of exposure) and item 7 (validity of measurements). Studies were classified as low quality (high risk of bias) and denoted with a negative (-) sign if the answer to the quality assessment questions was "yes" for three or less of the ten items. Conversely, studies were classified as high quality (low risk of bias) and denoted with a positive (+) sign if the answer was "yes" for a total of five or more items including all four important items. Studies were denoted with a neutral (Ø) sign if the answer was "yes" for a total of four or more items including three or less of the important items **(Supplemental Table 3.3)**. Quality was assessed and scored at a study level as follow: 1) for studies that involved two publications: the quality of each publication was rated separately. If the quality of the publications differed, the lowest quality among the publications was assigned to the study; 2) for publications that reported multiple studies, the quality assessment was done for each study separately. The overall grade for the strength of the evidence supporting the conclusion of the review was based on the grading definition of the American Dietetic Association's Evidence Analysis Manual as: 1) Good: evidence consists of results from studies of strong design for answering the question addressed, results with no serious concern about the generalizability, bias, and flaws in study design and consistent results at most; 2) Fair: evidence from studies of strong designs for answering the question addressed but inconsistencies among the results from different studies, or with consistent results from weaker designs for the question addressed only, or with minor concerns about the generalizability, bias, study design flaws, or adequacy of sample size; 3) Limited: evidence from a limited number of studies of weak design for answering the question addressed, with

inconclusive results due to lack of generalizability, bias, study design flaws, or inadequacy of sample size or unexplained inconsistency among results from different studies [189].

Statistical analysis

The data were analysed using SPSS Statistics 25 (IBM Corp, Armonk, NY). We assessed the agreement between the TSH marker and MUIC or goitre in school age children to classify iodine status in the populations as a primary outcome. An agreement is defined as having the same classification of iodine status of the populations while using TSH vs MUIC and/or TSH vs goitre. The agreement between these markers was assessed in two stages: firstly, the agreement between the markers in defining iodine sufficient and deficient; secondly, the agreement between the markers in defining the degree of iodine deficiency as mild, moderate or severe deficiency. The degree of iodine deficiency in the included studies is classified according to the WHO criteria to classify iodine status in populations [1, 12] (**Supplemental Table 3.4**). The overall agreement was assessed separately for studies that used neonatal TSH and those used cord blood TSH as the marker. In addition, subgroup analysis by age of neonatal blood sampling (before day 3, at 3-4 days, 3-5 days and those after 5 days of age), publication type (the markers were reported in the same publication vs. different publications) and country level iodine status (iodine deficiency vs. sufficiency countries) was also conducted to assess the agreement between markers. For the purpose of the subgroup analysis by the country level iodine status (deficient vs. sufficient), population iodine status was defined as follow: 1) based on the country level iodine status data from the WHO database if the data were available [156, 157]; 2) based on the Iodine Global Network if the data were not available in the WHO database [158, 159]; or 3) based on the authors' report when country level data were not reported in either the WHO or Iodine Global Network databases. We used MUIC as the primary marker to classify country level iodine status and used goitre as the secondary marker when MUIC data were not reported. For the studies reported goitre

prevalence by both palpation and ultrasound examination, results from the ultrasound examination were used as it is generally considered more accurate in the assessment of goitre in school age children [12, 34]. We also performed an exploratory analysis to assess the agreement between newborn TSH and MUIC of pregnant women in studies where the MUIC data for pregnant women were reported.

The secondary outcome was to assess the sensitivity and specificity of the TSH compared with MUIC and goitre prevalence. To calculate the sensitivity and specificity of the TSH marker in classifying the iodine status of populations compared with the MUIC or goitre prevalence, studies which consistently classified iodine deficiency using both the TSH and MUIC/goitre were categorised as true positives while those consistently diagnosing iodine sufficiency were considered as true negatives. Sensitivity indicates the ability of the TSH marker to correctly classify iodine deficient populations while specificity refers to the ability of the TSH marker to correctly classify iodine sufficient populations [190]. Youden's index [191] was also calculated ((Specificity + Sensitivity) - 1) to compare the performance of the markers [191]. The index has a value between zero (poor accuracy) and one (excellent accuracy). If the classification of population iodine status was the same using TSH and MUIC, or TSH and goitre, the index has a value of one. The receiver operating curve (ROC) analysis was performed and used in two ways: 1) to assess the predictive ability of newborn TSH to classify population iodine status based on MUIC and goitre prevalence; 2) to determine the optimal cut-off point for the proportion of TSH >5 mIU/L that gave the highest specificity and sensitivity when compared with MUIC and goitre prevalence [192]. Mann-Whitney statistics was used to determine the area under the ROC (AUC) [192, 193]. The AUC was used to compare the accuracy (both sensitivity and specificity) of the TSH to classify iodine status against MUIC and goitre prevalence: the larger the AUC the more

accurate the marker is. The accuracy of the TSH was categorised as excellent (AUC : 0.9-1.0), good (AUC : 0.8-0.9), poor (AUC : 0.7-0.8) and not good (AUC : 0.6-0.7) [194].

Results

Study characteristics

The search strategy identified 17,435 records and 30 eligible publications [17, 52, 54, 57, 60, 83-85, 116, 131, 179-184, 186-188, 195-205] were included in this systematic review **(Figure 3.1)**. TSH and UIC or goitre were reported in a same publication for 17/30 publications [60, 83-85, 181-184, 188, 195-199, 201, 203, 205]. For the remaining 13 publications, TSH and UIC or goitre were reported separately in eight (TSH only) [52, 54, 57, 116, 179, 186, 187, 200] and 5 (UIC only) publications [17, 131, 180, 202, 204].

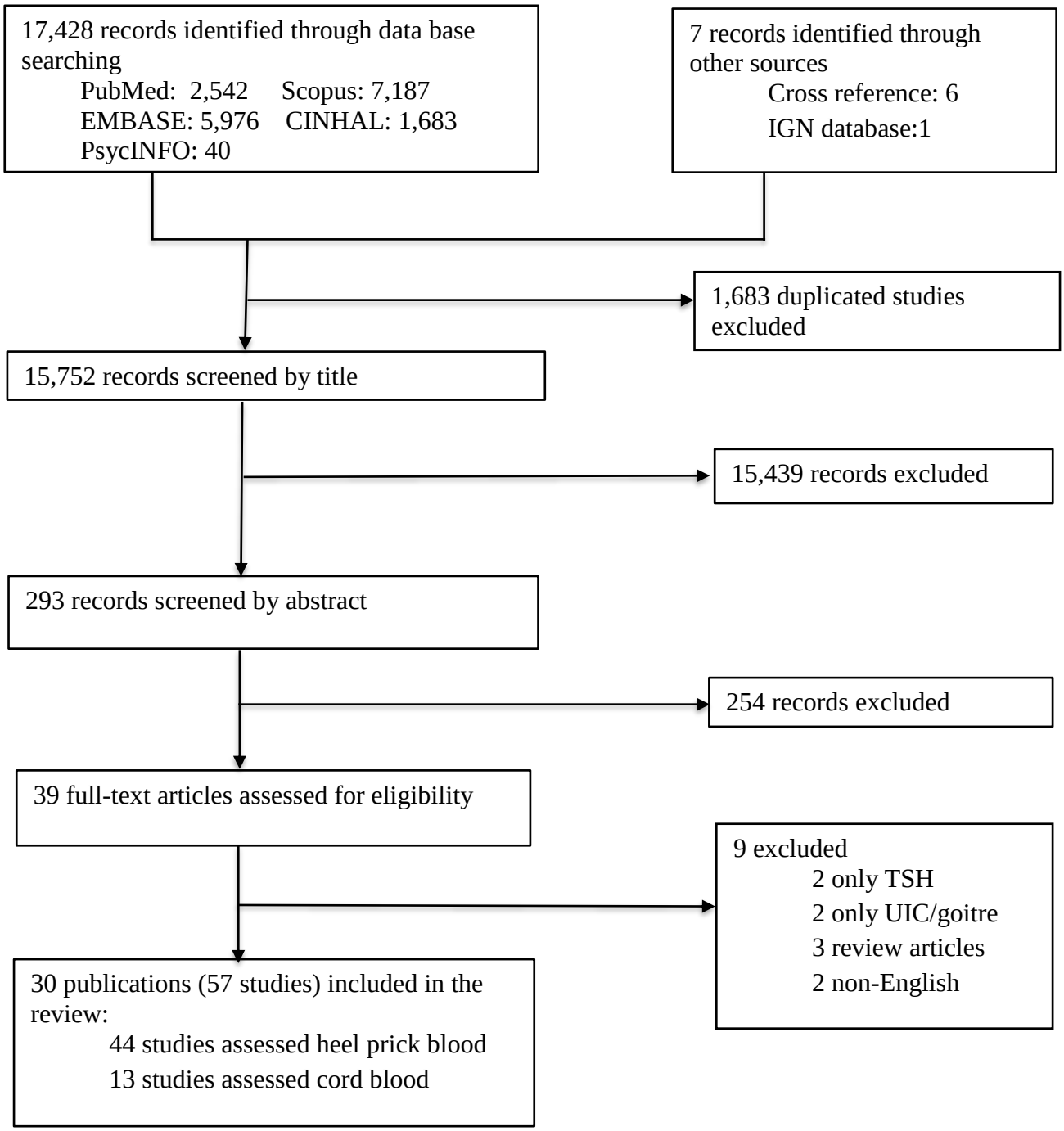


Figure 3. 1 Flowchart of publications for inclusion in the systematic review.
IGN, iodine global network; TSH, thyroid stimulating hormone; UIC, urinary iodine concentration.

Supplemental Table 3.5 summarizes the characteristics of the studies and **Supplemental Table 3.6** reports the iodine status of studies by different markers. Overall, a total of 57 studies from 30 publications were included in the analysis (**Supplemental Table 3.7**). Of the 57 studies, 56 studies reported TSH and MUIC of school age children (43 studies assessed TSH from infant blood samples and 13 studies assessed TSH from cord blood samples), and 36/57 studies reported TSH and goitre of school age children (23 studies assessed TSH from infant blood samples and 13 studies assessed TSH from cord blood samples) **(Tables 3.1 and 2)**. Of the 36 studies that assessed goitre prevalence, 35 of them assessed goitre by palpation [17, 60, 83, 85, 181-184, 186, 188, 195, 196, 198, 201-203], one study assessed goitre by ultrasound examination [198] and nine studies assessed goitre by both palpation and ultrasound examination [60, 182, 184, 198, 201] (**Supplemental Table 3.5**). Both methods consistently classified iodine status of the populations in all nine studies used both methods [60, 182, 184, 198, 201]. The included studies were conducted in 29 countries representing all WHO regions of the world. Fifty seven percent (32/56) of the studies were in countries classified as iodine deficient (**Tables 3.1 and 3.2**).

Table 3. 1 The agreement between newborn TSH and other markers to classify population iodine status[1]

PUBLICATION, COUNTRY	COUNTRY-LEVEL IODINE STATUS	STUDY[2]	AGREEMENT BETWEEN TSH & MUIC	AGREEMENT BETWEEN TSH & GOITRE
AGE AT TSH ASSESSMENT: 3-4 DAYS[3]				
Yamada, 2000, Mongolia [203]	Deficiency[4]	Study (Ulaanbaatar)	No	Yes
		Study (Uvurkhangai)	Yes	Yes
		Study (Huvsgol)	No	Yes
		Study (Dornod)	Yes	Yes
		Study (Hovd)	Yes	Yes
Zimmerman, 2005, Switzerland [84]	Sufficiency[4]	Study (Pre-intervention population)	Yes	NA
		Study (Post-intervention population)	Yes	NA
Clapin, 2014 [52]; ABS report, 2013, Western Australia [131]	Sufficiency[5]	Single study	No	NA
Evans 2014[179]; Vanderpump, 2011,Wales UK [180]	Deficiency[5]	Single study	No	NA
Rahman 2010 [200]; Li 2005, Victoria, Australia [17]	Deficiency[4]	Single study	Yes	NA
Fuse, 2003, Ulaan Baatar, Mongolia [198]	Deficiency[4]	Study (1996)	No	Yes
		Study (1996-1999)	No	Yes
Wilcken , 2011 [116]; Li, 2005, NSW, Australia [17]	Deficiency[5]	Single study	Yes	NA
AGE AT TSH ASSESSMENT: < 3 DAYS OR > 4 DAYS[6]				
Gruneiro- Papendieck, 2004, Argentina [183]	Sufficiency[7]	Single study	No	No
Al-Hosani, 2003, United Arab Emirates [181]	Deficiency[4]	Single study	No	No

PUBLICATION, COUNTRY	COUNTRY-LEVEL IODINE STATUS	STUDY[2]	AGREEMENT BETWEEN TSH & MUIC	AGREEMENT BETWEEN TSH & GOITRE
Wachter,1985, Tanzania [188]	Deficiency[7]	Study (Pre-intervention population)	No	No
		Study (Post intervention population)	No	NA
Evans, 2014, UK [179]; Vanderpump et al., 2011, UK[180]	Deficiency[4]	Single study	No	NA
AGE AT TSH ASSESSMENT: 3-4 DAYS AND < 3 DAYS OR > 4 DAYS[8]				
Mehran, 2016, Iran [205]	Sufficiency[5]	Single study	Yes	NA
Dalili, 2012, Iran [197]	Sufficiency[5]	Study (2006)	Yes	NA
		Study (2007)	Yes	
		Study (2008)	Yes	
		Study (2009)	Yes	
		Study (2010)	Yes	
Costan, 2002, Italy [196]	Deficiency[4]	Study (Inland territory)	Yes	Yes
		Study (Sea level territory)	Yes	Yes
Costan,1997, Italy [195]	Deficiency[4]	Single study	Yes	Yes
Burns, 2008 [54]; Vanderpump, 2011, Ireland [180]	Sufficiency[4]	Single study	No	NA
Mikelsaar, 1999, Estonia [57]	Deficiency[4]	Study (North):	Yes	NA
		Study (Central)	Yes	NA
		Study (South)	Yes	NA
Vandevijvere, 2012 [186] Vandevijvere, 2012, Belgium [202]	Sufficiency[4]	Study (2009-2011)	Yes	No
		Study (2010-2011)	No	Yes
		Study (2011-2011)	No	Yes
Simsek, 2003, Turkey[60]	Deficiency[4]	Study (Turkey)	NA	Yes
		Study (Bolu)	Yes	Yes
		Study (Duzce)	Yes	Yes

PUBLICATION, COUNTRY	COUNTRY-LEVEL IODINE STATUS	STUDY[2]	AGREEMENT BETWEEN TSH & MUIC	AGREEMENT BETWEEN TSH & GOITRE
		Study (Zongulda)	Yes	Yes
Konrade, 2014, Latvia [199]	Sufficiency[5]	Single study	No	NA
Svinaryov, 2003, Russia[201]	Deficiency[4]	Single study	Yes	Yes
Wachter,1985, Tanzania [188]	Deficiency[7]	Study (Pre-intervention population)	Yes	Yes
		Study (Post intervention population)	Yes	NA
Gruneiro- Papendieck, 2004, Argentina [183]	Sufficiency[7]	Single study	Yes	Yes
Gyurjyan, 2006, Latvia [187]; Selga, 2000, Latvia [204]	Deficiency[4]	Single study	Yes	NA

[1]NA, not applicable; TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration.

[2]Study conducted at different regions or different time periods within the same publication were classified as separate studies, with the region /time period in the brackets. The term 'Single study' denotes if the study was conducted in only one region or one time point within publication(s).

[3] Studies where the newborns' heel prick samples were collected within the World Health Organization's recommended age (at 3-4 days of infant's age).

[4]Data were obtained from the World Health Organization database to indicate iodine status of the countries where the studies were conducted.

[5]Data were obtained from the Iodine Global Network database to indicate iodine status of the countries where the studies were conducted.

[6] Studies where the newborns' heel prick samples were collected outside the World Health Organization's recommended age (at < 3 or > 4 days of infant's age).

[7]The iodine status of the population were classified based the authors' report when the population-level data are not reported in the World Health Organization or the Iodine Global Network databases.

[8]Studies where the newborns' heel prick samples were collected within (3-4 days of infant age) and outside (< 3 and > 4 days of infants age) the World Health Organization's recommended infant's age.

Quality of included studies

The quality of the included studies is reported in the **Supplemental Table 3.3**. Twenty-two studies had low risk of bias and the remaining studies were classified as neutral in quality. Overall, the strength of evidence was graded as "Limited" mainly due to lack of generalizability [52, 54, 57, 83, 116, 181, 184, 187, 200] or inadequate sample size [183, 188, 196] or bias in participant selections [57, 116, 183, 187, 188, 198], and inconsistent results from different studies [179, 180, 186, 202].

Agreement between newborn TSH and MUIC in defining iodine status of populations

Forty-three studies reported data on iodine status based on TSH and MUIC, data on iodine status from the two markers being compared were extracted from the same publication in 33 studies and from different publications in 10 studies. The agreement between the markers were 73% in the studies where the data were from the same publication and 40% in the studies from different publications (**Supplemental Table 3.8**). Overall, 65% (28/43) of the studies showed an agreement between MUIC and newborn TSH in classifying iodine status [57, 60, 84, 116, 183, 186-188, 195-197, 200, 201, 205] (**Supplemental Table 3.9**). The agreement in classifying iodine status of populations was observed in 69% (22/32) of the studies from iodine deficient populations and 56% (6/11) of the studies from iodine sufficient populations. Seventy eight percent (14/18) of the studies showed an agreement in classifying the severity of iodine deficiency. The agreement was 50% (7/14) in classifying mild iodine deficiency, 43% (6/14) in moderate iodine deficiency and 7% (1/14) in severe iodine deficiency **(Table 3.1, Supplemental Tables 3.5 and 3.6).** In the exploratory analysis, the agreement between newborn TSH and the MUIC of pregnant women in classifying iodine status was observed in 75% (6/8) of the studies (**Supplemental Table 3.10**).

Agreement between newborn TSH and goitre prevalence in defining iodine status of the population

Twenty-three studies reported data on iodine status based on TSH and goitre, data on iodine status from the two markers being compared were extracted from the same publication in 20 of the studies and from different publications in three of the studies. The agreement in classifying iodine status of populations was 85% in the studies where the data were from the same publication and 67% (only three studies) in the studies from different publications (**Supplemental Table 3.8**). Overall, 83% (19/23) of the studies showed an agreement between goitre and newborn TSH in classifying iodine status (**Supplemental Table 3.9**). Agreement in classifying iodine status of populations was observed in 89% (16/18) of the studies from iodine deficient populations and 80% (4/5) of the studies from iodine sufficient populations. Thirty-nine percent (7/18) [183, 186, 188, 195, 201-203] of the studies showed an agreement in classifying the degree of iodine deficiency (mild versus moderate versus severe) between the two markers **(Table 3.1, Supplemental Tables 3.5 and 3.6)**.

Newborn's age at heel prick blood sampling and the agreement between markers

Of the 13 studies assessing TSH concentration from samples collected at 3 to 4 days after birth according to the WHO recommendation, agreement in classifying iodine status of the populations was reported in 54% (7/13) [84, 116, 200, 203] of the studies that assessed iodine status by TSH and MUIC and 100% (7/7) [198, 203] of the studies that assessed iodine status by TSH and goitre. There was no agreement with MUIC or goitre for the studies that assessed newborn TSH at < 3 or > 4 days of infant's age [60, 179, 181, 183, 188] **(Table 3.1)**.

Agreement between cord blood TSH, MUIC and goitre prevalence in defining iodine status of populations

Iodine status of the populations was classified as iodine deficient in all 13 studies that used cord blood TSH as a marker. All 13 studies reported data on TSH, MUIC and goitre prevalence, 38% (5/13) [85, 182, 184] showed an agreement between TSH and MUIC and 77% (10/13) [83, 85, 182, 184] showed an agreement between TSH and goitre prevalence. No agreement was observed between the cord blood TSH and MUIC or goitre in classifying the degree of iodine deficiency (mild vs. moderate vs. severe) **(Table 3.2, Supplemental Tables 3.5 and 3.6)**.

Table 3. 2 The agreement between cord blood TSH and other markers to classify population iodine status[1]

Publication, country	Country-level iodine status	Study[2]	Agreement between TSH & MUIC	Agreement between TSH & goitre
Kapil, 2015, India [85]	Sufficiency[3]	Study (Kangra)	No	Yes
		Study (Kullu)	No	Yes
		Study (Solan)	Yes	Yes
Copeland, 2002, Bangladesh [182]	Deficiency[4]	Single study	Yes	Yes
Copeland, 2002, Guatemala [182]	Sufficiency[4]	Single study	No	Yes
Copeland, 2002, USA [182]	Sufficiency[4]	Single Study	No	No
Sareen, 2016, India [83]	Sufficiency[3]	Study (Udham Singh Nagar)	No	Yes
		Study (Nainital)	No	Yes
		Study (Pauri Garhwal)	No	Yes
Sullivan, 1997, Kyrgyzstan [184]	Deficiency[4]	Study (Bishkek)	Yes	Yes
		Study (OSH)	Yes	Yes
Sullivan,1997, Philippines [184]	Deficiency[4]	Single study	Yes	No
Gruneiro-Papendieck, 2004, Argentina [183]	Sufficiency[5]	Single study	No	No

[1]NA, not applicable; TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration.
[2]Study conducted at different regions within the same publication were classified as separate studies, with the region in the brackets. 'Single study' denotes if the study was conducted in only one region or one time point only within one publication or between two publications.
[3]Data were obtained from the Iodine Global Network database to indicate iodine status of the countries where the studies were conducted.
[4]Data were obtained from the World Health Organization database to indicate iodine status of the countries where the studies were conducted.
[5]The iodine status of the population were classified based the authors' report when the population-level data are not reported in the World Health Organization or Iodine Global Network databases.

Sensitivity and specificity of newborn TSH

Figures 3.2 and **3.3** show the ROC curves for the ability of newborn TSH in predicting iodine status compared with MUIC and goitre prevalence, respectively. The sensitivity, specificity and Youden's index was 0.75, 0.53 and 0.28, respectively for TSH compared to MUIC in classifying iodine status and 0.86, 0.50 and 0.36, respectively for TSH compared with goitre.

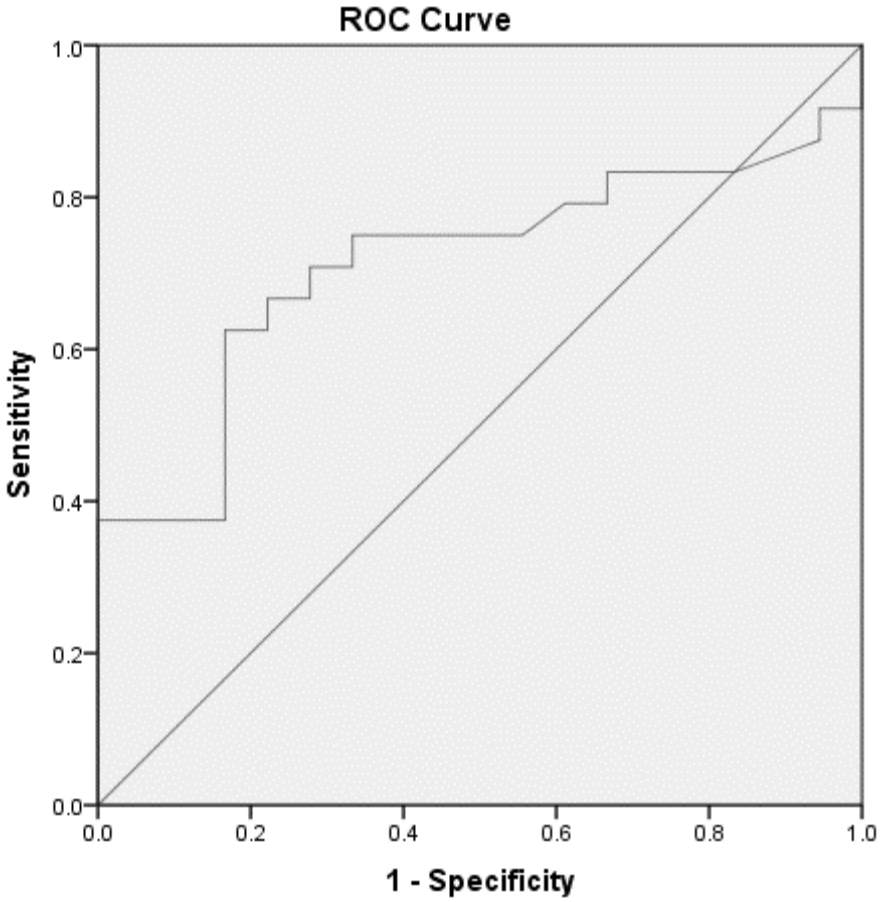


Figure 3. 2 The receiver operating curve (ROC) of the TSH marker to predict iodine status using the MUIC as a reference marker (Forty-three studies included in the analysis).
The area under the curve (AUC) was 0.71 (95% CI: 0.56, 0.87). TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration.

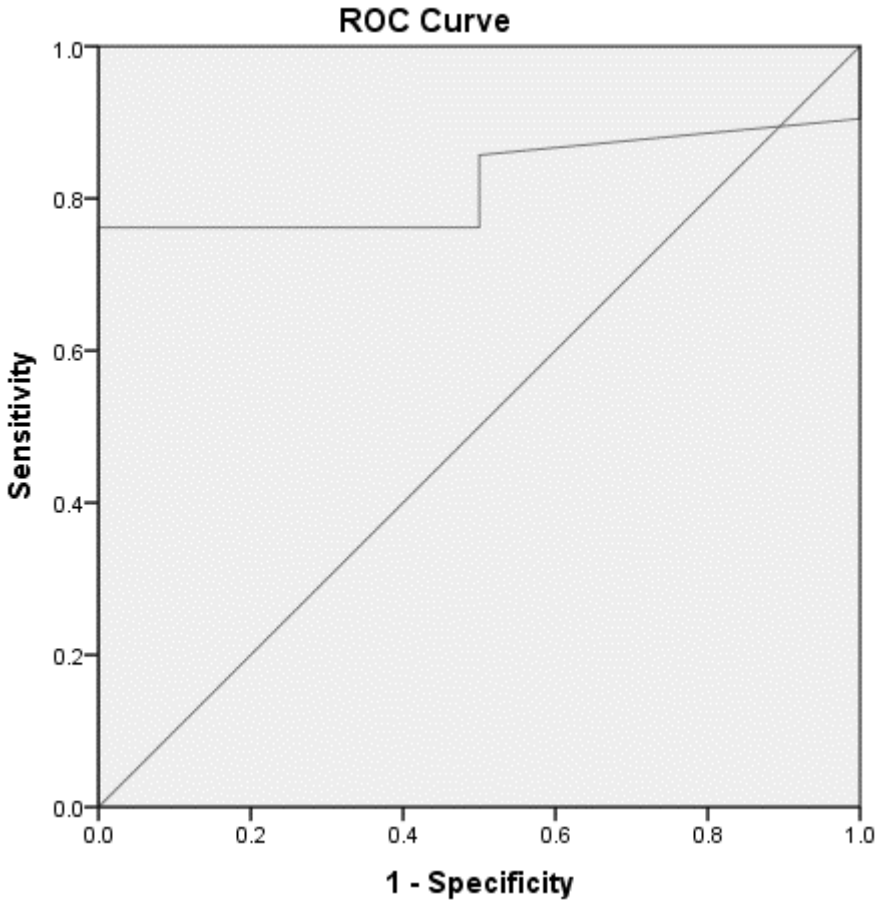


Figure 3. 3 The receiver operating curve (ROC) of the TSH marker to predict iodine status of populations using goitre prevalence as a reference marker (Twenty-three studies included in the analysis).

The area under curve (AUC) was 0.82 (95% CI: 0.65, 0.99). TSH, thyroid stimulating hormone.

Cut-off for the percentage of TSH >5 mIU/L from heel prick samples to define iodine sufficiency

The AUC was 0.71 (95% CI: 0.56, 0.87) for TSH vs MUIC and 0.82 (95% CI: 0.65, 0.99) for TSH vs goitre. Based on the coordinates of the ROC, the cut-off to define iodine sufficiency in populations using the percentage of newborn TSH >5 mIU/L was between 4.31% and 5.36% when using MUIC as a reference marker **(Supplemental Table 3.11)** and between 2.85% and 3.15 % when using goitre as a reference marker **(Supplemental Table 3.12)**.

Discussion

To the best of our knowledge, this is the first systematic review which compares the agreement between newborn TSH concentration and MUIC or goitre prevalence to classify iodine status in populations. We showed that based on the current criteria, TSH has a better agreement with goitre prevalence than MUIC in school age children. We also found that the current newborn TSH criteria (newborn TSH >5 mIU/L greater than 3%) in classifying iodine deficiency is supported when using goitre prevalence as a reference marker, but a higher cut-off of approximately 5% was indicated using MUIC in school age children as the reference marker.

MUIC of school age children from spot urine samples is recommended by the WHO as the first choice to assess iodine status in populations [12, 178]. It is a simple, convenient, inexpensive and widely used marker to classify iodine status of populations [178]. Revision of the TSH cut-off has been suggested by previous research [45, 52, 54, 179]. Based on the ROC analysis, we showed that a cut-off between 4.3% and 5.4% of newborn with TSH >5 mIU/L resulted in the highest specificity and sensitivity to define iodine status when the MUIC of school age children was used as a reference marker. MUIC, TSH, and goitre prevalence reflect iodine status over different periods. MUIC is a sensitive marker to identify short term changes in iodine intake of the populations unlike goitre prevalence which mainly reflects longer term low iodine intake [12]. Elevated neonatal TSH is an indicator of exposure to iodine deficiency during the fetal life while goitre indicates chronic iodine deficiency in school age children [12, 206]. Enlargement of the thyroid gland and elevation of newborn TSH concentration have also been linked to excessive iodine intake, they also have common causes other than chronic iodine deficiency including hypothalamic dysfunction and exposure to anti-thyroid compounds like Goitrogens and drugs [207-209]. While our results showed

that newborn TSH had a better agreement with goitre prevalence than MUIC, TSH or goitre prevalence may not be the appropriate markers to monitor short term effects of iodine interventions [1, 87]. Inter-observer variations in the classification of goitre by palpation or ultrasound examination, even among experienced examiners, make goitre prevalence a less-preferred marker for population iodine status [12].Nevertheless, both palpation and ultrasound examination methods consistently classified the iodine status of populations in the limited number of studies that used both methods to detect the presence of goitre.

A number of studies classified populations as iodine sufficient by MUIC but mildly iodine deficient by TSH concentration [47, 52, 183, 186, 198, 199, 202, 203]. However, many studies used newborn heel prick blood samples outside of the WHO's recommended time frame of 3-4 days after birth [54, 60, 181, 183, 187, 188]. The lack of an agreement between markers when newborn blood samples were collected outside of the WHO's recommendation suggests that TSH from newborn blood samples collected outside of 3-4 days is not suitable for classifying the iodine status of populations based on the current criteria. The current criteria may not be appropriate in countries where the newborn screening is performed before day 3 of birth due to early discharge from hospitals. Differences in the type of urine sample (spot urine or 24-hour urine) and the timing of urine sample collection may contribute to the low agreement between the MUIC and TSH [72, 210]. We are unable to conduct a sensitivity analysis based on different urine collection methods as approximately half of the included studies reported using spot urine samples while the remaining did not report what type of urine samples were collected for the UIC analysis.

In studies which assessed iodine status after one year of the introduction of an iodine intervention program, both TSH and MUIC showed an improvement in iodine status of populations, but the degree of iodine deficiency was classified differently using TSH vs.

MUIC [54, 84, 179, 180, 188, 197]. Currently there are no specific guidelines regarding which of the population markers would be the most appropriate for use following iodine intervention programs. A re-evaluation of the TSH criteria with clear guidelines may improve the sensitivity and specificity of the TSH to classify iodine status of populations. Application of the same TSH cut-off across all ethnic groups may also be a potential factor contributing to the disagreement among markers of population iodine status as it has been showed that differences in the newborn TSH concentration across different ethnic groups existed [211].

Our finding of a higher agreement between TSH and MUIC or goitre in iodine deficient than iodine sufficient populations also highlights the need for further research to formulate population specific guidelines in terms of which marker is the most appropriate to use in different settings. When disagreement occurs between markers of population iodine status, using MUIC in school age children as the first choice would facilitate the decision making process in determining appropriate population strategies to prevent iodine deficiency or monitor the efficacy and safety of any intervention programs [45]. In addition, iodine intake data from population surveys and from the vulnerable groups like pregnant women should also be considered in the decision-making process. The higher agreement between the newborn TSH and MUIC of the pregnant reflects the strong correlation between newborn TSH and maternal iodine status during pregnancy [57, 116, 198], suggesting that the newborn TSH marker may be a good indicator of the iodine status of pregnant women.

Our results indicate that the cord blood TSH is not a good marker based on the current cut-off. Of studies analysing cord blood TSH, there was no agreement between TSH and MUIC or goitre in classifying the degree of iodine deficiency [83, 85, 182, 183]. The discrepancy may be due to the use of the same cut-off in both heel prick and cord blood samples to classify iodine deficiency in populations because TSH concentration is higher in cord blood than in

newborn heel prick samples [178, 183]. The methods used to analyse newborn vs. cord blood TSH were also different. All studies analysed cord blood TSH using Enzyme Linked Immune Sorbent Assay (ELISA) while Dissociation Enhanced Fluroimmunoassay (DELFIA) was used for most of heel prick samples. Hence, the TSH assay methods used to measure the TSH concentration need to be considered in the interpretation of the results. A different TSH cut-off (cord blood vs. infant blood samples) may improve the agreement between TSH and MUIC or goitre in classifying iodine status in populations.

Our review has some limitations. Firstly, we retrieved some data on iodine status from different publications for the same country or region at the same time to assess the agreement between markers. Whether the study populations in the different publications matched was difficult to ascertain. Secondly, some studies assessed MUIC and/or goitre prevalence in school age children outside of the recommended age range of 6 - 12 years. Lastly, neonatal TSH is used as an indicator of population iodine status mainly in countries which have congenital hypothyroidism screening programs. Hence, our findings may not be applicable to those countries which do not have TSH screening programs currently.

In conclusion, this systematic review shows that neonatal TSH concentration >5 mIU/L greater than 3% as a population iodine status marker has a better agreement with goitre prevalence than MUIC in school age children. TSH concentration from heel prick samples outside of the WHO's recommendation of 3-4 days after birth or from cord blood samples is not suitable to classify iodine status of populations. Further research to determine the appropriate TSH cut-off as well as clear guidelines on which population markers is the most appropriate for use in different settings are recommended.

Acknowledgements

We thank Michael Draper (Librarian at University of Adelaide) for his assistance in building search strategies of databases.

Authors' contribution

MMW and SJZ designed the study. MMW performed literature search, data extraction, analysis and drafting the manuscript. SJZ oversaw the study conduct and critically reviewed the manuscript. PM contributed to the design, interpretation of the results and reviewed the manuscript. All authors approved the final manuscript.

Conflict of interest

All authors declare no conflict of interests.

CHAPTER 4 AGREEMENT BETWEEN URINARY IODINE MARKERS DERIVED FROM SPOT URINE SAMPLES TO CLASSIFY IODINE STATUS OF PREGNANT WOMEN

This chapter has been written based on the style of the *Journal of Nutrition.*

4.1 Manuscript

Wassie MM, Roberts, CT, & Zhou, S. Agreement between urinary iodine markers derived from spot urine samples to classify iodine status of pregnant women (To be submitted to the *Journal of Nutrition*).

Statement of Authorship

Title of Paper	Agreement between urinary iodine markers derived from spot urine samples during pregnancy
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wassie MM, Roberts CT, Zhou SJ. Agreement between urinary iodine markers derived from spot urine samples during pregnancy. Prepared based on the guidelines of Journal of Nutrition.

Principal Author

Name of Principal Author (Candidate)	Molla Mesele Wassie		
Contribution to the Paper	Design, laboratory analysis of urine samples, statistical analysis, interpretation of data, manuscript preparation and critical revision of the manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/03/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	*Claire T Roberts*

Contribution to the Paper	Design, supervise data collection, interpretation of data and critical revision of the manuscript		
Signature		Date	25/03/2020
Name of Co-Author	Shao Jia Zhou		
Contribution to the Paper	Conception, design, statistical analysis, interpretation of data and critical revision of the manuscript		
Signature		Date	25/03/2020

Please cut and paste additional co-author panels here as required.

Title Page

Agreement between urinary iodine markers derived from spot urine samples to classify iodine status of pregnant women

Author Names

Molla Mesele Wassie, Claire T Roberts, Shao Jia Zhou

Authors' affiliations

School of Agriculture Food and Wine, Faculty of Sciences, The University of Adelaide, Adelaide, Australia (MMW, SJZ).

Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia (MMW).

Adelaide Medical School, Faculty of Health and Medical Sciences, *The University of Adelaide, Adelaide, Australia (CR)*

Conflict of Interest and Funding Disclosure

MMW was supported by The University of Adelaide, Australian Government Research Training Program. The STOP study was supported by The University of Adelaide grant number U1111-1160-9802.

The word count for the entire manuscript

5919

Conflict of Interest and Funding Disclosure

None.

Corresponding Author

Shao Jia Zhou

School of Agriculture, Food & Wine

University of Adelaide, Waite Campus

PMB1, Glen Osmond, SA 5064

Phone: +61 8 8313 2065

Email: jo.zhou@adelaide.edu.au

The number of figures

Two.

The number of tables

Three.

Supplementary data submitted

Two.

Short running title

Agreement between markers of urinary iodine during pregnancy

Abbreviations

I/Cr, iodine-to-creatinine ratio; UIC~Cr, urinary iodine concentration corrected for creatinine; UIC, urinary iodine concentration; UIE, urinary iodine excretion and WHO, World Health Organization.

Abstract

Background: Multiple markers of iodine excretion from urine, including the 24-hours urinary iodine excretion (24-h UIE), urinary iodine concentration (UIC), iodine-to-creatinine ratio (I/Cr) and UIC-corrected for creatinine (UIC~Cr) have been used to assess iodine status of populations. There is limited evidence on the agreement between these markers in pregnant women.

Objectives: To assess the agreement between urinary iodine excretion measurements in pregnant women.

Methods: Spot UIC and urinary creatinine were measured in 534 pregnant women at 12 weeks of gestation. Iodine deficiency was defined based on the World Health Organization criteria of median UIC <150 µg/L. 24-h creatinine was estimated by using published equations by Mage et al., which was then used to calculate 24-h UIE. UIC~Cr was calculated from a regression of UIC over urine creatinine. Agreements between markers (estimated 24-h UIE, UIC, I/Cr and UIC~Cr) were assessed using the Bland-Altman plot and kappa coefficient (K).

Results: The median [Inter Quartile Range (IQR)] UIC was lower than estimated 24-h UIE [207.7 (121.7, 335.5) µg/L vs 261.0 (154.9, 466.3) µg/day, P <0.001], but did not differ from I/Cr or UIC~Cr. Based on the Bland-Altman plot, estimated 24-h UIE was 71.8, 76.8 and 86.0 units higher than I/Cr, UIC and UIC~Cr, respectively. I/Cr was 5.0 and 14.2 units higher than UIC and UIC~Cr, respectively. The agreement between the markers to classify iodine status in pregnant women was substantial between estimated 24-h UIE and I/Cr (k = 0.7) or UIC~Cr (K = 0.6), but fair (k = 0.3) between UIC and estimated 24-h UIE or between UIC and I/Cr.

Conclusions: All markers classified the study population as iodine sufficient. I/Cr has a better agreement with estimated 24-h UIE than UIC or UIC~Cr. Further study assessing which of these markers is the best surrogate marker of the actual 24-h UIE during pregnancy is recommended.

Key words: iodine, urine, creatinine, agreement, pregnancy

Introduction

Iodine deficiency or excess during pregnancy has been linked to impaired neurodevelopment and poor child growth [65, 212]. Hence, assessing iodine status in pregnant women using appropriate markers is an important public health measure [213]. There are several markers of iodine excretion from urine that have been used to assess iodine status of populations as approximately 90% of dietary iodine is excreted in the urine [214]. The 24-h urinary iodine excretion (UIE) in µg/day is viewed as the gold standard for assessing iodine status [12] because the original World Health Organization's (WHO) criteria established in 1994 to define iodine sufficiency were based on the lowest level of 24-h UIE associated with the absence of goitre, a functional outcome of iodine deficiency [37]. The study conducted in Central America, where the WHO's criteria were derived, showed that areas with 24-h UIE >100 µg/day had no goitre in adults and school age children (SAC) [41].

As collecting 24-h urine is a significant burden on participants and is not feasible in fieldwork, WHO subsequently recommended the use of spot urinary iodine concentration (UIC) expressed in µg/L [12] over 24-h UIE to assess iodine status of populations. This was based on the assumption that the average volume of 24-h urine is 1 L in SAC. The 24-h UIE and UIC would be comparable in SAC as their average 24-h urine volume is close to 1 L [215], but they may not be compatible in adults as their 24-h urine volume is approximately 1.5 L [216]. Hence, spot UIC is 60-65% lower than the 24-h UIE in adults [217]. Iodine-to-creatinine ratio (I/Cr) has been shown to be comparable with 24-h UIE when the 24-h urine creatinine excretion was 1 g and the 24-h urine volume was 1 L [214, 218]. However, the creatinine excretion level and 24-h urine volume vary across different age, sex, ethnicity and nutritional status, and these factors affect the compatibility of the 24-h UIE and I/Cr [68, 73, 219]. A recent systematic review showed that neither UIC nor I/Cr were good predictors of 24-h UIE although the included studies were mainly from adult populations with small

sample sizes (ranged from 29 to 188 participants) [220]. It has been shown that 24-h UIE calculated from spot UIC and 24-h creatinine excretion estimated using published equations had a better agreement with 24-h UIE in healthy adults [66, 69].

Recently, UIC corrected for creatinine by a regression method (UIC~Cr) was also used as a marker of iodine status [64]. In this method, UIC was regressed over the urinary creatinine excretion to compute residuals of the model, which represent the difference between the actual UIC and UIC predicted by urinary creatinine concentration, were added to the UIC [64, 221] and expressed as µg/L. Unlike I/Cr, the UIC~Cr value is uncorrelated with factors affecting urinary creatinine concentration. Due to this, UIC~Cr may be a better marker than I/Cr to define iodine status at an individual level.

Several studies have applied the WHO criteria to define iodine deficiency based on UIC (< 150 µg/L) to the 24-h UIE, UIC and I/Cr to define iodine status in pregnant women [65, 67, 71], but little is known about the agreement between these markers in pregnancy except in a small study (n = 143) by Li et al. [67] which showed a better agreement between I/Cr and 24-h UIE than UIC and 24-h UIE. While spot UIC is recommended by WHO as a marker of population iodine status, there is a scarcity of evidence to guide the choice of a suitable marker to assess individual iodine status in pregnancy. This study assessed the agreement between UIC, estimated 24-h UIE, I/Cr and UIC~Cr to define iodine status in pregnancy in a well-nourished iodine sufficient population.

Methods

Study design, setting and participants

A cross-sectional study. We used urine samples collected from pregnant women in South Australia participating in the Screening Tests to Identify Poor Outcome in Pregnancy study (STOP) [222]. Participants were recruited between 2015 and 2017. Nulliparous women with a

singleton pregnancy of less than 12 weeks of gestation who attended the first antenatal clinic were included. Women with known chronic disease were excluded [222]. A random spot urine sample was collected from participating pregnant women at 12 weeks of gestation and all samples have been stored at -80 degrees Celsius for batched UIC analysis. Data on maternal socio-economic and demographic characteristics were collected using a structured interviewer-administered questionnaire by research midwives. Height and weight measurement and urine sample collection was done by research midwives at the study entry. Six hundred pregnant women from the STOP study were randomly selected to assess iodine status (see Chapter 6) and 534 of them who had urinary creatinine excretion were included in this study.

Assessment of iodine status

Estimated 24-h UIE, UIC, I/Cr and UIC~Cr were assessed as markers of iodine status. The UIC was measured using a modified Sandell-Kolthoff spectrophotometric method (WHO's method B) in the spot urine samples [223]. The urinary creatinine concentration (mmol/L) was measured using the kinetic-Jaffe method at the South Australia Pathology Laboratory, an accredited diagnostic laboratory. Urine creatinine concentration in mmol/L was converted to g/L by dividing it by 8.840 [78]. I/Cr is the ratio of urinary iodine concentration in µg/L and urinary creatinine excretion in g/L. The 24-h UIE was calculated based on I/Cr and the 24-h urinary creatinine excretion as follows: I/Cr in µg/g x 24-h urinary creatinine excretion in g/day [214], the 24-h creatinine was estimated from the creatinine concentration in the spot urine sample. The estimated 24-h urinary creatinine value was calculated using two different predictive equations by Mage et al. [69] and by Kesteloot and Joosens [68]. While the estimated 24-h creatinine values were similar using both equations (**Supplemental table 4.1**), we used the estimated 24-h UIE value based on the Mage et al. method for comparison with other urinary excretion markers because it has been shown to have a better agreement with

24-h UIE [72]. The UIC~Cr was calculated by adding the residuals, which were created by regressing UIC against urinary creatinine concentration (both log-transformed) [221] to the UIC.

Statistical Analysis

Data were analysed using Stata 15 (Statacorp LP, College Station, Texas). Iodine status was reported as UIC in µg/L, estimated 24-h UIE in µg/day, I/Cr in µg/g and UIC~Cr in µg/L. The iodine status of women was categorised using the WHO criteria based on UIC: iodine deficiency (<150 µg/L), sufficiency (150-500 µg/L) and excess (≥500 µg/L) based on the WHO criteria [12]. We applied the same cut-offs to all markers irrespective of their units for comparisons with the literature. The median urinary iodine excretion values (using Wilcoxon rank-sum test) and the percentage of pregnant women in the iodine deficiency, iodine sufficiency or iodine excess category (using Pearson chi-square test) were compared across all markers. As the variables were not normally distributed, they were log-transformed for a pairwise correlation test.

Iodine status of the study population was classified using each of the markers. Agreement between markers in classifying iodine status of individual pregnant women was assessed using the 24-h UIE as a reference marker because it has been shown to better predict the 24-h UIE compared with UIC and I/Cr [66, 69]. Agreement between markers to define iodine status at individual-level was assessed in three ways. First, sensitivity and specificity of spot urinary iodine markers in classifying the iodine status of pregnant women compared with the estimated 24-h UIE was determined. Markers which classified iodine deficiency during pregnancy (<150 units) consistently with the estimated 24-h UIE marker were categorised as true positives while those consistently diagnosing iodine sufficiencies (≥150 units) were considered as true negatives. The sensitivity indicates the ability of the urinary iodine markers to correctly classify iodine deficient pregnant women while specificity refers to the ability of

the urinary iodine markers to correctly classify iodine sufficient pregnant women [190]. Second, the Kappa Coefficient (K) was used to assess the agreement between each of the marker and estimated 24-h UIE to classify iodine status as deficient or sufficient (as categorical variables). The agreement between markers was defined as poor if k <0.01, slight if K = 0.01-0.20, fair if K = 0.21-0.40, moderate if k = 0.41-0.60, substantial if K = 0.61-0.80, almost perfect if K = 0.81-1.0 by using the Kappa statistics [224]. Third, the Bland-Altman method was used to determine the mean difference (bias) and limits of agreement (LOA) between two markers [225]. The LOA are ± 1.96 SD of the mean difference of two markers. The agreement between other urinary iodine markers (I/Cr with UIC or UIC~Cr) was also explored. Based on the sample size (n = 534), we had greater than 95% power to compare the percentage of women with UIC <150 μg/L and I/Cr <150 μg/g at a 0.05 significance level; two-sided [81]. Statistical significance was set at a value of P <0.05.

Results

Participant characteristics

Table 4.1 presents the characteristics of 534 study participants. The mean (± SD) age and BMI of participants were 25.7 ± 5.0 years and 28.1 ± 7.3 kg/m^2, respectively. The mean (± SD) estimated 24-h urinary creatinine excretion was 1.3 ± 0.3 μg/day.

Table 4. 1 Characteristics of study participants[1]

Variables	STOP (n = 534)
Age, y	25.7 ± 5.0
Married	464 (87%)
Caucasian race	445 (83%)
Completed tertiary education	109 (18%)
SEI	903.9 ± 86.4
Drunk alcohol during first trimester	75 (14%)
Smoked during first trimester	98 (18%)
Took iodine-containing supplements during first trimester	400 (75%)
BMI, kg/m^2	28.1 ± 7.3
Urine creatinine, g/L	1.2 ± 0.8

[1]Values are presented as means ± SDs otherwise specified. BMI, body mass index; SD, standard deviation; SEI, Socio Economic Index for areas based on the Index of Relative Socio-Economic Disadvantage; STOP, Screening Tests to Identify Poor Outcome in Pregnancy; N, sample size; NA, not applicable.

[2]The mean SEI indicated low socio-economic status compared with other areas in Australia [226].

Iodine status in the study population

There was a significant difference in the median concentration between UIC and estimated 24-h UIE ($P < 0.001$) or between UIC and UIC~Cr ($P = 0.01$) but there was no significant difference between UIC and I/Cr ($P = 0.40$). The median estimated 24-h UIE was different from the median I/Cr or UIC~Cr while median I/Cr did not differ from median UIC~Cr ($P = 0.57$). The study population was classified as iodine sufficient based on all urinary iodine markers (median values >150). However**,** as shown in **Table 4.2**, there was a significant difference in the percentage of pregnant women with urinary iodine concentration >150 or >500 by all markers ($P < 0.001$).

Table 4. 2 Iodine status of participants using different markers (n=534)[1]

	UIC, µg/L	Estimated 24-h UIE, µg/day	I/Cr, µg/g	UIC~Cr µg/L
Median	207.7 (121.7, 335.5)	261.0 (154.9, 466.3)	214.0 (126.9, 343.1)	200.0 (131.8, 332.2)
Classification of iodine status				
<150 (deficiency)	168 (31%)	124 (23%)	177 (33%)	180 (34%)
150-499 (sufficiency)	308 (58%)	300 (56%)	300 (56%)	314 (59%)
≥500 (excess)	58 (11%)	110 (21%)	57 (11%)	40 (7%)

[1]Values are presented as medians (IQR) otherwise specified. I/Cr, iodine to creatinine ratio; SD, standard deviation; STOP, Screening Tests to Identify Poor Outcome in Pregnancy; UIC, urinary iodine concentration, UIC~Cr, UIC corrected for creatinine by regression method; UIE, urinary iodine excretion.

Agreement between markers of iodine status at an individual level

The sensitivity and specificity of the markers in classifying iodine status (<150 or ≥150) is presented in **Table 4.3**. When compared with the estimated 24-h UIE, UIC showed lower sensitivity and specificity than I/Cr and UIC~Cr. I/Cr had the best sensitivity and specificity when compared with the estimated 24-h UIE. A contingency table is presented in **Supplemental Table 4.2** for the number of pregnant women who were consistently categorised into iodine deficiency (<150 units) and iodine excess (≥500 units) by each marker.

Table 4. 3 Sensitivity and specificity of urinary iodine markers to define iodine status in pregnancy[1]

Reference markers	Test markers			
	UIC	Estimated 24-h UIE	I/Cr	UIC~Cr
Sensitivity and specificity of markers to define iodine deficiency <150 units[2]				
Estimated 24-h UIE as the reference				
Sensitivity	0.56	NA	0.98	0.88
Specificity	0.76	NA	0.87	0.83
I/Cr as the reference				
Sensitivity	0.49	0.69	NA	0.79
Specificity	0.77	0.99	NA	0.89
UIC as the reference				
Sensitivity	NA	0.42	0.52	0.71
Specificity	NA	0.85	0.75	0.88
Sensitivity and specificity of markers to define iodine excess ≥500 units[2]				
Estimated 24-h UIE as the reference				
Sensitivity	0.37	NA	0.49	0.35
Specificity	0.96	NA	0.99	1.00
I/Cr as the reference				
Sensitivity	0.42	0.00	NA	0.80
Specificity	0.93	0.79	NA	0.95
UIC as the reference				
Sensitivity	NA	0.71	0.41	0.55
Specificity	NA	0.86	0.93	0.98

[1]I/Cr, urinary iodine concentration to creatinine ratio; NA, not applicable; UIC, urinary iodine concentration; UIC~Cr, urinary iodine concentration corrected for creatinine; UIE, urinary iodine excretion. The sensitivity of a marker was calculated as a/(a + b) where 'a' is the number of pregnant women classified as iodine deficient by both the test and reference markers and 'b' is the number of pregnant women classified as iodine deficient by only the test marker. The specificity of a marker was calculated as d/(c + d) where 'd' is the number of pregnant

women classified as iodine sufficient by both the test and reference markers and ‘c’ is the number of pregnant women classified as iodine sufficient by only the test marker [190].

[2]The WHO criteria based on median UIC in µg/L was applied to all markers to define iodine deficiency (<150 units) and excess (≥500 units) including estimated 24-h UIE in µg/d, I/Cr in µg/g and UIC~Cr in µg/L [12].

Based on kappa coefficient, the agreement between markers to classify iodine status as <150 or ≥150 was 'substantial' for estimated 24-h UIE and I/Cr (k =0.74, P <0.001) or estimated 24-h UIE and UIC~Cr (k = 0.61, P <0.001) but 'fair' for estimated 24-h UIE and UIC (k = 0.30, P <0.001). Furthermore, the agreement between the markers was 'substantial' for I/Cr and UIC~Cr (k =0.68, P <0.001), 'Moderate' for UIC and UIC~Cr (k = 0.60, P <0.001) and "fair" for I/Cr and UIC (k = 0.27, P <0.001).

As shown in **Figure 4.1**, estimated 24-h UIE was 71.8 units (LOA: -118.8 to 262.4 units), 76.8 units (LOA: -298.2 to 451.8 units) and 86.0 units (LOA: -171.4 to 343.3 units) higher than I/Cr, UIC and UIC~Cr, respectively. I/Cr was 5.0 (LOA: -344.7 to 354.6) units higher than UIC and 14.2 (LOA: -184.7 to 213.0) units higher than UIC~Cr (**Figure 4.2**).

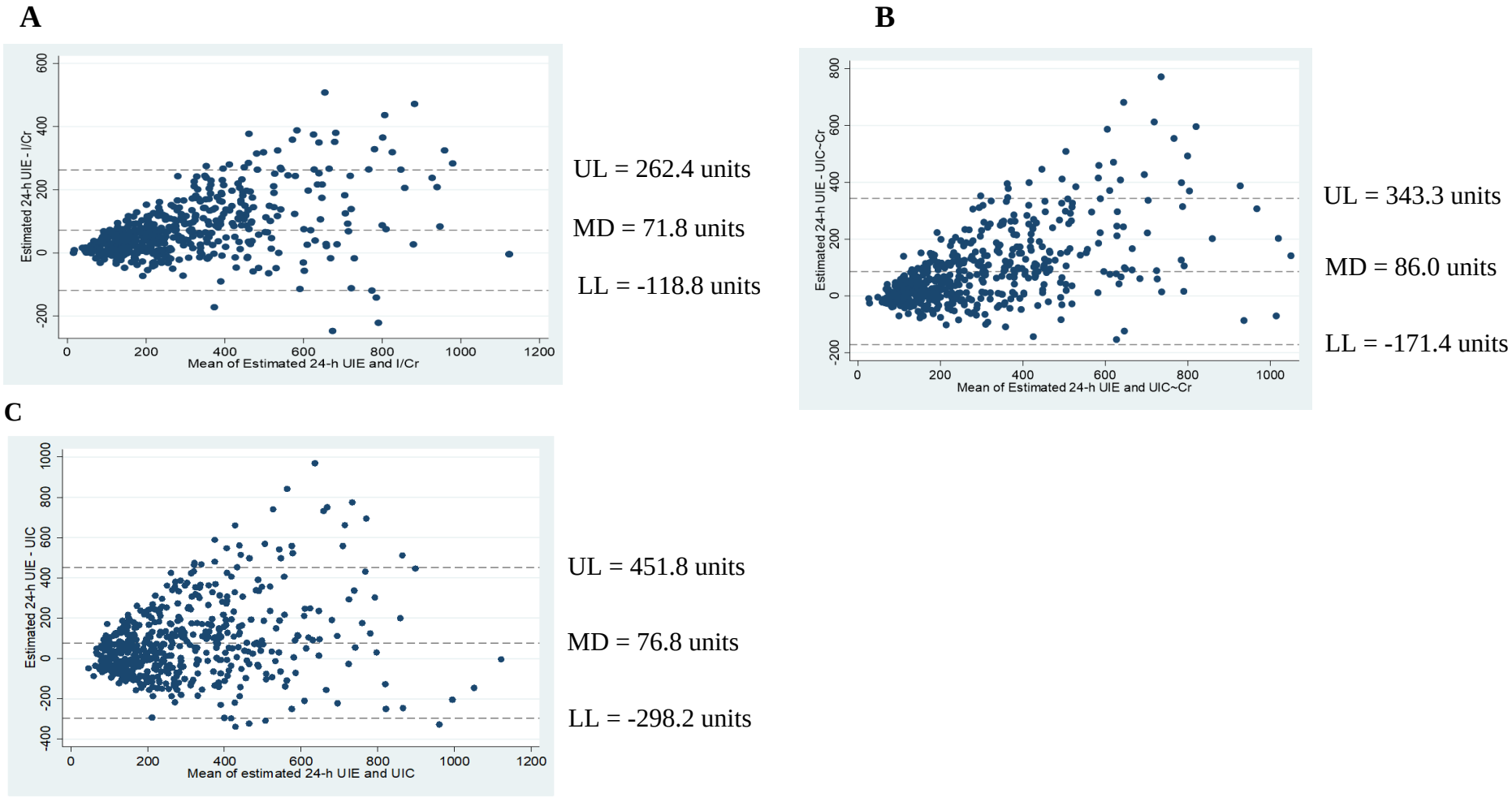


Figure 4. 1 The Bland-Altman plots for the agreement between estimated 24-h UIE and: A) I/Cr; B) UIC~Cr and C) UIC.

I/Cr, urinary iodine concentration to creatinine ratio; LL, lower limit of agreement; MD, mean difference; UIE, urinary iodine excretion; UIC, urinary iodine concentration; UIC~Cr, urinary iodine concentration corrected for creatinine; UL, upper limit of agreement.

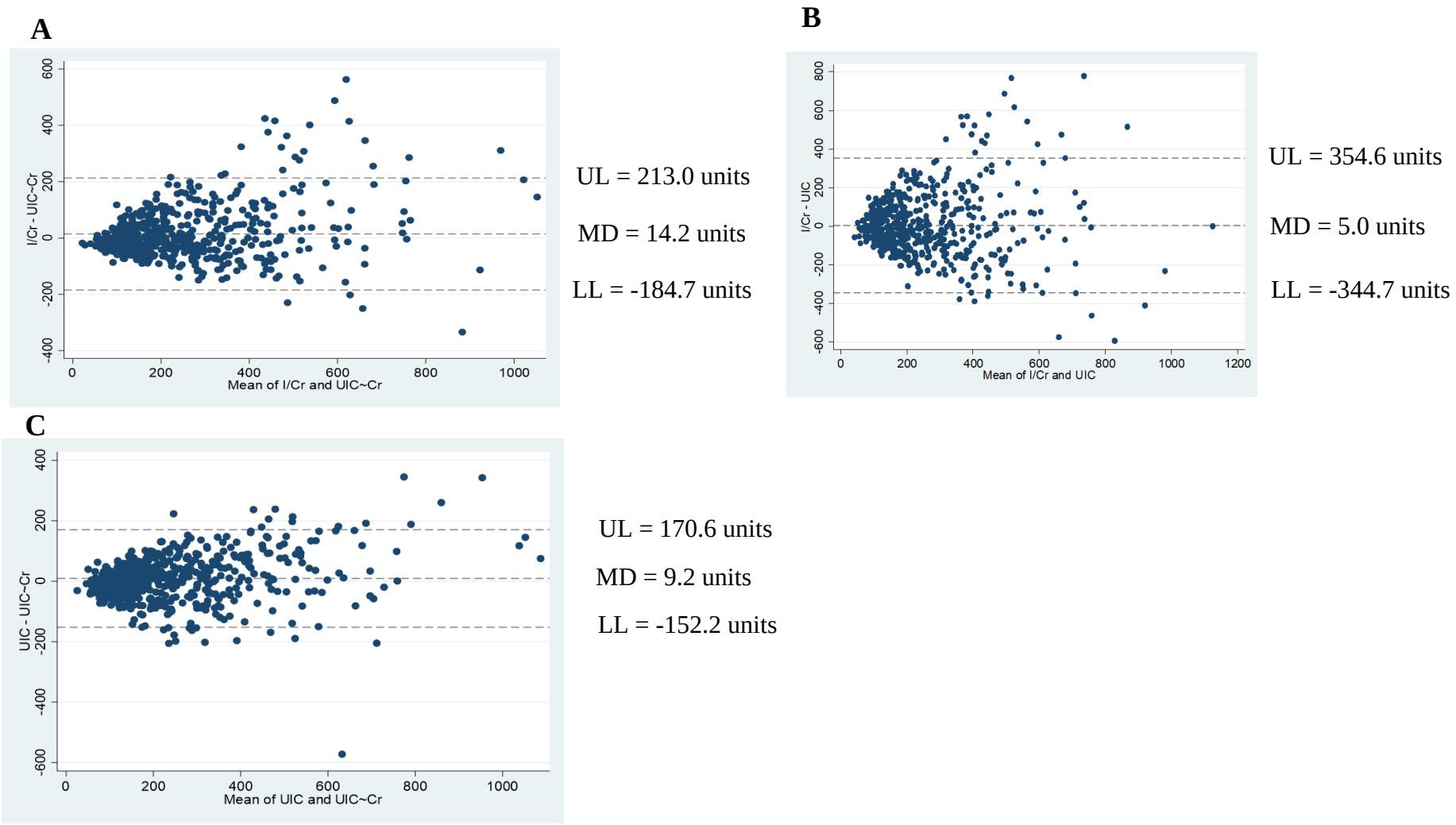


Figure 4. 2 The Bland-Altman plot for the agreement between A) I/Cr and UIC~Cr; B) I/Cr and UIC; C) UIC and UIC~Cr.

I/Cr, urinary iodine concentration to creatinine ratio; LL, lower limit of agreement; MD, mean difference; UIE, urinary iodine excretion; UIC, urinary iodine concentration; UIC~Cr, urinary iodine concentration corrected for creatinine; UL, upper limit of agreement.

The estimated 24-h UIE had a better correlation with I/Cr ($r = 0.94$, $P < 0.001$) than UIC~Cr ($r = 0.85$, $p < 0.001$) and UIC ($r = 0.57$, $P < 0.001$), while I/Cr had a better correlation with UIC~Cr ($r = 0.86$, $P < 0.001$) than UIC ($r = 0.52$, $P < 0.001$).

Discussion

The study population was defined as iodine sufficient when the WHO threshold of UIC >150 (µg/L) was applied to all markers. When considered as a marker of individual iodine status, UIC, I/Cr and UIC~Cr underestimated urinary iodine excretion compared with estimated 24-h UIE though I/Cr had a better agreement with estimated 24-h UIE compared with the other markers. However, the urine samples were obtained from iodine sufficient populations and the results may not be generalisable to iodine deficient populations.

Expressing urinary iodine excretion as µg/g of creatinine or as 24-h UIE in µg/day may not result in a more accurate assessment of iodine status of populations [74]. Our finding, that iodine status classified by spot UIC and I/Cr or UIC~Cr or estimated 24-h UIE was consistent in well-nourished pregnant women from iodine sufficient populations, suggests that the use of spot UIC as a marker would be appropriate in similar settings. Besides, the cost of measuring urine creatinine value that is required to estimate 24-h creatinine excretion would be a challenge for surveillance of iodine status by estimated 24-h UIE marker in large populations groups. Previous studies on populations of pregnant women also reported a consistent classification of iodine status using UIC and I/Cr [65, 67, 77, 78, 80] and using UIC and UIC~Cr [64]. However, using estimated 24-h UIE overestimated the percentage of pregnant women with iodine excess (≥500 units) and underestimated the percentage of iodine deficiency (<150 units) when compared with UIC and other spot urinary iodine markers (I/Cr or UIC~Cr). This finding of an overestimated percentage of iodine excess by 24-h UIE has also been observed in studies of healthy adults [61, 66, 72] and may be related to how the

current WHO UIC cut-off was established, which was derived from the 24-h UIE studies in SAC [41]. The 24-h urine volume is approximately 1.5 L in adults [216] which is higher than SAC and the 24-h UIE in adults would overestimate the UIC. There is limited data on the 24-h urine volume at 12 weeks of gestation, however a study in healthy nulliparous women showed a 24-h urine volume of >1.5 L during pre-pregnancy period and at 30-32 weeks of pregnancy [227]. Moreover, inconsistent classification of iodine status has also been reported by UIC vs I/Cr in populations of pregnant women [64, 71, 81]. The differences in UIC and I/Cr may be due to variation in 24-h creatinine excretion [64, 66] or the urine samples were collected at a different gestational age in the studies. Unlike the median I/Cr which showed an increasing trend from trimester one to trimester three, the median UIC decreased with increasing gestational age in several studies [67, 71, 77]. The UIC variations may be caused by physiological adaptations of the thyroid gland and an increase in glomerular filtration rate as gestational age increases [78, 228]. The increased maternal iodine requirement for thyroid hormone synthesis and maternal to fetal iodine transfer may lead to low UIC in late pregnancy [78]. Increase in glomerular filtration rate may lead to increased urine output which decreases the UIC across the trimesters [67]. Hence, as median UIC is lower in late pregnancy, application of the same UIC cut-off across all trimesters may contribute to incorrect classification of iodine status in pregnant women. Trimester specific spot UIC cut-points may be required to more accurately classify population-level iodine status during pregnancy [79].

Our results of a poor agreement between UIC and I/Cr or estimated 24-h UIE or UIC~Cr show that UIC is not a suitable marker of individual iodine status in pregnant women [12]. Expressing UIC per g of creatinine or estimating 24-h UIE from spot UIC reduced within-day variation in iodine excretion of individuals [66, 67, 220] but does not address the day-to-day variability in iodine intake, which needs a collection of ten or more repeated spot urine

samples [62, 63]. Comparison of each of the markers with a gold standard (24-h UIE) is required to identify the most appropriate marker to assess individual iodine status of pregnant women. Though estimated 24-h UIE showed a better agreement with 24-h UIE than I/Cr in adults, Li et al. suggested to use I/Cr over the estimated 24-h UIE as a suitable marker of individual iodine status of pregnant women [67]. Li et al. argued that estimating 24-h UIE from spot UIC is unnecessary as urine creatinine is not different between individuals in a specific population of pregnant women; i.e. reproductive age group and females [67]. As shown in our results, the daily estimated creatinine excretion was not the same between participants, the mean ± SD of 1.2 ± 0.8 g/L, and I/Cr and estimated 24-h UIE were not exchangeable. Besides, the urinary creatinine excretion may vary depending on nutritional status, age, and ethnicity [69] which affects the value of I/Cr or estimated 24-h UIE. The UIC~Cr marker, which is less affected by creatinine excretion [64], was compared with both I/Cr and estimated 24-h UIE in our study. Our results showed a better agreement between UIC~Cr and I/Cr than between UIC~Cr and estimated 24-h UIE. No previous study assessed agreement between UIC~Cr and 24-h UIE and further research is required to identify the best predictor of 24-h UIE in pregnancy among all markers obtained from spot urine samples.

The study has the following limitations: Firstly, we did not collect 24-h urine excretion so were unable to compare the agreement between markers from spot urine samples with the actual 24-h UIE. Secondly, we only assessed the urinary iodine markers at one time-point (12 weeks of gestation), and the results may not be generalisable in all trimesters during pregnancy [29, 71]. Thirdly, we did not collect data on the type and timing of the spot urine collection (first morning urine, fasting or not fasting) as timing and fasting may affect the UIC [72]. Lastly, both the Mage et al. and the Kesteloot and Joosens equations to estimate the 24-h urinary creatinine excretion were established based on healthy adult populations. The only factor in the equation by Mage et al. that differ between populations of pregnant and non-

pregnant women is the anthropometric score. Our study was conducted in pregnant women at 12 weeks of gestation and these women may not significantly differ from the non-pregnant women in terms of weight and body mass index.

In conclusion, all markers from spot urine samples defined the study population as iodine sufficient according to the WHO criteria for UIC, and correcting iodine for creatinine may not increase the accuracy to define population-wide iodine status in pregnancy. UIC shows a low agreement with estimated 24-h UIE or I/Cr as a marker of individual-level iodine status in pregnant women. Further study comparing the agreement between urinary iodine markers from spot urine samples, and UIE measured from 24-h urine samples to define individual-level iodine status in pregnant women is warranted.

Acknowledgements

We would like to acknowledge Dylan McCullough for assistance with urine sample selection and transporting the samples to the Waite Main Building, Waite Campus for analysis.

Statement of authors' contributions to the manuscript

MMW and SJZ designed the study; MMW analysed the UIC and performed the data analyses; and MMW drafted the paper with input from all authors. CR contributed to design, supervision of data collection and interpretation of results and reviewed the manuscript. SJZ oversight the conduct of the study. All authors read and approved the final manuscript.

Conflict of interest

Authors declare no conflict of interests.

CHAPTER 5 COMPARISON OF IODINE STATUS PRE- AND POST-MANDATORY IODINE FORTIFICATION OF BREAD IN SOUTH AUSTRALIA

This chapter has been published in *Public Health Nutrition* [229].

5.1 Publication

Wassie MM, Yelland, L., Smithers, L., Ranieri, E., & Zhou, S. (2019). Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: A population study using newborn thyroid-stimulating hormone concentration as a marker. *Public Health Nutrition*, **22**(16),3063-3072. DOI: https://doi.org/10.1017/S1368980019001915

To adapt to the format of the thesis, the form of the published version of the paper has been slightly modified in the following ways without any changes to the text content;

1. Re-numbered the consecutive page numbering
2. Re-numbered the consecutive reference numbering and used the numbered format of the references to be consistent with the rest of the chapters.
3. Re-numbered the tables and figures.
4. Embedded the tables and figures into the text.

Statement of Authorship

Title of Paper	Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: A population study using newborn thyroid-stimulating hormone concentration as a marker. Public Health Nutrition, 2019. doi: 10.1017/S1368980019001915
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wassie MM, Yelland LN, Smithers LG, Ranieri E, Zhou SJ. Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: a population study using newborn thyroid-stimulating hormone concentration as a marker. Public Health Nutr 2019:1-10. doi: 10.1017/S1368980019001915

Principal Author

Name of Principal Author (Candidate)	Molla Mesele Wassie		
Contribution to the Paper	Design, statistical analysis, interpretation of data, manuscript preparation and critical revision of the manuscript		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	13/03/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lisa Nicole Yelland		
Contribution to the Paper	Design, statistical analysis, interpretation of data and critical revision of the manuscript		
Signature		Date	18/3/20

Name of Co-Author	Lisa Gaye Smithers		
Contribution to the Paper	Design, interpretation of data and critical revision of the manuscript		
Signature		Date	16 March 2020

Name of Co-Author	Enzo Ranieri		
Contribution to the Paper	Design, supervise data collection, interpretation of data and critical revision of the manuscript		
Signature	(primary supervisor)	Date	25/03/2020

Name of Co-Author	Shao Jia Zhou		
Contribution to the Paper	Conception, design, statistical analysis, interpretation of data and critical revision of the manuscript		
Signature		Date	16/03/2020

Please cut and paste additional co-author panels here as required.

Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: a population study using newborn TSH concentration as a marker

Molla Mesele Wassie[1, 2], Lisa N Yelland[3,4], Lisa G Smithers[3], Enzo Ranieri[5], Shao Jia Zhou[1,6]*

Authors' affiliation

[1]School of Agriculture Food and Wine, Faculty of sciences, The University of Adelaide, Australia

[2]Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia

[3]School of Public Health, Faculty of Health and Medical Sciences, University of Adelaide, Australia

[4]South Australian Health and Medical Research Institute, Adelaide, Australia.

[5]South Australia Newborn Screening Centre, Women's and Children's Hospital, North Adelaide, Australia

[6]Robinson Research Institute, School of Paediatrics and Reproductive Health, The University of Adelaide, Adelaide, Australia

**Address Correspondence to*

Shao Jia Zhou

School of Agriculture, Food & Wine

University of Adelaide, Waite Campus

PMB1, Glen Osmond, SA 5064

Phone: 8313 2065

Email: jo.zhou@adelaide.edu.au.

Short title

Iodine status in South Australia.

Acknowledgements

We are grateful to the South Australia newborn screening centre for providing us the newborn screening data for infants born between 2005 and 2016.

Financial support

Molla Wassie was supported by The University of Adelaide Research Training Program scholarship. Lisa Yelland was supported by an Australian National Health and Medical Research Council Early Career Fellowship (ID 1052388).

Conflict of interest

None.

Authorship

The authors' responsibility is as follows: SJZ was responsible for the conception of the study and the overall conduct of the study; MMW, LY and SJZ were responsible for the design of the study; MMW and LY analysed and interpreted the data; LS and ER contributed to the design and interpretation of the results; MMW drafted the manuscript; All authors critically reviewed the manuscript. All authors approved the final manuscript.

Ethical Standards Disclosure

This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving human subjects were approved by the Human Research Ethics Committee (HREC) of the University of Adelaide (reference number 22274).

Abstract

Objective: This study aimed to evaluate the effect of mandatory iodine fortification of bread on the iodine status of South Australian populations using the newborn thyroid stimulating hormone (TSH) concentration as a marker.

Design: The study used an interrupted time series design.

Setting: TSH data collected between 2005 and 2016 (n=211,033) were extracted from the routine newborn screening program in South Australia for analysis. Iodine deficiency is indicated when more than 3% of newborns have TSH >5 mIU/l.

Subjects: Newborns were classified into three groups: the pre-fortification group (those born before October 2009); the transition group (born between October 2009 and June 2010); and the post-fortification group (born after June 2010).

Results: The percentage of newborns with TSH >5 mIU/l was 5.1%, 6.2% and 4.6% in the pre-fortification, transition and post-fortification groups, respectively. Based on a segmented regression model, newborns in the post-fortification period had a 10% lower risk of having TSH >5 mIU/l than newborns in the pre-fortification group [Incidence Rate Ratio (IRR): 0.90; 95% confidence interval (CI): 0.87, 0.94], while newborns in the transitional period had a 22% higher risk of having TSH >5 mIU/l compared with newborns in the pre-fortification period [IRR: 1.22; 95% CI: 1.13, 1.31].

Conclusion: Using TSH as a marker, South Australia would be classified as mild iodine deficiency post-fortification in contrast to iodine sufficiency using median urinary iodine concentration as a population marker. Re-evaluation of the current TSH criteria to define iodine status in populations is warranted in this context.

Key words: TSH, newborns, iodine deficiency, iodine fortification, population iodine status

Introduction

Iodine is an essential micronutrient required for synthesis of thyroid hormone [20], which is important for optimal growth and brain development [1, 2]. Iodine deficiency can affect thyroid hormone production, as indicated by altered levels of thyroxine (T_4), tri-iodothyronine (T_3) and thyroid stimulating hormone (TSH) [20]. Newborn TSH concentration is an indicator of fetal iodine nutrition prior to birth as iodine deficiency during fetal life can lead to an increase in the TSH concentration in the newborn [36, 137].

Methods of monitoring iodine status of populations include measuring *TSH* concentrations in newborns, median urinary iodine concentration (UIC) in school aged children, and prevalence of goitre in school aged children [230]. According to the World Health Organization (WHO), United Nations Children's Fund (UNICEF) and International Counsel for Control of Iodine Deficiency Disorders (ICCIDD) joint criteria, populations are classified as iodine deficient if more than 3% of newborns have TSH concentration >5 mIU/l in whole blood, or median UIC <100 μg/L, or goitre prevalence >5% in school aged children [230]. When TSH is used to classify population iodine status, the recent WHO recommendation [12] specified to use neonatal blood samples taken at 48-96 hours after birth because newborn TSH concentrations in the first 24 to 48 hours are higher due to the neonatal TSH surge [45] and are not recommended to be used to classify population iodine status.

In Australia, newborn TSH concentration is routinely assessed in whole blood spot samples taken by heel-prick as part of the Newborn Screening Test [231]. The test aims to detect disorders such as congenital hypothyroidism and newborn errors of metabolism. Iodine deficiency re-emerged in Australia [17, 116, 117, 200], and as a result, mandatory iodine fortification of bread was introduced in October, 2009 [16]. Pregnant and lactating women in

Australia are also recommended to take an iodine supplement of 150 μg per day since January, 2010 [107]. Following iodine fortification, the 2011-2012 National Health Survey and other surveys showed Australians are iodine sufficient as assessed by median UIC [29, 131, 232]. Data on iodine status of the population in Australia using newborn TSH as a marker are scarce. A recent study conducted in the Western Australia showed mild iodine deficiency using this marker [52] but the population was iodine sufficient based on UIC [131]. There is a lack of national data that assessed population iodine status using newborn TSH as a marker in the post-fortification period including South Australia.

Monitoring the iodine status of the population post-fortification using multiple markers is important to evaluate the efficacy and safety of iodine fortification program as disagreements between the TSH marker vs. median UIC or goitre prevalence in classifying iodine status of populations have been reported [45, 186]. In this study we use an interrupted time series (ITS) design to investigate newborn TSH levels before, during and after the implementation of mandatory iodine fortification program [233]. We also explore the impact of the intervention on the iodine status of the population using different TSH cut-offs to classify iodine status. Based on available evidence, we hypothesised that there will be a reduction in the proportion of newborns with TSH concentration >5 mIU/l following the fortification program.

Methods

Design, setting and participants: This was an ITS study using population-level data collected across the entire state of South Australia from 2005 to 2016 (n=254,156). The TSH concentration of all neonates who underwent newborn screening, a routine test in South Australia with a screening rate of 98% [234], were extracted from the newborn screening database in de-identified format. As part of the newborn screening process, newborn whole

blood samples were taken from neonates by heel-prick. Newborn TSH concentration was determined using Dissociation-Enhanced Lanthanide Fluorescence Immunoassay (DELFIA). The same method was used for the assessment of TSH over the entire study period. Newborns with congenital hypothyroidism, defined as newborn TSH concentration ≥13 mIU/l, were excluded from the study (n=388).

Outcome:

The primary outcome of this study was iodine status of the population as defined by the percentage of newborns with TSH >5 mIU/l based on the WHO criteria [230]. These criteria classify a population as iodine sufficient if the percentage of newborns with TSH >5 mIU/L is <3%, and as mildly, moderately and severely iodine deficient if the percentage of newborns with TSH >5 mIU/l is 3-19.9%, 20-39.9% and >40%, respectively [230]. Newborn TSH concentration declines over the first few days after birth [235, 236] and newborns sampled between 48-72 hours after birth had a higher median TSH concentrations than those sampled 72-96 hours after birth [52]. We considered the percentage of newborns with TSH >6 mIU/l as a secondary outcome [179], since the majority of the samples were collected early at 48-72 hours after birth when TSH concentrations are higher.

Exposure:

Newborns were classified into three groups based on their exposure to the mandatory iodine fortification during their fetal growth period assuming a full-term pregnancy of nine months. Since the fortification program was implemented in October 2009, newborns born before October 2009 formed the pre-fortification group. Newborns born between October 2009 and June 2010 formed the transitional group, as the fetal growth of these children occurred crossing both the pre- and post-fortification periods. Newborns born after June 2010 formed

the post-fortification group, as these children would have been conceived after the fortification program was mandated.

Time varying confounders:

In this study, time varying confounders are population characteristics that change over time. The timing of blood sample collection, birthweight and season at birth were adjusted in the analysis as potential time varying confounders. The timing of blood sample collection for TSH assessment has been shown to affect the newborn TSH concentration [43, 52, 183], with early blood samples collected in the first 24 hours of birth having higher TSH concentration due to the surge in neonatal TSH [45]. Low birthweight infants have been shown to have higher TSH concentrations [52, 237] and infants born in the spring and winter seasons had a higher TSH concentration due to cold atmospheric temperature [238]. Percentage of low birthweight infants may vary across time [239]. Use of iodine-containing antiseptics during delivery [45], gestational age at birth [47], sex of the newborn [47, 52], organochlorine pesticides exposure in perinatal period [50], and mode of delivery [45, 46] have been shown to affect TSH concentration, but there was little change in these factors over the study period in south Australia [240].

Statistical Analysis:

The data were analysed using Stata 14 (Statacorp LP, College Station, Texas). Median (interquartile range (IQR)) of TSH concentration and percentage of newborns with TSH concentration >5 mIU/L were calculated for the overall samples as well as separately by fortification group, newborn characteristics and year of sampling. Descriptive analyses were repeated for the percentage of newborns with TSH concentration >6 mlU/l. Chi-square tests were used to compare newborn characteristics (for categorical variables) and the percentage of newborns with TSH >5 mIU/l between fortification groups. The Kruskal-Wallis tests were

used to compare the median TSH concentration between subpopulations defined by newborn characteristics.

A segmented regression analysis [241] was employed to assess the effect of mandatory iodine fortification on the percentage of newborns with TSH concentration >5 mIU/l. Newborns with heel-prick blood sampled outside of the WHO recommendation of three to four days (48 to 96 hours) after birth were excluded from the segmented regression analysis. Newborns with missing information on time of blood sampling, or birthweight, or improbable values for birthweight, defined as <300 gm or >10 kg, were also excluded from the segmented regression analysis. Sex and gestational age were not included in the segmented regression analysis because data on sex were incomplete for post-fortification groups, while data on gestational age were incomplete for pre-fortification groups.

The primary analysis was an unadjusted analysis performed using a Poisson regression model with the number of newborns with TSH >5 per month as the outcome, the fortification groups as the primary predictor variable and the natural logarithm of the total number of newborns per month as an offset. Over-dispersion was handled by estimating the scale parameter by Pearson's chi-square statistic divided by the degrees of freedom. Several exploratory adjusted analyses were also performed. Firstly, birthweight and infants' age at sampling were adjusted in the model to control for possible time varying confounding. To allow for changes over time in birthweight we used the mean or median birthweight or the percentage of infants with low birthweight per month. Both linear and quadratic effects were considered. To examine potential changes over time in the infants' age at sampling we tested the median age at sampling or the percentage of infants sampled in each 6-hour period for each month. Akaike information criterion (AIC) was used to choose the best way to model the birthweight and time of sampling effects in these exploratory analyses, and the lowest value was taken for a

better model fit. Secondly, season of birth was included to model any seasonal trends in the data. Finally, time in months was included as a predictor to allow for a linear trend over time and an interaction between time in months and fortification group was included to test for possible slope changes between fortification periods.

Global tests were performed to assess whether there was any evidence of a jump in the percentage of newborns with TSH >5 mIU/l (fortification group effect) or, in exploratory analyses, a slope change (time in months by fortification group interaction effect) between the fortification periods. The effect of fortification group on the percentage of newborns with TSH concentration >5 mIU/l was reported as the incidence rate ratio (IRR) with a 95% confidence interval. A p- value less than 0.05 was considered statistically significant.

Results

Characteristics of the study population

Of the 254,156 newborns for whom TSH data were available, 388 were diagnosed with congenital *hypothyroidism* (15 per 10,000 births) and were excluded from the analysis. A heel prick blood sample was collected between 48 and 96 hours after birth in 213,642 (84.0%) newborns. Of these, 211,033 (98.8%) had data on birthweight and were included in the segmented regression analysis (**Figure 5.1**).

Table 5.1 shows the characteristics of the study population. The majority (91%) of newborns were sampled 48-72 hours after birth, both in the pre- and post-fortification periods. The percentage of newborns with TSH >5 mIU/l was higher in males than females, higher in lower birthweight than normal birthweight infants, higher in preterm than term infants, and higher in winter and spring than summer and autumn (**Table 5.2**). The percentage of

newborns with TSH >5 mIU/l and median TSH concentrations were highest in newborns blood sampled at two days of age and declined afterwards (**Figure 5.2**).

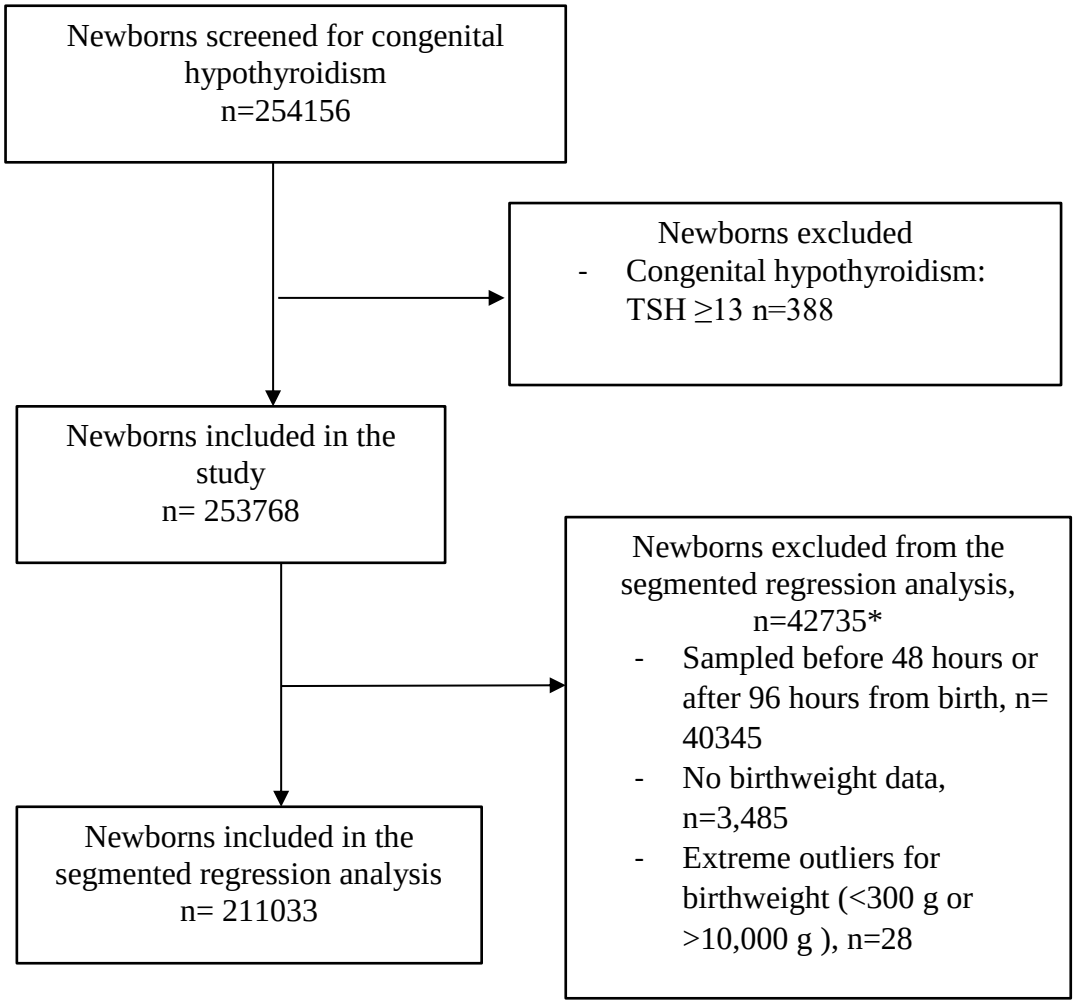


Figure 5. 1 Flow chart of participants
*Newborns could be excluded under multiple exclusion criteria. TSH, thyroid stimulating hormone.

Table 5. 1 Characteristics of the study population, 2005-2016, South Australia (n 211033)

Characteristic	n	Pre-fortification group n (%)	Transition group n (%)	Post-fortification group n (%)	P-value*
Neonatal age at blood sampling: 48-72 hrs	190997	75850 (90.7)	12200 (89.3)	102947 (90.5)	<0.01
Sex [a]: Males	30171	19432 (50.9)	3500 (52.0)	7239 (51.8)	<0.01
Birthweight: <2500 g	14720	5438 (6.5)	944 (6.9)	8318 (7.3)	<0.01
Gestational age [b]: <37 weeks	23456	2383 (15.3)	2021 (15.7)	19052 (17.5)	<0.01
Season					<0.01
Summer	51736	20115 (24.0)	4457 (32.6)	27164 (23.9)	
Autumn	53591	22037 (26.4)	4686 (34.3)	26868 (23.6)	
Winter	53166	22377 (26.8)	1515 (11.1)	29274 (25.7)	
Spring	52540	19078 (22.8)	3007 (22.1)	30455 (26.8)	

Abbreviations: n, number of samples; hrs, hours.

[a]The sample size for sex: n 58858;

[b] The sample size for gestational age: n 136958

*The p values were from the Chi-square test for all three groups.

Iodine status in the pre-fortification, transition and post-fortifications periods

The median TSH concentration was 2.0 mIU/l (IQR 1.3-2.9) in the post-fortification period, 2.0 mIU/l (IQR 1.2-3.1) in the transitional period and 1.7 mIU/l (IQR 1.0-2.7) in pre-fortification period. **Figure 5.3** presents the percentage of newborn with TSH concentration >5 mIU/l and >6 mIU/l. The percentage of newborns with TSH concentration >5 mIU/l was lower in the post-fortification (4.6%) period than the pre-fortification (5.1%) and transition (6.2%) periods (P value <0.01) (**Figure 5.3**). Using samples collected at 72-96 hours after birth only, the percentage of TSH >5 mIU/l was 2.9% (**Table 5.2**).

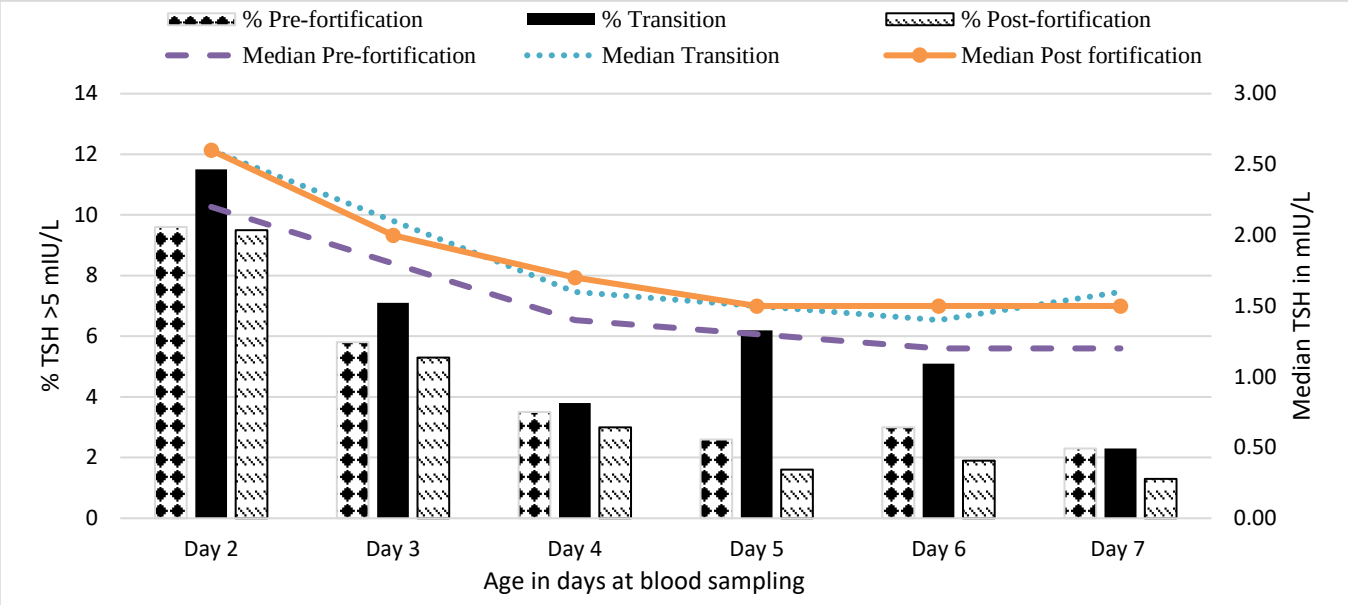


Figure 5. 2 Median TSH concentration and % of TSH concentration >5 mIU/L by neonatal age at blood sampling, 2005-2016, South Australia (n 239 182).

TSH, thyroid stimulating hormone.

Table 5. 2 Comparison of the percentage of TSH >5 mIU/L and median TSH concentration of newborns by potential time varying confounders, 2005-2016, South Australia (N 211033)

Variables	% TSH >5 mIU/L n (%)[a]	P-value *	Median TSH (mIU/L)	IQR (mIU/L)	P-value **
Neonatal age		<0.01			<0.01
48-72 hrs	9799 (5.1)		1.9	1.2, 2.9	
72-96 hrs	584 (2.9)		1.6	1.0, 2.4	
Sex		<0.01			<0.01
Males	1451 (4.8)[b]		1.8	1.1, 2.7	
Females	1065 (3.7)[b]		1.6	1.0, 2.5	
Birthweight		<0.01			<0.01
>2,500g	9345 (4.8)		1.9	1.2, 2.8	
≤2,500g	1308 (7.0)		2.0	1.3, 3.1	
Gestational age		0.01			0.73
≥37 weeks	5704 (5.0)[c]		2.0	1.3, 2.9	
<37 weeks	1282 (5.5)[c]		2.0	1.3, 2.9	
Season		<0.01			<0.01
Summer	2475 (4.8)		1.9	1.2, 2.8	
Autumn	2512 (4.7)		1.8	1.2, 2.8	
Winter	2741 (5.2)		1.9	1.2, 2.9	
Spring	2655 (5.0)		2.0	1.3, 2.9	

Abbreviations: TSH, thyroid stimulating hormone; IQR, inter quartile range; n, number of samples.

[a] Denominator was N=211,033 unless otherwise specified.

[b] The Denominator was N=58858

[c] The Denominator was N= 136958

* The p values were from the Chi square test.

** The p values were from the Kruskal-Wallis test.

Effect of mandatory iodine fortification of bread on iodine status of the population

In the primary unadjusted segmented regression analysis, newborns in the post-fortification period had a 10% lower risk of having TSH concentration >5 mIU/L than newborns in the pre-fortification group [IRR: 0.90, 95% confidence interval (CI): 0.87, 0.94]. Newborns in the transitional period had a 22% higher risk of having TSH concentration >5 mIU/L compared with newborns in the pre-fortification period [IRR: 1.22; 95% CI: 1.13, 1.31] (**Table 5.3**).

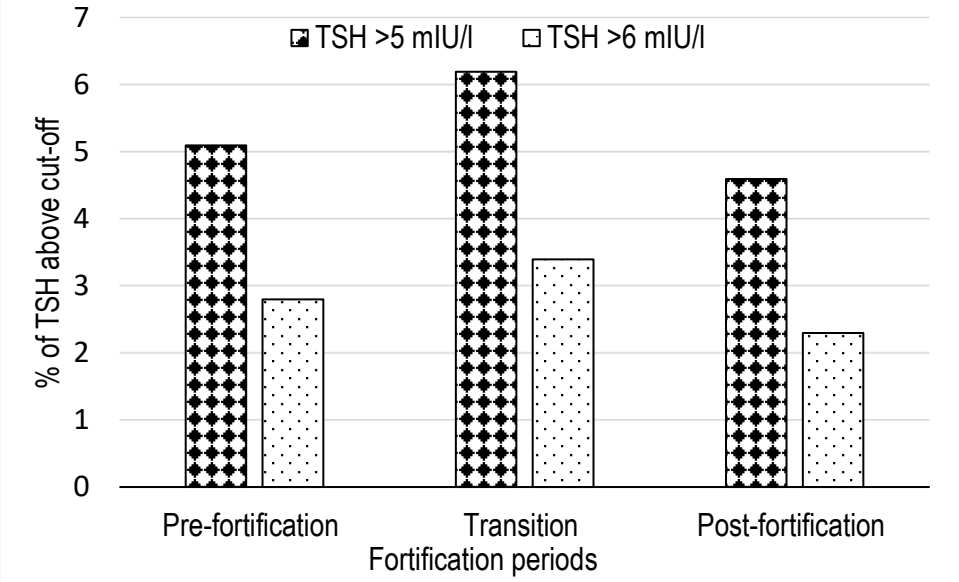


Figure 5. 3 Percentage of newborns with TSH concentration exceeding different cut-offs, 2005-2016, South Australia (n 211033).
TSH, thyroid stimulating hormone.

Table 5. 3 The effect of fortification on the proportion of newborns with TSH concentration >5 mIU/L using segmented regression analysis, 2005-2016, South Australia (n 211033)

	Pre- fortification	Transition		Post-fortification	
	IRR(95% CI)	IRR(95% CI)	P-value	IRR(95% CI)	P-value
Unadjusted Model	Reference	1.22 (1.13, 1.31)	<0.01	0.90 (0.87, 0.94)	<0.01
Adjusted Model 1	Reference	1.17 (0.96, 1.43)	0.12	0.88 (0.78, 0.99)	0.04
Adjusted Model 2	Reference	1.13 (0.92, 1.37)	0.24	0.93 (0.81, 1.07)	0.30
Adjusted Model 3	Reference	1.24 (1.01, 1.52)	0.04	0.90 (0.81, 1.01)	0.07
Adjusted Model 4	Reference	1.07 (0.88, 1.30)	0.49	0.88 (0.77, 1.01)	0.06
Adjusted Model 5	Reference	1.45 (1.15, 1.84)	<0.01	1.43 (1.10, 1.85)	0.01

Abbreviations: IRR, incidence rate ratio; CI, confidence interval.

Model 1: adjusted for only the linear and quadratic effects for the percentage of infants with low birthweight

Model 2: adjusted for only the linear and quadratic effects for the percentage of bloods sampled in each six-hour period

Model 3: adjusted for only season at birth

Model 4: adjusted for season, linear and quadratic effects for the percentage of bloods sampled in each six-hour period and linear and quadratic effects for the percentage of infants with low birthweight

Model 5: The adjusted model that includes time in months and also adjusted for season, linear and quadratic effects for the percentage of infants' blood sampled in each six-hour period from birth and linear and quadratic effects for the percentage of infants with low birthweight

Effect estimates (IRR) were similar in the adjusted exploratory models (models 1-4) except model 5 when time in months was included in the adjustment, which changed the direction of the effect in the post-fortification period and substantially increased the CIs (**Table 5.3**). There was no evidence of an interaction between the fortification periods and time in months (interaction p value= 0.35) and hence the interaction term (fortification groups and time) was not included in the final adjusted models.

Based on the predicted values of the exploratory segmented regression including fortification group and time in months in the model, there was a jump in the proportion of TSH >5 mIU/L at the commencement of the iodine fortification, followed by a decreasing trend throughout the post-fortification period which was similar to the trend seen in the pre-fortification period (**Figure 5.4**). Based on the counterfactual model, the proportion of TSH >5 mIU/L would also decrease in the post-fortification period even if the mandatory iodine fortification of bread was not introduced in October 2009 (**Figure 5.4**).

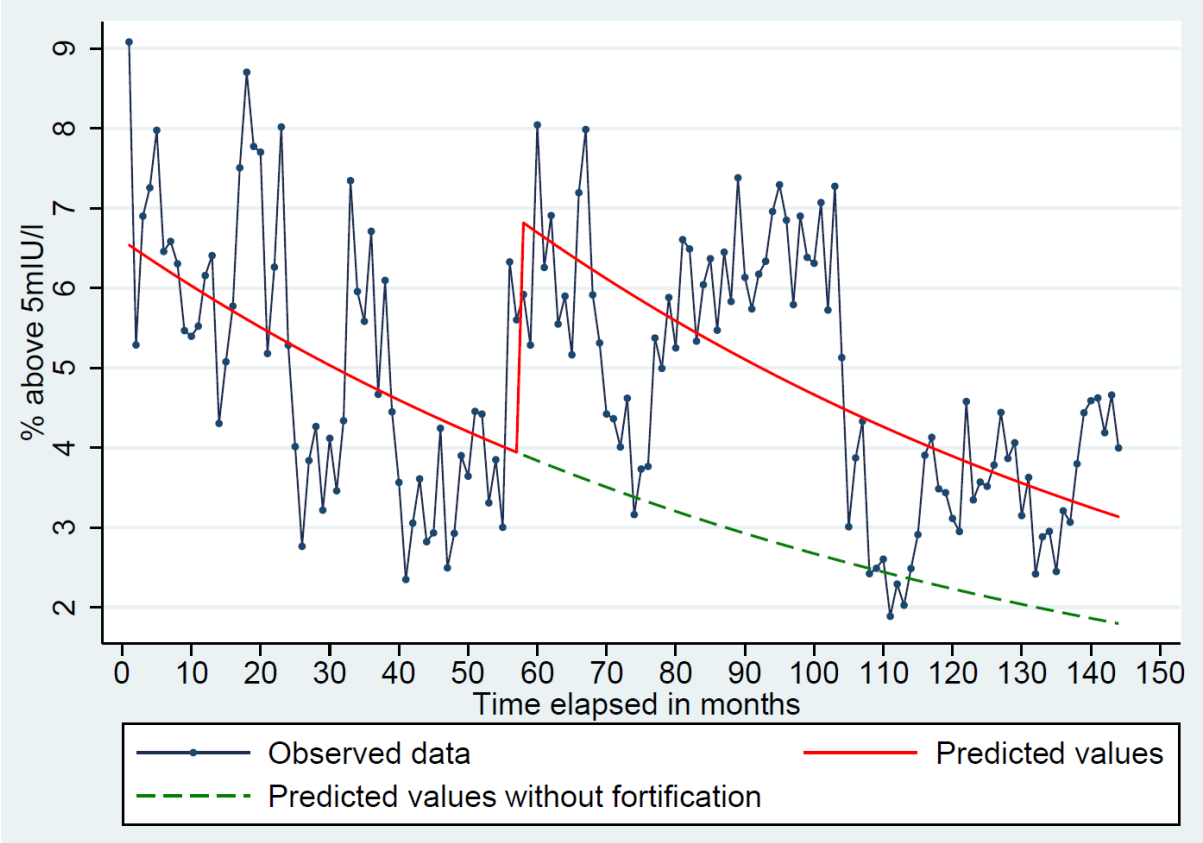


Figure 5. 4 Percentage of TSH >5 mIU/L across time in months based on a segmented regression model, 2005-2016, South Australia.
Dotted line, observed data; solid line, predicted values; dashed line, predicted trend without introduction of fortification based on an unadjusted regression model including fortification group and time. TSH, thyroid stimulating hormone.

Discussion

The present study assessed the effect of the mandatory iodine fortification program in the South Australian population. Using one of the recommended markers of population iodine status, the percentage of newborns with TSH concentration >5 mIU/L, South Australia would be classified as mild iodine deficiency post-fortification. Our finding of mild iodine deficiency in South Australia using newborn TSH as a marker is in contrast to iodine sufficiency defined by median UIC among school aged children [131], pregnant women [29] and breast-milk iodine concentration [232] in South Australia post-fortification. The majority of newborn blood samples were collected between 48 and 72 hours after birth and this may

increase the percentage of newborns with TSH concentration >5 mIU/L and contribute to the discrepancy in the classification of iodine status between the two markers of population iodine status of UIC and newborns TSH. The WHO criteria specify using newborn TSH assessed between 48 and 96 hours after birth to classify iodine status of populations. Our data showed that TSH concentrations of samples collected 48-72 hours after birth were higher than samples collected 72-96 hours after birth. The population became iodine sufficient if samples collected between 72 and 96 hours after birth were used. A similar finding has also been reported in Western Australia, where the population was iodine sufficient based on the median UIC [131] but mildly iodine deficient using TSH as a marker in the post-fortification period [52]. In the Western Australian study, the percentage of newborns with TSH concentration >5 mIU/L was higher in samples collected 48-72 hours after birth [52] similar to our study. Discrepancies in classifying population iodine status using newborns TSH concentration versus median UIC have also been reported elsewhere [43, 179, 186]. Evans et al [179] and Vandevijvere et al. [186] reported iodine deficiency by median UIC [43, 186] though a population study using newborn TSH showed iodine sufficiency. However, TSH was measured from blood sampled between 72-120 hours of age, outside the WHO recommendation, which may partly explain the difference. Furthermore, the acute perinatal stress during the early neonatal period may increase the TSH concentration in a healthy newborns with blood sample taken prior to 72 hours compared with newborns sampled after 72 hours of birth [236, 237].

Many countries have shifted to earlier blood sampling for neonatal screening due to earlier discharge from hospitals. This change in the practice may affect the newborns TSH value and the percentage with TSH >5 mIU/L at a population-level even that blood samples were collected within the WHO recommended 48 to 96 hours' time frame. In countries with earlier discharge, applying the 3% TSH cut-off suggested for samples collected 48-96 hours after

birth may not be appropriate since early samples have a higher TSH concentration and this may lead to incorrect classification of iodine status. When a higher cut-off of 6 mIU/L for TSH was used, we found that South Australia was iodine sufficient post-fortification, consistent with finding from the 2011-2012 National Health Survey which indicated iodine sufficiency in South Australia [131]. In contrast, Evans et al [179] called for WHO to modify its criteria and showed that using a cut-off of 2 mIU/L and 3 mIU/L for samples collected from 72 hours after birth was consistent with UIC. Together, these data suggest that in countries with earlier blood sampling for newborn TSH, increasing either the cut-off TSH value of 5 mIU/L, or the percentage of newborns with TSH concentration >5 mIU/L used to define iodine deficiency, or applying specify cut-offs for TSH depending on when the newborn blood was sampled may improve the consistency between TSH and median UIC in the classification of iodine status. In such countries like Australia, the newborn TSH concentration >5 mIU/L at its current cut-off of >3% may not be appropriate marker for classifying iodine status of populations.

Different cut-offs have been applied to newborn TSH for estimating population iodine status in the literature. Most studies use the 5 mIU/L TSH cut-off [179, 186, 200] while others use values rounded to one decimal place (5.0 mIU/L) [52, 54]. The WHO recommends using TSH concentration >5 mIU/L [12, 230], however recent sensitive TSH assays make reporting of TSH levels to one decimal place possible. It has been shown that using TSH results rounded to the nearest whole number and applying a 5 mIU/L cut-off decreased the percentage of newborns with elevated TSH concentration compared with using TSH results rounded to one decimal place and applying a 5.0 mIU/L cut-off [52]. This discrepancy is due to rounding a TSH results between 5.0 mIU/L and 5.5 mIU/L to a whole number (5 mIU/L) and applying a 5 mIU/L cut-off. *As* the recent TSH assay methods are more sensitive and able to measure

TSH concentration to more than one decimal place [52, 242], an update to the TSH cut-off value with one decimal place is suggested.

There have been debates over whether the current WHO criteria of newborn TSH concentration >5 mIU/L above 3% is appropriate to define iodine deficiency [45, 52, 54, 179, 186]. Due to the discrepancy between TSH and UIC to classify population iodine status, Burns et al [54] and Vandevijvere et al. [186] suggested to use the trend of percentage of TSH concentration >5 mIU/L over time to see the change in the iodine status of populations. Though the trends are important to observe changes in iodine status over time, the cut-off on the percentage of newborn TSH concentration >5 mIU/L is required to classify iodine status in populations using this marker.

The post-fortification group had a 10% lower risk of newborn TSH concentration >5 mIU/L compared with the pre-fortification group. This indicates the success of both the mandatory iodine fortification of bread and the recommendation for routine iodine supplementation in pregnancy in Australia to lower the risk of iodine deficiency. This finding is consistent with studies in Ireland [54], Armenia [243] and Switzerland [84] that showed a declining trend in the percentage of newborn TSH concentration >5 mlU/l after introduction of iodine fortification or salt iodisation. However, an exploratory analysis controlling for timing of blood sample collection, birthweight, season at birth and the underlying time trend resulted in a reversing of the effect estimate and a loss of precision. These seemingly contradictory results are due in part to the underlying time trend, which suggests that the percentage of newborns with TSH concentration >5 mIU/L was decreasing over time even before fortification was introduced. Although the actual percentage of TSH above 5 mIU/L in the post-fortification period is lower than the pre-fortification period, it is higher than the predicted percentage when taking into consideration the decreasing time trend. There is no

clear explanation for the decreasing time trend in the pre-fortification periods. Increased awareness of the re-emergence of iodine deficiency in Australia from the 2006 national iodine nutrition survey may have led to increasing iodised salt use in place of non-iodised salt in the population, which may partly explain the decreasing time trend observed in the pre-fortification [244].

Surprisingly, the percentage of TSH concentration >5 mIU/L and the median TSH concentration increased during the transitional period despite that TSH was measured using a same method over the entire study period. Our finding of a transient increase in the TSH concentration in the transitional period could be explained by the acute Wolf Chaikoff effect [26], which refers to a transient inhibition of thyroid function when iodine deficient populations exposed to a sudden increase in iodine intake, like in Australia from a combination of iodine supplementation in pregnancy and mandatory iodine fortification of bread during the transition period. A recent study reported that the mean iodine intake of pregnant women in South Australia who took iodine supplements at the recommended dose of 150 µg/day or above was 377 µg/day, which was well above the Recommended Dietary Intake (RDI) for iodine in pregnancy [29]. In addition, based on the median UIC, 20% of Australian children aged 2-3 years [245] and 24% of South Australian infants aged 3 months [246] had iodine intake above the upper level of safe intake in the post-fortification period [245]. A national bread consumption survey [245] conducted in 2010 also indicated that iodine content varied in the transitional period, with some as high as 270 µg/100g of bread in contrast to the target level of 46 µg/100g of bread [16]. Implementation of both the mandatory iodine fortification of bread in October 2009 [247] followed by the recommendation for routine iodine supplementation in pregnancy in January 2010 [107] in Australia may have resulted in a transient inhibitory effect on thyroid function due to a sudden increase in iodine intake in pregnant women. The fetal thyroid system is more susceptible to both inadequate

and excess iodine. Elevated newborn TSH concentration of above 17 mIU/L was reported in an iodine deficient population exposed to increased intake above requirements, suggesting a transient inhibitory effect on the synthesis of thyroid hormones [248]. A transient inhibitory effect from a sudden increase in iodine intake during the implementation of the fortification and supplementation programs may contribute to a sharp increase in the percentage of TSH above 5 mIU/L in the transitional period [26]. In the post-fortification period, the population (including childbearing aged women) became iodine sufficient [29, 131]. The same level of iodine intake in the post-fortification period is unlikely to cause the acute Wolf Chaikoff effect as women entering pregnancy were no longer iodine deficient.

Our study reported seasonal variation in the percentage of newborns with TSH concentration >5 mIU/L in the study period and this was slightly higher in winter and spring compared with summer and autumn. A similar seasonal pattern was observed in Belgium [238]. Seasonality in TSH concentration may due to variation in dietary iodine intake (mainly fish and iodine fortified foods) [238] but there are no data on the seasonality of bread consumption in Australia. Cold weather conditions can increase thyroid hormone secretion, and this may contribute to the higher TSH concentrations observed in winter and spring [249, 250].

The current study has a number of limitations. Firstly, iodine supplementation for pregnant women was recommended by National Health and Medical Research Council in January 2010, shortly after the implantation of the fortification program. Therefore, it was not possible to examine the effect of the fortification separately from the supplementation program. Secondly, gestational age was not controlled in the final model due to incomplete information but we controlled for birthweight as a proxy which is highly correlated with gestational age [239]. Thirdly, we did not adjust the caesarean section rate because the data was not collected from the neonatal screening program. However, the caesarean section rate was steady over the

study period and is unlikely to affect the percentage of TSH >5 mIU/L over the study period [240]. Fourthly, due to a move towards earlier discharge of mothers and newborns from hospital, a majority of newborns were sampled 48 to 72 hours after birth for newborn screening. As a result, there is a difference in the mean timing of blood sampling between the three fortification periods which may increase the percentage of newborns with TSH concentration >5 mIU/L in the post-fortification period. Although we attempted to control for this variation in the timing of blood sample collection in the analysis by considering the percentage of infants sampled in each 6-hour period, there may be a residual effect.

Conclusions

Using newborn TSH as a marker, South Australia would be classified as mild iodine deficiency post-fortification in contrast to iodine sufficiency based on the median UIC, the most widely used population marker of iodine status. Our study together with the literature suggest the newborn TSH concentration >5 mIU/L at its current cut-off of >3% may not be a reliable marker for defining the iodine status of populations. There has been a change in clinical practice worldwide towards early discharge from hospital and newborn bloods are collected early for the newborn screening. Re-evaluation of the current WHO criteria on TSH as a marker of population iodine status is warranted in this context.

Abbreviations

AIC, Akaike Information Criterion; CI, Confidence Interval; DELFIA, Dissociation-Enhanced Lanthanide Fluorescence Immunoassay; ICCIDD, International Council for Control of Iodine Deficiency Disorders; IDD, Iodine Deficiency Disorder; SA, South Australia; TSH, Thyroid Stimulating Hormone; UIC, Urinary Iodine Concentration; WHO, World Health Organization; UNICEF, United Nations Children's Fund.

CHAPTER 6 IODINE STATUS OF PREGNANT WOMEN FROM LOW SOCIO-ECONOMIC STATUS BEFORE AND AFTER MANDATORY IODINE FORTIFICATION OF BREAD IN AUSTRALIA

This chapter has been written based on the style of the *Food Policy* journal.

6.1 Manuscript

Wassie MM, Roberts, CT, & Zhou, S. Iodine status of pregnant women from low socio-economic status before and after mandatory iodine fortification of bread in Australia (To be submitted to *Food Policy*).

Statement of Authorship

Title of Paper	Iodine status of pregnant women from low socio-economic status before and after mandatory iodine fortification of bread in Australia
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wassie MM, Roberts CT, Zhou SJ. Iodine status of pregnant women from low socio-economic status before and after mandatory iodine fortification of bread in Australia. Prepared based on the guidelines of Food Policy Journal.

Principal Author

Name of Principal Author (Candidate)	Molla Mesele Wassie
Contribution to the Paper	Design, laboratory analysis of urine samples, statistical analysis, interpretation of data, manuscript preparation and critical revision of the manuscript
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 25/03/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	*Claire T Roberts*

Contribution to the Paper	Design, supervise data collection, interpretation of data and critical revision of the manuscript		
Signature		Date	25/03/2020
Name of Co-Author	Shao Jia Zhou		
Contribution to the Paper	Conception, design, statistical analysis, interpretation of data and critical revision of the manuscript		
Signature		Date	25/03/2020

Please cut and paste additional co-author panels here as required.

Title Page

Iodine status of pregnant women from low socio-economic status before and after mandatory iodine fortification of bread in Australia.

Author names *and affiliations*

Molla Mesele Wassie [a,b], Claire T Roberts [c], Shao Jia Zhou [a]

[a] *School of Agriculture Food and Wine, Faculty of Sciences, The University of Adelaide, Adelaide, Australia.*

[b] *Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia.*

[c] Adelaide Medical School, Faculty of Health and Medical Sciences, *The University of Adelaide, Adelaide, Australia.*

Corresponding Author

Shao Jia Zhou

School of Agriculture, Food & Wine

University of Adelaide, Waite Campus

PMB1, Glen Osmond, SA 5064

Phone: +61 8 8313 2065

Email: jo.zhou@adelaide.edu.au.

Abstract

Background: Iodine deficiency re-emerged in Australia and that led to the implementation of mandatory iodine fortification of bread in October 2009. There is no study that compared the iodine status of pregnant women pre- and post- iodine fortification in pregnant women from low socio-economic status (SES).

Objectives: To assess and compare iodine status before and after the mandatory iodine fortification of bread in pregnant women from low SES in South Australia.

Methods: A comparative study using spot urine samples collected (between 12 to15 weeks of gestation) from two pregnancy cohort studies conducted before (n = 600) and after (n = 600) the mandatory fortification of bread in pregnant women from low SES. SES was determined based on a relative social economic index for areas. Urinary iodine concentration (UIC) was measured using a modified Sandell-Kolthoff (S-K) spectrophotometric method. Iodine deficiency during pregnancy was defined based on the WHO criteria of median UIC <150 µg/L.

Results: The median [Inter Quartile Range (IQR)] UIC of pregnant women was 103 (65 163) µg/L in the pre-fortification cohort. In the post-fortification cohort, the median (IQR) UIC were 210 (125, 338) µg/L for all pregnant women and 162 (106, 232) µg/L for pregnant women who did not take iodine-containing supplements.

Conclusions: Iodine status of pregnant women from low SES improved from deficiency pre-fortification to sufficiency post-fortification period regardless of whether iodine-containing supplements were consumed. National iodine survey on pregnant women is warranted to confirm iodine sufficiency in other states of Australia.

Key words: Iodine status, iodine fortification, pregnant women, low socio-economic status

1. Introduction

Iodine is an essential micronutrient needed for the synthesis of thyroid hormones, which are important for proper growth and development [3]. Iodine requirements increase during pregnancy to fulfil the feto-placental demand for thyroid hormone [251]. Iodine deficiency during pregnancy can lead to adverse pregnancy outcomes [252] and impaired neurodevelopment in children [65, 212]. According to the World Health Organization (WHO), pregnant women are classified as iodine deficient if the median UIC is less than 150 µg/L in a representative sample of pregnant women [178].

Iodine deficiency re-emerged in Australia [17] and that led to the implementation of mandatory fortification of salt used in bread making in October 2009 [16]. Dietary modelling suggested that the increased intake from fortified bread was not enough to meet the recommended dietary intake in pregnancy and breastfeeding [253]. Hence, the Australian National Health and Medical Research Council recommends that pregnant and breastfeeding women take iodine supplements of 150 µg/day followed soon after the mandatory iodine fortification [107]. The mandatory iodine fortification of bread has been successful in preventing iodine deficiency in the general population and reproductive-age women in Australia [131], but the impact on the iodine status of pregnant women is uncertain due to a lack of nationally representative data. There is only one study that met the WHO recommended sample size to assess iodine status in populations of pregnant women [29]. That study (n = 783) was conducted in South Australia and showed iodine sufficiency. However, one small study in low socio-economic status (SES) (n = 196) in South Australia suggested that pregnant women remained iodine deficient post-fortification [39].

It is well recognised that most individuals from low SES have a low dietary quality with low intake in key micronutrients and a higher risk of deficiency [254]. There are limited data on the

iodine status of populations, particularly in low SES. Pregnant women from low SES may also have a higher risk of iodine deficiency as studies showed low compliance with iodine supplementation recommendation [255, 256] and low intake of fish and seafood due to the cost [254]. Monitoring of iodine status will help to ensure either deficient or excess levels of iodine do not widely occur in vulnerable groups of the population. Hence, this study aimed to assess and compare the iodine status of pregnant women from low SES in South Australia before and after the mandatory iodine fortification of bread.

2. Methods

2.1 Study design, setting and participants

A comparative study. We used urine samples collected from pregnant women participating in two pregnancy cohort studies: the SCOPE (screening for pregnancy end points) [257] and STOP (screening tests to identify poor outcome in pregnancy) [222]. All pregnant women were recruited from the Lyell McEwin Hospital in Elizabeth Vale, northern Adelaide, South Australia, and were from the most disadvantaged suburbs in South Australia based on a relative social economic index for areas that measures the relative disadvantage [258]. Although the relative social economic index for areas summarises the socio-economic conditions of population groups within an area, it application to classify SES at individual level is limited [226]. Women in the SCOPE study recruited between 2006 and 2008 formed the pre-fortification cohort and women in the STOP study recruited between 2015 and 2017 formed the post-fortification cohort. All women were nulliparous, and the mean gestational age was less than 15 weeks at study entry. Women with multiple pregnancy, major fetal anomalies, ≥3 terminations of pregnancy or ≥3 miscarriages and known chronic disease were excluded in both studies [222, 257]. A random spot urine sample was collected from all pregnant women at study entry and all samples were stored at -80 degrees Celsius for analysis. Data on socio-economic and demographic characteristics were collected at study entry using a structured interviewer-administered questionnaire by research midwives. In

addition, information on dietary supplement intake at study entry was also collected in the post-fortification cohort. Maternal measurements of height and weight, as well as urine sample collection, were performed by research midwives. A total of 1,165 nulliparous women participated in the SCOPE study in Adelaide, South Australia and 1,373 women in the STOP study (**Figure 6.1**).

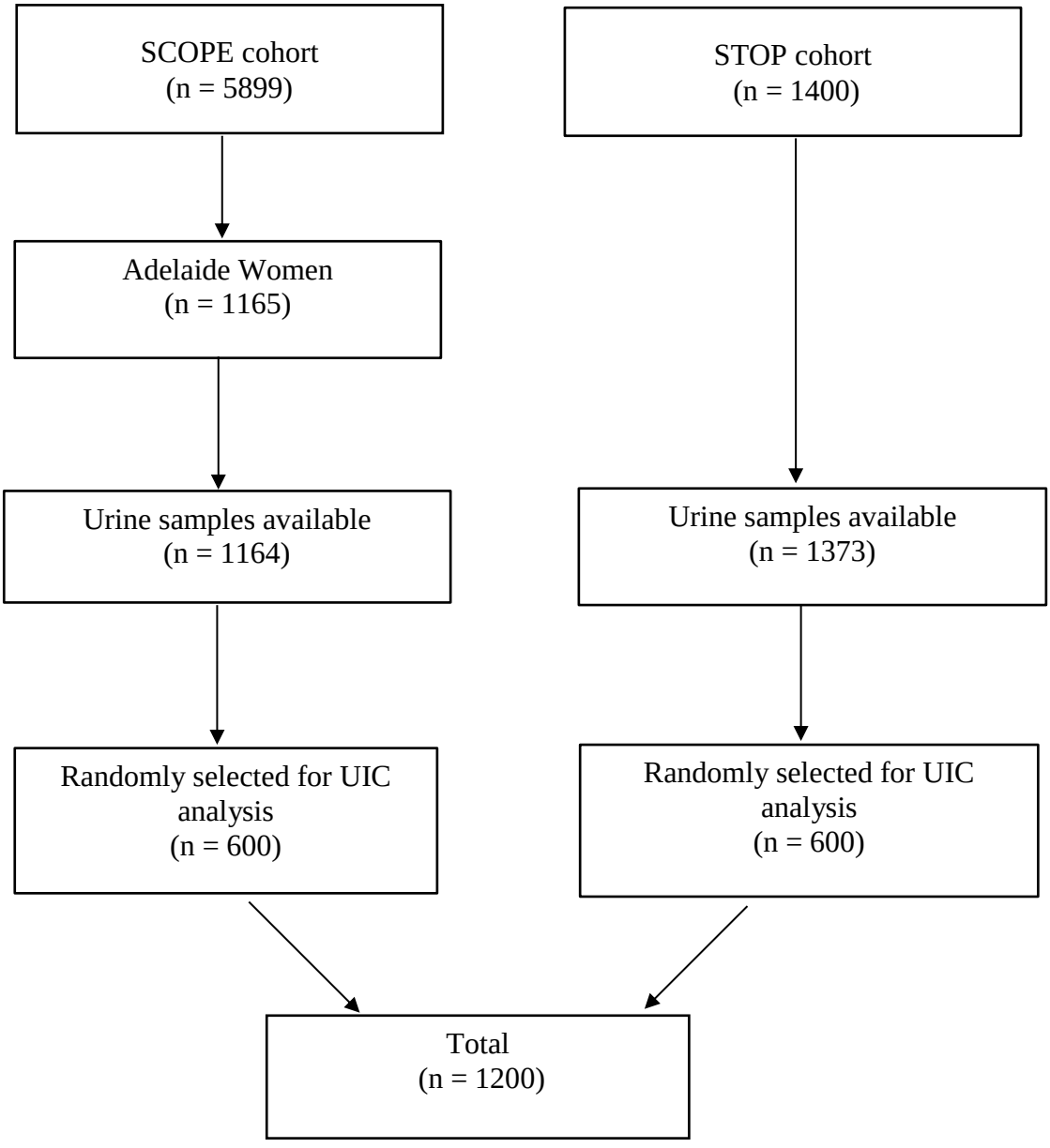


Figure 6. 1 Selection of participants of the SCOPE and STOP cohorts

Notes: SCOPE, screening for pregnancy end points; n, number of participants; STOP, screening tests to identify poor outcome in pregnancy; UIC, urinary iodine concentration.

Based on the WHO recommendation, a sample size of 600 to 900 is sufficient, in most instances, to have a reasonable confidence interval around the percentage estimate for iodine deficiency [12]. Andersen et al. [63] also suggested the number of spot urine samples required to estimate the iodine status in a population with 95% confidence and a 5% precision was about 500. Six hundred women were randomly selected from each cohort with a total of 1,200 pregnant women included in this study.

2.2 Outcome

Urinary Iodine Concentration (UIC): Iodine status of pregnant women assessed by median UIC. A modified Sandell-Kolthoff spectrophotometric method (WHO's method B) [223] was used to analyse the iodine concentration in the spot urine samples. The SeronormTM (SERO) Trace Elements Urine L-2 (LOT 1403081) was used as the external reference. The mean iodine concentration of the SeronormTM obtained was 299 (SD 12) μg/L compared with the certified reference iodine concentration of 297 (95% CI: 237-356) μg/L. The lower limit of detection was 10 μg/L and the limit of quantification of the assay was 30 μg/L. The intra-assay and inter-assay coefficient of variation (CV) was 5.4% and 3.2%, respectively.

2.3 Statistical Analysis

Data were analysed using Stata 15 (Statacorp LP, College Station, Texas). Iodine status was reported as median UIC in μg/L. The iodine status of pregnant women was categorised into iodine deficiency (<150 μg/L), sufficiency (150-249 μg/L), more than adequate (250-500 μg/L) and excess (≥500 μg/L) based on the WHO criteria [12]. Pearson chi-square test was used to compare the percentage of pregnant women with UIC <150 μg/L between the cohorts. As UIC is not normally distributed, Kruskal wails test, a non-parametric test, was used to compare the median UIC between the two cohorts and across maternal characteristics in each cohort. UIC was compared between maternal age ≤19 years (young maternal age), 20-34 years and ≥35 years (advanced maternal age) because maternal age was associated with adverse birth

outcomes [259]. A comparison was also made between women who took iodine-containing supplements vs. those who did not take the supplements in the post-fortification cohort where data on intake of supplements were collected. Statistical significance was set at a value of $P < 0.05$.

3. Results

3.1 Participant characteristics

Table 6.1 presents the characteristics of the study participants. The mean ± SD age of participants was 25 ± 5 years. More women were Caucasian, smoked, consumed alcohol and were obese at study entry in the pre-fortification cohort compared with women in the post-fortification cohort, while more women completed tertiary education in the post-fortification cohort compared with women in the pre-fortification cohort.

Table 6. 1 Characteristics of study participants

Variables	Overall	SCOPE (Pre-fortification)	STOP (Post-fortification)
n	1200	600	600
Age, y	24.6 ± 5.1	23.4 ± 4.9	25.7 ± 5.0
Age category, y			
≤19	205 (17.1)	141 (23.5)	64 (10.7)
20-34	945 (78.8)	442 (73.8)	503 (83.8)
≥35	49 (4.1)	16 (2.7)	33 (5.5)
Married, n (%)	1034 (86.2)	511 (85.3)	523 (87.2)
Caucasian race, n (%)	1056 (88.1)	553 (92.3)	503 (83.8)
Completed tertiary education, n (%)	160 (13.3)	51 (8.5)	109 (18.2)
Consumed alcohol during first trimester, n (%)	303 (25.3)	213 (35.6)	90 (15.0)
Smoked during first trimester, n (%)	334 (27.9)	228 (38.1)	106 (17.7)
Took iodine-containing supplements during first trimester, n (%)	NA	NA	450 (75)
BMI, kg/m^2	27.2 ± 6.9	26.6 ± 6.4	27.9 ± 7.2
Obese, n (%)	332 (27.7)	145 (24.2)	187 (31.2)

Notes: Values are presented as frequencies (percentages) or means ± SDs. BMI, body mass index; SD, standard deviation; SCOPE, screening for pregnancy end points; STOP, screening tests to identify poor outcome in pregnancy; N, sample size; NA, not applicable.

3.2 Iodine status in pregnant women pre- and post- mandatory iodine fortification of bread

As shown in **Table 6.2**, pregnant women in the pre-fortification cohort were iodine deficient while those in the post-fortification cohort were iodine sufficient based on the median UIC. The median UIC was higher (210 µg/L vs. 103 µg/L, p <0.001) and the percentage of pregnant women with UIC <150 µg/L was lower (31.0% vs. 71.7%, p <0.001) in the post-fortification cohort than the pre-fortification cohort. However, the percentage of pregnant women with UIC >500 µg/L was higher (11.5% vs. 2.0%, P <0.001) in the post-fortification cohort than the pre-fortification cohort. In the post-fortification cohort where data on the intake of supplements were collected, the median (IQR) UIC was 242 (142, 377) µg/L for women who took iodine-containing supplements and 162 (106, 232) µg/L for those who did not take supplements, indicating iodine sufficiency regardless of whether women took iodine-containing supplements.

Table 6. 2 Iodine status of participants in pre- and post-fortification cohorts

	SCOPE (Pre-fortification)	STOP (Post-fortification)	P value [a]
n	600	600	
Median UIC (IQR), µg/L	103 (65, 163)	210 (125, 338)	0.0001[b]
UIC <50 µg/L, n (%)	99 (16.5)	14 (2.3)	<0.0001[c]
UIC: 50-149 µg/L, n (%)	331 (55.2)	172 (28.7)	
UIC: 150-299 µg/L, n (%)	117 (19.5)	168 (28.0)	
UIC: 300-499 µg/L, n (%)	41 (6.8)	177 (29.5)	
UIC >500 µg/L, n (%)	12 (2)	69 (11.5)	

Notes: Values are presented as frequencies (percentages) or medians (IQR). IQR, inter-quartile range; SCOPE, screening for pregnancy end points; SD, standard deviation; STOP, screening tests to identify poor outcome in pregnancy; UIC, urinary iodine concentration.

[a] The p values were for comparison of UIC and iodine status between SCOPE and STOP studies.

[b] Kruskal wails test.

[c] Pearson chi-square test.

3.3 Comparison of iodine status in pregnant women by maternal characteristics

As depicted in **Table 6.3**, the median UIC and percentage of women with UIC <150 µg/L were not different across maternal characteristics except age and obesity status in both SCOPE and STOP. Younger pregnant women aged ≤19 years had lower median UIC in STOP when compared with women aged 20-34 years of age or with advanced maternal age (≥35). Obese women had a higher median UIC when compared with non-obese women in both SCOPE and STOP.

Table 6. 3 Comparison of the median urinary iodine concentration by maternal characteristics

Variables	SCOPE (Pre-fortification)				STOP (Post-fortification)			
	Median UIC (IQR), µg/L	P-value [a]	n (%) for UIC <150 µg/L	P-value [b]	Median UIC (IQR), µg/L	P-value [a]	N (%) for UIC <150 µg/L	P-value [b]
n	600				600			
Age, years								
≤19	109 (66, 155)	0.43	104 (17.4)	0.79	168 (119)	0.05	23 (3.8)	0.62
20-34	102 (65, 163)		314 (52.4)		217 (128, 339)		152 (25.3)	
≥35	71 (43, 164)		12 (2.0)		268 (125, 503)		11 (1.8)	
Obese								
Yes	126 (81, 180)	<0.001	87 (14.5)	<0.001	237 (142, 377)	0.04	49 (8.2)	0.09
No	98 (60, 149)		343 (57.3)		200 (119, 322)		137 (22.8)	
Ethnicity								
Caucasian	103 (65, 162)	0.56	400 (66.8)	0.30	202 (125, 322)	0.09	158 (26.3)	0.62
Others	106 (77, 173)		30 (5.0)		255 (132, 380)		28 (4.7)	
Completed tertiary education								
No	103 (65, 162)	0.78	397 (66.3)	0.24	202 (122, 321)	0.27	157 (26.2)	0.05
Yes	112 (51, 182)		33 (5.5)		245 (146, 391)		29 (4.8)	
Consumed alcohol during 1st trimester								
No	106 (66, 162)	0.30	278 (46.4)	0.86	204 (122, 339)	0.43	164 (27.4)	0.14
Yes	97 (60.7, 165)		152 (25.4)		237 (152, 336)		22 (3.7)	
Smoked during 1st trimester								
No	103 (66, 162)	0.80	266 (44.4)	0.95	216 (132, 342)	0.12	145 (24.2)	0.06
Yes	103 (63, 163)		164 (27.4)		185 (113, 323)		41 (6.8)	
Took iodine-containing supplement at study entry								
Yes	NA				242 (142, 377)	0.00	123 (20.5)	0.001

Variables	SCOPE (Pre-fortification)				STOP (Post-fortification)			
	Median UIC (IQR), µg/L	P-value [a]	n (%) for UIC <150 µg/L	P-value [b]	Median UIC (IQR), µg/L	P-value [a]	N (%) for UIC <150 µg/L	P-value [b]
n	600				600			
No	NA				162 (106, 232)		63 (10.5)	

Notes: BMI, body mass index; IQR, interquartile range; UIC, urinary iodine concentration.

[a] Kruskal wails test was used to compare median UIC across maternal characteristics in the combined cohort.

[b] Pearson chi-square test was used to compare the percentage of women with <150 µg/L across maternal characteristics in the combined cohort.

[c] Data on iodine-containing supplement intake were obtained only from the post-fortification cohort.

4. Discussion

Our study is the only study in Australia with a sufficient sample size to assess and compare the iodine status of pregnant women from low SES pre- and post-mandatory iodine fortification of bread. The study showed that pregnant women from low SES in South Australia were iodine deficient in the pre-fortification period but became iodine sufficient in the post-fortification period regardless of whether women took iodine-containing supplements.

Iodine deficiency in the pre-fortification cohort was consistent with previous studies conducted in pregnant women from other states of Australia before the introduction of the mandatory iodine fortification of bread [40, 119, 129]. Our finding of iodine sufficiency in pregnant women in the post-fortification cohort was consistent with a large study conducted in the post-fortification period in South Australia (n = 783) [29] and a small study conducted in New South Wales (n = 95) [119] although both studies were not focused on pregnant women from low SES. Contrary to our finding of iodine sufficiency post-fortification, a small study conducted in pregnant women from low SES in northern Adelaide (n = 196) [39] reported iodine deficiency but included pregnant women in the transitional period. Two other small studies in the Northern Territory (n = 24) [126] and Victoria's Gippsland region (n = 86) [130] also reported iodine deficiency. Findings from these small studies are unlikely to be representatives of the iodine status of pregnant women post-fortification. As the studies in northern Adelaide and Gippsland were conducted in the transitional period after the introduction of mandatory iodine fortification of bread, awareness of the importance of adequate iodine intake might be low. The study in the Northern Territory [126], conducted between 2013 and 2015, reported a median UIC of 93 µg/L in remote indigenous and non-indigenous pregnant women which was higher than the pre-fortification value of 48 µg/L in a

small sample (n = 22) of pregnant women in the same region. Mandatory iodine fortification of bread and iodine supplementation during pregnancy may be less commonly practised in remote communities compared with major cities. The differences in iodine status between the pre- and post-fortification cohorts are less likely to be reflective of the changing demographic of the general population [260-262] and a changing dietary habit due to difference in the ethnic makeup between the two cohorts.

Together with the results of iodine sufficiency from the National Survey in SAC [131] and a large pregnancy study in South Australia [29], our finding of iodine sufficiency in pregnant women from low SES in the post-fortification cohort during the early pregnancy periods, indicated the success of the mandatory fortification of bread. The national survey showed bread as the major source of dietary iodine in Australia including in women from low SES post-fortification [263]. Although there are no national data on bread intake by SES in pregnant women in Australia, reproductive-age women [125, 263] and children [263] in the lowest SES had higher iodine intake than those from higher SES. Besides, pregnant women in Australia are recommended to follow the National Health and Medical Research Council's recommendation to take iodine supplements of 150 μg/day during pregnancy. Based on data from a subgroup of women in our study (n = 150) and the other South Australia study (n = 148), pregnant women before 20 weeks of gestation who did not take iodine-containing supplements were iodine sufficient. Both our study and the other South Australian study [29] suggested iodine intake of pregnant women from low SES is adequate regardless of whether they are taking iodine supplements in pregnancy. However, the number of pregnant women not taking iodine-containing supplements was small in both studies. Further study assessing the iodine status of pregnant women who did not take iodine-containing supplements from a large representative sample is warranted.

The urine samples were collected before 15 weeks of gestation in both cohorts, but whether iodine status changes over trimesters of pregnancy is uncertain. There is some evidence suggesting a lower median UIC in the second and third trimester compared to first trimester [29, 71, 127]. This may be explained by hemodilution and increased iodine requirement in late trimesters. Therefore, trimester-specific UIC cut-off may be required to define iodine status during pregnancy [79].

The difference in median UIC or percentage of women <150 µg/L in younger and older pregnant women may be due to differences in dietary practice. Obese women had a higher median UIC compared with non-obese women. This finding was consistent with the National Health Survey on adults [131] and the median UIC difference by obesity status may be reflective of the differences in dietary patterns of obese and non-obese women. However, this finding may not be representative as the number of obese pregnant women was small and assessing the iodine status of these women with a larger sample size is warranted.

In conclusion, pregnant women from low SES in South Australia are iodine deficient pre-fortification but iodine sufficient post-fortification. The mandatory iodine fortification of bread improved the iodine status of pregnant women from low SES in South Australia irrespective of iodine-containing supplement intake. A representative survey is warranted in other states to confirm iodine sufficiency as geographical differences exist in the iodine status of pregnant women in Australia.

Abbreviations

UIC, Urinary Iodine Concentration; SES, Socio-Economic Status; SCOPE, screening for pregnancy end points; STOP, screening tests to identify poor outcome in pregnancy; WHO, World Health Organization.

Acknowledgements

We would like to acknowledge Dylan McCullough for assisting us in urine sample selection and transporting the samples to the Waite Main Building, Waite Campus for analysis.

Statement of authors' contributions to manuscript

MMW and SJZ designed the study; MMW analysed the UIC and performed data analysis; and MMW drafted the paper with input from all authors. CR contributed to the interpretation of results and reviewed the manuscript. SJZ oversight the conduct of the study. All authors read and approved the final manuscript.

Funding Disclosure

MMW was supported by The University of Adelaide, Australian Government Research Training Program. The SCOPE study was supported by the South Australia Premier's Science and Research Fund. South Australian Government. The STOP study was supported by The University of Adelaide grant number U1111-1160-9802.

Conflict of Interest

None.

CHAPTER 7 ASSOCIATIONS BETWEEN NEWBORN THYROID STIMULATING HORMONE CONCENTRATION AND NEURODEVELOPMENT AND GROWTH OF CHILDREN AT 18 MONTHS

This chapter has been written based on the style of *The American Journal of Clinical Nutrition.*

7.1 Manuscript

Wassie MM, Smithers, L., Yelland, L., Makrides, M & Zhou, S. Associations between newborn Thyroid Stimulating Hormone concentration and neurodevelopment and growth of children at 18 months (To be submitted to *The American Journal of Clinical Nutrition*).

Statement of Authorship

Title of Paper	Associations between newborn TSH concentration and neurodevelopment and growth of children at 18 months
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wassie MM, Smithers LG, Yelland LN, Makrides M, Zhou SJ. Associations between newborn TSH concentration and neurodevelopment and growth of children at 18 months. Prepared based on the guidelines of the American Journal of Clinical Nutrition.

Principal Author

Name of Principal Author (Candidate)	Molla Mesele Wassie		
Contribution to the Paper	Design, statistical analysis, interpretation of data, manuscript preparation and critical revision of the manuscript		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	12/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lisa Gaye Smithers
Contribution to the Paper	Design, interpretation of data and critical revision of the manuscript

Signature		Date	16 Mar 2020

Name of Co-Author	Lisa Nicole Yelland		
Contribution to the Paper	Design, interpretation of data and critical revision of the manuscript		
Signature		Date	18/3/20

Name of Co-Author	Maria Makrides		
Contribution to the Paper	Design, interpretation of data and critical revision of the manuscript		
Signature		Date	17/3/20

Name of Co-Author	Shao Jia Zhou		
Contribution to the Paper	Conception, design, statistical analysis, interpretation of data and critical revision of the manuscript		
Signature		Date	16/03/2020

Please cut and paste additional co-author panels here as required.

Title Page

Associations between newborn Thyroid Stimulating Hormone concentration and neurodevelopment and growth of children at 18 months

Author Names

Molla Mesele Wassie, Lisa Gaye Smithers, Lisa Nicole Yelland, Maria Makrides, Shao Jia Zhou

Authors' affiliations

School of Agriculture Food and Wine, Faculty of Sciences, The University of Adelaide, Adelaide, Australia (MMW, SJZ).

Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia (MMW).

School of Public Health, Faculty of Health and Medical Sciences, *The University of Adelaide, Adelaide, Australia (LGS, LNY).*

Robinson Research Institute, The University of Adelaide, Adelaide, Australia (LGS, SJZ)

Adelaide Medical School, Faculty of Health and Medical Sciences, *The University of Adelaide, Adelaide, Australia (MM).*

South Australian Health and Medical Research Institute, *Adelaide, Australia (LNY, MM)*

Authors' last names

Wassie, Smithers, Yelland, Makrides, Zhou

Disclaimers

None.

Corresponding Author

Shao Jia Zhou

School of Agriculture, Food & Wine

University of Adelaide, Waite Campus

PMB1, Glen Osmond, SA 5064

Phone: +61 8 8313 2065

Email: jo.zhou@adelaide.edu.au.

Sources of Support

MMW was supported by The University of Adelaide, Australian Government Research Training Program. LNY and MM were supported by Australian National Health and Medical Research Council Fellowships (IDs 1052388 and 1061704, respectively). The DOMInO and PINK studies were funded by the Australian National Health and Medical Research Council (IDs 349301 and 626800, respectively).

Short running head

Associations between newborn TSH and neurodevelopment and growth.

Abbreviations

Bayley-III, Bayley Scales of Infant and Toddler Development third edition; DOMInO, Docosahexaenoic acid to Optimize Mother Infant Outcome; HCZ, head circumference-for-age; LAZ, length-for-age; PINK, Pregnancy Iodine and Neurodevelopment in Kids; RR, Relative Risk; TSH, Thyroid Stimulating Hormone; WAZ, weight-for-age; WHO, World Health Organization; WLZ, weight-for-length.

Abstract

Background: Evidence of an association between newborn thyroid stimulating hormone concentration (TSH), a marker of iodine nutrition in early life, and childhood neurodevelopment and growth is limited, and findings from previous studies are inconsistent.

Objective: To assess the associations between newborn TSH and childhood neurodevelopment and growth using data collected from two pregnancy studies, one in a borderline iodine deficient setting and one in an iodine sufficient setting.

Design: TSH data were obtained from routine newborn screening and were analysed as a continuous variable or in quartiles. Neurodevelopment was assessed at 18 months using the Bayley Scales of Infant and Toddler Development third edition (Bayley-III). Weight, height and head circumference were measured at 18 months and anthropometric z scores were calculated using the WHO's Child Growth Standards. Key confounders of neurodevelopment and growth were controlled in the analyses.

Results: In total, 1520 children were included in the analysis. Comparing the highest with the lowest TSH quartile, the Mean Differences (MDs) in the Bayley-III scores ranged from -2.0 (95% CI: -4.7, 0.8) to -2.5 (95% CI: -5.9, 0.9) points in a borderline iodine deficient setting and 1.2 (95% CI: -1.4, 3.7) to 1.4 (95% CI: -2.0, 4.9) points in an iodine sufficient setting in the cognitive, language and motor scales; the MDs in the anthropometric z scores ranged from 0.04 (95% CI: -0.2, 0.2) to -0.4 (95% CI: -0.8, -0.1) in studies from both settings.

Conclusion: A null association between newborn TSH and neurodevelopment or growth was observed but we cannot rule out a different pattern of association in borderline iodine deficient and iodine sufficient settings. Longitudinal studies that utilise newborn TSH data

and examine neurodevelopmental outcomes at later ages are warranted, as neurodevelopmental assessments in older children are more predictive of later achievement.

Key words: thyroid stimulating hormone, iodine, growth, development

Introduction

Iodine is an integral part of thyroid hormone, which in turn is required for normal brain development and physical growth during the early critical periods of life [264]. Insufficient iodine intake in pregnancy has been linked to poor growth [22, 265] and impaired neurodevelopment in children [22, 70, 266] and can result in cretinism if iodine deficiency in pregnancy is severe [22]. Because of the concern of potential adverse effects on offspring, iodine supplementation in pregnancy has been recommended in many countries, including countries with sustained programs to combat iodine deficiency such as Australia with mandatory iodine fortification of bread [107]. However, the World Health Organization (WHO) does not recommend iodine supplementation during pregnancy in iodine sufficient populations with sustained programs to combat iodine deficiency [267]. Evidence from randomised controlled trials (RCT) suggested iodine supplementation during pregnancy improved child neurodevelopment and growth in severe iodine deficiency settings [174, 265] but that there is a lack of benefit in mild iodine deficiency to iodine sufficiency settings based on limited data [265, 268, 269]. There is a concern of excess iodine intake in pregnant women from iodine sufficient settings who took iodine-containing supplements [132]. Excessive iodine intake in pregnancy has been linked with adverse health outcomes and cognitive deficits [10].

Both iodine deficiency [83, 85, 178] and excess [270-272] in pregnancy were associated with elevated newborn thyroid stimulating hormone (TSH) concentration. Newborn TSH has been suggested as a marker of iodine nutrition in pregnancy [36] and it may be useful to identify children at risk of neurodevelopmental or growth delay in countries where newborn screening is routinely practised.

There are several studies that assessed the association between newborn TSH concentration and childhood neurodevelopment. Studies in iodine deficient populations showed that newborns with higher TSH concentrations of ≥5 mIU/L [86, 166, 167] had an increased risk of developmental delay, but the association between newborn TSH and neurodevelopmental scores were inconsistent in iodine sufficient populations [142, 161, 162, 273]. Most published studies had methodological limitations such as inadequate adjustment for potential confounders [86, 167], small sample sizes of <30 participants [163], or subjective assessment of the outcome [142, 273]. Only a few studies have examined the association between newborn TSH concentration and childhood growth [141, 142, 173]. A cohort study from an iodine deficient population showed no correlation between newborn TSH and infant growth [173]. In iodine sufficient populations, a higher risk of low birthweight was observed in newborns whose TSH was in the highest quartile compared with newborns in the lower three TSH quartiles [141], but Cuestas et al. [142] reported no difference in childhood growth between newborns with high (≥10 mIU/L) or low (<10 mIU/L) TSH concentrations.

The quality of the studies examining the relationship between newborn TSH concentration and neurodevelopment or growth of children was limited, and the findings were inconsistent. The objective of the current study was to investigate the associations between newborn TSH concentration and neurodevelopment and growth of children at 18 months of age utilising data collected from two large pregnancy studies in Australia, one in a borderline iodine deficient setting before the implementation of the mandatory iodine fortification of bread, and the other in an iodine sufficient setting after fortification, with adjustment for key confounding variables.

Subjects and Methods

Study design, setting and participants:

Participants were children from two pregnancy studies conducted in the same area of Australia: the Docosahexaenoic acid to Optimize Mother Infant Outcome (DOMInO) trial [274] and the Pregnancy Iodine and Neurodevelopment in Kids (PINK) study [70]. Children in the DOMInO study were born before the implementation of the mandatory iodine fortification of bread (born between 2005-2008), when the population was defined as borderline iodine deficient, and children in the PINK study were conceived and born after the implementation of the mandatory fortification (born between 2011-2013), when the population was defined as iodine sufficient. Detailed descriptions of DOMInO and PINK are reported elsewhere [70, 274]. Briefly, DOMInO was a multicenter randomised controlled trial that investigated the effect of docosahexaenoic acid (DHA) supplementation during pregnancy on postnatal depression and childhood cognitive and language development. Neurodevelopment and growth were assessed in a subset of children at 18 months of age (n = 726), consisting of a random sample of children born at term (n = 630) and all children born preterm (n = 96) from Adelaide, South Australia. The overall results of the trial showed DHA supplementation during pregnancy had no effect on the neurodevelopmental scores of children at 18 months, although the DHA treatment group had a lower risk of cognitive delay [274]. PINK was a prospective cohort study investigating the relationship between maternal iodine intake during pregnancy and the neurodevelopment of children at 18 months of age. The results of the PINK study showed that maternal iodine intake during pregnancy in the lowest or highest quartiles was associated with poorer Bayley-III scores when compared with maternal iodine intake in the middle quartiles [70]. Women were <21 weeks of gestation at enrolment in both studies. Women with known fetal abnormalities or the following attributes were excluded in both studies: drug or alcohol abuse, English not spoken at home or unable to give informed consent. Mothers with known thyroid disease were also excluded in the PINK

study. The study protocols were approved by the institutional review boards at each participating centre in both DOMInO (REC1657/12/2007) and PINK (REC2230/12/15 and 076.10) and each participant provided informed consent.

Assessment of Outcomes:

Neurodevelopment of children

Neurodevelopment of children was assessed at 18 months of age using the Bayley Scales of Infant and Toddler Development third edition (Bayley-III) in both studies [70, 274]. The Bayley-III Scale consists of cognitive, language, motor, adaptive behaviour and social emotional scales. The cognitive scale assesses sensorimotor development, exploration, manipulation, object relatedness, concept formation, memory and other aspects of cognitive processing. The language scale assesses both receptive and expressive communications. Both fine and gross motor skills were assessed using the motor scale. The adaptive behaviour and social emotional scales assessed parent reported child behaviours [275].

Growth of Children

Weight, length and head circumference were measured at the time of Bayley assessment using the standard procedures of the WHO in both studies [276]. Length was measured in a lying/supine position using a measuring board and was recorded to the nearest 0.1 centimetre. Weight was measured by placing the child on a calibrated weighing scale and was recorded to the nearest 10 g. Head circumference (HC) was measured using a non-stretching tape positioned above the eyebrows anteriorly and olecranon fossa posteriorly and was recorded to the nearest 0.1 centimetre [274]. Weight-for-age z scores (WAZ), weight-for-length z scores (WLZ), length-for-age z scores (LAZ) and head circumference-for-age z scores (HCZ) were calculated using the WHO's Child Growth Standard [277].

Assessment of Exposure:

Newborn TSH concentration data were extracted from the routine newborn screening database in the South Australian Neonatal Screening Centre [278]. The TSH concentration was determined from newborn whole blood samples taken between 3-4 days of infant's age by heel-prick. The method used to measure newborn TSH was consistent between the two studies.

Assessment of Confounders:

The potential confounders of child neurodevelopment and growth were identified based on the literature and are presented in **Supplemental Figures 7.1 and 7.2**. Socio-economic status data, including parental education and occupation, were collected using a structured self-report questionnaire. The Home Screening Questionnaire (HSQ) was completed when a parent attended the 18-month Bayley-III appointment to assess the level of cognitive, social and emotional support available to the child in the home environment [279]. The HSQ score was also used as a proxy for overall maternal care and support during pregnancy [280]. Pregnancy related background data on smoking, alcohol consumption, multivitamin intake and history of depression were collected during pregnancy by self-report. The mother's body mass index (BMI) was calculated using maternal weight and height measured at enrolment. Information on caesarean-section delivery, gestational age at birth, 5-minute Apgar score (a proxy for perinatal stress) and birthweight were retrieved from a case note audit at birth. Birthweight-for-gestational age z score was calculated using the Australian national birthweight percentiles by gestational age data [239].

Statistical analysis

The data were analysed using Stata 15 (Statacorp LP, College Station, Texas). Multiple imputation (MI) using chained equations was performed to handle missing information. The

pattern of missingness was non-monotonic and the data were assumed to be missing at random [281]. A total of twenty imputed datasets were generated. The MI model included the exposure variable, outcome variables, auxiliary variables (age at TSH assessment, hospital of enrolment, maternal height and weight) and potential confounders (parity, ethnicity, occupation, education, smoking, alcohol consumption, supplement intake, depression, mode of delivery, mothers age, gestational age, sex of the child, Apgar score, HSQ score and birthweight-for-gestational age z score). The newborn TSH in quartiles, and the binary variables (developmental delay, growth delay and dichotomised TSH as ≤5 mIU/L vs. >5 mIU/L) were created before the imputation. We compared these variables descriptively between the observed dataset and imputed dataset in the DOMInO and PINK studies, and they were similar in both studies (**Supplemental Table 7.1**). All analyses were performed on the imputed datasets.

The primary analyses were performed on the DOMInO and PINK data separately for several reasons. First, the two studies were conducted over different time periods where the iodine status of the population differed. DOMInO was conducted in the borderline iodine deficient setting before the implementation of mandatory iodine fortification of bread and recommendation of routine iodine supplementation for pregnant women, while PINK was conducted in the iodine sufficient setting after those national iodine interventions. Second, the two studies may not be directly compatible due to possible changes in the demographic and economic characteristics of the general population or the type of antenatal care over time. Third, while the participants were from the same geographic area in Australia, the design of the studies differed (DOMInO was a RCT while PINK was a cohort study). Finally, the direction of the association between TSH and neurodevelopment differed in the two studies. Meta-analysis was performed to examine the overall effect combining both studies.

To assess the association between newborn TSH and Bayley-III and growth outcomes, the exposure variable newborn TSH was treated as a continuous variable and categorised into quartiles in separate analyses. The newborn TSH distribution differed between DOMInO and PINK studies and hence different TSH quartile cut-offs were applied in each study. Sensitivity analyses were conducted to examine the robustness of the main findings by examining newborn TSH in tertiles, or quintiles, or applying the DOMInO TSH quartile cut points to PINK and vice versa, and using the TSH corrected for an infant's age at blood sampling (TSH~age) because infants in the highest TSH quartile were sampled earlier on average than infants in the lowest TSH quartile. It is well recognised that the age at sampling affects newborn TSH concentration due to the physiological surge in newborn TSH concentration for early infant blood sampling [45]. The TSH~age variable was created by regressing TSH on age at blood sampling. The residuals of the models, which represent the differences between each infant's actual TSH and the TSH concentration predicted by age at sampling, were added to the mean TSH. TSH and age at blood sampling were log-transformed for the analysis and back transformed to the original scales [64, 221]. TSH was also dichotomised based on the commonly reported categories in literature (>5 mIU/L vs. ≤5 mIU/L) to facilitate comparison with previous studies [12].

There are no well-designed studies on which to base power calculations for the assessment of childhood neurodevelopment and growth outcomes between children in the lowest and highest TSH quartiles. However, a five point difference on average in the cognitive and psychomotor development between children with adequate vs. deficient iodine or iron intake was considered clinically significant [70, 154]. With the sample sizes available in DOMInO (n = 726; all singleton) and PINK (n = 794; 11 twins), a mean Bayley-III score of 100 and a SD of 15, the studies have 89% and 91% power to detect a difference of five points in the Bayley-III scale between children in the highest and lowest TSH quartiles.

The neurodevelopment and growth outcomes of children were analysed as both continuous and categorical outcomes. For the categorical outcomes, developmental delay was defined as Bayley-III scores below 85 and growth delay was defined as anthropometric z scores below -1 (i.e. more than 1 SD below the mean).

A linear regression model was used to analyse continuous outcomes and a log Poisson regression model with robust variance estimation was used to analyse binary outcomes [282]. Both unadjusted and adjusted analyses were performed, with adjusted analyses considered to be the primary analysis. In DOMInO, we adjusted for DHA treatment group to investigate potential confounding as DHA may influence child development [283], though there is no known biological mechanism for an effect of DHA supplementation on iodine metabolism and newborn TSH concentration. The following confounders were controlled in the adjusted analyses for neurodevelopmental and growth outcomes in both DOMInO and PINK: parity, ethnicity, parents' occupation, parents' education, maternal smoking and alcohol consumption during and before pregnancy (Yes or No), maternal multivitamin intake (Yes or No) during pregnancy, maternal depression during and before pregnancy (Yes or No), mother's BMI, HSQ score, delivery by caesarean section (Yes or No), mother's age, gestational age at birth, 5 minute Apgar score, sex and birthweight-for-gestational age z score. Twins in the PINK study were not excluded in the analysis as the number of twins was small (n = 11) and the impact of clustering would not be a concern [284]. The mean differences (MDs) in Bayley-III scores with the 95% confidence intervals (CIs) estimated from the linear regression model and the Relative Risks (RRs) and 95% CIs estimated from the log Poisson regression model were reported to indicate the magnitude and direction of the associations. In line with current best practice [285-288], we have not used p values to interpret the associations. The STATA 'metan' command was used to perform the meta-analysis [289]. I^2 was used to evaluate

heterogeneity between the two studies. Heterogeneity between the results was classified as low (I^2 : <25%), moderate (I^2: 25-75%) and high (I^2: >75%) [147, 148]. A random effects model was applied due to the heterogeneity between the two studies. The pooled effect sizes were calculated to see the magnitude and direction of the associations between newborn TSH and neurodevelopment and growth. The pooled effect sizes were reported using forest plots.

Results

Maternal and infant characteristics

A total of 726 children in DOMInO and 794 children in PINK were included in the analysis (**Figure 7.1**). **Table 7.1** shows the characteristics of the DOMInO and PINK participants. The two studies differed in some socio-demographic characteristics. There were more Caucasian women, smokers and alcohol consumers in DOMInO compared with PINK. As a result of the study design that included a subset of randomly selected term infants and all preterm infants [274], there was a higher percentage of preterm infants in DOMInO. More PINK mothers completed secondary education and took multivitamin supplements during pregnancy compared with DOMInO mothers. The two studies had similar mean HSQ scores and birthweight-for-gestational age z scores.

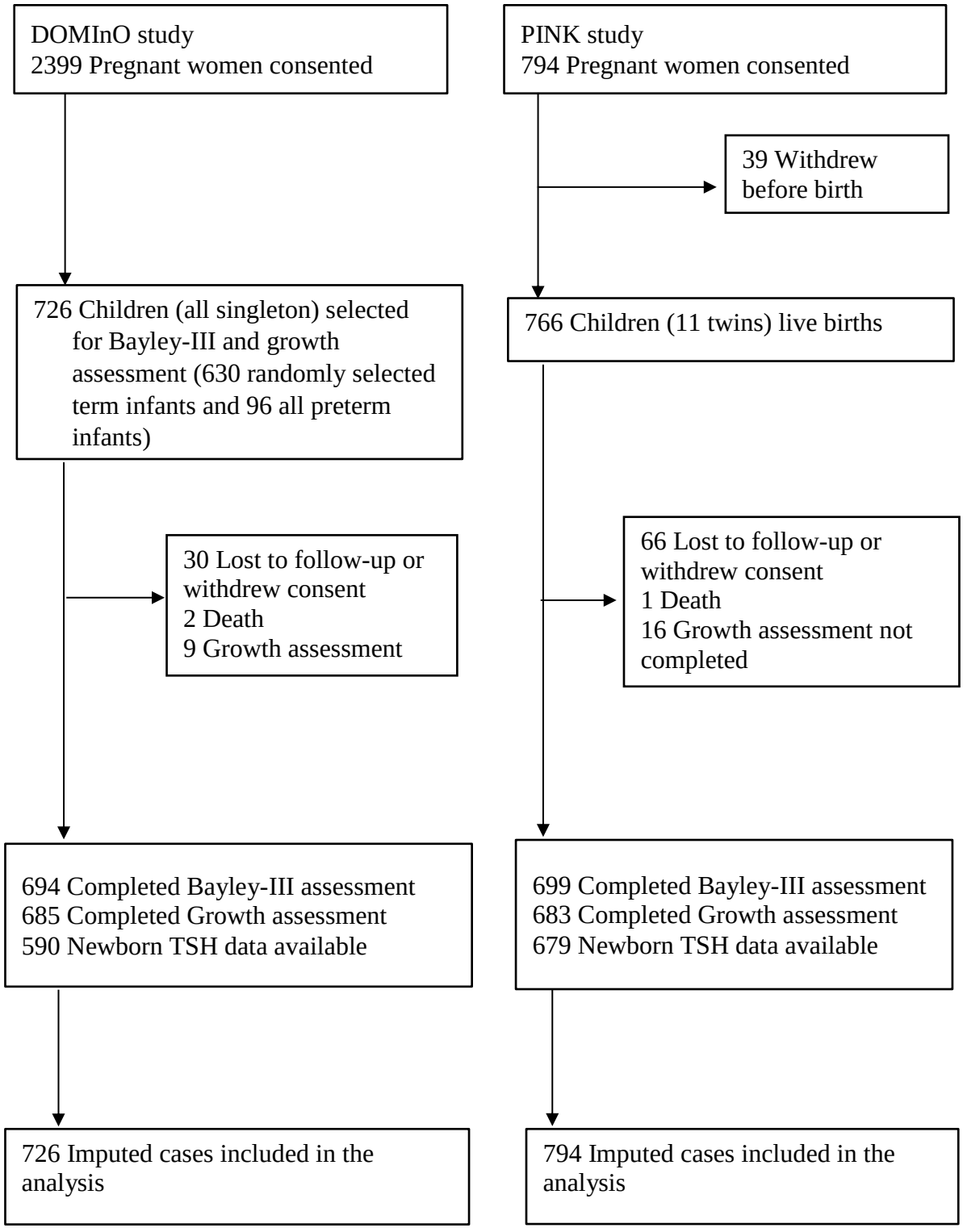


Figure 7. 1 Selection of participants from the DOMInO and PINK studies

Table 7. 1 Characteristics of the participants[1]

	DOMInO (n = 726)	PINK (n = 794)
Mother characteristics		
Caucasian ethnicity	663 (91.3)	668 (84.1)
Age, y	28.8 ± 5.8	30.0 ± 5.1
BMI, kg/m^2	27.3 ± 5.9	26.1 ± 5.4
Nulliparous	405 (55.8)	427 (53.8)
Completed secondary education	450 (62.0)	659 (83.0)
HSQ score	33.4 ± 3.6	33.0 ± 2.9
Smoking during second trimester	113 (15.6)	45 (5.7)
Previous smoking	227 (31.3)	102 (12.9)
Alcohol consumption during second trimester	76 (10.5)	41 (5.2)
Previous alcohol consumption	445 (61.3)	526 (66.3)
Supplement intake during second trimester	352 (48.5)	708 (89.2)
Depression during second trimester	48 (6.6)	48 (6.1)
Previous depression	159 (21.9)	169 (21.3)
Caesarean Section	203 (28.0)	222 (28.0)
Child characteristics		
Female sex	361 (49.7)	382 (48.1)
Gestational age, wk	38.7 ± 2.2	38.9 ± 1.8
Birthweight-for-gestational age z score	0.2 ± 1.0	0.0 ± 1.0
Preterm	90 (12.4)	61 (7.7)
5 minute APGAR score	9.0 ± 0.9	8.9 ± 0.7
Newborn TSH[2], mIU/L	1.7 (1.0, 2.9)	2.2 (1.4, 3.3)
Age at TSH assessment, h	58.7 ± 9.2	57.5 ± 9.0
Newborns with TSH >5 mIU/L	33 (4.6)	43 (5.4)
Age at Bayley assessment, mo	18.8 ± 1.8	19.5 ± 4.8

[1]Data are mean ± SD or n (%) unless otherwise specified. DOMInO, DHA to Optimize Mother Infant Outcome; HSQ, home screening questionnaire; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; TSH, thyroid-stimulating hormone.
[2]Median (Inter Quartile Range).

As shown in **Table 7.1**, the median (IQR) newborn TSH concentration was 1.7 (1.0, 2.9) mIU/L in DOMInO and 2.2 (1.4, 3.3) mIU/L in PINK. The percentage of newborns with TSH concentration >5 mIU/L was lower in DOMInO (4.6%) than PINK (5.4%). Newborn heel prick blood samples were collected between 48-96 hours of age in both studies and the median age at the TSH assessment was 57.5 hours in PINK, which was 1.2 hours earlier than DOMInO. In DOMInO, the median age at sampling was 1.5 hours earlier in preterm than term infants, but the median newborn TSH was comparable between these subgroups (1.8 mIU/L in preterm vs. 1.7 mIU/L in term infants).

TSH concentration and Bayley-III scores at 18 months of age

Table 7.2 reports the Bayley-III scores of children in the DOMInO and PINK studies. The mean Bayley-III scores were 3 to 6 points higher in DOMInO than PINK across all domains except the adaptive behaviour domain, which had similar mean scores in both studies. For example, the mean ± SD cognitive scores were 102 ± 12 points in DOMInO and 96 ± 12 points in PINK. The percentage of children with cognitive or language delay was lower in DOMInO than PINK. Children in the highest TSH quartile had the lowest mean Bayley-III score in the DOMInO study but the highest mean score in PINK study when compared with the other TSH quartiles.

Table 7. 2 Bayley-III outcomes at 18 months of age by newborn TSH concentration in quartiles[1]

	DOMInO			PINK		
	n	Bayley-III Score	Developmental delay	n	Bayley-III Score	Developmental delay
Cognitive scale						
All children	726	101.7 ± 11.9	36 (4.9)	794	96.3 ± 11.9	80 (10.0)
Quartile 1	184	103.3 ± 12.2	7 (4.0)	200	96.2 ± 11.7	21 (10.0)
Quartile 2	176	101.6 ± 11.5	6 (4.0)	196	96.1 ± 12.2	20 (10.0)
Quartile 3	185	101.8 ± 12.2	10 (6.0)	201	95.0 ± 11.8	25 (13.0)
Quartile 4	181	100.2 ± 11.9	12 (7.0)	197	97.8 ± 11.9	14 (7.0)
Language scale						
All children	726	97.2 ± 14.7	127 (17.4)	794	93.9 ± 18.1	234 (29.4)
Quartile 1	184	98.8 ± 14.8	23 (13.0)	200	93.7 ± 18.1	56 (28.0)
Quartile 2	176	98.2 ± 14.6	29 (17.0)	196	92.5 ± 6.3	56 (28.0)
Quartile 3	185	96.9 ± 14.9	36 (19.0)	201	94.0 ± 19.5	63 (31.0)
Quartile 4	181	94.7 ± 14.4	38 (21.0)	197	95.5 ± 18.2	59 (30.0)
Motor scale						
All children	726	102.5 ± 11.2	37 (5.1)	794	98.9 ± 11.2	35 (4.4)
Quartile 1	184	103.8 ± 10.6	6 (3.0)	200	99.5 ± 9.9	11 (5.0)
Quartile 2	176	102.6 ± 10.6	8 (5.0)	196	98.2 ± 11.6	8 (4.0)
Quartile 3	185	102.6 ± 11.4	8 (4.0)	201	97.6 ± 13.1	10 (5.0)
Quartile 4	181	101.0 ± 12.2	15 (8.0)	197	100.4 ± 9.6	6 (3.0)
Social emotional scale						
All children	726	106.8 ± 17.8	36 (4.9)	794	103.7 ± 14.7	44 (5.5)
Quartile 1	184	107.9 ± 18.4	10 (5.0)	200	102.7 ± 14.3	9 (5.0)
Quartile 2	176	107.5 ± 18.0	8 (5.0)	196	103.4 ± 15.2	13 (7.0)
Quartile 3	185	107.5 ± 18.0	8 (5.0)	201	103.9 ± 15.4	13 (6.0)
Quartile 4	181	104.6 ± 16.9	9 (5.0)	197	104.8 ± 13.7	9 (4.0)
Adaptive behaviour scale						
All children	726	100.0 ± 14.4	87 (12.1)	794	100.1 ± 14.0	87 (11.0)
Quartile 1	184	100.3 ± 15.1	19 (10.0)	200	100.3 ± 13.9	18 (9.0)
Quartile 2	176	100.2 ± 14.4	21 (12.0)	196	98.4 ± 14.3	29 (15.0)
Quartile 3	185	100.2 ± 13.6	21 (11.0)	201	100.6 ± 14.5	23 (11.0)
Quartile 4	181	99.1 ± 14.7	27 (15.0)	197	101.0 ± 13.3	17 (9.0)

[1]Data are mean ± SD or n (%). Developmental delay was defined as Bayley-III scores below 85. TSH was categorised into quartiles in both DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L.
DOMInO, DHA to Optimize Mother Infant Outcome; n, number of participants; PINK, Pregnancy Iodine and Neurodevelopment in Kids; TSH, thyroid-stimulating hormone.

There was a null association between newborn TSH and neurodevelopment. For example, when TSH was treated as a continuous exposure (**Table 7.3**), a 1 mIU/L increase in TSH was associated with a -0.3 point (95% CI: -1.0, 0.3) change in the mean cognitive score in DOMInO, compared with a 0.2 point (95% CI: -0.4, 0.7) change in PINK, and a RR of 1.1 (95% CI: 0.9, 1.3) for cognitive delay in DOMInO compared with a RR of 0.9 (95% CI: 0.8, 1.0) in PINK. When TSH was categorised into quartiles, the MD for children in the highest TSH quartile compared with the lowest quartile was -2.2 points (95% CI: -5.0, 0.7) for the cognitive score in DOMInO compared with 1.2 points (95% CI: -1.4, 3.7) in PINK, and the RR was 2.0 (95% CI: 0.6, 6.1) for cognitive delay in DOMInO compared with 0.7 (95% CI: 0.3, 1.4) in PINK (**Table 7.3**).

Table 7. 3 Adjusted associations between newborn TSH concentration and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley-III Score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Cognitive scale						
1 mIU/L increase in TSH[5]	726	-0.3 (-1.0, 0.3)	1.1 (0.9, 1.3)	794	0.2 (-0.4, 0.7)	0.9 (0.8, 1.0)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-1.3 (-4.0, 1.4)	1.1 (0.3, 4.3)	196	-0.7 (-3.1, 1.8)	1.1 (0.5, 2.1)
Quartile 3	185	-1.6 (-4.4, 1.2)	2.3 (0.7, 7.2)	201	-0.9 (-3.4, 1.5)	1.1 (0.6, 2.1)
Quartile 4	181	-2.2 (-5.0, 0.7)	2.0 (0.6, 6.1)	197	1.2 (-1.4, 3.7)	0.7 (0.3, 1.4)
Language scale						
1 mIU/L increase in TSH[5]	726	-0.4 (-1.2, 0.3)	1.1 (1.0, 1.2)	794	0.5 (-0.2, 1.3)	1.0 (1.0, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-3.4, 3.2)	1.3 (0.7, 2.2)	196	-2.0 (-5.4, 1.3)	1.1 (0.8, 1.5)
Quartile 3	185	-1.4 (-4.8, 2.0)	1.5 (0.9, 2.7)	201	0.6 (-3.2, 4.5)	1.1 (0.8, 1.6)
Quartile 4	181	-2.5 (-5.9, 0.9)	1.4 (0.8, 2.4)	197	1.4 (-2.0, 4.9)	1.1 (0.7, 1.5)
Motor scale						
1 mIU/L increase in TSH[5]	726	-0.3 (-0.9, 0.4)	1.2 (1.0, 1.4)	794	0.3 (-0.2, 0.8)	0.9 (0.7, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-1.0 (-3.5, 1.6)	1.6 (0.5, 5.3)	196	-1.3 (-3.6, 1.0)	0.7 (0.2, 2.0)
Quartile 3	185	-1.4 (-4.1, 1.3)	1.6 (0.5, 5.2)	201	1.5 (-3.9, 1.0)	0.8 (0.3, 2.2)
Quartile 4	181	-2.0 (-4.7, 0.8)	2.3 (0.8, 7.1)	197	1.3 (-0.7, 3.4)	0.5 (0.1, 1.8)
Social emotional scale						
1 mIU/L increase in TSH[5]	726	0.0 (-1.0, 1.0)	0.9 (0.7, 1.2)	794	0.4 (-0.2, 1.1)	1.0 (0.8, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.7 (-4.8, 3.4)	1.2 (0.4, 3.6)	196	0.7 (-2.5, 4.0)	1.6 (0.6, 4.6)
Quartile 3	185	-0.2 (-4.4, 4.0)	0.8 (0.2, 2.5)	201	1.6 (-1.8, 4.9)	1.8 (0.6, 5.0)
Quartile 4	181	-2.5 (-6.8, 1.7)	0.9 (0.3, 2.6)	197	2.9 (-0.3, 6.1)	1.1 (0.3, 3.9)
Adaptive behaviour scale						

	DOMInO			PINK		
	n	Bayley-III Score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
1 mIU/L increase in TSH[5]	726	0.0 (-0.8, 0.8)	1.0 (0.9, 1.2)	794	0.1 (-0.6, 0.8)	1.0 (0.8, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	0.0 (-3.4, 3.3)	1.2 (0.6, 2.2)	196	-2.0 (-5.1, 1.1)	1.6 (0.8, 3.0)
Quartile 3	185	-0.1 (-3.3, 3.5)	1.0 (0.5, 2.1)	201	1.1 (-2.0, 4.1)	0.9 (0.4, 1.9)
Quartile 4	181	-0.5 (-4.0, 3.1)	1.3 (0.7, 2.5)	197	0.9 (-2.0, 3.8)	0.8 (0.4, 1.7)

[1]The regression models were adjusted for sex, parity, ethnicity, occupation, education, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risk; TSH, thyroid-stimulating hormone.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable log Poisson regression model with robust variance estimation.

[5]TSH is modelled continuously.

[6]TSH was categorised into quartiles separately in DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L.

A similar pattern of null association was observed between newborn TSH and neurodevelopment in both studies regardless of whether TSH was categorised into tertiles (**Supplemental Table 7.2**), quintiles (**Supplemental Table 7.3**) or dichotomised at 5 mIU/L (**Supplemental Table 7.4**), or when the quartile cut points from the PINK study were applied to DOMInO and vice versa (**Supplemental Table 7.5**), or when TSH was corrected for age at sampling (**Supplemental Table 7.6**). Null association was also observed in the unadjusted analyses (**Supplemental Table 7.7**).

TSH concentration and growth at 18 months of age

As shown in **Table 7.4,** the mean WLZ, LAZ, WAZ and HCZ scores were mostly greater than zero for both DOMInO and PINK, indicating children were larger on average than the populations used to define the WHO growth standards. The percentage of children with z scores <-1 ranged from 3.5% to 15.6% in the DOMInO and 4.2% to 15.5% in the PINK studies.

Table 7. 4 Growth outcomes at 18 months of age by newborn TSH concentration in quartiles[1]

	DOMInO			PINK		
	n	Anthropometric z score	Growth delay	n	Anthropometric z score	Growth delay
WLZ						
All children	726	0.7 ± 1.0	25 (3.5)	794	0.6 ± 1.3	41 (5.2)
Quartile 1	184	0.7 ± 1.0	6 (3.0)	200	0.8 ± 1.1	7 (3.0)
Quartile 2	176	0.6 ± 1.0	5 (3.0)	196	0.6 ± 1.1	10 (5.0)
Quartile 3	185	0.6 ± 1.1	10 (5.0)	201	0.6 ± 1.1	13 (6.0)
Quartile 4	181	0.7 ± 1.0	5 (3.0)	197	0.6 ±1.8	12 (6.0)
LAZ						
All children	726	0.1 ± 1.1	113 (15.6)	794	-0.1 ± 0.2	123 (15.5)
Quartile 1	184	0.2 ± 1.1	24 (13.0)	200	0.0 ± 1.0	23 (11.0)
Quartile 2	176	0.1 ± 1.0	26 (15.0)	196	0.1 ± 1.2	27 (14.0)
Quartile 3	185	0.1 ± 1.1	26 (14.0)	201	-0.1 ± 0.3	42 (21.0)
Quartile 4	181	-0.1 ± 1.2	37 (20.0)	197	-0.2 ± 0.4	31 (16.0)
WAZ						
All children	726	0.5 ± 1.0	38 (5.3)	794	0.5 ± 1.4	56 (7.0)
Quartile 1	184	0.6 ± 1.0	12 (6.0)	200	0.8 ± 1.7	9 (4.0)
Quartile 2	176	0.5 ± 1.0	5 (3.0)	196	0.6 ± 1.6	12 (6.0)
Quartile 3	185	0.5 ± 1.1	7 (4.0)	201	0.3 ± 1.1	18 (9.0)
Quartile 4	181	0.5 ± 1.0	15 (8.0)	197	0.3 ± 1.1	17 (9.0)
HCZ						
All children	726	0.7 ± 1.0	30 (4.1)	794	0.9 ± 1.9	33 (4.2)
Quartile 1	184	0.8 ± 1.0	5 (3.0)	200	1.0 ± 1.3	7 (7.0)
Quartile 2	176	0.6 ± 0.9	5 (3.0)	196	0.9 ± 1.7	8 (10.0)
Quartile 3	185	0.7 ± 1.0	9 (5.0)	201	0.7 ± 1.4	12 (13.0)
Quartile 4	181	0.6 ± 1.0	10 (6.0)	197	0.9 ± 2.8	6 (12.0)

[1]Data are mean ± SD or n (%). Growth delay was defined as z scores below 1 SD. TSH was categorised into quartiles in both DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/l; quartile 2: 1.1-1.7 mIU/l; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L. cm, centimetre;

DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; n, number of participants; PINK, Pregnancy Iodine and Neurodevelopment in Kids; SD, standard deviation; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

The associations between newborn TSH and growth outcomes were similar in both PINK and DOMInO. Null associations were observed between newborn TSH and growth outcomes in both studies regardless of whether TSH was treated as a continuous variable or categorised into quartiles (**Table 7.5**), tertiles or quintiles (data not shown) or when TSH was dichotomised at 5 mIU/L (**Supplemental Table 7.8**), or when TSH was corrected for age at sampling (**Supplemental Table 7.9**). A similar pattern of null association was also observed in the unadjusted analyses (**Supplemental Table 7.10**)**.**

Table 7. 5 Adjusted associations between newborn TSH and growth outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Anthropometric z score	Growth delay[2]	n	Anthropometric z score	Growth delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
WLZ						
1 mIU/L increase in TSH[5]	726	0.01 (-0.04, 0.05)	0.8 (0.6, 1.1)	794	-0.02 (-0.09, 0.03)	1.0 (0.9, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.2 (-0.4, 0.0)	1.6 (0.3, 10.0)	196	-0.2 (-0.4, 0.1)	1.3 (0.4, 4.4)
Quartile 3	185	-0.1 (-0.3, 0.1)	2.2 (0.5, 9.8)	201	-0.2 (-0.5, 0.0)	1.4 (0.4, 4.7)
Quartile 4	181	-0.04 (-0.2, 0.2)	0.6 (0.1, 4.6)	197	-0.1 (-0.5, 0.2)	1.2 (0.4, 4.1)
LAZ						
1 mIU/L increase in TSH[5]	726	-0.01 (-0.05, 0.06)	1.0 (0.9, 1.1)	794	-0.01 (-0.07, 0.04)	1.0 (0.9, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-0.3, 0.1)	1.2 (0.7, 2.1)	196	0.1 (-0.1, 0.4)	1.1 (0.6, 2.1)
Quartile 3	185	-0.1 (-0.3, 0.1)	1.1 (0.6, 2.0)	201	0.0 (-0.3, 0.2)	1.5 (0.8, 2.6)
Quartile 4	181	-0.2 (-0.4, 0.1)	1.4 (0.8, 2.4)	197	-0.1 (-0.3, 0.2)	1.1 (0.6, 2.1)
WAZ						
1 mIU/L increase in TSH[5]	726	0.01 (-0.04, 0.05)	1.0 (0.8, 1.2)	794	-0.09 (-0.16, -0.01)	1.1 (0.9, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.2 (-0.4, 0.0)	0.6 (0.2, 2.5)	196	-0.2 (-0.5, 0.1)	1.2 (0.4, 3.7)
Quartile 3	185	-0.1 (-0.3, 0.1)	0.8 (0.2, 2.6)	201	-0.4 (-0.8, -0.1)	1.8 (0.7, 4.6)
Quartile 4	181	-0.1 (-0.3, 0.1)	1.2 (0.5, 3.3)	197	-0.4 (-0.8, -0.1)	1.5 (0.5, 4.2)
HCZ						
1 mIU/L increase in TSH[5]	726	-0.01 (-0.06, 0.03)	1.1 (0.9, 1.3)	794	0.00 (-0.12, 0.06)	0.8 (0.7, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.2 (-0.4, 0.1)	1.8 (0.4, 7.4)	196	-0.1 (-0.4, 0.3)	1.2 (0.3, 4.6)
Quartile 3	185	-0.1 (-0.3, 0.1)	2.8 (0.8, 9.6)	201	-0.2 (-0.6, 0.1)	1.3 (0.4, 5.0)
Quartile 4	181	-0.2 (-0.4, 0.0)	2.4 (0.7, 8.1)	197	0.0 (-0.5, 0.5)	0.6 (0.1, 2.5)

[1]The regression models were adjusted for sex, parity, ethnicity, occupation, education, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

[2]Growth delay was defined as z scores below 1 SD

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable log Poisson regression model with robust variance estimation.

[5]TSH is modelled continuously.

[6]TSH was categorised into quartiles in both DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5- 2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L.

Meta-analysis combined the DOMInO and PINK studies

As shown in **Supplemental Figures 7.3 and 7.4**, there was high heterogeneity between the DOMInO and PINK studies in the developmental outcomes, but low heterogeneity in the growth outcomes (**Supplemental Figures 7.5 and 7.6)** with null associations in all meta-analyses.

Discussion

Our study showed a null association between newborn TSH and both childhood neurodevelopment and growth in the DOMInO and PINK studies. However, children in the highest TSH quartile had the lowest mean Bayley-III score in the DOMInO study but the highest mean score in PINK study. We cannot rule out a different pattern of association between newborn TSH and neurodevelopmental outcomes in the two studies where the iodine status of the study populations differed because of the wide CIs around the effect estimates, as well as possible residual confounding. Similarly, we cannot rule out a poorer growth outcome of children in the highest TSH quartile compared with children in the lowest quartile in both studies.

Newborn TSH as an indicator of iodine nutrition during pregnancy may depend on maternal iodine intake, as either too little [83, 85, 178] or too much [270, 271] iodine during pregnancy may lead to high TSH. Excess iodine intake may also lead to lower TSH in iodine induced hyperthyroidism during pregnancy [290]. DOMInO was conducted in a borderline iodine deficient population before the mandatory iodine fortification of bread in Australia while PINK was conducted in an iodine sufficient population after the introduction of fortification. In studies from iodine deficient populations, like DOMInO, high newborn TSH may be due to inadequate maternal iodine intake during pregnancy. Iodine deficiency during pregnancy can lead to lower maternal free thyroxine concentration [291] and shifts the intra-thyroidal formation of thyroxine to the more active metabolite tri-iodothyronine [270, 292]. This leads

to a lower transfer of thyroid hormones, primarily thyroxine, to the foetus and may cause impaired neurodevelopment in early childhood. Although a majority of studies conducted in iodine deficient populations showed a negative association between newborn TSH and childhood neurodevelopment [86, 167, 169, 170], all except one study [86], had a small sample size (ranged from 61 to 250 participants), and did not account for key confounders including maternal socio-economic status, gestational age at birth and home environment. Our finding of a null association in DOMInO was consistent with a large study (n = 691) in a mildly iodine deficient population that controlled for key confounders including maternal socio-economic status and infants age at blood sampling for TSH assessment [160]. Similar to DOMInO, Murcia et al. [160] used Bayley scales to assess neurodevelopment of offspring under two years of age. Participants in DOMInO and Murcia et al. [160] were healthy children from population-based cohorts unlike a study in Italy from hospitalised children [169] where the cause for a higher TSH may differ. Furthermore, outcome assessors were not blinded to the exposure variable and maternal characteristics in a prospective cohort study by Costeira et al. [170], which may lead to bias. All of the studies in iodine deficient populations assessed neurodevelopment in toddlers and pre-schoolers except a large Australian study [86] that assessed educational performance at a median age of 10 years. In that Australian study [86], infants without congenital hypothyroidism whose TSH concentration was above the 99th percentile (TSH 12.4-20 mIU/L) had a higher risk of poorer school achievement than infants with TSH <75th percentile (TSH concentration ≤4 mIU/L). Neurodevelopmental assessments at later ages may be more stable and a better indication of cognitive function [293].

In contrast to the DOMInO study, the higher TSH concentration in PINK may be due to a high iodine intake as a result of mandatory iodine fortification and routine iodine supplementation in a previously iodine deficient population [82]. This is partially explained by the acute Wolff-Chaikoff effect, that excess iodine intake temporarily inhibits synthesis of

thyroid hormones and resulted in elevated TSH [270]. Our finding of a null association between newborn TSH and Bayley-III scores in PINK study is consistent with the PsychoTSH study (n = 315) [161, 162, 273] in an iodine sufficient population in Belgium, which is a prospective cohort study designed to assess the association between newborn TSH and neurodevelopmental outcomes in pre-schoolers. Similar to PINK, children who participated in the PsychoTSH study were born during a transitional period when Belgium introduced a voluntary program to use iodised salt in bread-making in 2009 [202], and the higher TSH may be due to the transient excess in iodine intake. Although the PsychoTSH study reported no association between TSH level and most of the neurodevelopmental scales [161, 162, 273], it showed higher mean verbal IQ scores in children with TSH concentration in the range 5-9 mIU/L compared with children whose TSH concentration <5 mIU/L [161]. We cannot rule out a higher neurodevelopmental score in children with TSH in the highest quartile (median TSH of 4.5 mIU/L) in the PINK study due to the wide CIs around the estimates. In contrast, though not directly compatible to our study, a smaller (n = 250) prospective cohort study by Cuestas et al. [142] in the iodine sufficient population in Argentina showed a higher odds of developmental delay in SAC with newborn TSH ≥10 mIU/L compared with TSH <10 mIU/L at birth, however, the developmental outcome assessment was by parents' evaluation of developmental status, and TSH was sampled earlier at 2-3 days of age outside the WHO recommendation. Prospective cohorts or population data linkage studies using exposure data from newborn screening and cognitive function at later ages may help to evaluate long-term effect of the newborn TSH concentration on cognition.

The newborn TSH level is influenced by the timing of blood sampling due to physiological surge within 48 hours after birth. Higher TSH level may also be a reflection of earlier blood sampling in newborn screening [45]. Although the median age at blood sampling was earlier in PINK than DOMINO overall, children in the highest TSH quartile had an earlier blood

sampling compared with the other quartiles in both the DOMInO and PINK studies. Therefore, age at the newborn TSH assessment is unlikely to explain the different patterns of associations observed between newborn TSH and neurodevelopmental outcomes in the two studies.

There is only one study (n = 250) [142], which examined the association between newborn TSH in full-term infants and childhood growth, which was conducted in Argentina, in an iodine sufficient population. It showed no difference in the weight and height at 6 years of age between children with newborn TSH ≥10 mIU/L and those with newborn TSH <10 mIU/L at birth though no adjustment was made for potential confounding variables. While the Argentina study is not directly compatible to our study, both suggested no difference in the growth of children between newborns with high or low TSH in populations of mild iodine deficiency or iodine sufficiency.

The WHO defines iodine deficiency in populations when >3% of newborns have TSH concentration >5 mIU/L [12]. Several studies examining the association between newborn TSH and developmental outcomes dichotomised TSH into ≤5 mIU/L or >5 mIU/L [167, 273], but it is important to note that TSH >5 mIU/L alone is not a criterion in classifying iodine deficiency either in populations or individuals. Further research is warranted to establish a newborn TSH cut-off associated with impaired neurodevelopment to identify children at risk of developmental delay for monitoring and early intervention as appropriate.

Our results should be interpreted with caution because our study is a secondary analysis utilising data from the DOMInO and PINK studies that differed in the distribution of several confounders. In the DOMInO study, for instance, there was an over-representation of preterm infants due to the study design. Although preterm birth was adjusted for in all models, and

Bayley-III scores were assessed based on the child's corrected age, preterm birth may increase TSH concentration and confound the association. The timing of newborn screening was earlier in PINK, as there was a shift towards earlier discharge from hospitals after birth. We attempted to address this through correcting TSH for an infant's age at blood sampling but residual confounding may still present.

Conclusion

Our study shows a null association between newborn TSH and both childhood neurodevelopment and growth. However, we cannot exclude the possibility of poorer neurodevelopment or growth in infants with high TSH in a borderline iodine deficient setting, and better neurodevelopment in infants with high TSH in an iodine sufficient setting, due to the wide CIs around the estimated effects. Follow-up studies that utilise newborn TSH data obtained from routine newborn screening and examines neurodevelopmental outcomes at a later age are warranted to evaluate whether newborn TSH predicts long-term neurodevelopmental outcome.

Acknowledgements

We thank the participants of both the DOMInO and PINK studies; and the staff at the University of Adelaide who participated in data collection and management in both studies.

Authors' contribution

MMW, LGS, LNY, MM and SJZ designed the study. MMW performed the analysis and drafted the manuscript. SJZ oversaw the study conduct and critically reviewed the manuscript. LGS contributed to the data analysis, interpretation of the results and reviewed the manuscript. LNY and MM contributed to interpretation of results and reviewed the manuscript. All authors approved the final manuscript.

Conflict of interest

Authors declare no conflict of interests.

CHAPTER 8 GENERAL DISCUSSION

8.1 Discussion of key findings and overall significance of the work

Iodine deficiency remains a public health problem worldwide [7]. There is a re-emergence of iodine deficiency even in countries that were previously iodine sufficient like Australia. Therefore, it would seem prudent to have regular monitoring of iodine status using appropriate markers of iodine status to prevent iodine deficiency or excess in general populations and at-risk population such as pregnant women.

My thesis is the first to systematically investigate the agreement between different markers of iodine status in populations. I have shown there is a discrepancy between these markers in classifying population iodine status. The iodine status of South Australians was sufficient by the UIC marker but classified as mildly iodine sufficient by the TSH marker. This suggests a need to use multiple markers of iodine status in populations as markers are conducted in different population groups. The WHO, UNICEF and ICCIDD jointly recommended using the median UIC in SAC, goitre prevalence in SAC, or proportion of newborns with TSH concentration >5 mIU/L to define iodine status in populations [12].

Choosing an appropriate marker is the important first step in monitoring iodine status of populations and early identification of iodine deficiency or excess in high-risk populations such as neonates and pregnant women. However, there is not a gold standard marker to define iodine status in populations that can be applied in all settings. The results from the systematic review presented in this thesis showed that these markers might not consistently classify the iodine status of populations [15]. The results showed that a population that is iodine sufficient by one marker may be classified as iodine deficient by another marker. In South Australia, there has been an inconsistent classification of iodine status when using different markers of population iodine status. The national health survey showed iodine sufficiency by UIC in

SAC [118] but iodine deficiency by the TSH marker [229]. This kind of discrepancy may affect the decision-making process towards monitoring and controlling IDD.

Several factors may contribute to the observed discrepancies among different markers to define the iodine status of populations. For instance, the length of time when populations were exposed to a low or high iodine environment may affect the choice of a marker. Markers of population iodine status differ in their application to monitor iodine status as the UIC marker may be preferable to assess iodine status shortly after an iodine supplementation or fortification program, while goitre and newborn TSH may be used to monitor long-term changes following iodine interventions [12]. Therefore, using multiple markers may give better representation for the iodine status of populations. Furthermore, the cut-off applied for different markers may not be appropriate and could be a source of disagreement in classifying iodine status. The current UIC cut-offs for SAC, adults and pregnant women, and the current newborn TSH cut-off recommended by the WHO have been criticised. The concern about the UIC marker is that it may overestimate iodine intake in SAC [294, 295] while underestimating iodine intake in adults [217]. In the recent WHO recommendation of UIC >100 µg /L to define iodine sufficiency [12], it was assumed that SAC had an average urine volume of 1 L irrespective of their age, though urine volume may be less than 1 L in many SAC [295] and UIC would be higher than 24-h UIE [296]. In adults, however, their daily urinary excretion 1.5 L and UIC would be lower than the 24-h UIE. As measuring 24-h UIE in a population survey is impractical, prone to errors in urine collection, and imposes burdens to participants [12], revision of the UIC criteria to define iodine status in adults needs to be considered. As the original WHO classification of iodine status using the median UIC was based from a functional assessment of goitre in SAC, the current newborn TSH cut-off also needs to be re-evaluated by either a functional outcome like childhood neurodevelopment or with the iodine status of SAC by 24-h UIE. Besides, consideration of the timing of iodine

intervention programs, re-evaluation of the current cut-offs for UIC and TSH markers has been suggested to define the iodine status of populations correctly.

The main concern with the newborn TSH marker is due to a recent move towards early discharge of mothers and newborns from hospitals, as the age at which TSH sampling for newborn screening is younger in recent surveys [52]. Early discharge results in artificially higher newborn TSH concentrations and higher proportions of newborns with TSH >5 mIU/L. Using the median UIC as a reference marker, the systematic review presented in my thesis suggested iodine sufficiency in populations when more than 5% of newborns have TSH >5 mIU/L [15]. The increased percentage cut-off in the proposed criterion may be due to the increasing number of studies with early blood sampling before three days of age for TSH assessment. This criterion is likely to be applicable in countries like Australia with the early discharge of infants that leads to early blood sampling for TSH assessment. Re-evaluation of the current TSH cut-off may improve the sensitivity and accuracy of the TSH marker to define and monitor the iodine status of populations. However, using the TSH marker is feasible only in countries that practise newborn screening for CH.

Exposure to a deficient or an excess iodine environment during pregnancy has been associated with adverse birth outcomes and neurodevelopmental deficits in childhood. Although UIC is not recommended for defining the iodine status of individuals due to a day-to-day or within-day variability in the iodine intake and hydration [12], measurement of urinary iodine remained a better biomarker to assess exposure to poor iodine nutrition compared with goitre and thyroid hormones [36]. However, discrepancies in urinary iodine markers (UIC, I/Cr and UIE) to classify individual iodine status in pregnancy, is a challenge to evaluate the associations between iodine nutrition in pregnancy and health, and neurodevelopmental outcomes in the offspring. Although median UIC is primarily designed

as a marker of population-level iodine status, the dichotomised median UIC (<150 vs ≥150 µg/L) from spot urine samples has been widely used to assess iodine status in individual pregnant women. However, UIC is highly variable during pregnancy due to hemodilution and iodine excretion [297, 298]. Whilst the 24-h UIE reduces within-day variation in iodine excretion, assessing the actual 24-h UIE is not practical and also does not address the day-to-day variation in iodine excretion. I/Cr and the estimated 24-h UIE, which are estimated from spot UIC and estimated 24-h urinary creatinine excretion, have been used as surrogates of actual 24-h UIE in healthy adults to assess iodine status in individuals [72, 214]. My thesis is the first to assess the agreement between spot UIC and other urinary iodine markers (estimated 24-h UIE, I/Cr and UIC~Cr) in pregnancy from iodine sufficient populations. The results showed that the spot UIC was not comparable with the estimated 24-h UIE nor I/Cr as the spot UIC underestimated iodine excess and overestimated iodine deficiency in pregnancy.

The results of my study that assessed the agreement between the urinary iodine markers during pregnancy showed that I/Cr and UIC~Cr had a better agreement with the estimated 24-h UIE than spot UIC. However, the agreement between each of these markers and the actual 24-h UIE is not known during pregnancy, and this could be the subject of future research. If the estimated 24-h UIE is a good predictor of actual 24-h UIE, obtaining socio-demographic variables (maternal age, ethnicity, height, weight and BMI) required to estimate 24-h creatinine excretion [69] would not be a challenge as these variables are commonly reported in many cohort studies during pregnancy. However, these variables may not be available in population-wide surveys that are designed to monitor the iodine status of populations. Therefore, continued use of the current UIC criterion >150 µg/L to define iodine sufficiency has been suggested to monitor iodine status in populations of pregnant women.

This thesis also examines the iodine status of general populations and populations of pregnant women who were from areas of low SES in South Australia. Based on findings from this thesis, the mandatory iodine fortification of bread implemented in October 2009 was successful in improving the iodine status of pregnant women from low SES areas in South Australia. Although the South Australians would be classified as mildly iodine deficient post-fortification using TSH as a marker in contrast to iodine sufficiency by UIC, the decreasing trend in the percentage of TSH >5 mIU/L in this large population study indicated improvement in iodine status of populations following the mandatory fortification of bread [229]. Moreover, the iodine status of the general population in South Australia would be improved from mild iodine deficiency to iodine sufficiency if the new proposed cut-off at the percentage of TSH >5 mIU/L greater than 5% from this thesis has been applied in classifying iodine status in the post-fortification period.

The results of my study on the iodine status of pregnant women showed that pregnant women in low SES areas of South Australia were iodine sufficient irrespective of iodine supplement intake. The WHO recommends iodine supplements in pregnancy when the median UIC was <100 µg/L in SAC [267] with sustained iodine prevention program. The median UIC was greater than 100 µg/L in SAC in the Australian Health Survey conducted three to four years post-fortification [118]. Therefore, based on the WHO's recommendation, routine iodine supplementation may not be required in Australia, for an iodine sufficient population [267, 299]. The routine iodine supplementation for pregnant and breastfeeding women was introduced in Australia without evidence from double-blind RCTs on the value of iodine supplementation in borderline iodine deficient or iodine sufficient population. Instead, pre-emptive recommendations were mainly based on the increasing concerns of negative consequences on the offspring following mild iodine deficiency in pregnancy. The implementation of both mandatory iodine fortification of bread and routine iodine

supplementation in pregnancy in an iodine sufficient population may lead to excessive iodine intake [70]. Iodine excess during pregnancy may be associated with thyroid hormone disorders [13, 132, 300] and subsequent impaired neurodevelopment in children [70]. Therefore, a double-blind RCT on the impacts of iodine supplementation on the iodine status of pregnant women, as well as on the general health outcomes, is warranted in iodine sufficient settings. However, conducting such a trial would not be practical in Australia, as routine iodine supplementation is recommended for pregnancy [269]. However, a national survey of iodine status in a representative sample of pregnant women would be informative in helping to understand the impact of both mandatory iodine fortification and routine iodine supplementation in improving the iodine status of pregnant women. Moreover, the supplementation strategy may be expensive and not a feasible intervention for women from low SES, as they may not be able to adhere to the recommendation. The fact that both iodine deficiency and excess are associated with impaired neurodevelopment further underscores the need for an iodine monitoring program, particularly among high-risk populations like pregnant women in Australia.

Even though the mandatory iodine fortification of bread was successful in improving iodine status among SAC and pregnant women in South Australia, public health strategies to reduce salt intake may be an issue for ongoing success in the prevention of iodine deficiency [301, 302]. For example, there was a significant reduction in the sodium content of bread from 2009 to 2013 in Australia [303], which may have a significant impact on the iodine content of bread. Whilst salt reduction is a very important measure to reduce cardiovascular disease risk, the implications of low iodine intake of individuals on impaired growth and development should not be overlooked even in iodine sufficient populations. Therefore, the two public health measures need to be balanced [106, 304] with a shared goal of achieving salt reduction at the same time as improving iodine intake. Moreover, as bread is one of the major dietary

sources of iodine and a major contributor of calories in Australians, including pregnant women, the public health message on bread choices and quantity to prevent iodine deficiency needs to be advocated.

This thesis also examined whether newborn TSH is a suitable biomarker to identify children at risk of impaired childhood neurodevelopment in both borderline iodine deficient and iodine sufficient settings. The results showed a null association between newborn TSH and neurodevelopment in both settings. However, previous studies showed a lower neurodevelopmental score in infants with high TSH in iodine deficient population [86, 165-167, 169, 170] and no association in iodine sufficient population [162, 168]. Therefore, the associations between newborn TSH and child neurodevelopment might vary depending on the iodine status of populations. However, most of the previous studies collected newborn TSH samples outside of the recommended age range of 3-4 days after birth, did not control key confounders and had assessed neurodevelopmental outcomes at early ages [167, 169, 170]. Neurodevelopmental assessment at early ages may not be conclusive for long-term neurodevelopmental consequences [293]. There was no study that examined the association between newborn TSH and neurodevelopmental outcomes assessed at later ages using standard scales from a large population-based data in both iodine deficient and sufficient settings. Assessment of neurodevelopmental outcomes in older children has been suggested as more precise, reliable and a better predictor of long-term mental function and adulthood productivity [293]. Therefore, long-term follow up studies that assess the associations between newborn TSH and neurodevelopment outcomes in later ages are warranted to use newborn TSH as a screening tool to identify children at risk of neurodevelopmental deficits.

As thyroid hormones are essential for normal physical growth in early life, newborn TSH could be applied as a marker to identify children at risk of growth delay. This thesis has

examined the association between newborn TSH and childhood anthropometry in healthy well-nourished populations using an adequate sample size and appropriate age at TSH sampling that controlled for key confounders of child growth. The results showed a null association between newborn TSH and anthropometry in both borderline iodine deficient and iodine sufficient settings. However, there were only a few children with anthropometric z scores <-1 SD with a mean anthropometric z score of 0 SD in participating children. A study in healthy infants from iodine sufficient populations also showed no association between newborn TSH and child anthropometry though the study did not adjust for key confounders of child growth [142]. It might not be worthwhile using TSH as a marker of child anthropometry in borderline iodine deficient or iodine sufficient well-nourished populations.

The use of newborn TSH as a biomarker to screen children at increased risk of impaired neurodevelopment or growth has multiple benefits. First, as newborn TSH data is available from routine newborn screenings, there will be no additional cost for the determination of the biomarker. Second, newborn TSH can identify children exposed to both iodine deficiency and iodine excess in pregnancy. Third, as the marker is collected at birth, children identified as having a higher risk of impaired neurodevelopment are able to access the available treatments as early as possible.

8.2 Potential future works

The following areas of future research work have been suggested based on the results of this thesis.

- Re-evaluation of the current TSH cut-off to improve the accuracy and sensitivity of the TSH marker to define iodine status in populations appropriately.
- A double-blind RCT on the health and neurodevelopmental impacts of iodine supplementation during pregnancy on iodine sufficient populations.

- A program for monitoring iodine status in pregnant women across all states in Australia as both iodine deficiency and excess may be associated with impaired neurodevelopment.
- Ways of integrating public health programs that recommend cutting salt intake with the national program that encourages to increase iodine in bread-making. Whether increasing the level of iodine in the salt, from its current range of 25-65 mg of iodine per kg of salt to higher acceptable thresholds that will align with a lower amount of salt, is technically achievable and safe remains an area of further research.
- Assessment of the agreement between actual 24-h UIE and estimated 24-h UIE as well as the agreement between spot UIC and I/Cr with actual 24-h UIE in pregnant women from all settings.
- Population-based data linkage studies that assess the associations between newborn TSH and neurodevelopment outcomes at later ages.
- Continuous monitoring and appropriate intervention for infants with mildly elevated TSH >5 mIU/L at birth to prevent later childhood cognitive impairment.
- Utilising newborn TSH data from newborn screening as a potential biomarker to identify children at risk of impaired development or growth in different settings.

8.3 Conclusions

Regular monitoring of iodine status using an appropriate measure of population-level and individual-level iodine status helps to identify populations or individuals at increased risk of IDD such as impaired neurodevelopment. The disagreement between markers to classify iodine status in South Australia may affect measures to prevent and control iodine deficiency. Re-evaluation of the current TSH criteria for classifying iodine status in populations, but continued use of median UIC from spot UIC in populations of pregnant women, is suggested. Mandatory iodine fortification of bread was successful in improving the iodine status of

pregnant women in South Australia regardless of iodine supplement intake. However, more data is required on the iodine status of pregnant women by way of a national representative survey. Newborn TSH may be a potential marker for identifying children at risk of poor neurodevelopmental scores at 18 months because infants with higher TSH may have poor cognitive scores, but future research should focus on neurodevelopmental outcomes at a later age for evaluating long-term neurodevelopmental consequences in countries where routine newborn screening is practised.

REFERENCES

1. WHO, UNICEF, and ICCIDD, *Ellimination of iodine deficiency disorders: a manual for health workers* in *EMRO Technical Publications Series*. 2008.
2. Bernal, J., *Thyroid hormones and brain development.* Vitam Horm, 2005. **71**: p. 95-122.
3. Hetzel, B.S., *Iodine deficiency disorders (IDD) and their eradication.* Lancet, 1983. **2**(8359): p. 1126-9.
4. Velasco, I., S.C. Bath, and M.P. Rayman, *Iodine as Essential Nutrient during the First 1000 Days of Life.* Nutrients, 2018. **10**(3).
5. Qian, M., et al., *The effects of iodine on intelligence in children: a meta-analysis of studies conducted in China.* Asia Pac J Clin Nutr, 2005. **14**(1): p. 32-42.
6. *World Summit for Children*, in *Goals for Children and Development in the 1990s* 1990, United Nations: New York
7. Iodine Global Network, *Global scorecard of iodine nutrition in 2019 based on median urinary iodine concentration (mUIC) in school-age children (SAC).* 2019, Iodine Global Network: Zurich, Switzerland.
8. Bath, S.C., et al., *A multi-centre pilot study of iodine status in UK schoolchildren, aged 8-10 years.* Eur J Nutr, 2016. **55**(6): p. 2001-9.
9. Li, M. and C.J. Eastman, *The changing epidemiology of iodine deficiency.* Nat Rev Endocrinol, 2012. **8**(7): p. 434-40.
10. Pearce, E.N., et al., *Consequences of iodine deficiency and excess in pregnant women: an overview of current knowns and unknowns.* Am J Clin Nutr, 2016. **104 Suppl 3**: p. 918S-23S.
11. Gartner, R., *Recent data on iodine intake in Germany and Europe.* J Trace Elem Med Biol, 2016. **37**: p. 85-89.
12. WHO, UNICEF, and ICCIDD, *Assessment of iodine deficiency disorders and monitoring their elimination-a guide for programme managers* 2007, World Health Organization
13. Burgi, H., *Iodine excess.* Best Pract Res Clin Endocrinol Metab, 2010. **24**(1): p. 107-15.
14. Monahan, M., et al., *Costs and benefits of iodine supplementation for pregnant women in a mildly to moderately iodine-deficient population: a modelling analysis.* Lancet Diabetes Endocrinol, 2015. **3**(9): p. 715-22.
15. Wassie, M.M., P. Middleton, and S.J. Zhou, *Agreement between markers of population iodine status in classifying iodine status of populations: a systematic review.* Am J Clin Nutr, 2019. **110**(4): p. 949-958.
16. Food Standards Australia New Zealand, *Australia New Zealand Food Standards Code – Standard 2.1.1 – Cereal and cereal products* 2009, Federal Register of Legislative Instruments.
17. Li, M., et al., *Are Australian children iodine deficient? Results of the Australian National Iodine Nutrition Study.* Med J Aust, 2006. **184**(4): p. 165-9.
18. *Turkish Association for Psychopharmacology (TAP) 5th International Congress on Psychopharmacology & International Symposium on Child and Adolescent Psychopharmacology October 30-November 3, 2013 Cornelia Diamond Hotel, Antalya, Turkey Abstracts.* Klinik Psikofarmakoloji Bulteni-Bulletin of Clinical Psychopharmacology, 2013. **23**: p. S1-S292.
19. Adachi, M., et al., *Mass screening of newborns for congenital hypothyroidism of central origin by free thyroxine measurement of blood samples on filter paper.* Eur J Endocrinol, 2012. **166**(5): p. 829-38.
20. Rousset, B., et al., *Chapter 2 Thyroid Hormone Synthesis And Secretion*, in *Endotext*, K.R. Feingold, et al., Editors. 2000: South Dartmouth (MA).

21. Porterfield, S.P., *Vulnerability of the developing brain to thyroid abnormalities: environmental insults to the thyroid system.* Environ Health Perspect, 1994. **102 Suppl 2**: p. 125-30.
22. Zimmermann, M.B., *The role of iodine in human growth and development.* Semin Cell Dev Biol, 2011. **22**(6): p. 645-52.
23. Bernal, J. and J. Nunez, *Thyroid hormones and brain development.* Eur J Endocrinol, 1995. **133**(4): p. 390-8.
24. Delange, F., *The role of iodine in brain development.* Proc Nutr Soc, 2000. **59**(1): p. 75-9.
25. Zimmermann, M.B., P.L. Jooste, and C.S. Pandav, *Iodine-deficiency disorders.* Lancet, 2008. **372**(9645): p. 1251-62.
26. Leung, A.M. and L.E. Braverman, *Consequences of excess iodine.* Nat Rev Endocrinol, 2014. **10**(3): p. 136-42.
27. National Health and Medical Research Council, *Nutrient Reference Values for Australia and New Zealand Including Recommended Dietary Intakes* 2006.
28. Charlton, K., Y. Probst, and G. Kiene, *Dietary Iodine Intake of the Australian Population after Introduction of a Mandatory Iodine Fortification Programme.* Nutrients, 2016. **8**(11).
29. Condo, D., et al., *Iodine status of pregnant women in South Australia after mandatory iodine fortification of bread and the recommendation for iodine supplementation.* Matern Child Nutr, 2017. **13**(4).
30. Food Standards Australia New Zealand, *Australia New Zealand Food Standards Code - Standard 2.10.2 - Salt and salt products*. 2009: cambera
31. Santos, J.A.R., et al., *Iodine fortification of foods and condiments, other than salt, for preventing iodine deficiency disorders.* Cochrane Database Syst Rev, 2019. **2**: p. CD010734.
32. WHO, U., *Reaching Optimal Iodine Nutrition in Pregnant and Lactating Women and Young Children* 2007, WHO: Geneva.
33. Public Health Committee of the American Thyroid, A., et al., *Iodine supplementation for pregnancy and lactation-United States and Canada: recommendations of the American Thyroid Association.* Thyroid, 2006. **16**(10): p. 949-51.
34. Zimmermann, M.B., et al., *New reference values for thyroid volume by ultrasound in iodine-sufficient schoolchildren: a World Health Organization/Nutrition for Health and Development Iodine Deficiency Study Group Report.* Am J Clin Nutr, 2004. **79**(2): p. 231-7.
35. Andersson, M., V. Karumbunathan, and M.B. Zimmermann, *Global iodine status in 2011 and trends over the past decade.* J Nutr, 2012. **142**(4): p. 744-50.
36. Rohner, F., et al., *Biomarkers of nutrition for development--iodine review.* J Nutr, 2014. **144**(8): p. 1322S-1342S.
37. WHO, ICCIDD, and UNICEF, *Indicators for assessing iodine deficiency disorders and their control through salt iodization*. 1994, Geneva : World Health Organization.
38. Peterson, S., et al., *Classification of thyroid size by palpation and ultrasonography in field surveys.* Lancet, 2000. **355**(9198): p. 106-10.
39. Clifton, V.L., et al., *The impact of iodine supplementation and bread fortification on urinary iodine concentrations in a mildly iodine deficient population of pregnant women in South Australia.* Nutr J, 2013. **12**: p. 32.
40. Travers, C.A., et al., *Iodine status in pregnant women and their newborns: are our babies at risk of iodine deficiency?* Med J Aust, 2006. **184**(12): p. 617-20.
41. Ascoli W, A.G., *Epidemiologia el bocio endémico en Centro América. Relación entre prevalencia y excreción urinaria de yodo [Epidemiology of endemic goiter in Central America. Association between prevalence and urinary iodine excretion]*. . Arch Latinoam Nutr, 1970. **20**: p. 309–320.

42. Chanoine, J.P., et al., *Smoking during pregnancy: a significant cause of neonatal thyroid enlargement.* Br J Obstet Gynaecol, 1991. **98**(1): p. 65-8.
43. Vanderpump, M., *Thyroid and iodine nutritional status: a UK perspective.* Clin Med (Lond), 2014. **14 Suppl 6**(6): p. s7-11.
44. Domene, H.M., et al., *The C105fs114X is the prevalent thyrotropin beta-subunit gene mutation in Argentinean patients with congenital central hypothyroidism.* Horm Res, 2004. **61**(1): p. 41-6.
45. Li, M. and C.J. Eastman, *Neonatal TSH screening: is it a sensitive and reliable tool for monitoring iodine status in populations?* Best Pract Res Clin Endocrinol Metab, 2010. **24**(1): p. 63-75.
46. Lao, T.T. and N.S. Panesar, *Neonatal thyrotrophin and mode of delivery.* Br J Obstet Gynaecol, 1989. **96**(10): p. 1224-7.
47. Korada, M., et al., *TSH levels in relation to gestation, birth weight and sex.* Horm Res, 2009. **72**(2): p. 120-3.
48. Nazeri, P., et al., *Can postpartum maternal urinary iodine be used to estimate iodine nutrition status of newborns?* Br J Nutr, 2016. **115**(7): p. 1226-31.
49. Brucker-Davis, F., et al., *Neurotoxicant exposure during pregnancy is a confounder for assessment of iodine supplementation on neurodevelopment outcome.* Neurotoxicol Teratol, 2015. **51**: p. 45-51.
50. Freire, C., et al., *Prenatal exposure to organochlorine pesticides and TSH status in newborns from Southern Spain.* Sci Total Environ, 2011. **409**(18): p. 3281-7.
51. Moreno-Reyes, R., et al., *ETA Abstracts 2012.* European Thyroid Journal, 2012. **1**(s1): p. 75-208.
52. Clapin, H., et al., *Factors influencing neonatal thyroid-stimulating hormone concentrations as a measure of population iodine status.* J Pediatr Endocrinol Metab, 2014. **27**(1-2): p. 101-6.
53. Anjos, T., et al., *Nutrition and neurodevelopment in children: focus on NUTRIMENTHE project.* Eur J Nutr, 2013. **52**(8): p. 1825-42.
54. Burns, R., et al., *Can neonatal TSH screening reflect trends in population iodine intake?* Thyroid, 2008. **18**(8): p. 883-8.
55. Zimmermann, L.J., et al., *Surfactant metabolism in the neonate.* Biol Neonate, 2005. **87**(4): p. 296-307.
56. Caylan, N., et al., *Neonatal Thyroid-Stimulating Hormone Screening as a Monitoring Tool for Iodine Deficiency in Turkey.* J Clin Res Pediatr Endocrinol, 2016. **8**(2): p. 187-91.
57. Mikelsaar, R.V. and M. Viikmaa, *Neonatal thyroid-stimulating hormone screening as an indirect method for the assessment of iodine deficiency in Estonia.* Horm Res, 1999. **52**(6): p. 284-6.
58. Simsek, E., M. Karabay, and K. Kocabay, *Neonatal screening for congenital hypothyroidism in West Black Sea area, Turkey.* Int J Clin Pract, 2005. **59**(3): p. 336-41.
59. Dalili, S., et al., *1547 Can we Compare Indicators of Iodine Deficiency Disorder in Neonate With School- Aged Children?* Archives of Disease in Childhood, 2012. **97**(Suppl 2): p. A438-A438.
60. Simsek, E., et al., *Congenital hypothyroidism and iodine status in Turkey: a comparison between the data obtained from an epidemiological study in school-aged children and neonatal screening for congenital hypothyroidism in Turkey.* Pediatr Endocrinol Rev, 2003. **1 Suppl 2**: p. 155-61.
61. Als, C., et al., *Quantification of urinary iodine: a need for revised thresholds.* Eur J Clin Nutr, 2003. **57**(9): p. 1181-8.

62. Konig, F., et al., *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**(11): p. 2049-54.
63. Andersen, S., et al., *Reliability of studies of iodine intake and recommendations for number of samples in groups and in individuals.* Br J Nutr, 2008. **99**(4): p. 813-8.
64. Murcia, M., et al., *Iodine intake from supplements and diet during pregnancy and child cognitive and motor development: the INMA Mother and Child Cohort Study.* J Epidemiol Community Health, 2018. **72**(3): p. 216-222.
65. Bath, S.C., et al., *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**(9889): p. 331-7.
66. Knudsen, N., et al., *Age- and sex-adjusted iodine/creatinine ratio. A new standard in epidemiological surveys? Evaluation of three different estimates of iodine excretion based on casual urine samples and comparison to 24 h values.* Eur J Clin Nutr, 2000. **54**(4): p. 361-3.
67. Li, C., et al., *The Urine Iodine to Creatinine as an Optimal Index of Iodine During Pregnancy in an Iodine Adequate Area in China.* J Clin Endocrinol Metab, 2016. **101**(3): p. 1290-8.
68. Kesteloot, H. and J.V. Joossens, *On the determinants of the creatinine clearance: a population study.* J Hum Hypertens, 1996. **10**(4): p. 245-9.
69. Mage, D.T., R.H. Allen, and A. Kodali, *Creatinine corrections for estimating children's and adult's pesticide intake doses in equilibrium with urinary pesticide and creatinine concentrations.* J Expo Sci Environ Epidemiol, 2008. **18**(4): p. 360-8.
70. Zhou, S.J., et al., *Association Between Maternal Iodine Intake in Pregnancy and Childhood Neurodevelopment at Age 18 Months.* Am J Epidemiol, 2019. **188**(2): p. 332-338.
71. Snart, C.J.P., et al., *Maternal Iodine Status and Associations with Birth Outcomes in Three Major Cities in the United Kingdom.* Nutrients, 2019. **11**(2).
72. Perrine, C.G., et al., *Comparison of population iodine estimates from 24-hour urine and timed-spot urine samples.* Thyroid, 2014. **24**(4): p. 748-57.
73. Konno, N., et al., *Clinical evaluation of the iodide/creatinine ratio of casual urine samples as an index of daily iodide excretion in a population study.* Endocr J, 1993. **40**(1): p. 163-9.
74. Charlton, K.E., et al., *Iodine Status Assessment in South African Adults According to Spot Urinary Iodine Concentrations, Prediction Equations, and Measured 24-h Iodine Excretion.* Nutrients, 2018. **10**(6): p. 736.
75. Remer, T. and F. Manz, *The inadequacy of the urinary iodine-creatinine ratio for the assessment of iodine status during infancy, childhood and adolescence.* J Trace Elem Electrolytes Health Dis, 1994. **8**(3-4): p. 217-9.
76. Ohira, S., et al., *Creatinine adjustment of spot urine samples and 24 h excretion of iodine, selenium, perchlorate, and thiocyanate.* Environ Sci Technol, 2008. **42**(24): p. 9419-23.
77. Fuse, Y., et al., *Iodine status of pregnant and postpartum Japanese women: effect of iodine intake on maternal and neonatal thyroid function in an iodine-sufficient area.* J Clin Endocrinol Metab, 2011. **96**(12): p. 3846-54.
78. Brander, L., et al., *Urinary iodine concentration during pregnancy in an area of unstable dietary iodine intake in Switzerland.* J Endocrinol Invest, 2003. **26**(5): p. 389-96.
79. Bath, S.C., et al., *Gestational changes in iodine status in a cohort study of pregnant women from the United Kingdom: season as an effect modifier.* Am J Clin Nutr, 2015. **101**(6): p. 1180-7.

80. Ji, C., et al., *Determination of Reference Intervals of Ratios of Concentrations of Urinary Iodine to Creatinine and Thyroid Hormone Concentrations in Pregnant Women Consuming Adequate Iodine in Harbin, Heilongjiang Province.* Biol Trace Elem Res, 2020. **193**(1): p. 36-43.
81. Gunton, J.E., et al., *Iodine deficiency in ambulatory participants at a Sydney teaching hospital: is Australia truly iodine replete?* Med J Aust, 1999. **171**(9): p. 467-70.
82. Nohr, S.B. and P. Laurberg, *Opposite variations in maternal and neonatal thyroid function induced by iodine supplementation during pregnancy.* J Clin Endocrinol Metab, 2000. **85**(2): p. 623-7.
83. Sareen, N., et al., *Iodine nutritional status in Uttarakhand State, India.* Indian J Endocrinol Metab, 2016. **20**(2): p. 171-6.
84. Zimmermann, M.B., et al., *Increasing the iodine concentration in the Swiss iodized salt program markedly improved iodine status in pregnant women and children: a 5-y prospective national study.* Am J Clin Nutr, 2005. **82**(2): p. 388-92.
85. Kapil, U., et al., *Iodine nutritional status in Himachal Pradesh state, India.* Indian J Endocrinol Metab, 2015. **19**(5): p. 602-7.
86. Lain, S.J., et al., *Association between borderline neonatal thyroid-stimulating hormone concentrations and educational and developmental outcomes: a population-based record-linkage study.* Lancet Diabetes Endocrinol, 2016. **4**(9): p. 756-765.
87. Zimmermann, M.B., *Methods to assess iron and iodine status.* Br J Nutr, 2008. **99 Suppl 3**(S3): p. S2-9.
88. Zimmermann, M.B., et al., *Thyroglobulin is a sensitive measure of both deficient and excess iodine intakes in children and indicates no adverse effects on thyroid function in the UIC range of 100-299 mug/L: a UNICEF/ICCIDD study group report.* J Clin Endocrinol Metab, 2013. **98**(3): p. 1271-80.
89. Ma, Z.F. and S.A. Skeaff, *Thyroglobulin as a biomarker of iodine deficiency: a review.* Thyroid, 2014. **24**(8): p. 1195-209.
90. ICCIDD/UNICEF/WHO, *Global Prevalence of Iodine Deficiency Disorders*, in *MDIS working paper #1*, M.D.I. System, Editor. 1993, World Health Organization Geneva.
91. Andersson, M., et al., *Current global iodine status and progress over the last decade towards the elimination of iodine deficiency.* Bull World Health Organ, 2005. **83**(7): p. 518-25.
92. Aburto N, A.M., Candeias V, Wu T. , *Effect and safety of salt iodization to prevent iodine deficiency disorders: a systematic review with meta-analyses*, in *WHO eLibrary of Evidence for Nutrition Actions (eLENA)*. 2014, World Health Organization Geneva.
93. Dold, S., et al., *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**(4): p. 587-598.
94. Doggui, R., et al., *Adequacy Assessment of a Universal Salt Iodization Program Two Decades after Its Implementation: A National Cross-Sectional Study of Iodine Status among School-Age Children in Tunisia.* Nutrients, 2016. **9**(1).
95. Jaiswal, N., et al., *The iodized salt programme in Bangalore, India provides adequate iodine intakes in pregnant women and more-than-adequate iodine intakes in their children.* Public Health Nutr, 2015. **18**(3): p. 403-13.
96. Taylor, P.N., et al., *Therapy of endocrine disease: Impact of iodine supplementation in mild-to-moderate iodine deficiency: systematic review and meta-analysis.* Eur J Endocrinol, 2014. **170**(1): p. R1-R15.
97. Sukkhojaiwaratkul, D., et al., *Effects of maternal iodine supplementation during pregnancy and lactation on iodine status and neonatal thyroid-stimulating hormone.* J Perinatol, 2014. **34**(8): p. 594-8.
98. Andersen, S.L., et al., *Iodine deficiency in Danish pregnant women.* Dan Med J, 2013. **60**(7): p. A4657.

99. Anaforoglu, I., et al., *Iodine status among pregnant women after mandatory salt iodisation.* Br J Nutr, 2016. **115**(3): p. 405-10.
100. Randremanana, R.V., et al., *First national iodine survey in Madagascar demonstrates iodine deficiency.* Matern Child Nutr, 2019. **15**(2): p. e12717.
101. Perrine, C.G., et al., *Iodine Status of Pregnant Women and Women of Reproductive Age in the United States.* Thyroid, 2019. **29**(1): p. 153-154.
102. Caldwell, K.L., et al., *Iodine status in pregnant women in the National Children's Study and in U.S. women (15-44 years), National Health and Nutrition Examination Survey 2005-2010.* Thyroid, 2013. **23**(8): p. 927-37.
103. Sun, D., et al., *Eliminating Iodine Deficiency in China: Achievements, Challenges and Global Implications.* Nutrients, 2017. **9**(4).
104. Bath, S.C. and M.P. Rayman, *Iodine deficiency in the U.K.: an overlooked cause of impaired neurodevelopment?* Proc Nutr Soc, 2013. **72**(2): p. 226-35.
105. Leung, A.M., *What's the best measure of population iodine status? It's not a simple answer.* Am J Clin Nutr, 2019. **110**(4): p. 797-798.
106. WHO, *Salt reduction and iodine fortification strategies in public health: report of a joint technical meeting convened by the World Health Organization and The George Institute for Global Health in collaboration with the International Council for the Control of Iodine Deficiency Disorders Global Network*. 2013: Sydney, Australia.
107. National Health and Medical Research Council, *NHMRC Public Statement: Iodine Supplementation for Pregnant and Breastfeeding Women*. 2010.
108. National Health and Medical Research Council, *Literature Review: Iodine Supplementation During Pregnancy and Lactation*. 2009, National Health and Medical Research Council.
109. Australian Population Health Development Principal Committee, *The prevalence and severity of iodine deficiency in Australia*. 2007, Australian Health Ministers Advisory Committee.
110. McDonnell, C.M., M. Harris, and M.R. Zacharin, *Iodine deficiency and goitre in schoolchildren in Melbourne, 2001.* Med J Aust, 2003. **178**(4): p. 159-62.
111. Guttikonda, K., et al., *Iodine deficiency in urban primary school children: a cross-sectional analysis.* Med J Aust, 2003. **179**(7): p. 346-8.
112. Li, M., et al., *Re-emergence of iodine deficiency in Australia.* Asia Pacific Journal of Clinical Nutrition, 2001. **10**(3): p. 200-203.
113. Eastman, C.J., *Where has all our iodine gone?* Med J Aust, 1999. **171**(9): p. 455-6.
114. Hynes, K.L., et al., *Persistent iodine deficiency in a cohort of Tasmanian school children: associations with socio-economic status, geographical location and dietary factors.* Aust N Z J Public Health, 2004. **28**(5): p. 476-81.
115. Kazi, T.G., et al., *Evaluation of iodine, iron, and selenium in biological samples of thyroid mother and their newly born babies.* Early Hum Dev, 2010. **86**(10): p. 649-55.
116. Wilcken, B.M. and V.C. Wiley, *Increased iodine deficiency in Victoria, Australia: analysis of neonatal thyroid-stimulating hormone data, 2001 to 2006.* Med J Aust, 2011. **194**(4): p. 209-10.
117. McElduff, A., et al., *Neonatal thyroid-stimulating hormone concentrations in northern Sydney: further indications of mild iodine deficiency?* Med J Aust, 2002. **176**(7): p. 317-20.
118. Australian Bureau of Statistics, *Australian Health Survey: Biomedical Results for Nutrients, 2011-2012*. 2013, Australian Bureau of Statistics cambera.
119. Charlton, K.E., et al., *Improvement in iodine status of pregnant Australian women 3years after introduction of a mandatory iodine fortification programme.* Preventive Medicine, 2013. **57**(1): p. 26-30.
120. Li, M., et al., *Re-emergence of iodine deficiency in Australia.* Asia Pac J Clin Nutr, 2001. **10**(3): p. 200-3.

121. Charlton, K.E., et al., *Suboptimal iodine status of Australian pregnant women reflects poor knowledge and practices related to iodine nutrition.* Nutrition, 2010. **26**(10): p. 963-8.
122. Burgess, J.R., et al., *A case for universal salt iodisation to correct iodine deficiency in pregnancy: another salutary lesson from Tasmania.* Med J Aust, 2007. **186**(11): p. 574-6.
123. Hamrosi, M.A., E.M. Wallace, and M.D. Riley, *Iodine status in pregnant women living in Melbourne differs by ethnic group.* Asia Pac J Clin Nutr, 2005. **14**(1): p. 27-31.
124. Mackerras, D.E., G.R. Singh, and C.J. Eastman, *Iodine status of Aboriginal teenagers in the Darwin region before mandatory iodine fortification of bread.* Med J Aust, 2011. **194**(3): p. 126-30.
125. Australian Institute of Health and Welfare, *Monitoring the Health Impacts of Mandatory Folic Acid and Iodine Fortification*. 2016, Australian Institute of Health and Welfare: Canberra, Australia.
126. Singh, G.R., et al., *Iodine status of Indigenous and non-Indigenous young adults in the Top End, before and after mandatory fortification.* Med J Aust, 2019. **210**(3): p. 121-125.
127. Hynes, K.L., et al., *Women Remain at Risk of Iodine Deficiency during Pregnancy: The Importance of Iodine Supplementation before Conception and Throughout Gestation.* Nutrients, 2019. **11**(1): p. 172.
128. Charlton, K.E., et al., *Improvement in iodine status of pregnant Australian women 3 years after introduction of a mandatory iodine fortification programme.* Prev Med, 2013. **57**(1): p. 26-30.
129. Blumenthal, N., K. Byth, and C.J. Eastman, *Iodine Intake and Thyroid Function in Pregnant Women in a Private Clinical Practice in Northwestern Sydney before Mandatory Fortification of Bread with Iodised Salt.* J Thyroid Res, 2012. **2012**: p. 798963.
130. Rahman, A., et al., *Urinary iodine deficiency in Gippsland pregnant women: the failure of bread fortification?* Med J Aust, 2011. **194**(5): p. 240-3.
131. Australian Health Survey. Biomedical Results for Nutrients. 2011-2012 [cited 2019 01 February]; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.006Chapter1202011-12.
132. Shi, X., et al., *Optimal and safe upper limits of iodine intake for early pregnancy in iodine-sufficient regions: a cross-sectional study of 7190 pregnant women in China.* J Clin Endocrinol Metab, 2015. **100**(4): p. 1630-8.
133. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
134. Wassie, M.M., L.G. Smithers, and S.J. Zhou, *Association of newborn thyroid stimulating hormone (TSH) concentration with child growth and neurodevelopment: a systematic review and meta-analysis*. 2020, PROSPERO
135. Chiamolera, M.I. and F.E. Wondisford, *Minireview: Thyrotropin-releasing hormone and the thyroid hormone feedback mechanism.* Endocrinology, 2009. **150**(3): p. 1091-6.
136. Smallridge, R.C. and P.W. Ladenson, *Hypothyroidism in pregnancy: consequences to neonatal health.* J Clin Endocrinol Metab, 2001. **86**(6): p. 2349-53.
137. Nazeri, P., et al., *Neonatal thyrotropin concentration and iodine nutrition status of mothers: a systematic review and meta-analysis.* Am J Clin Nutr, 2016. **104**(6): p. 1628-1638.
138. LaFranchi, S., *Congenital hypothyroidism: etiologies, diagnosis, and management.* Thyroid, 1999. **9**(7): p. 735-40.

139. Liu, Y., et al., *The Association Between Maternal Subclinical Hypothyroidism and Growth, Development, and Childhood Intelligence: A Meta-analysis.* J Clin Res Pediatr Endocrinol, 2018. **10**(2): p. 153-161.
140. European Society for Paediatric Endocrinology, *Revised guidelines for neonatal screening programmes for primary congenital hypothyroidism. Working Group on Neonatal Screening of the European Society for Paediatric Endocrinology.* Horm Res, 1999. **52**(1): p. 49-52.
141. Korzeniewski, S.J., et al., *Are preterm newborns who have relative hyperthyrotropinemia at increased risk of brain damage?* J Pediatr Endocrinol Metab, 2014. **27**(11-12): p. 1077-88.
142. Cuestas, E., M.I. Gaido, and R.H. Capra, *Transient neonatal hyperthyrotropinemia is a risk factor for developing persistent hyperthyrotropinemia in childhood with repercussion on developmental status.* Eur J Endocrinol, 2015. **172**(4): p. 483-90.
143. Lain, S., et al., *Are lower TSH cutoffs in neonatal screening for congenital hypothyroidism warranted?* Eur J Endocrinol, 2017. **177**(5): p. D1-D12.
144. Delange, F., *Neonatal thyroid screening as a monitoring tool for the control of iodine deficiency.* Acta Paediatr Suppl, 1999. **88**(432): p. 21-4.
145. Bougma, K., et al., *Iodine and mental development of children 5 years old and under: a systematic review and meta-analysis.* Nutrients, 2013. **5**(4): p. 1384-416.
146. Sterne, J.A., et al., *ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions.* BMJ, 2016. **355**: p. i4919.
147. Holger Schünemann, et al., *GRADE Handbook*. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. 2013: The GRADE Working Group.
148. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* BMJ, 2003. **327**(7414): p. 557-60.
149. Hedges, L.V., *Distribution Theory for Glass's Estimator of Effect size and Related Estimators.* Journal of Educational Statistics, 2016. **6**(2): p. 107-128.
150. Cumming, G., *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. 2013: Routledge.
151. Altman, D.G. and J.M. Bland, *How to obtain the confidence interval from a P value.* BMJ, 2011. **343**: p. d2090.
152. Cohen, J., *Statistical power analysis for the behavioral sciences*. 2013: Routledge.
153. Zhou, S.J., et al., *Effect of iron supplementation during pregnancy on the intelligence quotient and behavior of children at 4 y of age: long-term follow-up of a randomized controlled trial.* Am J Clin Nutr, 2006. **83**(5): p. 1112-7.
154. Walter, T., et al., *Iron deficiency anemia: adverse effects on infant psychomotor development.* Pediatrics, 1989. **84**(1): p. 7-17.
155. Gaudino, R., et al., *Proportion of various types of thyroid disorders among newborns with congenital hypothyroidism and normally located gland: a regional cohort study.* Clin Endocrinol (Oxf), 2005. **62**(4): p. 444-8.
156. WHO, *Iodine status worldwide : WHO Global Database on Iodine Deficiency*, Bruno de Benoist, et al., Editors. 2004: Geneva.
157. Bruno de Benoist, et al., *Iodine deficiency in 2007: Global progress since 2003.* Food and Nutrition Bulletin, vol. 29, no. 3 © 2008, The United Nations University., 2008. **29**(3).
158. Iodine Global Network, *Global Scorecard of Iodine Nutrition in 2017 in the general population and in pregnant women (PW)*. 2017, Iodine Global Network Zurich, Switzerland.
159. Iodine Global Network, *Global Iodine Nutrition Scorecard* 2015, Iodine Global Network: Zurich, Switzerland.

160. Murcia, M., et al., *Effect of iodine supplementation during pregnancy on infant neurodevelopment at 1 year of age.* Am J Epidemiol, 2011. **173**(7): p. 804-12.
161. Trumpff, C., et al., *Thyroid-Stimulating Hormone (TSH) Concentration at Birth in Belgian Neonates and Cognitive Development at Preschool Age.* Nutrients, 2015. **7**(11): p. 9018-32.
162. Trumpff, C., et al., *Neonatal thyroid-stimulating hormone concentration and psychomotor development at preschool age.* Arch Dis Child, 2016. **101**(12): p. 1100-1106.
163. Calaciura, F., et al., *Childhood IQ measurements in infants with transient congenital hypothyroidism.* Clin Endocrinol (Oxf), 1995. **43**(4): p. 473-7.
164. Azizi, F., et al., *Effects of transient neonatal hyperthyrotropinemia on intellectual quotient and psychomotor performance.* Int J Vitam Nutr Res, 2001. **71**(1): p. 70-3.
165. Choudhury, N. and K.S. Gorman, *Subclinical prenatal iodine deficiency negatively affects infant development in Northern China.* J Nutr, 2003. **133**(10): p. 3162-5.
166. Freire, C., et al., *Newborn TSH concentration and its association with cognitive development in healthy boys.* Eur J Endocrinol, 2010. **163**(6): p. 901-9.
167. Riano Galan, I., et al., *Psycho-intellectual development of 3 year-old children with early gestational iodine deficiency.* J Pediatr Endocrinol Metab, 2005. **18 Suppl 1**: p. 1265-72.
168. Trumpff, C., et al., *No Association between Elevated Thyroid-Stimulating Hormone at Birth and Parent-Reported Problem Behavior at Preschool Age.* Front Endocrinol (Lausanne), 2016. **7**: p. 161.
169. Belcari, F., et al., *Thyroid-stimulating hormone levels in the first days of life and perinatal factors associated with sub-optimal neuromotor outcome in pre-term infants.* J Endocrinol Invest, 2011. **34**(10): p. e308-13.
170. Costeira, M.J., et al., *Psychomotor development of children from an iodine-deficient region.* J Pediatr, 2011. **159**(3): p. 447-53.
171. Leviton, A., et al., *Systemic inflammation on postnatal days 21 and 28 and indicators of brain dysfunction 2years later among children born before the 28th week of gestation.* Early Hum Dev, 2016. **93**: p. 25-32.
172. Kuban, K.C., et al., *Circulating Inflammatory-Associated Proteins in the First Month of Life and Cognitive Impairment at Age 10 Years in Children Born Extremely Preterm.* J Pediatr, 2017. **180**: p. 116-123 e1.
173. Shields, B.M., et al., *Fetal thyroid hormone level at birth is associated with fetal growth.* J Clin Endocrinol Metab, 2011. **96**(6): p. E934-8.
174. Zhou, S.J., et al., *Effect of iodine supplementation in pregnancy on child development and other clinical outcomes: a systematic review of randomized controlled trials.* Am J Clin Nutr, 2013. **98**(5): p. 1241-54.
175. Levie, D., et al., *Association of Maternal Iodine Status With Child IQ: A Meta-Analysis of Individual Participant Data.* J Clin Endocrinol Metab, 2019. **104**(12): p. 5957-5967.
176. Ghassabian, A., et al., *Maternal urinary iodine concentration in pregnancy and children's cognition: results from a population-based birth cohort in an iodine-sufficient area.* BMJ Open, 2014. **4**(6): p. e005520.
177. Velasco, I., et al., *Effect of iodine prophylaxis during pregnancy on neurocognitive development of children during the first two years of life.* J Clin Endocrinol Metab, 2009. **94**(9): p. 3234-41.
178. WHO, ICCIDD, and UNICEF, *Indicators for assessing iodine deficiency disorders and their control through salt iodization*. 1994, Geneva: World Health Organization.
179. Evans, C., et al., *Neonatal blood TSH concentration in Wales (UK): an indicator of iodine sufficiency.* Clin Endocrinol (Oxf), 2014. **81**(4): p. 606-9.

180. Vanderpump, M.P., et al., *Iodine status of UK schoolgirls: a cross-sectional survey.* Lancet, 2011. **377**(9782): p. 2007-12.
181. Al-Hosani, H., et al., *Prevalence of iodine deficiency disorders in the United Arab Emirates measured by raised TSH levels.* East Mediterr Health J, 2003. **9**(1-2): p. 123-30.
182. Copeland, D.L., et al., *Comparison of neonatal thyroid-stimulating hormone levels and indicators of iodine deficiency in school children.* Public Health Nutr, 2002. **5**(1): p. 81-7.
183. Gruneiro-Papendieck, L., et al., *Neonatal TSH levels as an index of iodine sufficiency: differences related to time of screening sampling and methodology.* Horm Res, 2004. **62**(6): p. 272-6.
184. Sullivan, K.M., et al., *Use of thyroid stimulating hormone testing in newborns to identify iodine deficiency*. J Nutr, 1997. **127**(1): p. 55-8.
185. Ristic-Medic, D., et al., *Methods of assessment of iodine status in humans: a systematic review.* Am J Clin Nutr, 2009. **89**(6): p. 2052S-2069S.
186. Vandevijvere, S., et al., *Neonatal thyroid-stimulating hormone concentrations in Belgium: a useful indicator for detecting mild iodine deficiency?* PLoS One, 2012. **7**(10): p. e47770.
187. Gyurjyan, R.H., et al., *Newborn thyrotropin screening confirms iodine deficiency in Latvia.* Eur J Clin Nutr, 2006. **60**(5): p. 688-90.
188. Wachter, W., et al., *Iodine deficiency, hypothyroidism, and endemic goitre in southern Tanzania. A survey showing the positive effects of iodised oil injections by TSH determination in dried blood spots.* J Epidemiol Community Health, 1985. **39**(3): p. 263-70.
189. *Evidence analysis manual: Steps in the ADA evidence analysis process* A.D.A. Research and Strategic Business Development, Editor. 2011, Americal Dietetic Association.
190. Lalkhen, A.G. and A. McCluskey, *Clinical tests: sensitivity and specificity.* Bja Education, 2008. **8**(6): p. 221-223.
191. Youden, W.J., *Index for rating diagnostic tests.* Cancer, 1950. **3**(1): p. 32-5.
192. Kumar, R. and A. Indrayan, *Receiver operating characteristic (ROC) curve for medical researchers.* Indian Pediatr, 2011. **48**(4): p. 277-87.
193. Greiner, M., D. Pfeiffer, and R.D. Smith, *Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests.* Prev Vet Med, 2000. **45**(1-2): p. 23-41.
194. Zhu, W., N. Zeng, and N. Wang, *Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations*, in *Health Care and Life Sciences*. 2010, Northeast SAS User Group proceedings, Section of Health Care and Life Sciences: Baltimore, Maryland.
195. Costante, G., et al., *The statistical analysis of neonatal TSH results from congenital hypothyroidism screening programs provides a useful tool for the characterization of moderate iodine deficiency regions.* J Endocrinol Invest, 1997. **20**(5): p. 251-6.
196. Costante, G., et al., *Iodine deficiency in Calabria: characterization of endemic goiter and analysis of different indicators of iodine status region-wide.* J Endocrinol Invest, 2002. **25**(3): p. 201-7.
197. Dalili, S., et al., *The prevalence of iodine deficiency disorder in two different populations in northern province of Iran: a comparison using different indicators recommended by WHO.* Acta Med Iran, 2012. **50**(12): p. 822-6.
198. Fuse, Y., et al., *Epidemiological survey of thyroid volume and iodine intake in schoolchildren, postpartum women and neonates living in Ulaan Baatar.* Clin Endocrinol (Oxf), 2003. **59**(3): p. 298-306.

199. Konrade, I., et al., *A cross-sectional survey of urinary iodine status in Latvia.* Medicina (Kaunas), 2014. **50**(2): p. 124-9.
200. Rahman, A., et al., *Increased iodine deficiency in Victoria, Australia: analysis of neonatal thyroid-stimulating hormone data, 2001 to 2006.* Med J Aust, 2010. **193**(9): p. 503-5.
201. Svinaryov, M. and V. Aranovich, *Iodine deficiency disorders in the Saratov province in Russia.* J Endocrinol Invest, 2003. **26**(2 Suppl): p. 16-9.
202. Vandevijvere, S., et al., *Fortification of bread with iodized salt corrected iodine deficiency in school-aged children, but not in their mothers: a national cross-sectional survey in Belgium.* Thyroid, 2012. **22**(10): p. 1046-53.
203. Yamada, C., et al., *Current status of iodine deficiency in Mongolia in 1998-1999.* Asia Pac J Public Health, 2000. **12**(2): p. 79-84.
204. Selga, G., M. Sauka, and G.Gerasimov, *Status of Iodine Deficiency in Latvia Reconsidered: Results of Nation-wide Survey of 587 Schoolchildren in the Year 2000.*, in *IDD newsletter* 2000, International Council for the Control of Iodine Deficiency Disorders (ICCIDD).
205. Mehran, L., et al., *The Impact of Iodine Status on the Recall Rate of the Screening Program for Congenital Hypothyroidism: Findings from Two National Studies in Iran.* Nutrients, 2017. **9**(11).
206. Zimmermann, M.B., et al., *Thyroid size and goiter prevalence after introduction of iodized salt: a 5-y prospective study in schoolchildren in Cote d'Ivoire.* Am J Clin Nutr, 2003. **77**(3): p. 663-7.
207. Gaitan, E., *Goitrogens.* Baillieres Clin Endocrinol Metab, 1988. **2**(3): p. 683-702.
208. Zimmermann, M.B., et al., *High thyroid volume in children with excess dietary iodine intakes.* Am J Clin Nutr, 2005. **81**(4): p. 840-4.
209. Momotani, N., et al., *Effects of iodine on thyroid status of fetus versus mother in treatment of Graves' disease complicated by pregnancy.* J Clin Endocrinol Metab, 1992. **75**(3): p. 738-44.
210. Vejbjerg, P., et al., *Estimation of iodine intake from various urinary iodine measurements in population studies.* Thyroid, 2009. **19**(11): p. 1281-6.
211. Peters, C., et al., *Defining the Newborn Blood Spot Screening Reference Interval for TSH: Impact of Ethnicity.* J Clin Endocrinol Metab, 2016. **101**(9): p. 3445-9.
212. Hynes, K.L., et al., *Mild iodine deficiency during pregnancy is associated with reduced educational outcomes in the offspring: 9-year follow-up of the gestational iodine cohort.* J Clin Endocrinol Metab, 2013. **98**(5): p. 1954-62.
213. Liu, P., et al., *Should urinary iodine concentrations of school-aged children continue to be used as proxy for different populations? Analysis of data from Chinese national surveys.* Br J Nutr, 2016. **116**(6): p. 1068-76.
214. Vought, R.L., et al., *Reliability of Estimates of Serum Inorganic Iodine and Daily Fecal and Urinary Iodine Excretion from Single Casual Specimens.* J Clin Endocrinol Metab, 1963. **23**: p. 1218-28.
215. Wong, G.W., et al., *Childhood goitre and urinary iodine excretion in Hong Kong.* Eur J Pediatr, 1998. **157**(1): p. 8-12.
216. Manz, F., et al., *Water balance throughout the adult life span in a German population.* Br J Nutr, 2012. **107**(11): p. 1673-81.
217. Rasmussen, L.B., L. Ovesen, and E. Christiansen, *Day-to-day and within-day variation in urinary iodine excretion.* Eur J Clin Nutr, 1999. **53**(5): p. 401-7.
218. Frey, H.M., B. Rosenlund, and J.P. Torgersen, *Value of single urine specimens in estimation of 24 hour urine iodine excretion.* Acta Endocrinol (Copenh), 1973. **72**(2): p. 287-92.

219. Haddow, J.E., et al., *Urine iodine measurements, creatinine adjustment, and thyroid deficiency in an adult United States population.* J Clin Endocrinol Metab, 2007. **92**(3): p. 1019-22.
220. Ji, C., et al., *Systematic review of studies evaluating urinary iodine concentration as a predictor of 24-hour urinary iodine excretion for estimating population iodine intake.* Revista Panamericana De Salud Publica-Pan American Journal of Public Health, 2015. **38**(1): p. 73-81.
221. Willett, W.C., G.R. Howe, and L.H. Kushi, *Adjustment for total energy intake in epidemiologic studies.* Am J Clin Nutr, 1997. **65**(4 Suppl): p. 1220S-1228S; discussion 1229S-1231S.
222. Roberts, C., *Screening Tests to identify poor Outcomes in Pregnancy (STOP) Study*, in *Australian New Zealand Clinical Trials Registry*. 2014, Australian New zealand clinical trials registry
223. Amir A. Makhmudov and K.L. Caldwell, *The Challenge of Iodine Deficiency Disorder: A Decade of CDC's Ensuring the Quality of Urinary Iodine Procedures Program*, in *EQUIP 10 year Anniversary* 2011.
224. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
225. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.
226. Australian Bureau of Statistics, *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011* 2013: Canberra
227. Phillips, J.K., et al., *Examination of Prepregnancy and Pregnancy Urinary Protein Levels in Healthy Nulliparous Women.* Reprod Sci, 2017. **24**(3): p. 407-412.
228. Davison, J.M. and W. Dunlop, *Renal hemodynamics and tubular function normal human pregnancy.* Kidney Int, 1980. **18**(2): p. 152-61.
229. Wassie, M.M., et al., *Comparison of iodine status pre- and post-mandatory iodine fortification of bread in South Australia: a population study using newborn thyroid-stimulating hormone concentration as a marker.* Public Health Nutr, 2019. **22**(16): p. 3063-3072.
230. World Health Organization, I., UNICEF, *Indicators for assessing iodine deficiency disorders and their control through salt iodization*. 1994, Geneva : World Health Organization.
231. Human Genetics Society Of Australasia, *Newborn bloodspot screening Policy*. 2017: NSW, Australia.
232. Huynh, D., et al., *Comparison of breast-milk iodine concentration of lactating women in Australia pre and post mandatory iodine fortification.* Public Health Nutr, 2017. **20**(1): p. 12-17.
233. Soumerai, S.B., D. Starr, and S.R. Majumdar, *How Do You Know Which Health Care Effectiveness Research You Can Trust? A Guide to Study Design for the Perplexed.* Prev Chronic Dis, 2015. **12**: p. E101.
234. Metz, M.P., et al., *Newborn screening in South Australia: is it universal?* Med J Aust, 2003. **179**(8): p. 412-5.
235. Fisher, D.A. and A.H. Klein, *Thyroid development and disorders of thyroid function in the newborn.* N Engl J Med, 1981. **304**(12): p. 702-12.
236. Lee, S.Y., *Perinatal factors associated with neonatal thyroid-stimulating hormone in normal newborns.* Ann Pediatr Endocrinol Metab, 2016. **21**(4): p. 206-211.
237. Rashmi, et al., *Effect of perinatal factors on cord blood thyroid stimulating hormone levels.* J Pediatr Endocrinol Metab, 2007. **20**(1): p. 59-64.
238. Trumpff, C., et al., *Neonatal thyroid-stimulating hormone level is influenced by neonatal, maternal, and pregnancy factors.* Nutr Res, 2015. **35**(11): p. 975-81.

239. Dobbins, T.A., et al., *Australian national birthweight percentiles by sex and gestational age, 1998-2007.* Med J Aust, 2012. **197**(5): p. 291-4.
240. SA Health, *Pregnancy Outcome in South Australia, 2016*. 2018, Pregnancy Outcome Unit, Prevention and Population Health Branch, SA Health, Government of South Australia: Adelaide.
241. Bernal, J.L., S. Cummins, and A. Gasparrini, *Interrupted time series regression for the evaluation of public health interventions: a tutorial.* Int J Epidemiol, 2017. **46**(1): p. 348-355.
242. Spencer, C.A., *Assay of Thyroid Hormones and Related Substances,* in *Endotext*, K.R. Feingold, et al., Editors. 2000: South Dartmouth (MA).
243. Hutchings, N., et al., *Neonatal thyrotropin (TSH) screening as a tool for monitoring iodine nutrition in Armenia.* Eur J Clin Nutr, 2019. **73**(6): p. 905-909.
244. Li, M., et al., *Can even minimal news coverage influence consumer health-related behaviour? A case study of iodized salt sales, Australia.* Health Educ Res, 2008. **23**(3): p. 543-8.
245. Food Standards Australia New Zealand (FSANZ), *Monitoring the Australian population's intake of dietary iodine before and after mandatory fortification*. 2016.
246. Huynh, D., et al., *Iodine status of postpartum women and their infants in Australia after the introduction of mandatory iodine fortification.* Br J Nutr, 2017. **117**(12): p. 1656-1662.
247. Food Standards Australia New Zealand, *Australia New Zealand Food Standards,* in *Australia New Zealand Food Standards Code - Standard 2.10.2 - Salt and salt products*. 2009: Canberra
248. Nishiyama, S., et al., *Transient hypothyroidism or persistent hyperthyrotropinemia in neonates born to mothers with excessive iodine intake.* Thyroid, 2004. **14**(12): p. 1077-83.
249. Ruppert, F., et al., *Thyrotropin and prolactin response to ambient temperature in newborn infants.* Acta Paediatr Acad Sci Hung, 1982. **23**(2): p. 189-94.
250. Fregly, M.J., *Activity of the hypothalamic-pituitary-thyroid axis during exposure to cold.* Pharmacol Ther, 1989. **41**(1-2): p. 85-142.
251. Glinoer, D., *The importance of iodine nutrition during pregnancy.* Public Health Nutr, 2007. **10**(12A): p. 1542-6.
252. Zimmermann, M.B., *The effects of iodine deficiency in pregnancy and infancy.* Paediatr Perinat Epidemiol, 2012. **26 Suppl 1**: p. 108-17.
253. Mackerras, D.E. and C.J. Eastman, *Estimating the iodine supplementation level to recommend for pregnant and breastfeeding women in Australia.* Med J Aust, 2012. **197**(4): p. 238-42.
254. Darmon, N. and A. Drewnowski, *Does social class predict diet quality?* Am J Clin Nutr, 2008. **87**(5): p. 1107-17.
255. El-mani, S., et al., *Limited knowledge about folic acid and iodine nutrition in pregnant women reflected in supplementation practices.* Nutrition & Dietetics, 2014. **71**(4): p. 236-244.
256. Malek, L., et al., *Poor adherence to folic acid and iodine supplement recommendations in preconception and pregnancy: a cross-sectional analysis.* Aust N Z J Public Health, 2016. **40**(5): p. 424-429.
257. Kenny, L.C., et al., *Early pregnancy prediction of preeclampsia in nulliparous women, combining clinical risk and biomarkers: the Screening for Pregnancy Endpoints (SCOPE) international cohort study.* Hypertension, 2014. **64**(3): p. 644-52.
258. Statistics, A.B.o., *Adelaide's advantaged and disadvantaged subrubs* in *SA Stats*. 2008: Canberra.
259. Fall, C.H.D., et al., *Association between maternal age at childbirth and child and adult outcomes in the offspring: a prospective study in five low-income and middle-*

income countries (COHORTS collaboration). The Lancet Global Health, 2015. **3**(7): p. e366-e377.
260. Australian Bureau of Statistics, *2006 Census QuickStats*. 2006, ABS: Canberra.
261. Australian Bureau of Statistics, *2016 Census QuickStats*. 2016, ABS: Canberra.
262. SA Health, *Pregnancy Outcome in South Australia 2016* 2018, Pregnancy Outcome Unit, Prevention and Population Health Branch, SA Health, Government of South Australia: Adelaide.
263. Charlton, K., Y. Probst, and G. Kiene, *Dietary Iodine Intake of the Australian Population after Introduction of a Mandatory Iodine Fortification Programme.* Nutrients, 2016. **8**(11): p. 701.
264. Zimmermann, M.B., *Iodine deficiency in pregnancy and the effects of maternal iodine supplementation on the offspring: a review.* Am J Clin Nutr, 2009. **89**(2): p. 668S-72S.
265. Farebrother, J., et al., *Effects of Iodized Salt and Iodine Supplements on Prenatal and Postnatal Growth: A Systematic Review.* Adv Nutr, 2018. **9**(3): p. 219-237.
266. Abel, M.H., et al., *Suboptimal Maternal Iodine Intake Is Associated with Impaired Child Neurodevelopment at 3 Years of Age in the Norwegian Mother and Child Cohort Study.* J Nutr, 2017. **147**(7): p. 1314-1324.
267. Secretariat, W.H.O., et al., *Prevention and control of iodine deficiency in pregnant and lactating women and in children less than 2-years-old: conclusions and recommendations of the Technical Consultation.* Public Health Nutr, 2007. **10**(12A): p. 1606-11.
268. Gowachirapant, S., et al., *Effect of iodine supplementation in pregnant women on child neurodevelopment: a randomised, double-blind, placebo-controlled trial.* Lancet Diabetes Endocrinol, 2017. **5**(11): p. 853-863.
269. Zhou, S.J., et al., *The effect of iodine supplementation in pregnancy on early childhood neurodevelopment and clinical outcomes: results of an aborted randomised placebo-controlled trial.* Trials, 2015. **16**: p. 563.
270. Markou, K., et al., *Iodine-Induced hypothyroidism.* Thyroid, 2001. **11**(5): p. 501-10.
271. Chen, W., et al., *Neonatal thyroid function born to mothers living with long-term excessive iodine intake from drinking water.* Clin Endocrinol (Oxf), 2015. **83**(3): p. 399-404.
272. Sait, H., et al., *Association Between Neonatal Thyroid Stimulating Hormone Status and Maternal Urinary Iodine Status.* Indian Pediatr, 2019. **56**(6): p. 472-475.
273. Trumpff, C., et al., *No Association between Elevated Thyroid-Stimulating Hormone at Birth and Parent-Reported Problem Behavior at Preschool Age.* Front Endocrinol (Lausanne), 2016. **7**(DEC): p. 161.
274. Makrides, M., et al., *Effect of DHA supplementation during pregnancy on maternal depression and neurodevelopment of young children: a randomized controlled trial.* JAMA, 2010. **304**(15): p. 1675-83.
275. Bayley, N., *Bayley scales of infant and toddler development. Third edition*. 2006.
276. WHO, *Physical status: The use and interpretation of anthropometry: report of a WHO expert committee* 1995, World Health Organization: Geneva
277. WHO, *World Health Organization: Child growth standards* in *Anthro and macro*. 2011, World Health Organization.
278. South Australian Neonatal Screening Center, *Screening tests for your new baby: helping to ensure the health of your child*. 2010, Women's and Children's Hospital: Adelaide, South Australia.
279. Frankenburg, W.K. and C.E. Coons, *Home Screening Questionnaire: its validity in assessing home environment.* J Pediatr, 1986. **108**(4): p. 624-6.
280. Iltus, S., *Paper commissioned for the EFA Global Monitoring Report 2007, Strong foundations: early childhood care and education*, in *Significance of home environments as proxy indicators for early childhood care and education*. 2006.

281. Sterne, J.A., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* BMJ, 2009. **338**: p. b2393.
282. Zou, G., *A modified poisson regression approach to prospective studies with binary data.* Am J Epidemiol, 2004. **159**(7): p. 702-6.
283. Makrides, M., *DHA supplementation during the perinatal period and neurodevelopment: Do some babies benefit more than others?* Prostaglandins Leukot Essent Fatty Acids, 2013. **88**(1): p. 87-90.
284. Hibbs, A.M., et al., *Accounting for multiple births in neonatal and perinatal trials: systematic review and case study.* J Pediatr, 2010. **156**(2): p. 202-8.
285. Amrhein, V., S. Greenland, and B. McShane, *Scientists rise up against statistical significance.* Nature, 2019. **567**(7748): p. 305-307.
286. Greenland, S., et al., *Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations.* Eur J Epidemiol, 2016. **31**(4): p. 337-50.
287. Wasserstein, R.L. and N.A. Lazar, *The ASA's Statement on p-Values: Context, Process, and Purpose.* American Statistician, 2016. **70**(2): p. 129-131.
288. Wasserstein, R.L., A.L. Schirm, and N.A. Lazar, *Moving to a World Beyond "p < 0.05".* The American Statistician, 2019. **73**(sup1): p. 1-19.
289. Harris, R.J., et al., *metan: fixed- and random-effects meta-analysis.* Stata Journal, 2008. **8**(1): p. 3-28.
290. Parveen, S., et al., *Iodized salt induced thyrotoxicosis: Bangladesh perspective.* Mymensingh Med J, 2009. **18**(2): p. 165-8.
291. Glinoer, D., *The regulation of thyroid function during normal pregnancy: importance of the iodine nutrition status.* Best Pract Res Clin Endocrinol Metab, 2004. **18**(2): p. 133-52.
292. Chan, S.Y., E. Vasilopoulou, and M.D. Kilby, *The role of the placenta in thyroid hormone delivery to the fetus.* Nat Clin Pract Endocrinol Metab, 2009. **5**(1): p. 45-54.
293. Dietrich, K.N., et al., *Principles and practices of neurodevelopmental assessment in children: lessons learned from the Centers for Children's Environmental Health and Disease Prevention Research.* Environ Health Perspect, 2005. **113**(10): p. 1437-46.
294. Chen, W., et al., *24-Hour Urine Samples Are More Reproducible Than Spot Urine Samples for Evaluation of Iodine Status in School-Age Children.* J Nutr, 2016. **146**(1): p. 142-6.
295. Beckford, K., *Iodine intakes and food sources of iodine in Victorian Schoolchildren* in *School of Exercise and Nutrition Sciences*. 2018, Deakin University p. 182.
296. Beckford, K., et al., *Iodine Intakes of Victorian Schoolchildren Measured Using 24-h Urinary Iodine Excretion.* Nutrients, 2017. **9**(9): p. 961.
297. Glinoer, D., *The regulation of thyroid function in pregnancy: pathways of endocrine adaptation from physiology to pathology.* Endocr Rev, 1997. **18**(3): p. 404-33.
298. Delange, F., *Iodine requirements during pregnancy, lactation and the neonatal period and indicators of optimal iodine nutrition.* Public Health Nutr, 2007. **10**(12A): p. 1571-80; discussion 1581-3.
299. Andersen, S.L. and P. Laurberg, *Iodine Supplementation in Pregnancy and the Dilemma of Ambiguous Recommendations.* Eur Thyroid J, 2016. **5**(1): p. 35-43.
300. Rebagliato, M., et al., *Iodine intake and maternal thyroid function during pregnancy.* Epidemiology, 2010. **21**(1): p. 62-9.
301. Trieu, K., et al., *Salt Reduction Initiatives around the World - A Systematic Review of Progress towards the Global Target.* PLoS One, 2015. **10**(7): p. e0130247.
302. Webster, J., et al., *The development of a national salt reduction strategy for Australia.* Asia Pac J Clin Nutr, 2009. **18**(3): p. 303-9.
303. National Heart Foundation of Australia, *Report on the Evaluation of the nine Food Categories for which reformulation targets were set under the Food and Health Dialogue* 2016.

304. Webster, J., et al., *Reducing dietary salt intake and preventing iodine deficiency: towards a common public health agenda.* Med J Aust, 2014. **201**(9): p. 507-8.

APPENDIXES

Appendix I Supplemental Tables

Supplemental Table 2. 1 Search strategy in different databases

PUBMED:			
Iodine [mh:noexp] OR iodides [TIAB] OR iodates [TIAB] OR Iodine status [TW] OR Urinary iodi* [TW] OR iodine nutritional status [TW] OR iodide status [TW] OR goiter [Mesh] OR Iodine Deficienc*[TW] OR	Thyroid function [tiab] OR Thyroid stimulating hormone [TW] OR TSH [TIAB] OR Thyrotropin [MH:noexp] OR Thyrotropin [TIAB] OR Thyrotrophin [TW] OR Thyreotropin [TW] OR Hypothyroidism [TIAB] OR hypothyroxin* [TW] OR thyroid hormones [TIAB] OR thyroxine [TW] OR Triiodothyronine [TW]	child development [mh:noexp] OR cognition [mh:noexp] OR cognition [TIAB] OR intelligence [TIAB] OR motor skills [TIAB] OR psychomotor performance [TIAB] OR intelligence quotient [TIAB] OR IQ [TIAB] OR neurological development [TIAB] OR cognition disorders [TIAB] OR cognitive function [TIAB] OR neural development [TIAB] OR intellectual disability [TW] OR post-natal development [TIAB] OR memory [TIAB] OR executive function [TIAB] OR mental process [TIAB]	physical development [TIAB] OR growth [TIAB] OR Head circumference [TIAB] OR Length [TIAB] OR Height [TIAB] OR stature [TIAB] OR Weight [TIAB] OR Mid upper arm circumference [TIAB] OR MUAC [TIAB] OR Anthropometric status [TIAB] OR Z-score [TW] OR Body mass index [TW] OR BMI [TW] OR Underweight [TIAB] OR Obesity [TIAB] OR Stunting [TIAB] OR wasting [TIAB] OR Child growth rate [TIAB]

		OR learning [TIAB] OR child behaviour [Mesh] OR brain development [TIAB] OR language [TIAB] OR language development [mh:noexp] OR child language [mh:noexp] OR child language [TIAB] OR language development [TIAB]	
(((Iodine [mh:noexp] OR iodides [TIAB] OR iodates [TIAB] OR Iodine status [TW] OR Urinary iodi* [TW] OR iodine nutritional status [TW] OR iodide status [TW] OR goiter [Mesh] OR Iodine Deficienc*[TW] OR Thyroid function [tiab] OR Thyroid stimulating hormone [TW] OR TSH [TIAB] OR Thyrotropin [MH:noexp] OR Thyrotropin [TIAB] OR Thyrotrophin [TW] OR Thyreotropin [TW] OR Hypothyroidism [TIAB] OR hypothyroxin* [TW] OR thyroid hormones [TIAB] OR thyroxine [TW] OR Triiodothyronine [TW])) AND (child development [mh:noexp] OR cognition [mh:noexp] OR cognition [TIAB] OR intelligence[TIAB] OR motor skills [TIAB] OR psychomotor performance [TIAB] OR intelligence quotient [TIAB] OR IQ [TIAB] OR neurological development[TIAB] OR cognition disorders [TIAB] OR cognitive function [TIAB] OR neural development [TIAB] OR intellectual disability [TW] OR post-natal development [TIAB] OR memory [TIAB] OR executive function [TIAB] OR mental process [TIAB] OR learning [TIAB] OR child behaviour [Mesh] OR brain development [TIAB] OR language [TIAB] OR language development [mh:noexp] OR child language [mh:noexp] OR child language [TIAB] OR language development [TIAB] OR physical development [TIAB] OR growth [TIAB] OR Head circumference [TIAB] OR Length [TIAB] OR Height [TIAB] OR stature [TIAB] OR Weight [TIAB] OR Mid upper arm circumference [TIAB] OR MUAC [TIAB] OR Anthropometric status [TIAB] OR Z-score [TW] OR Body mass index [TW] OR BMI [TW] OR Underweight [TIAB] OR Obesity [TIAB] OR Stunting [TIAB] OR wasting [TIAB] OR Child growth rate [TIAB])) AND (Infant [Mesh] OR Neonat*[TW] OR Newborn*[TW] OR Child*[TW] OR			

adolescent[Mesh] OR baby[TW] OR babies[TW] or toddler[TW] or preschool children[TW] OR School age [TW] children [TW] OR school child [TW] OR teenager [TW])			
Embase:			
Iodine OR iodides OR iodates OR 'Iodine status' OR 'Urinary iodi*' OR 'iodine nutritional status' OR 'iodide status' OR goiter OR 'Iodine Deficienc*' OR 'Thyroid function' OR 'Thyroid stimulating hormone' OR TSH OR Thyrotropin OR Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin* OR 'thyroid hormones' OR thyroxine OR Triiodothyronine	'child development' OR cognition OR intelligence OR 'motor skills' OR 'psychomotor performance' OR 'intelligence quotient' OR IQ OR 'neurological development' OR 'cognition disorders' OR 'cognitive function' OR 'neural development' OR 'intellectual disability' OR 'post-natal development' OR memory OR 'executive function' OR 'mental process' OR learning OR 'child behaviour' OR 'brain development' OR language OR 'language development' OR 'child language' OR 'language development' OR 'physical development' OR growth OR 'Head circumference' OR Length OR Height	Infant OR Neonat* OR Newborn* OR Child* OR adolescent OR baby OR babies or toddler or 'preschool children' OR 'School age children' OR 'school child' OR teenager	Humans

	OR stature OR Weight OR 'Mid upper arm circumference' OR MUAC OR 'Anthropometric status' OR 'Z-score' OR 'Body mass index' OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR 'Child growth rate'		
('thyroid stimulating hormone' OR thyrotropin OR tsh OR thyrotrophin OR thyreotropin OR hypothyroidism OR 'iodine deficienc*' OR Goiter OR 'Thyroid volume' OR hypothyroxin* OR 'iodine/deficiency' OR 'iodine status' OR 'urinary iodi*' OR 'iodine nutritional status' OR 'iodide status') AND (infant OR newborn OR neonat* OR newborn*) AND (child*) AND humans			
Scopus:			
A	B	C	D
Iodine OR iodides OR iodates OR "Iodine status" OR "Urinary iodi*" OR "iodine nutritional status" OR "iodide status" OR goiter OR "Iodine Deficienc*" OR "Thyroid function" OR "Thyroid stimulating hormone" OR TSH OR Thyrotropin OR Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin* OR "thyroid	"child development" OR cognition OR intelligence OR "motor skills" OR "psychomotor performance" OR "intelligence quotient" OR IQ OR "neurological development" OR "cognition disorders" OR "cognitive function" OR "neural development" OR "intellectual disability" OR "postnatal development" OR memory OR "executive function" OR "mental	Infant OR Neonat* OR Newborn* OR Child* OR adolescent OR baby OR babies or toddler or "preschool children" OR "School age children" OR "school child" OR teenager	Title

hormones" OR thyroxine OR Triiodothyronine	process" OR learning OR "child behaviour" OR "brain development" OR language OR "language development" OR "child language" OR "physical development" OR growth OR "Head circumference" OR Length OR Height OR stature OR Weight OR "Mid upper arm circumference" OR MUAC OR "Anthropometric status" OR "Z score" OR "Body mass index" OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR "Child growth rate"		
Title (A AND B AND C)			
CINAHL:			
Iodine OR iodides OR iodates OR "Iodine status" OR "Urinary iodi*" OR "iodine nutritional status" OR "iodide status" OR goiter OR "Iodine Deficienc*" OR "Thyroid function" OR "Thyroid stimulating hormone" OR TSH OR Thyrotropin OR	("child development" OR cognition OR intelligence OR "motor skills" OR "psychomotor performance" OR "intelligence quotient" OR IQ OR "neurological development" OR "cognition disorders" OR "cognitive function" OR "neural development" OR "intellectualdisability" OR "post-	(Infant OR Neonat* OR Newborn* OR Child* OR adolescent OR baby OR babies or toddler or "preschool children" OR "School age children" OR "school child" OR teenager)	

Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin* OR “thyroid hormones” OR thyroxine OR Triiodothyronine) AND	natal development” OR memory OR “executive function” OR “mental process” OR learning OR “child behaviour” OR “brain development” OR language OR “language development” OR “child language” OR “language development” OR “physical development” OR growth OR “Head circumference” OR Length OR Height OR stature OR Weight OR “Mid upper arm circumference” OR MUAC OR “Anthropometric status” OR “Z-score” OR “Body mass index” OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR “Child growth rate”)	
(Iodine OR iodides OR iodates OR “Iodine status” OR “Urinary iodi*” OR “iodine nutritional status” OR “iodide status” OR goiter OR “Iodine Deficienc*” OR “Thyroid function” OR “Thyroid stimulating hormone” OR TSH OR Thyrotropin OR Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin* OR “thyroid hormones” OR thyroxine OR Triiodothyronine) AND (“child development” OR cognition OR intelligence OR “motor skills” OR “psychomotor performance” OR “intelligence quotient” OR IQ OR “neurological development” OR “cognition disorders” OR “cognitive function” OR “neural development” OR “intellectualdisability” OR “post-natal development” OR memory OR “executive function” OR “mental process” OR learning OR “child behaviour” OR “brain development” OR language OR “language development” OR “child language” OR “language development” OR “physical development” OR		

growth OR “Head circumference” OR Length OR Height OR stature OR Weight OR “Mid upper arm circumference” OR MUAC OR “Anthropometric status” OR “Z-score” OR “Body mass index” OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR “Child growth rate”) AND (Infant OR Neonat* OR Newborn* OR Child* OR adolescent OR baby OR babies or toddler or “preschool children” OR “School age children” OR “school child” OR teenager)
PsycINFO:
(Iodine OR iodides OR iodates OR Iodine status OR Urinary iodine OR iodine nutritional status OR iodide status OR goiter OR Iodine Deficiency OR Thyroid function OR Thyroid stimulating hormone OR TSH OR Thyrotropin OR Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin OR thyroid hormones OR thyroxine OR Triiodothyronine) AND (child development OR cognition OR intelligence OR motor skills OR psychomotor performance OR intelligence quotient OR IQ OR neurological development OR cognition disorders OR cognitive function OR neural development OR intellectual disability OR post-natal development OR memory OR executive function OR mental process OR learning OR child behaviour OR brain development OR language OR language development OR child language OR language development OR physical development OR growth OR Head circumference OR Length OR Height OR stature OR Weight OR Mid upper arm circumference OR MUAC OR Anthropometric status OR Z-score OR Body mass index OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR Child growth rate) AND (Infant OR Neonate OR Newborn OR Child OR Children OR adolescent OR baby OR babies or toddler or preschool children OR School age children OR school child OR teenager)
Web of Science
(Title (Iodine OR iodides OR iodates OR Iodine status OR Urinary iodine OR iodine nutritional status OR iodide status OR goiter OR Iodine Deficiency OR Thyroid function OR Thyroid stimulating hormone OR TSH OR Thyrotropin OR Thyrotropin OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR hypothyroxin OR thyroid hormones OR thyroxine OR Triiodothyronine) AND (child development OR cognition OR intelligence OR motor skills OR psychomotor performance OR intelligence quotient OR IQ OR neurological development OR cognition disorders OR cognitive function OR neural development OR intellectual disability OR post-natal development OR memory OR executive function OR mental process OR learning OR child behaviour OR brain development OR language OR language development OR child language OR language development OR physical development OR growth OR Head circumference OR Length OR Height OR stature

OR Weight OR Mid upper arm circumference OR MUAC OR Anthropometric status OR Z-score OR Body mass index OR BMI OR Underweight OR Obesity OR Stunting OR wasting OR Child growth rate) AND (Infant OR Neonate OR Newborn OR Child OR Children OR adolescent OR baby OR babies or toddler or preschool children OR School age children OR school child OR teenager))

Supplemental Table 2. 2 The risk of bias in non-randomized studies of interventions (ROBINS-I) assessment tool for assessing quality of the studies[1]

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered		**Y / PY** / PN / N
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3.		NA / Y / PY / **PN / N** / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / **PN / N** / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / **PN / N** / NI
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / **PN / N** / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / **Y / PY** / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / **PN / N** / NI
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / **PN / N** / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI

Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4		**Y / PY** / PN / N / NI
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / **Y / PY** / PN / N / NI
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / **Y / PY** / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY / **PN / N** / NI
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NA / Y / PY / **PN / N** / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?		Y / PY / **PN / N** / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY / **PN / N** / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		**Y / PY** / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI

Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable
Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		**Y / PY** / PN / N / NI
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / **Y / PY** / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		Y / PY / **PN / N** / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY / **PN / N** / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY / **PN / N** / NI
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA / Y / PY / **PN / N** / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y / PY / **PN / N** / NI
5.2 Were participants excluded due to missing data on intervention status?		**Y / PY** / PN / N / NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		**Y / PY** / PN / N / NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		NA / Y / PY / **PN / N** / NI

5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY **/ PN / N** / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		**Y / PY** / PN / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		**Y / PY** / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY / **PN / N** / NI
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		**Y / PY** / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from...		
7.1. ... multiple outcome *measurements* within the outcome domain?		**Y / PY** / PN / N / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		**Y / PY** / PN / N / NI
7.3 ... different *subgroups*?		**Y / PY** / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

[1]Responses underlined in green are potential markers for low risk of bias, and responses in red/bold are potential markers for a risk of bias. No formatting is used for questions 1.2 and 1.3 because these questions are only sign posts to other questions. NA, not applicable; NI, no information; N, no; PN, partially no; PY, partially yes and Y, yes.

Supplemental Table 2. 3 Quality of the included studies assessed by ROBINS-I tool[1]

Study	Confounding	Selection of participants	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of the reported result	Overall risk of bias
Riano Galan, 2005 [167]	S	L	L	L	L	L	M	S
Murcia, 2011 [160]	M	L	L	L	L	M	M	M
Trumpff, 2015 [161]	M	L	M	L	L	L	M	M
Trumpff, 2016 [168]	M	L	L	L	L	S	L	S
Trumpff, 2016 [162]	M	L	L	L	M	M	L	M
Belcari, 2011 [169]	M	M	S	L	L	M	L	S
Costeira, 2011 [170]	S	M	M	M	M	M	L	S
Lain, 2016 [86]	L	M	L	M	M	L	M	M
Cuestas, 2015 [142]	S	M	M	M	L	S	L	S
Korzeniewski, 2014 [141]	S	S	M	L	L	L	L	S
Leviton, 2016 [171]	S	M	S	L	NI	L	M	S
Kuban, 2017 [172]	S	L	M	M	L	L	M	S
Calaciura, 1995 [163]	M	L	L	L	L	L	M	M
Azizi, 2001 [164]	M	L	M	L	L	L	L	M
Choudhury, 2003 [165]	M	L	M	L	L	L	M	M
Freire, 2010 [166]	M	L	M	L	L	L	M	M
Shields, 2011 [173]	S	L	M	L	L	M	M	S

[1]L, low risk of bias; M, medium risk of bias; NI, no information and S, serious risk of bias.

Supplemental Table 3. 1 Search strategy in different databases

PUBMED:		
Thyroid stimulating hormone[TW] OR Thyrotropin[MH:noexp] OR Thyrotropin[TIAB] OR TSH[TW] OR Thyrotrophin[TW] OR Thyreotropin[TW] OR Hypothyroidism[TW] OR Iodine Deficienc*[TW] OR Goiter[TW] OR Thyroid volume [TW] OR Hypothyroxin* [TW] OR "Iodine/deficiency"[Mesh] OR Iodine status[TW] OR Urinary iodi* [TW] OR iodine nutritional status [TW] OR iodide status[TW]	Infant, Newborn[MH:noexp] OR Neonat*[TW] OR Newborn*[TW]	Child*[TW]
(Thyroid stimulating hormone[TW] OR Thyrotropin[MH:noexp] OR Thyrotropin[TIAB] OR TSH[TW] OR Thyrotrophin[TW] OR Thyreotropin[TW] OR Hypothyroidism[TW] OR Iodine Deficienc*[TW] OR hypothyroxin* [TW] OR "Iodine/deficiency"[Mesh] OR Goiter[TW] OR Thyroid volume [TW] OR Iodine status[TW] OR Urinary iodi* [TW] OR iodine nutritional status [TW] OR iodide status[TW]) AND (Infant, Newborn[MH:noexp] OR Neonat*[TW] OR Newborn*[TW]) AND Child*[TW]		
Embase:		
“Thyroid stimulating hormone” OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR “Iodine Deficienc*” OR	Infant OR Newborn OR Neonat* OR Newborn*	Child*

hypothyroxin* OR "Iodine/deficiency" OR Goiter OR “Thyroid volume” OR “Iodine status” OR “Urinary iodi*” OR “iodine nutritional status” OR “iodide status”
('thyroid stimulating hormone' OR thyrotropin OR tsh OR thyrotrophin OR thyreotropin OR hypothyroidism OR 'iodine deficienc*' OR Goiter OR 'Thyroid volume' OR hypothyroxin* OR 'iodine/deficiency' OR 'iodine status' OR 'urinary iodi*' OR 'iodine nutritional status' OR 'iodide status') AND (infant OR newborn OR neonat* OR newborn*) AND (child*) AND humans
Scopus:

A	B	C	D
“Thyroid stimulating hormone” OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR “Iodine Deficiency” OR Goiter OR “Thyroid volume” OR hypothyroxine OR "Iodine deficiency" OR “Iodine status” OR “Urinary iodine” OR “iodine nutritional status” OR “iodide status”	Infant OR Newborn OR Neonate OR Newborn	Child	Humans
ABS (“Thyroid stimulating hormone” OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR “Iodine Deficiency” OR Goiter OR “Thyroid volume” OR hypothyroxine OR "Iodine deficiency" OR “Iodine status” OR			

“Urinary iodine” OR “iodine nutritional status” OR “iodide status”) AND (Infant OR Newborn OR Neonate OR Newborn) AND (Child) AND (humans) ABS (A AND B AND C AND D)		
CINAHL:		
“Thyroid stimulating hormone” OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR “Iodine Deficiency” OR Goiter OR “Thyroid volume” OR hypothyroxine OR "Iodine deficiency" OR “Iodine status” OR “Urinary iodine” OR “iodine nutritional status” OR “iodide status”	Infant OR Newborn OR Neonate OR Newborn	Child or children or “school age children”
PsycINFO:		
Thyroid stimulating hormone OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR Iodine Deficiency OR Goiter OR Thyroid volume OR hypothyroxine OR Iodine deficiency OR Iodine status OR Urinary iodine OR iodine nutritional status OR iodide status	Infant OR Newborn OR Neonate OR Newborn	Child or children or school age children

(Thyroid stimulating hormone OR Thyrotropin OR Thyrotropin OR TSH OR Thyrotrophin OR Thyreotropin OR Hypothyroidism OR Iodine Deficiency OR hypothyroxine OR Iodine deficiency OR Goiter OR Thyroid volume OR Iodine status OR Urinary iodine OR iodine nutritional status OR iodide status) AND (Infant OR Newborn OR Neonate OR Newborn) AND (Child or children or school age children)

Supplemental Table 3. 2 Check list for assessing quality of the studies[1]

REVIEWER_____ DATE_____
AUTHOR_____ YEAR_____ RECORD NUMBER_____

ITEMS	QUALITY ASSESSMENT QUESTIONS	YES	NO	UNCLEAR	NA
1	Was the research question clearly stated? 1.1 Was the outcome clearly indicated? 1.2 Were the target population and setting specified?				
2[2]	Was the selection of study subjects free from bias? 2.1 Were inclusion/exclusion criteria specified and with sufficient detail and without omitting criteria critical to the study? 2.2 were health, demographics and other characteristics of subjects described? 2.3 Were the subjects/patients a representative sample of the relevant population?				
3[2]	Were study groups comparable? 3.1 Were groups comparable on important confounding factors, period of outcome (TSH vs UIC vs goitre) assessment and/or were pre-existing differences accounted for by using appropriate adjustments in statistical analysis?				
4	Was methods of handling withdrawals described? 4.1 Was the number, characteristics of response rate described for each group? 4.2 Were all enrolled subjects/patients (in the original sample) accounted for?				
5	Was blinding used to prevent introduction of bias? 5.1 Were data collectors blinded for outcome assessment? (If outcome is measured using an objective test, these criteria is assumed to be met) 5.2 were measurements of outcomes and risk factors blinded?				
6[2]	Were exposure factor described in detail? 6.1 was the study setting and exposure factor described?				
7[2]	Were outcomes clearly defined and the measurements valid and reliable? 7.1 Were the observations and measurements based on standard, valid, and reliable data collection instruments/tests/procedures?				
8	Was the statistical analysis appropriate for the study design and type of outcome indicators? 8.1 were the statistical analysis adequately described? 8.2 were correct statistical tests were used and assumptions of the test not violated?				

	8.3 were statistics reported with levels of significance and or confidence intervals?
9	Are conclusions supported by the results with biases and limitations taken into consideration? 9.1 Is there a discussion of findings? 9.2 Are biases and study limitations identified and discussed?
10	Is bias due to study's funding or sponsorship unlikely? 10.1 were sources of funding and investigators affiliations described? 10.2 was there no apparent conflict of interest? Rate of the overall quality

[1]The checklist was modified by authors from American Dietetic Association (ADA) quality assessment checklist to be used for cross sectional studies[189].

[2]The important items to assess the quality of cross-sectional studies.

Supplemental Table 3. 3 Quality of the included studies

Studies	Clear research question	Selection free from bias[1]	Comparability [1]	No missing outcome	Blinding	Clearly defined exposures [1]	Valid and reliable measurement tools[1]	Appropriate statistical analysis	Conclusions based on results	Funding	Study quality [2]
Al-Hosani [181]	Yes	Yes	No	Yes	Unclear	No	Yes	Yes	Yes	Yes	Ø
Simsek [60]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Simsek [60]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Simsek [60]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Simsek [60]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Costante [196]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Costante [196]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Dalili [197]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Dalili [197]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Dalili [197]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Dalili [197]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Dalili [197]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Fuse [198]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø
Fuse[198]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø
Konrade [199]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Svinaryov [201]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Wachter [188]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø
Wachter [188]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø
Wachter [188]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø

Studies	Clear research question	Selection free from bias[1]	Comparability[1]	No missing outcome	Blinding	Clearly defined exposures[1]	Valid and reliable measurement tools[1]	Appropriate statistical analysis	Conclusions based on results	Funding	Study quality[2]
Wachter [188]	Yes	No	No[3]	Yes	No	Yes	Yes	Yes	Yes	Yes	Ø
Yamada [203]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Yamada [203]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Yamada [203]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Yamada [203]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Yamada [203]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Zimmerman [84]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Zimmerman [84]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Costante [195]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Gruneiro-Papendieck [183]	Yes	No	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Gruneiro-Papendieck [183]	Yes	No	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Kapil [85]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Kapil [85]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Kapil [85]	Yes	Yes	Yes[3]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Copeland [182]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+

Studies	Clear research question	Selection free from bias[1]	Comparability[1]	No missing outcome	Blinding	Clearly defined exposures[1]	Valid and reliable measurement tools[1]	Appropriate statistical analysis	Conclusions based on results	Funding	Study quality[2]
Copeland [182]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Copeland [182]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Sareen [83]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Sareen [83]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Sareen [83]	Yes	Yes	Yes[3]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Sullivan [184]	Yes	Yes	Unclear[3]	Unclear	Yes	Yes	No	Yes	Yes	Yes	Ø
Sullivan [184]	Yes	Yes	Unclear[3]	Unclear	Yes	Yes	No	Yes	Yes	Yes	Ø
Sullivan [184]	Yes	Yes	Unclear[3]	Unclear	Yes	Yes	No	Yes	Yes	Yes	Ø
Burns [54] and Vanderpump [180]	Yes	No	No[4]	No	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Burns [54] and Vanderpump [180]	Yes	No	No[4]	No	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Clapin [52] and ABS report[131]	Yes	Yes	Yes[4]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Evans [179] and Vanderpump [180]	Yes	No	No[4]	No	Yes	Yes	Yes	Yes	Yes	Yes	Ø

Studies	Clear research question	Selection free from bias[1]	Comparability [1]	No missing outcome	Blinding	Clearly defined exposures [1]	Valid and reliable measurement tools[1]	Appropriate statistical analysis	Conclusions based on results	Funding	Study quality [2]
Evans [179] and Vanderpump [180]	Yes	No	No[4]	No	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Gyurjyan [187] and Selga [204]	Yes	No	Yes[4]	No	Yes	Yes	No	Yes	Yes	Yes	Ø
Wilcken [116] and Li [17]	Yes	No	Yes[4]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Mikelsaar [57]	Yes	No	No	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Mikelsaar [57]	Yes	No	No	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Mikelsaar [57]	Yes	No	No	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Ø
Rahman [200] and Li [17]	Yes	Yes	Yes[4]	Yes	Yes	Yes	No	Yes	Yes	Yes	Ø
Vandevijvere [186] and Vandevijvere [202]	Yes	Yes	Yes[4]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Vandevijvere [186] and Vandevijvere [202]	Yes	Yes	Yes[4]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+
Vandevijvere [186] and	Yes	Yes	Yes[4]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+

Studies	Clear research question	Selection free from bias[1]	Comparability [1]	No missing outcome	Blinding	Clearly defined exposures [1]	Valid and reliable measurement tools[1]	Appropriate statistical analysis	Conclusions based on results	Funding	Study quality [2]
Vandevijvere [202]											
Mehran [205]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	+

[1]The important items to assess the quality of cross-sectional studies.

[2]The quality of included studies was assessed based on the American Dietetic Association's study quality analysis manual [189]. Studies with + sign were considered as high quality/low risk of bias, studies with Ø sign were considered as neutral and studies with – sign were considered as low quality/high risk of bias.

[3]Two or more studies were reported in a single publication.

[4]The Data on iodine status based on the different markers were extracted from different publications of the same country/region conducted in the same time period to determine the agreement between markers.

Supplemental Table 3. 4 The World Health Organization's criteria in classifying population iodine status[1]

Population Markers	Iodine Status[2]			
	Sufficiency	Mild-deficiency	Moderate-deficiency	Sever-deficiency
Newborn TSH >5 mIU/L, %	0-3	> 3-19.9	20-39.9	≥40
MUIC in SAC, µg/L	≥100	50-99	20-49	<20
Goitre in SAC, %	0-4.9	5-19.9	20-29.9	≥30

[1]TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration; SAC, school age children.

[2]Iodine status is defined based on the WHO criteria [12, 178].

Supplemental Table 3. 5 Characteristics of the included studies[1]

PUBLICATIONS, COUNTRY	STUDY[2]	% TSH >5 mIU/L (N)	MUIC µg/L (N)	% GOITRE BY PALPATION (N)	% TVOL > 97TH PERCENTILE (N)
AGE AT TSH ASSESSMENT:3-4 DAYS[3]					
Yamada, 2000, Mongolia [203]	Study (Ulaanbaatar)	18.0 (128)	160.4 (210)	30.7 (1,209)	NA
	Study (Uvurkhangai)	30.4 (112)	45.0 (300)	19 (1,200)	NA
	Study (Huvsgol)	13.3 (207)	126 (140)	29.5 (516)	NA
	Study (Dornod)	12.2 (368)	40.6 (110)	36.9 (428)	NA
	Study (Hovd)	28.3 (99)	39.1 (150)	28.5 (302)	NA
Zimmerman, 2005, Switzerland [84]	Study (Pre-intervention population)	2.9 (259,035)	115 (610)	NA	NA
	Study (Post-intervention population)	1.7 (218,665)	141 (386)	NA	NA
Clapin, 2014, [52]; ABS report, 2013, Western Australia [131]	Single study	5.7	261.3 (NA)	NA	NA
Evans, 2014, [179]; Vanderpump, 2011, Wales, UK [180]	Single study	1.5 (535)	80.1 (737)[4]	NA	NA
Rahman, 2010, [200]; Li, 2005, Victoria, Australia [17]	Single study	9.26 (NA)	73.5 (348)[4]	NA	NA
Fuse, 2003, Ulaan Baatar, Mongolia [198]	Study (1996), 1996	17.8 (129)	152.5 (300)	49.4 (1,203)	43.3 (300)
	Study (1996-1999)	17.8 (129)	160.4 (198)	NA	29.8 (205)
Wilcken, 2011, [116]; Li , 2005, NSW, Australia [17]	Single study	5.02 (NA)	89.0 (427)[4]	NA	NA
AGE AT TSH ASSESSMENT: < 3 DAYS OR > 4 DAYS[5]					
Gruneiro- Papendieck, 2004, Argentina [183]	Single study	3.6 (112)	143 (100)	4.5(500)	NA
Al-Hosani, 2003, United Arab Emirates [181]	Single study	2.6 (23,116)	90.5 (NA)	40.4(NA)	NA

PUBLICATIONS, COUNTRY	STUDY[2]	% TSH >5 mIU/L (N)	MUIC µg/L (N)	% GOITRE BY PALPATION (N)	% TVOL > 97TH PERCENTILE (N)
Wachter,1985, Tanzania [188]	Study (Pre-intervention population)	2.7 (150)	17.6 (23)[6]	91.6 (3031)	NA
Wachter,1985, Tanzania [188]	Study (Post intervention population)	0 (74)	93.3 (13)[6]	NA	NA
Evans,2014 [179]; Vanderpump, 2011, Wales, UK [180]	Single study	0.9 (41,809)	80.1 (737)[4]	NA	NA
AGE AT TSH ASSESSMENT: 3-4 DAYS AND < 3 DAYS OR > 4 DAYS[7]					
Mehran, 2016, Iran [205]	Single study	2.8 (128,8237)	145 (11280)	NA	NA
Dalili, 2012, Iran [197]	Study (2006)	1.7 (9,251)	270.2 (240)	NA	NA
	Study (2007)	1.4 (23,529)	200.4 (240)	NA	NA
	Study (2008)	2.1 (27,427)	200.1 (240)	NA	NA
	Study (2009)	1.8 (29,511)	200 (240)	NA	NA
	Study (2010)	1.9 (29,983)	200.2 (240)	NA	NA
Costante, 2002, Italy [196]	Study (Inland territory)	14.8 (9,957)	53.8 (191)	41.3 (10243)	NA
	Study (Sea level territory)	14.1 (11,121)	89.6 (93)	17.4 (3741)	NA
Costante,1997, Italy [195]	Single study	14.4 (21,078)	65.6 (284)	37 (13,984)	
Burns, 2008 [54]; Vanderpump, 2011, Ireland [180]	Single study	2.35	80.1 (737)	NA	NA
Mikelsaar, 1999; Not cited, Estonia,[57]	Study (North)	16.4	61.5[4]	NA	NA
	Study (Central)	21	71[4]	NA	NA
	Study (South)	17.2	53.5[4]	NA	NA
Vandevijvere, 2012 [186]; Vandevijvere, 2012, Belgium [202]	Study (2009-2011)	2.6 (81,107)	113 (1,541)[4]	7.2 (1,541)	NA
	Study (2010-2011)	3.0 (92,961)	113 (1,541)[4]	7.2 (1,541)	NA
	Study (2011)	3.3 (81,511)	113 (1,541)[4]	7.2 (1,541)	NA
Simsek, 2003, Turkey [60]	Study (Turkey)	26.7 (18,606)	NA	52 (1,046)	NA
	Study (Bolu)	32.1 (5,210)	31 (342)	76 (342)	79 (342)
	Study (Duzce)	26 (6,140)	49 (384)	67 (384)	71 (384)

PUBLICATIONS, COUNTRY	STUDY[2]	% TSH >5 mIU/L (N)	MUIC µg/L (N)	% GOITRE BY PALPATION (N)	% TVOL > 97TH PERCENTILE (N)
	Study (Zongulda)	23.2 (7,256)	23.2 (320)	9 (320)	17 (320)
Konrade, 2014, Latvia [199]	Single study	8.2 and 9.3 (31,274)	107.3 (913)[6]	NA	NA
Svinaryov, 2003, Russia [201]	Single study	37.4 (95,657)	38.2 (654)	35.8 (2,040)	25.8 (2040)
Wachter, 1985, Tanzania [188]	Study (Pre-intervention population)	45 (69)	17.6 (23)[6]	91.6 (3,031)	NA
Wachter, 1985, Tanzania [188]	Study (Post-intervention population)	21 (117)	93.3 (13)[6]	NA	NA
Gruneiro- Papendieck,, 2004, Argentina [183]	Single study	2.7 (1500)	143 (100)	4.5 (500)	NA
Gyurjyan, 2006[187]; Selga, 2000, Latvia [204]	Single study	16.5 (NA)	59 (NA)[4]	NA	NA
TSH ASSESSED IN CORD BLOOD SAMPLES					
Kapil, 2015, India [85]	Study (Kangra)	73.4 (613)	200 (463)	15.8 (1,864)	NA
	Study (Kullu)	79.8 (638)	175 (532)	23.4 (1,986)	
	Study (Solan)	63.2 (683)	62.5 (513)	15.4 (1,898)	
Copeland, 2002, Bangladesh [182]	Single study	84 (208)	73 (400)	27 (3,99)	26 (351)
Copeland, 2002, Guatemala [182]	Single study	58 (141)	181 (458)	15 (518)	NA
Copeland, 2002, USA [182]	Single study	82 (243)	282 (284)	2 (305)	0 (300)
Sareen, 2016, India[83]	Study (Udham Singh Nagar)	53.3 (649)	150 (1,807)	13.2 (NA)	NA
	Study (Nainital)	76.4 (670)	125 (2,269)	15.9 (NA)	
	Study (Pauri Garhwal)	72.8 (694)	115 (2,067)	16.8 (NA)	
Sullivan, 1997, Kyrgyzstan [184]	Study (Bishkek)	74 (90)	30 (NA)	49 (NA)	39 (NA)
	Study (OSH)	47 (92)	30 (NA)	49 (NA)	39 (NA)
Sullivan,1997, Philippines [184]	Single study	32 (750)	40 (NA)	2 (NA)	NA
Gruneiro-Papendieck, 2004, Argentina [183]	Single study	11.3 (186)	143 (100)	4.5 (500)	

[1]NA, not applicable; N, number of children; TSH, thyroid stimulating hormone; Tvol, thyroid volume; MUIC, median urinary iodine concentration

[2]Studies conducted at different regions or different periods within the same publication were classified as separate studies, with the region /time period in the brackets. “Single study” denotes that the publication(s) reported a study that was conducted in only one region or one time point.

[3]Studies where the newborns’ heel prick samples were collected within the World Health Organization’s recommended age (at 3-4 days of infant’s age).

[4]MUIC or goitre results from the same country/region conducted in the same time period were extracted from different publications.

[5]Studies where the newborns’ heel prick samples were collected outside the World Health Organization’s recommended age (at < 3 or > 4 days of infant’s age).

[6]Urinary iodine concertation as µg iodine per gram of creatinine.

[7]Studies where the newborns’ heel prick samples were collected within (3-4 days of infant age) and outside (< 3 and > 4 days of infants age) the World Health Organization’s recommended infant’s age.

Supplemental Table 3. 6 Classification of the iodine status of the included studies based on TSH, MUIC and goitre prevalence[1]

PUBLICATIONS, COUNTRY	STUDY[2]	IODINE STATUS BY POPULATION MARKERS[3]			
		TSH	MUIC	GOITRE BY PALPATION	TVOL BY ULTRASOUND
AGE AT TSH ASSESSMENT:3-4 DAYS[4]					
Yamada, 2000, Mongolia [203]	Study (Ulaanbaatar)	Mild	Sufficiency	Severe	NA
	Study (Uvurkhangai)	Moderate	Moderate	Moderate	NA
	Study (Huvsgol)	Mild	Sufficiency	Moderate	NA
	Study (Dornod)	Mild	Moderate	Severe	NA
	Study (Hovd)	Moderate	Moderate	Moderate	NA
Zimmerman, 2005, Switzerland [84]	Study (Pre-intervention population)	Sufficiency	Sufficiency	NA	NA
	Study (Post-intervention population)	Sufficiency	Sufficiency	NA	NA
Clapin, 2014, [52]; ABS report, 2013, Western Australia [131]	Single study	Mild	Sufficiency	NA	NA
Evans, 2014, [179]; Vanderpump, 2011, Wales, UK [180]	Single study	Sufficiency	Mild	NA	NA
Rahman, 2010, [200]; Li, 2005, Victoria, Australia [17]	Single study	Mild	Mild	NA	NA
Fuse, 2003, Ulaan Baatar, Mongolia [198]	Study (1996), 1996	Mild	Sufficiency	Severe	Severe
	Study (1996-1999)	Mild	Sufficiency	NA	Moderate
Wilcken, 2011, [116]; Li , 2005, NSW, Australia [17]	Single study	Mild	Mild	NA	NA
AGE AT TSH ASSESSMENT: < 3 DAYS OR > 4 DAYS[5]					
Gruneiro- Papendieck, 2004, Argentina [183]	Single study	Mild	Sufficiency	Sufficiency	NA
Al-Hosani, 2003, United Arab Emirates [181]	Single study	Sufficiency	Mild	Severe	NA

PUBLICATIONS, COUNTRY	STUDY[2]	IODINE STATUS BY POPULATION MARKERS[3]			
		TSH	MUIC	GOITRE BY PALPATION	TVOL BY ULTRASOUND
Wachter,1985, Tanzania [188]	Study (Pre-intervention population)	Sufficiency	Severe	Severe	NA
Wachter,1985, Tanzania [188]	Study (Post intervention population)	Sufficiency	Mild	NA	NA
Evans,2014 [179]; Vanderpump, 2011, Wales, UK [180]	Single study	Sufficiency	Mild	NA	NA
AGE AT TSH ASSESSMENT: 3-4 DAYS AND < 3 DAYS OR > 4 DAYS[6]					
Mehran, 2016, Iran [205]	Single study	Sufficiency	Sufficiency	NA	NA
Dalili, 2012, Iran [197]	Study (2006)	Sufficiency	Sufficiency	NA	NA
	Study (2007)	Sufficiency	Sufficiency	NA	NA
	Study (2008)	Sufficiency	Sufficiency	NA	NA
	Study (2009)	Sufficiency	Sufficiency	NA	NA
	Study (2010)	Sufficiency	Sufficiency	NA	NA
Costante, 2002, Italy [196]	Study (Inland territory)	Mild	Mild	Severe	NA
	Study (Sea level territory)	Mild	Mild	Mild	NA
Costante,1997, Italy [195]	Single study	Mild	Mild	Severe	
Burns, 2008 [54]; Vanderpump, 2011, Ireland [180]	Single study	Sufficiency	Mild	NA	NA
Mikelsaar, 1999; Not cited, Estonia,[57]	Study (North)	Mild	Mild	NA	NA
	Study (Central)	Mild	Mild	NA	NA
	Study (South)	Mild	Mild	NA	NA
Vandevijvere, 2012 [186]; Vandevijvere, 2012, Belgium [202]	Study (2009-2011)	Sufficiency	Sufficiency	Mild	NA
	Study (2010-2011)	Mild	Sufficiency	Mild	NA
	Study (2011)	Mild	Sufficiency	Mild	NA
Simsek, 2003, Turkey [60]	Study (Turkey)	Moderate	NA	Severe	NA
	Study (Bolu)	Moderate	Moderate	Severe	Severe

PUBLICATIONS, COUNTRY	STUDY[2]	IODINE STATUS BY POPULATION MARKERS[3]			
		TSH	MUIC	GOITRE BY PALPATION	TVOL BY ULTRASOUND
	Study (Duzce)	Moderate	Moderate	Severe	Severe
	Study (Zongulda)	Moderate	Moderate	Mild	Mild
Konrade, 2014, Latvia [199]	Single study	Mild	Sufficiency	NA	NA
Svinaryov, 2003, Russia [201]	Single study	Moderate	Moderate	Severe	Moderate
Wachter, 1985, Tanzania [188]	Study (Pre-intervention population)	Severe	Severe	Severe	NA
Wachter, 1985, Tanzania [188]	Study (Post-intervention population)	Moderate	Mild	NA	NA
Gruneiro- Papendieck,, 2004, Argentina [183]	Single study	Sufficiency	Sufficiency	Sufficiency	NA
Gyurjyan, 2006[187]; Selga, 2000, Latvia [204]	Single study	Mild	Mild	NA	NA
TSH ASSESSED IN CORD BLOOD SAMPLES					
Kapil, 2015, India [85]	Study (Kangra)	Severe	Sufficiency	Mild	NA
	Study (Kullu)	Severe	Sufficiency	Moderate	
	Study (Solan)	Severe	Mild	Mild	
Copeland, 2002, Bangladesh [182]	Single study	Severe	Mild	Moderate	Moderate
Copeland, 2002, Guatemala [182]	Single study	Severe	Sufficiency	Mild	NA
Copeland, 2002, USA [182]	Single study	Severe	Sufficiency	Sufficiency	Sufficiency
Sareen, 2016, India[83]	Study (Udham Singh Nagar)	Severe	Sufficiency	Mild	NA
	Study (Nainital)	Severe	Sufficiency	Mild	
	Study (Pauri Garhwal)	Severe	Sufficiency	Mild	
Sullivan, 1997, Kyrgyzstan [184]	Study (Bishkek)	Severe	Moderate	Severe	Severe
	Study (OSH)	Severe	Moderate	Severe	Severe
Sullivan,1997, Philippines [184]	Single study	Moderate	Moderate	Sufficiency	NA

PUBLICATIONS, COUNTRY	STUDY[2]	IODINE STATUS BY POPULATION MARKERS[3]			
		TSH	MUIC	GOITRE BY PALPATION	TVOL BY ULTRASOUND
Gruneiro-Papendieck, 2004, Argentina [183]	Single study	Mild	Sufficiency	Sufficiency	

[1]NA, not applicable; TSH, thyroid stimulating hormone; Tvol, thyroid volume; MUIC, median urinary iodine concentration

[2]Studies conducted at different regions or different periods within the same publication were classified as separate studies, with the region /time period in the brackets. "Single study" denotes that the publication(s) reported a study that was conducted in only one region or one time point.

[3]The Iodine status was defined according to the WHO criteria [12, 178] using the results of MUIC, percentage of TSH >5 mIU/L and goitre prevalence reported in the included studies.

[4]Studies where the newborns' heel prick samples were collected within the World Health Organization's recommended age (at 3-4 days of infant's age).

[5]Studies where the newborns' heel prick samples were collected outside the World Health Organization's recommended age (at < 3 or > 4 days of infant's age).

[6]Studies where the newborns' heel prick samples were collected within (3-4 days of infant age) and outside (< 3 and > 4 days of infants age) the World Health Organization's recommended infant's age.

Supplemental Table 3. 7 Summary of the number of studies and publications included in the systematic review

References	Number of publications	Number of studies per publication	Total number of studies
Publication contained > 1 study			
Dalili [197], Yamada [203]	Two	5	10
Simsek [60], Wachter [188]	Two	4	8
Copeland [182], Gruneiro-Papendieck [183], Kapil [85], Mikelsaar [57], Sareen [83], Sullivan [184]	Six	3	18
Costante [196], Evans [179] Fuse [198], Zimmermann [84]	Four	2	8
Al-Hosani [181], Costante [195], Konrade [199], Rahman [200], Svinaryov [201], Mehran [205]	Seven	1	6
One study from two publications			
Burns [54] Clapin [52], Gyurjyan [187], Li[17], Vanderpump [180] Wilcken [116], Selga [204] Australian Health Survey [131]	Eight		4
Three studies from two publications			
Vandevijvere [186], Vandevijvere [202]	Two		3
Total	30		57

Supplemental Table 3. 8 Comparison of the agreement between markers from the same or different publications[1]

	Agreement between TSH and MUIC			Agreement between TSH and goitre		
	Yes	No	Total	Yes	No	Total
Studies from the same publication	24 (73%)	9 (27%)	33 (100%)	17 (85%)	3 (15%)	20 (100%)
Studies from the different publications	4 (40%)	6 (60%)	10 (100%)	2 (67%)	1 (33%)	3 (100%)

[1] N, number of studies; TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration. Data are n(%).

Supplemental Table 3. 9 A contingency table of the iodine status of the included studies based on different markers[1]

	Iodine deficiency by MUIC[2]		Iodine deficiency by goitre[2]	
Iodine deficiency by TSH[2]	No	Yes	No	Yes
No	10	6	1	3
Yes	9	18	1	18

[1]TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration

[2]Iodine deficiency is defined as ≥ 3% of the population with newborn TSH >5 mIU/L, or MUIC < 100 µg/L or prevalence of goitre > 5% [12, 178].

Supplemental Table 3. 10 The agreement between newborn TSH and MUIC in pregnant women in classifying population iodine status[1]

Publications, country	Study	% TSH >5 mIU/l (N)	MUIC µg/L (N)	Agreement between TSH and MUIC
Sareen, 2016, India [83]	Udham Singh Nagar: 2013-2014 [3,4]	53.3(649)	124(532)	Yes
	Nainital: 2013-2014 [3,4]	76.4(670)	117.5(468)	Yes
	Pauri Garhwal: 2013-2014 [3,4]	72.8(694)	110(404)	Yes
	1992-1998[3],1999[4]	2.9 (259,035)	138(511)	No
Zimmerman, 2005, Switzerland [84]	1999-2004[3], 2004[4]	1.7(218,665)	249(279)	Yes
Kapil, 2015, India [85]	Kangra: 2014[43,4]	73.4 (613)	200 (368)	No
	Kullu:2014[3,4]	79.8 (638)	149 (439)	Yes
	Solan:2014[3,4]	63.2 (683)	130 (311)	Yes

[1]TSH, thyroid stimulating hormone; MUIC, median urinary iodine concentration.

[2]Agreement between newborn TSH and MUIC to classify iodine status. Iodine status is defined as deficiency if ≥ 3% of the population with newborn TSH >5 mIU/L, or MUIC < 150µg/L during pregnancy [12, 178].

[3]Year of data collection for newborn TSH concentration.

[4]Year of data collection for MUIC in pregnant women.

Supplemental Table 3. 11 Coordinates of the receiver operating curve for TSH >5 mIU/L vs UIC <100 µg/L[1]

Test result variable: proportion of TSH>5 mIU/L State variable: median UIC <100 µg/L		
Iodine sufficient if proportion greater than or equal to[2]	Sensitivity	1 - Specificity
-1.0000	1.000	1.000
0.4500	0.958	1.000
1.1500	0.917	1.000
1.4500	0.917	0.947
1.6000	0.875	0.947
1.7500	0.833	0.842
1.8500	0.833	0.789
2.0000	0.833	0.737
2.2250	0.833	0.684
2.4750	0.792	0.684
2.6500	0.792	0.632
2.7500	0.750	0.579
2.8500	0.750	0.526
2.9500	0.750	0.474
3.1500	0.750	0.421
3.4500	0.750	0.368
4.3100	0.750	0.316
5.3600	0.708	0.316
7.4800	0.708	0.263
9.2800	0.667	0.263
10.7500	0.667	0.211
12.7500	0.625	0.211
13.7000	0.625	0.158
14.2500	0.583	0.158
14.6000	0.542	0.158
15.6000	0.500	0.158
16.4500	0.458	0.158
16.8500	0.417	0.158
17.5000	0.375	0.158
17.9000	0.375	0.053
19.5000	0.375	0.000
22.1000	0.292	0.000
24.6000	0.250	0.000
27.1500	0.208	0.000
29.3500	0.167	0.000
31.2500	0.125	0.000
34.7500	0.083	0.000
41.2000	0.042	0.000

[1]TSH, thyroid stimulating hormone; UIC, urinary iodine concentration

[2]The smallest cut-off value is the minimum observed test value minus 1, and the largest cut-off value is the maximum observed test value plus 1. All the other cut-off values are the averages of two consecutive ordered observed test values.

Supplemental Table 3. 12 Coordinates of the receiver operating curve for TSH >5 mIU/L vs prevalence of goitre >5%[1]

Test result variable(s): proportion of TSH >5 mIU/L State variable: prevalence of goitre >5%		
Iodine sufficient if proportion greater than or equal to[2]	Sensitivity	1 - Specificity
0.7000	1.000	1.000
2.1500	0.952	1.000
2.6500	0.905	1.000
2.8500	0.857	0.500
3.1500	0.810	0.500
3.4500	0.762	0.500
7.9000	0.762	0.000
12.7500	0.714	0.000
13.7000	0.667	0.000
14.2500	0.619	0.000
14.6000	0.571	0.000
16.3000	0.524	0.000
17.9000	0.429	0.000
20.6000	0.381	0.000
24.6000	0.333	0.000
26.3500	0.286	0.000
27.5000	0.238	0.000
29.3500	0.190	0.000
31.2500	0.143	0.000
34.7500	0.095	0.000
41.2000	0.048	0.000
46.0000	0.000	0.000

[1]TSH, thyroid stimulating hormone

[2]The smallest cut-off value is the minimum observed test value minus 1, and the largest cut-off value is the maximum observed test value plus 1. All the other cut-off values are the averages of two consecutive ordered observed test values.

Supplemental Table 4. 1 Iodine status of participants classified using different markers (n=534)[1].

	Estimated 24-h UIE, µg/day	
	Mage et al.	Kesteloot and Joosens
Median	261.0 (154.9, 466.3)	259.0 (153.6, 415.8)
Broad classification of iodine status		
<150 (deficiency)	124 (23.2)	128 (24.0)
150-499 (sufficiency)	300 (56.2)	315 (59.0)
≥500 (excess)	110 (20.6)	91 (17.0)

[1] Values are presented as frequencies (percentages) unless otherwise specified. UIE, urinary iodine excretion.

[2]Medians (Inter Quartile Range).

Supplemental Table 4. 2 A contingency table of the number of pregnant women with urinary iodine value <150 (yes or no) or ≥500 (yes or no) based on estimated 24-h UIE and other urinary iodine markers[1]

Reference markers	Test markers							
	UIC		Estimated 24-h UIE		I/Cr		UIC~Cr	
	Yes	No	Yes	No	Yes	No	Yes	No
Comparison of markers to define iodine deficiency at <150 units[2]								
Estimated 24-h UIE as the reference								
Yes	70	54	NA	NA	122	2	109	15
No	98	312	NA	NA	55	355	71	339
I/Cr as the reference								
Yes	87	90	122	55	NA	NA	140	37
No	81	276	2	355	NA	NA	40	317
UIC as the reference								
Yes	NA	NA	70	98	87	81	127	41
No	NA	NA	54	312	90	276	53	313
Comparison of markers to define iodine excess at ≥500 units[2]								
Estimated 24-h UIE as the reference								
Yes	41	69	NA	NA	54	56	38	72
No	17	407	NA	NA	3	421	2	422
I/Cr as the reference								
Yes	24	33	0	3	NA	NA	32	8
No	34	443	110	421	NA	NA	25	469
UIC as the reference								
Yes	NA	NA	41	17	24	34	32	26
No	NA	NA	69	407	33	443	8	468

[1]I/Cr, urinary iodine concentration to creatinine ratio; NA, not applicable; UIC, urinary iodine concentration; UIC~Cr, urinary iodine concentration corrected for creatinine; UIE, urinary iodine excretion.

[2]The WHO criteria based on median UIC in µg/L was applied in all markers to define iodine deficiency (<150) or excess (≥500) including I/Cr in µg/g and UIC~Cr in µg/L and estimated 24-h UIE in µg/d [12].

Supplemental Table 7. 1 Comparison of observed vs. imputed values[1]

Variables	DOMInO				PINK			
	n	Observed cases	n	Imputed cases[2]	n	Observed cases	n	Imputed cases[2]
Mother characteristics								
Age at enrolment, y	726	28.8 ± 5.8	726	NA	792	30.0 ± 5.1	794	30.0 ± 5.1
Weight, kg	720	73.7 ± 16.3	726	73.7 ± 16.3	789	70.8 ± 16.0	794	70.8 ± 16.0
Height, cm	726	1.6 ± 0.1	726	NA	788	1.6 ± 0.1	794	1.6 ± 0.0
BMI, kg/m^2	720	27.3 ± 5.9	726	27.3 ± 5.9	787	26.1 ± 5.4	794	26.1 ± 0.4
HSQ score	687	33.7 ± 3.6	726	33.4 ± 3.6	685	33.0 ± 2.9	794	33.0 ± 2.9
Hospital	726		726		794		794	
FMC		257 (35.4)		NA		402 (50.6)		NA
WCH		469 (64.6)		NA		392 (49.4)		NA
Parity	726		726		794		794	
0		405 (55.8)		NA		427 (53.8)		NA
≥1		321 (44.2)		NA		367 (46.2)		NA
Race	726		726		794		794	
Caucasian		663 (91.3)		NA		668 (84.1)		NA
Others		63 (8.7)		NA		126 (15.9)		NA
Occupation	726		726		794		794	
Managers		165 (22.7)		NA		44 (5.5)		NA
Professionals		165 (22.7)		NA		266 (33.5)		NA
Sales, trades person, clerical service		201 (27.7)		NA		333 (41.9)		NA
Labourers		0		NA		16 (2.0)		NA
Unemployed/pension		24 (3.3)		NA		14 (1.8)		NA
Student		29 (4.0)		NA		25 (3.2)		NA
Home duties		142 (19.6)		NA		96 (12.1)		NA
Secondary education	726	450 (62.0)	726	NA	794	659 (83.0)	794	NA
Qualification	726		726		794		794	
Certificate or Diploma		317 (43.7)		NA		274 (34.5)		NA
Degree		142 (19.6)		NA		330 (41.6)		NA
Higher Degree		25 (3.4)		NA		79 (9.9)		NA
None		242 (33.3)		NA		111 (14.0)		NA
Smoking during second trimester	726	113 (15.6)	726	NA	794	45 (5.7)	794	NA

Variables	DOMInO				PINK			
	n	Observed cases	n	Imputed cases[2]	n	Observed cases	n	Imputed cases[2]
Previous smoking	726	227 (31.3)	726	NA	794	102 (12.9)	794	NA
Alcohol consumption during second trimester	726	76 (10.5)	726	NA	794	41 (5.2)	794	NA
Previous alcohol consumption	726	445 (61.3)	726	NA	794	526 (66.3)	794	NA
Supplement intake during second trimester	726	352 (48.5)	726	NA	794	708 (89.2)	794	NA
Previous depression	726	159 (21.9)	726	NA	794	169 (21.3)	794	NA
Depression during second trimester	726	48 (6.6)	726	NA	794	48 (6.1)	794	NA
Caesarean section	726	203 (28.0)	726	NA	765	213 (27.8)	794	222 (28.0)
Father characteristics								
Secondary education	726	389 (53.6)	726	NA	785	587 (74.8)	794	592 (74.5)
Qualification	691		726		793		794	
Certificate/Diploma		313 (45.3)		327 (45.0)		315 (39.7)		315 (39.7)
Degree		107 (15.5)		112 (15.4)		218 (27.5)		218 (27.5)
Higher Degree		27 (3.9)		29 (4.0)		74 (9.3)		74 (9.3)
None		244 (35.3)		258 (35.6)		186 (23.5)		187 (23.5)
Occupation	694		726		784		794	
Managers		127 (18.3)		131 (18.1)		92 (11.7)		93 (11.7)
Professionals		300 (43.2)		313 (43.1)		234 (29.9)		235 (29.7)
Sales, trades person, clerical service		207 (29.8)		218 (30.0)		352 (44.9)		357 (45.0)
Labourers		2 (0.3)		2 (0.3)		75 (9.8)		77 (9.7)
Unemployed/pension		33 (0.8)		35 (4.9)		16 (2.0)		17 (2.1)
Student		21 (3.0)		23 (3.1)		9 (1.2)		9 (1.2)
Home duties		4 (0.6)		4 (0.6)		6 (0.8)		6 (0.8)
Child characteristics								
Sex	726		726		765		794	
Female		361 (49.7)		NA		371 (48.5)		382 (48.1)
Male		365 (50.3)		NA		394 (51.5)		412 (51.9)
Gestational age at birth, weeks	726	38.7 ± 2.2	726	NA	765	39.0 ± 1.8	794	39.0 ± 1.8
Birthweight, gm	725	3392.2 ± 607.2	726	3391.5 ± 607.1	764	3380.4 ± 548.7	794	3379.5 ± 549.0
Birthweight-for-gestational age z score	725	0.2 ± 1.0	726	0.2 ± 1.0	764	0.0 ± 1.0	794	0.0 ± 1.0
Birth Length, cm	722	49.5 ± 3.0	726	49.5 ± 3.0	744	49.8 ± 2.6	794	49.8 ± 2.6
Birth Head circumference, cm	717	34.5 ± 2.0	726	34.5 ± 2.0	737	34.5 ± 1.6	794	34.5 ± 1.6
Age at outcome assessment, months	695	18.8 ± 1.7	726	18.8 ± 1.8	679	19.7 ± 2.2	794	19.5 ± 4.8

Variables	DOMInO				PINK			
	n	Observed cases	n	Imputed cases[2]	n	Observed cases	n	Imputed cases[2]
1 minute APGAR score	721	8.2 ± 1.4	726	8.2 ± 1.4	760	8.2 ± 1.4	794	8.2 ± 1.4
5 minute APGAR score	721	9.0 ± 0.9	726	9.0 ± 0.9	760	8.9 ± 0.7	794	8.9 ± 0.7
Newborn TSH[3], mIU/L	590	1.7 (1.0, 2.9)	726		679	2.2 (1.4, 3.3)	794	2.2 (1.4, 3.3)
Newborn TSH quartiles	590		726		679		794	
First		148 (25.1)		184 (25.3)		170 (25.0)		200 (25.2)
Second		147 (24.9)		176 (24.3)		170 (25.0)		196 (24.7)
Third		148 (25.1)		185 (25.5)		170 (25.0)		201 (25.3)
Fourth		147 (24.9)		181 (24.9)		169 (24.9)		197 (24.9)
Newborn TSH category	590		726		679		794	
>5 mIU/L		29 (4.9)		33 (4.6)		37 (5.5)		43 (5.4)
≤5 mIU/L		561 (95.1)		693 (95.4)		642 (94.6)		751 (94.6)
Weight, kg	684	11.5 ± 1.5	726	11.5 ± 1.5	683	11.6 ± 1.5	794	11.6 ± 0.5
Length, cm	685	82.4 ± 3.4	726	82.5 ± 3.4	672	83.0 ± 3.9	794	83.0 ± 4.0
Head circumference, cm	685	47.8 ± 1.5	726	47.8 ± 1.5	660	48.2 ± 2.1	794	48.2 ± 2.2
WLZ	682	0.7 ± 1.0	726	0.7 ± 1.0	669	0.6 ± 1.2	794	0.6 ± 1.3
LAZ	685	0.1 ± 1.1	726	0.1 ± 1.1	670	-0.1 ± 1.2	794	-0.1 ± 1.2
WAZ	684	0.5 ± 1.0	726	0.5 ± 1.0	681	0.5 ± 1.4	794	0.5 ± 1.4
HCZ	685	0.7 ± 1.0	726	0.7 ± 1.0	657	0.9 ± 1.4	794	0.9 ± 1.9
WLZ < - 1	682	23 (3.4)	726	25 (3.5)	669	31 (4.6)	794	41 (5.2)
LAZ < - 1	685	104 (15.2)	726	113 (15.6)	670	100 (14.9)	794	123 (15.5)
WAZ < - 1	684	35 (5.1)	726	38 (5.3)	681	43 (6.3	794	56 (7.0)
HCZ < - 1	685	28 (4.1)	726	30 (4.1)	657	21 (3.2)	794	33 (4.2)
Cognitive score	694	102.0 ± 11.8	726	101.7 ± 12.0	699	96.6 ± 11.6	794	96.3 ± 11.9
Language score	692	97.4 ± 14.5	726	97.2 ± 14.7	698	94.5 ± 17.8	794	93.9 ± 18.1
Motor score	694	102.6 ± 11.1	726	102.5 ± 11.2	698	99.4 ± 10.6	794	98.9 ± 11.2
Social emotional score	681	107.0 ± 17.5	726	106.8 ± 17.8	698	104.1 ± 14.3	794	103.7 ± 14.7
Adaptive behaviour score	680	100.0 ± 14.3	726	100.0 ± 14.4	699	100.3 ± 13.6	794	100.1 ± 14.0
Cognitive delay (<85)	694	31 (4.5)	726	36 (4.9)	699	63 (9.0)	794	80 (10.0)
Language delay (<85)	692	118 (17.1)	726	127 (17.4)	698	196 (28.1)	794	234 (29.4)
Motor delay (<85)	694	33 (4.8)	726	37 (5.1)	698	24 (3.4)	794	35 (4.4)
Social emotional delay (<85)	681	31 (4.6)	726	36 (4.9)	698	34 (4.9)	794	44 (5.5)
Adaptive behaviour delay (<85)	680	80 (11.8)	726	87 (12.1)	699	70 (10.0)	794	87 (11.0)

[1]Data are mean ± SD or n (%) unless otherwise specified. Developmental delay was defined as Bayley-III scores below 85 and low growth indices were defined as z scores below 1 SD. BMI, body mass index; DOMInO, DHA to Optimize Mother Infant Outcome; HCZ, head circumference-for-age z score; HSQ, home screening questionnaire; LAZ, length-for-age z score; n, total sample size; NA, not applicable; PINK, Pregnancy Iodine and Neurodevelopment in Kids; SD, standard deviation; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

[2]NA when data were complete.

[3]Median (IQR).

Supplemental Table 7. 2 Adjusted associations between newborn TSH tertiles and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley-III score	Developmental delay[2]	n	Bayley-II score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Cognitive scale						
Tertile 1	278	ref	ref	257	ref	ref
Tertile 2	210	-1.8 (-4.1, 0.5)	1.8 (0.7, 4.9)	256	-0.3 (-2.5, 1.9)	0.8 (0.5, 1.5)
Tertile 3	238	-1.8 (-4.3, 0.6)	1.6 (0.6, 4.3)	281	1.0 (-1.2, 3.1)	0.6 (0.4, 1.1)
Language scale						
Tertile 1	278	ref	ref	257	ref	ref
Tertile 2	210	-1.0 (-3.8, 1.8)	1.4 (0.9, 2.3)	256	0.5 (-2.7, 3.6)	1.0 (0.7, 1.3)
Tertile 3	238	-1.6 (-4.4, 1.2)	1.4 (0.9, 2.2)	281	2.5 (-0.5, 5.4)	1.0 (0.7, 1.4)
Motor scale						
Tertile 1	278	ref	ref	257	ref	ref
Tertile 2	210	-1.6 (-4.0, 0.7)	2.6 (1.0, 6.2)	256	-1.1 (-3.4, 1.1)	1.0 (0.4, 2.4)
Tertile 3	238	-1.6 (-4.1, 0.8)	2.1 (0.8, 5.6)	281	1.0 (-0.9, 3.0)	0.6 (0.2, 1.7)
Social emotional scale						
Tertile 1	278	ref	ref	257	ref	ref
Tertile 2	210	-0.8 (-4.2, 2.6)	0.8 (0.3, 2.2)	256	1.3 (-1.6, 4.2)	1.2 (0.5, 2.9)
Tertile 3	238	-1.3 (-5.0, 2.3)	0.8 (0.3, 2.2)	281	2.3 (-0.5, 5.1)	1.1 (0.4, 3.0)
Adaptive behaviour scale						
Tertile 1	278	ref	ref	257	ref	ref
Tertile 2	210	-0.9 (-3.7, 1.8)	1.5 (0.9, 2.6)	256	0.6 (-2.1, 3.4)	0.9 (0.5, 1.6)
Tertile 3	238	-0.2 (-3.3, 2.8)	1.2 (0.6, 2.2)	281	1.7 (-0.8, 4.2)	0.7 (0.4, 1.3)

[1] TSH was categorised into tertiles in both DOMInO and PINK. DOMInO tertile 1 (Lowest): <1.3 mIU/L; tertile 2: 1.4-2.5 mIU/L; tertile 3 (Highest): ≥2.6 mIU/L. PINK tertile 1 (Lowest): <1.6 mIU/L; tertile 2: 1.7-2.7 mIU/L; tertile 3 (Highest): ≥2.8 mIU/L. The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening

questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable Poisson regression model.

Supplemental Table 7. 3 Adjusted associations between newborn TSH quintile and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley-III score	Developmental delay[2]	n	Bayley-III score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Cognitive scale						
Quintile 1	167	ref	ref	173	ref	ref
Quintile 2	149	-0.8 (-3.6, 2.0)	1.0 (0.2, 4.6)	140	-1.2 (-4.0, 1.7)	1.4 (0.7, 3.0)
Quintile 3	114	-1.9 (-5.1, 1.4)	2.2 (0.6, 7.7)	143	-2.0 (-4.8, 0.9)	1.1 (0.5, 2.4)
Quintile 4	143	-1.3(-4.4, 1.8)	1.4 (0.3, 5.3)	173	-1.0 (-3.7, 1.7)	1.1 (0.5, 2.3)
Quintile 5	153	-2.4 (-5.4, 0.6)	2.0 (0.6, 7.0)	165	0.6 (-2.2, 3.4)	0.6 (0.3, 1.5)
Language scale						
Quintile 1	167	ref	ref	173	ref	ref
Quintile 2	149	-0.3 (-3.7, 3.0)	1.2 (0.7, 2.2)	140	-3.6 (-7.4, 0.2)	1.3 (0.9, 1.9)
Quintile 3	114	-1.3 (-5.1, 2.5)	1.5 (0.9, 2.8)	143	-1.2 (-5.2, 2.9)	1.0 (0.7, 1.5)
Quintile 4	143	-0.6 (-4.2, 3.1)	1.4 (0.8, 2.6)	173	0.4 (-3.6, 4.3)	1.2 (0.8, 1.8)
Quintile 5	153	-2.7 (-6.2, 0.8)	1.4 (0.8, 2.6)	165	1.6 (-2.0, 5.4)	1.1 (0.7, 1.5)
Motor scale						
Quintile 1	167	ref	ref	173	ref	ref
Quintile 2	149	-0.5 (-3.0, 2.0)	1.1 (0.3, 4.0)	140	-0.8 (-3.5, 1.8)	0.6 (0.1, 2.3)
Quintile 3	114	-2.4 (-5.5, 0.7)	2.6 (0.9, 7.7)	143	-2.2 (-5.3, 0.9)	0.8 (0.3, 2.5)
Quintile 4	143	-1.1 (-4.2, 2.0)	1.0 (0.2, 4.2)	173	-0.6 (-3.0, 1.9)	0.7 (0.2, 2.1)
Quintile 5	153	-2.3(-5.3, 0.6)	2.7 (0.9, 8.1)	165	1.3 (1.0, -3.8)	0.4 (0.1, 1.8)
Social emotional scale						
Quintile 1	167	ref	ref	173	ref	ref
Quintile 2	149	0.7 (-2.8, 4.3)	1.0 (0.3, 3.1)	140	-1.0 (-5.0, 2.9)	2.1 (0.6, 6.8)
Quintile 3	114	-1.1 (-4.6, 2.6)	0.7 (0.2, 2.6)	143	2.1 (-1.8, 5.9)	1.3 (0.3, 4.9)
Quintile 4	143	-0.2 (-3.9, 3.6)	0.7 (0.2, 2.1)	173	0.1 (-3.5, 3.6)	2.7 (0.8, 9.0)
Quintile 5	153	0.0 (-3.8, 3.8)	0.7 (0.2, 2.5)	165	2.8 (-0.6, 6.2)	0.9 (0.2, 4.0)
Adaptive behaviour scale						
Quintile 1	167	ref	ref	173	ref	ref
Quintile 2	149	-0.9 (-5.2, 3.3)	0.9 (0.4, 1.8)	140	-3.5 (-6.9, -0.1)	1.7 (0.8, 3.7)
Quintile 3	114	-2.4 (-7.2, 2.3)	1.5 (0.7, 2.9)	143	0.0 (-3.7, 3.8)	1.0 (0.4, 2.3)
Quintile 4	143	-0.3 (-4.7, 4.0)	1.0 (0.5, 2.0)	173	0.0(-3.1, 3.2)	1.0 (0.5, 2.2)
Quintile 5	153	-2.4 (-7.1, 2.3)	1.1 (0.5, 2.4)	165	0.2 (-3.0, 3.4)	0.8 (0.4, 1.9)

[1]TSH was categorised into quintiles in both DOMInO and PINK. DOMInO quintile 1 (Lowest): <0.9 mIU/L; quintile 2: 1.0-1.5 mIU/L; quintile 3: 1.6-2.1 mIU/L; quintile 4 : 2.2-3.0 mIU/L, quintile 5 (Highest): ≥3.1 mIU/L. PINK quintile 1 (Lowest): <1.3 mIU/L; quintile 2: 0.4-1.8 mIU/L; quintile 3: 1.9-2.4 mIU/L; quintile 4 : 2.5-3.4 mIU/L, quintile 5 (Highest): ≥3.5 mIU/L. The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable Poisson regression model.

Supplemental Table 7. 4 Adjusted associations between newborn TSH dichotomised at 5 mIU/L and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
		Bayley-III score	Developmental delay[2]		Bayley-III score	Developmental delay[2]
	n	MD (95% CI)[3]	RR (95% CI)[4]	n	MD (95% CI)[3]	RR (95% CI)[4]
Cognitive scale						
TSH ≤5 mIU/L	693	ref	ref	751	ref	
TSH >5 mIU/L	33	0.2 (-4.6, 5.0)	0.8 (0.1, 7.2)	43	-0.3 (-3.6, 3.0)	0.3 (0.1, 1.6)
Language scale						
TSH ≤5 mIU/L	693	ref	ref	751	ref	
TSH >5 mIU/L	33	-1.6 (-7.2, 4.0)	1.3 (0.6, 2.5)	43	0.8 (-4.5, 6.1)	1.1 (0.7, 1.6)
Motor scale						
TSH ≤5 mIU/L	693	ref	ref	751	ref	
TSH >5 mIU/L	33	0.7 (-5.4, 6.7)	3.4 (1.0, 13.9)	43	1.9 (-1.4, 5.2)	0.8 (0.1, 4.8)
Social emotional scale						
TSH ≤5 mIU/L	693	ref	ref	751	ref	
TSH >5 mIU/L	33	-0.1 (-7.0, 6.7)	0.7 (0.2, 3.0)	43	-0.1 (-4.2, 4.0)	1.0 (0.2, 4.8)
Adaptive behaviour scale						
TSH ≤5 mIU/L	693	ref	ref	751	ref	
TSH >5 mIU/L	33	0.2 (-6.2, 6.6)	1.0 (0.3, 3.3)	43	-1.3 (-6.4, 3.9)	1.4 (0.6, 3.6)

[1]The TSH classification was based on the World Health Organization‘s recommendation to classify population iodine status based on the percentage of TSH >5 mIU/L. The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD; mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with Univariable and multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with Univariable and multivariable log Poisson regression model with robust variance estimation.

Supplemental Table 7. 5 Adjusted associations between newborn TSH quartiles, Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley –III score	Developmental delay[2]	n	Bayley –III score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CI)[3]	RR (95% CI)[4]
Cognitive scale						
Quartile 1	127	ref	ref	127	ref	ref
Quartile 2	155	-1.3 (-3.9, 1.3)	1.1 (0.2, 4.3)	155	-0.9 (-4.0, 2.0)	1.9 (0.8, 4.4)
Quartile 3	262	-1.3 (-4.1, 1.5)	2.3 (0.7, 7.2)	262	-1.5 (-4.1, 1.2)	1.3 (0.6, 2.9)
Quartile 4	250	-2.0 (-4.5, 0.7)	1.9 (0.6, 6.1)	250	0.4 (-2.2, 3.1)	0.9 (0.4, 2.0)
Language scale						
Quartile 1	127	ref	ref	127	ref	ref
Quartile 2	155	-1.0 (-4.1, 2.0)	1.4 (0.9, 2.2)	155	-2.3 (-6.7, 2.1)	1.7 (1.1, 2.6)
Quartile 3	262	-1.1 (-4.1, 2.0)	1.2 (0.7, 2.0)	262	-0.7 (-4.7, 3.3)	1.3 (0.9, 1.9)
Quartile 4	250	-2.3 (-5.5, 1.0)	1.3 (0.8, 2.2)	250	1.8 (-2.1, 5.7)	1.3 (0.8, 1.9)
Motor scale						
Quartile 1	127	ref	ref	127	ref	ref
Quartile 2	155	-2.4 (-5.0, 0.1)	3.0 (1.2, 7.3)	155	-0.5 (-3.2, 2.1)	1.2 (0.4, 4.0)
Quartile 3	262	-0.9 (-3.6, 1.7)	0.9 (0.2, 3.3)	262	-1.4 (-4.0, 1.3)	1.0 (0.3, 3.2)
Quartile 4	250	-1.8 (-4.6, 1.0)	2.3 (0.9, 6.3)	250	0.7 (-1.7, 3.2)	0.6 (0.2, 2.2)
Social emotional scale						
Quartile 1	127	ref	ref	127	ref	ref
Quartile 2	155	-1.5 (-5.6, 2.6)	0.7 (0.2, 2.0)	155	-1.2 (-5.0, 2.7)	2.1 (0.6, 7.1)
Quartile 3	262	-0.8 (-4.9, 3.2)	0.6 (0.2, 1.8)	262	1.0 (-2.4, 4.5)	1.5 (0.5 , 4.7)
Quartile 4	250	-1.4 (-5.7, 2.8)	0.8 (0.3, 2.3)	250	1.4 (-2.0, 4.9)	1.5 (0.4, 5.3)
Adaptive behaviour scale						
Quartile 1	127	ref	ref	127	ref	ref
Quartile 2	155	-1.2 (-4.3, 1.8)	1.6 (0.9, 2.8)	155	-3.3 (-6.9, 0.3)	1.6 (0.7, 34)
Quartile 3	262	-0.7 (-4.1, 2.5)	1.1 (0.6, 2.0)	262	-0.6 (-4.0, 2.8)	1.0 (0.5, 2.2)
Quartile 4	250	0.0 (-3.7, 3.6)	1.2 (0.6, 2.4)	250	0.1 (-3.1, 3.3)	0.8 (0.4, 1.8)

[1]TSH in the DOMInO study was categorised into quartiles based on the PINK quartile cut-points (quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L). TSH in the PINK study was categorised into quartiles based on the

DOMInO quartile cut-points (quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L). The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score, birthweight-for-gestational age z score and treatment group in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable Poisson regression model with robust variance estimation.

Supplemental Table 7. 6 Adjusted associations between newborn TSH~age and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley-III score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Cognitive scale						
1 mIU/L increase in TSH~age[5]	726	-0.3 (-0.8, 0.3)	1.0 (0.8, 1.3)	794	0.2 (-0.4, 0.7)	0.9 (0.7, 1.0)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-1.1 (-3.8, 1.6)	1.6 (0.5, 5.0)	179	0.0 (-2.5, 2.5)	0.9 (0.4, 1.8)
Quartile 3	173	-1.3 (-3.9, 1.4)	1.5 (0.5, 4.7)	205	-1.5 (-4.1, 0.9)	1.1 (0.6, 2.0)
Quartile 4	191	-2.1 (-4.7, 0.4)	1.6 (0.5, 5.3)	205	1.2 (-1.4, 3.8)	0.6 (0.3, 1.2)
Language scale						
1 mIU/L increase in TSH~age[5]	726	-0.3 (-1.0, 0.5)	1.0 (0.9, 1.1)	794	0.5 (-0.2, 1.2)	1.0 (1.0, 1.1)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	0.3 (-2.7, 3.4)	1.1 (0.7, 1.9)	179	-2.2 (-5.5, 1.1)	1.1 (0.8, 1.6)
Quartile 3	173	-1.1 (-4.3, 2.2)	1.4 (0.8, 2.4)	205	0.4 (-3.1, 4.0)	1.1 (0.8, 1.5)
Quartile 4	191	-1.7 (-4.9, 1.4)	1.4 (0.8, 2.2)	205	2.1 (-1.2, 5.5)	1.1 (0.8, 1.5)
Motor scale						
1 mIU/L increase in TSH~age[5]	726	-0.2 (-0.8, 0.4)	1.2 (1.0, 1.4)	794	0.2 (-0.3, 0.7)	0.9 (0.7, 1.2)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-0.6 (-3.0, 1.7)	2.1 (0.7, 6.6)	179	-1.2 (-3.5, 1.1)	0.9 (0.3, 2.6)
Quartile 3	173	-1.6 (-4.4, 1.1)	1.8 (0.6, 5.9)	205	-1.8 (-4.2, 0.7)	1.0 (0.4, 2.8)
Quartile 4	191	-1.6 (-4.3, 1.0)	2.4 (0.8, 7.5)	205	0.6 (-1.5, 2.8)	0.5 (0.1, 2.0)
Social emotional scale						

	DOMInO			PINK		
	n	Bayley-III score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
1 mIU/L increase in TSH~age[5]	726	0.3 (-0.7, 1.3)	0.9 (0.7, 1.2)	794	0.5 (-0.2, 1.1)	1.0 (0.8, 1.2)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-1.3 (-5.4, 2.8)	1.0 (0.3, 3.0)	179	0.8 (-2.8, 4.4)	1.8 (0.6, 5.5)
Quartile 3	173	-1.4 (-5.6, 2.7)	0.9 (0.3, 2.4)	205	1.9 (-1.4, 5.2)	1.4 (0.5, 4.1)
Quartile 4	191	0.1 (-4.0, 4.2)	0.7 (0.2, 2.1)	205	2.8 (-0.4, 6.1)	1.4 (0.4, 4.8)
Adaptive behaviour scale						
1 mIU/L increase in TSH~age[5]	726	0.1 (-0.6, 0.8)	1.0 (0.9, 1.1)	794	0.1 (-0.7, 0.8)	1.0 (0.8, 1.2)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-0.4 (-3.7, 2.9)	1.0 (0.5, 2.0)	179	-2.6 (-5.7, 0.5)	1.3 (0.7, 2.6)
Quartile 3	173	-1.7 (-5.0, 1.6)	1.4 (0.8, 2.7)	205	0.9 (-2.0, 4.0)	0.8 (0.4, 1.5)
Quartile 4	191	0.3 (-3.1, 3.8)	1.0 (0.5, 2.0)	205	0.7 (-2.1, 3.6)	0.7 (0.3, 1.6)

[1]The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data.CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone; TSH~age, thyroid-stimulating hormone corrected for age at blood sampling.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable Poisson regression model with robust variance estimation.

[5]TSH~age is modelled continuously.

[6]TSH~age was categorised into quartiles in both DOMInO and PINK. TSH~age was created by regressing TSH on age at blood sampling [64, 221]. DOMInO quartile 1 (Lowest): <1.0 mIU/L; quartile 2: 1.1-1.8 mIU/L; quartile 3: 1.9-2.8 mIU/L; quartile 4 (Highest): ≥2.9 mIU/L. PINK quartile 1 (Lowest): <1.4 mIU/L; quartile 2: 1.4-2.1 mIU/L; quartile 3: 2.2-3.1 mIU/L; quartile 4 (Highest): ≥3.2 mIU/L.

Supplemental Table 7. 7 Unadjusted associations between newborn TSH concentration and Bayley-III outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Bayley-III Score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Cognitive scale						
1 mIU/L increase in TSH[5]	726	-0.5 (-1.1, 0.1)	1.1 (0.9, 1.3)	794	0.1 (-0.4, 0.7)	0.9 (0.8, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-1.7 (-4.5, 1.1)	0.9 (0.3, 3.3)	196	-0.2 (-2.8, 2.4)	1.0 (0.5, 1.9)
Quartile 3	185	-1.6 (-4.5, 1.4)	1.5 (0.5, 4.5)	201	-1.2 (-3.7, 1.2)	1.2 (0.7, 2.3)
Quartile 4	181	-3.1 (-6.0, -0.2)	1.7 (0.6, 5.1)	197	1.5 (-1.0, 4.1)	0.7 (0.3, 1.4)
Language scale						
1 mIU/L increase in TSH[5]	726	-0.8 (-1.5, 0.0)	1.1 (1.0, 1.2)	794	0.4 (-0.4, 1.3)	1.0 (1.0, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.6 (-4.1, 2.8)	1.3 (0.7, 2.3)	196	-1.1 (-4.8, 2.6)	1.0 (0.7, 1.5)
Quartile 3	185	-1.9 (-5.4,1.6)	1.5 (0.9, 2.7)	201	0.4 (-3.7, 4.3)	1.1 (0.8, 1.6)
Quartile 4	181	-4.1 (-7.6, -0.7)	1.7 (1.0, 2.8)	197	1.8 (-2.0, 5.6)	1.1 (0.8, 1.5)
Motor scale						
1 mIU/L increase in TSH[5]	726	-0.4 (-1.1, 0.2)	1.2 (1.0, 1.4)	794	0.1 (-0.4, 0.6)	0.9 (0.7, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-1.2 (-3.6, 1.2)	1.4 (0.4, 4.5)	196	-1.3 (-3.7, 1.1)	0.7 (0.3, 2.0)
Quartile 3	185	-1.2 (-3.8, 1.4)	1.3 (0.4, 4.3)	201	-1.9 (-4.4, 0.7)	1.0 (0.4, 2.5)
Quartile 4	181	-2.8 (-5.5, -0.1)	2.5 (0.9, 7.4)	197	0.9 (-1.1, 3.0)	0.6 (0.2, 1.8)
Social emotional scale						
1 mIU/L increase in TSH[5]	726	-0.3 (-1.3, 0.6)	1.0 (0.8, 1.2)	794	0.2 (-0.4, 0.8)	1.0 (0.8, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.4 (-4.5, 3.6)	0.9 (0.3, 2.5)	196	0.7 (-2.6, 3.9)	1.5 (0.6, 3.9)
Quartile 3	185	-0.8 (-4.9, 3.4)	0.9 (0.3, 2.4)	201	1.2 (-2.1, 4.5)	1.5 (0.5, 3.9)

	DOMInO			PINK		
	n	Bayley-III Score	Developmental delay[2]	n	Bayley-III Score	Developmental delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
Quartile 4	181	-3.4 (-7.7, 1.0)	1.0 (0.4, 2.6)	197	2.1 (-1.0, 5.2)	1.0 (0.3, 3.0)
Adaptive behaviour scale						
1 mIU/L increase in TSH[4]	726	-0.2 (-1.0, 0.5)	1.0 (0.9, 1.2)	794	0.0 (-0.7, 0.8)	1.0 (0.8, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-3.5, 3.3)	1.2 (0.6, 2.2)	196	-1.9 (-5.1, 1.4)	1.6 (0.8, 3.0)
Quartile 3	185	-0.1 (-3.5, 3.4)	1.1 (0.5, 2.2)	201	0.3 (-2.8, 3.5)	0.9 (0.4, 1.8)
Quartile 4	181	-1.2 (-4.9, 2.4)	1.4 (0.7, 2.8)	197	0.7 (-2.3, 3.7)	0.8 (0.4, 1.6)

[1]CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference; RR, relative risks; TSH, thyroid-stimulating hormone; SD, standard deviation.

[2]Developmental delay was defined as Bayley-III scores below 85.

[3]The Mean Differences (95% CIs) were estimated using Univariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated using Univariable Poisson regression model with robust variance estimation.

[5]TSH is modelled continuously.

[6]TSH was categorised into quartiles in both DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L.

Supplemental Table 7. 8 Adjusted associations between newborn TSH dichotomised at 5 mIU/L and growth outcomes at 18 months of age[1]

	DOMInO			PINK		
		Anthropometric z score	Growth delay[2]		Anthropometric z score	Growth delay[2]
	n	MD (95% CI)[3]	RR (95% CI)[4]	n	MD (95% CI)[3]	RR (95% CI)[4]
WLZ						
TSH ≤5 mIU/L	693	ref		751	ref	ref
TSH >5 mIU/L	33	-0.2 (-0.5, 0.1)	[5]	43	0.1 (-0.2, 0.5)	0.7 (0.2, 2.5)
LAZ						
TSH ≤5 mIU/L	693	ref	ref	751	ref	ref
TSH >5 mIU/L	33	0.1 (-0.3, 0.4)	1.0 (0.4, 2.5)	43	0.2 (-0.2, 0.6)	0.4 (0.1, 1.2)
WAZ						
TSH ≤5 mIU/L	693	ref	ref	751	ref	ref
TSH >5 mIU/L	33	-0.1 (-0.4, 0.2)	1.0 (0.2, 4.7)	43	0.1 (-0.3, 0.4)	0.5 (0.1, 2.0)
HCZ						
TSH ≤5 mIU/L	693	ref		751	ref	ref
TSH >5 mIU/L	33	-0.1 (-0.4, 0.2)	[5]	43	0.0 (-0.5, 0.5)	0.5 (0.1, 3.2)

[1]The TSH classification was based on the World Health Organization‘s recommendation to classify population iodine status based on the percentage of TSH >5 mIU/L. The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD; mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

[2]Growth delay was defined as z scores below 1 SD.

[3]The Mean Differences (95% CIs) were estimated with Univariable and multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with Univariable and multivariable log Poisson regression model with robust variance estimation.

[5]The Relative Risks (95% CIs) and the confidence intervals are not presented because the model did not converge.

Supplemental Table 7. 9 Adjusted associations between newborn TSH~age and growth outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Anthropometric z core	Growth delay[2]	n	Anthropometric z score	Growth delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CI)[3]	RR (95% CI)[4]
WLZ						
1 mIU/L increase in TSH~age[5]	726	0.0 (0.0, 0.1)	0.8 (0.6, 1.1)	794	0.0 (-0.1, 0.0)	1.0 (0.9, 1.2)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	0.0 (-0.2, 0.2)	1.1 (0.2, 5.7)	179	-0.1 (-0.3, 0.1)	0.9 (0.3, 3.2)
Quartile 3	173	0.0 (-0.2, 0.2)	1.1 (0.3, 4.0)	205	-0.1 (-0.4, 0.1)	1.0 (0.3, 3.3)
Quartile 4	191	0.1 (-0.1, 0.3)	0.5 (0.1, 2.5)	205	-0.1 (-0.5, 0.2)	1.2 (0.4, 3.2)
LAZ						
1 mIU/L increase in TSH~age[5]	726	0.0 (-0.1, 0.1)	1.0 (0.9, 1.1)	794	0.0 (-0.1, 0.0)	1.0 (0.9, 1.1)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-0.1 (-0.3, 0.1)	1.3 (0.7, 2.4)	179	0.0 (-0.2, 0.3)	1.2 (0.6, 2.3)
Quartile 3	173	-0.2 (-0.4, 0.0)	1.2 (0.7, 2.1)	205	0.0 (-0.2, 0.3)	1.3 (0.7, 2.3)
Quartile 4	191	-0.2 (-0.4, 0.1)	1.3 (0.7, 2.1)	205	-0.2 (-0.4, 0.1)	1.4 (0.8, 2.5)
WAZ						
1 mIU/L increase in TSH~age[5]	726	0.0 (0.0, 0.1)	1.0 (0.8, 1.2)	794	-0.1 (-0.2, 0.0)	1.1 (0.9, 1.2)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	0.0 (-0.2, 0.2)	0.5 (0.1, 2.6)	179	-0.2 (-0.5, 0.2)	1.1 (0.3, 3.7)
Quartile 3	173	-0.1 (-0.3, 0.1)	1.3 (0.5, 3.6)	205	-0.3 (-0.7, -0.2)	1.4 (0.5, 3.7)
Quartile 4	191	0.0 (-0.2, 0.2)	1.1 (0.4, 2.9)	205	-0.4 (-0.8, -0.1)	1.9 (0.7, 5.0)
HCZ						
1 mIU/L increase in TSH~age[5]	726	0.0 (-0.1, 0.0)	1.1 (0.9, 1.3)	794	0.0 (-0.1, 0.1)	0.8 (0.7, 1.1)
TSH~age in quartiles[6]						
Quartile 1	198	ref	ref	198	ref	ref
Quartile 2	164	-0.2 (-0.4, 0.1)	2.2 (0.6, 8.6)	179	-0.1 (-0.4, 0.3)	1.4 (0.4, 4.7)
Quartile 3	173	-0.2 (-0.4, 0.1)	3.3 (1.0, 10.8)	205	-0.2 (-0.5, 0.2)	1.0 (0.2, 4.1)
Quartile 4	191	-0.1 (-0.3, 0.1)	2.8 (0.8, 8.7)	205	0.0 (-0.5, 0.5)	0.8 (0.2, 2.9)

[1]The regression models were adjusted for sex, parity, race, occupation, education, qualification, smoking during second trimester, previous smoking, alcohol consumption during second trimester, previous alcohol consumption, supplement intake during second trimester, previous depression, depression during second trimester, mode of delivery, mothers age, gestational age at birth, 5 minute Apgar score, mothers BMI, home screening questionnaire score and birthweight-for-gestational age z score in both DOMInO and PINK studies. In addition, treatment group in the DOMInO study was also added to the adjusted models when analysing the DOMInO data. CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD, mean difference; n, number of participants, PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference category; RR, relative risks; TSH, thyroid-stimulating hormone; TSH~age, thyroid stimulating hormone corrected for age at blood sampling; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

[2]Growth delay was defined as z scores below 1 SD.

[3]The Mean Differences (95% CIs) were estimated with multivariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with multivariable Poisson regression model with robust variance estimation.

[5]TSH~age is modelled continuously.

[6]TSH~age was categorised into quartiles in both DOMInO and PINK. TSH~age was created by regressing TSH on age at blood sampling [64, 221]. DOMInO quartile 1 (Lowest): <1.0 mIU/L; quartile 2: 1.1-1.8 mIU/L; quartile 3: 1.9-2.8 mIU/L; quartile 4 (Highest): ≥2.9 mIU/L. PINK quartile 1 (Lowest): <1.4 mIU/L; quartile 2: 1.4-2.1 mIU/L; quartile 3: 2.2-3.1 mIU/L; quartile 4 (Highest): ≥3.2 mIU/L.

Supplemental Table 7. 10 Unadjusted associations between newborn TSH concentration and growth outcomes at 18 months of age[1]

	DOMInO			PINK		
	n	Anthropometric z score	Growth delay[2]	n	Anthropometric z score	Growth delay[2]
		MD (95% CI)[3]	RR (95% CI)[4]		MD (95% CIs)[3]	RR (95% CI)[4]
WLZ						
1 mIU/L increase in TSH[5]	726	0.0 (-0.1, 0.0)	0.9 (0.7, 1.2)	794	-0.1 (-0.1, 0.0)	1.2 (1.0, 1.4)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-0.3, 0.2)	0.8 (0.2, 3.3)	196	-0.2 (-0.4, 0.0)	1.5 (0.5, 5.0)
Quartile 3	185	-0.1 (-0.3, 0.2)	1.6 (0.5, 4.7)	201	-0.3 (-0.5, 0.0)	1.9 (0.6, 6.3)
Quartile 4	181	0.0 (-0.3, 0.2)	0.8 (0.2, 3.5)	197	-0.3 (-0.6, 0.1)	1.8 (0.5, 6.1)
LAZ						
1 mIU/L increase in TSH[5]	726	0.0 (-0.1, 0.0)	1.1 (0.9, 1.2)	794	-0.1 (-0.1, 0.0)	1.0 (0.9, 1.1)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	0.0 (-0.3, 0.2)	1.1 (0.6, 2.0)	196	0.1 (-0.2, 0.3)	1.2 (0.6, 2.3)
Quartile 3	185	-0.1 (-0.3, 0.2)	1.1 (0.6, 1.9)	201	-0.1 (-0.4, 0.1)	1.9 (1.1, 3.3)
Quartile 4	181	-0.2 (-0.5, 0.0)	1.6 (0.9, 2.7)	197	-0.2 (-0.5, 0.1)	1.4 (0.7, 2.6)
WAZ						
1 mIU/L increase in TSH[5]	726	0.0 (-0.1, 0.0)	1.0 (0.8, 1.3)	794	-0.1 (-0.2, 0.1)	1.2 (1.0, 1.3)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-0.3, 0.2)	0.4 (0.1, 1.5)	196	-0.2 (-0.6, 0.2)	1.4 (0.5, 4.3)
Quartile 3	185	-0.1 (-0.3, 0.2)	0.6 (0.2, 1.7)	201	-0.4 (-0.8, 0.1)	2.1 (0.8, 5.2)
Quartile 4	181	-0.1 (-0.4, 0.1)	1.3 (0.6, 2.9)	197	-0.5 (-0.8, 0.2)	2.1 (0.8, 5.6)
HCZ						
1 mIU/L increase in TSH[5]	726	0.0 (-0.1, 0.0)	1.1 (0.9, 1.3)	794	-0.1 (-0.1, 0.0)	0.9 (0.7, 1.2)
TSH in quartiles[6]						
Quartile 1	184	ref	ref	200	ref	ref
Quartile 2	176	-0.1 (-0.3, 0.1)	1.1 (0.3, 4.4)	196	-0.1 (-0.5, 0.2)	1.2 (0.3, 4.2)
Quartile 3	185	-0.1 (-0.3, 0.1)	1.8 (0.6, 6.0)	201	-0.3 (-0.6, 0.1)	1.7 (0.6, 5.3)
Quartile 4	181	-0.2 (-0.4, 0.1)	2.0 (0.7, 6.5)	197	-0.1 (-0.6, 0.4)	0.8 (0.2, 3.1)

[1]CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD; mean difference; N, number of participants, p, p-value; PINK, Pregnancy Iodine and Neurodevelopment in Kids; ref, reference; RR, relative risks; TSH, thyroid-stimulating hormone; SD, standard deviation; WAZ, weight-for-age z score; WLZ, weight-for-length z score.

[2]Growth delay was defined as z scores below 1 SD.

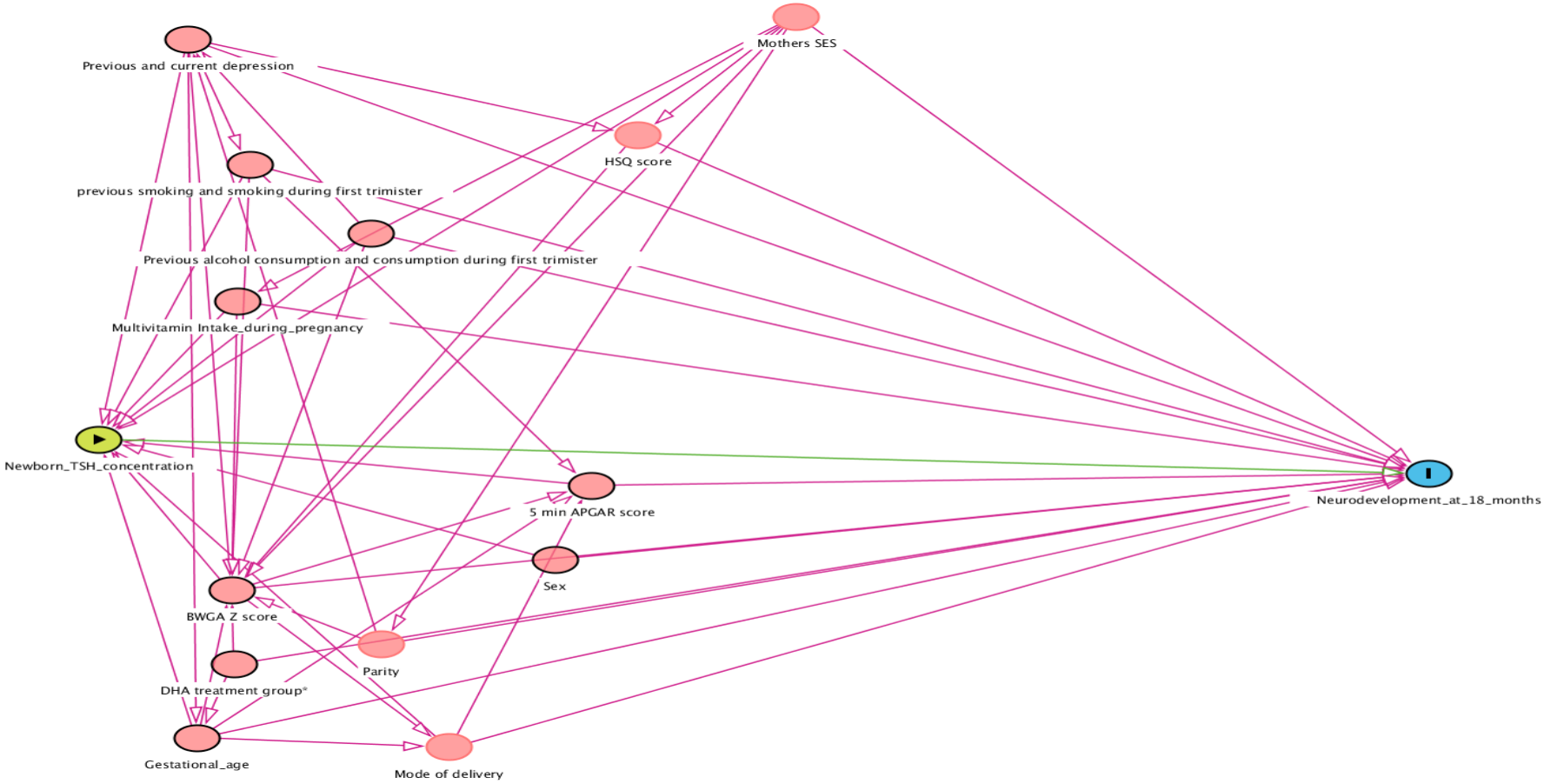
[3]The Mean Differences (95% CIs) were estimated with Univariable linear regression model.

[4]The Relative Risks (95% CIs) were estimated with Univariable Poisson regression model with robust variance estimation.

[5]TSH~age is modelled continuously.

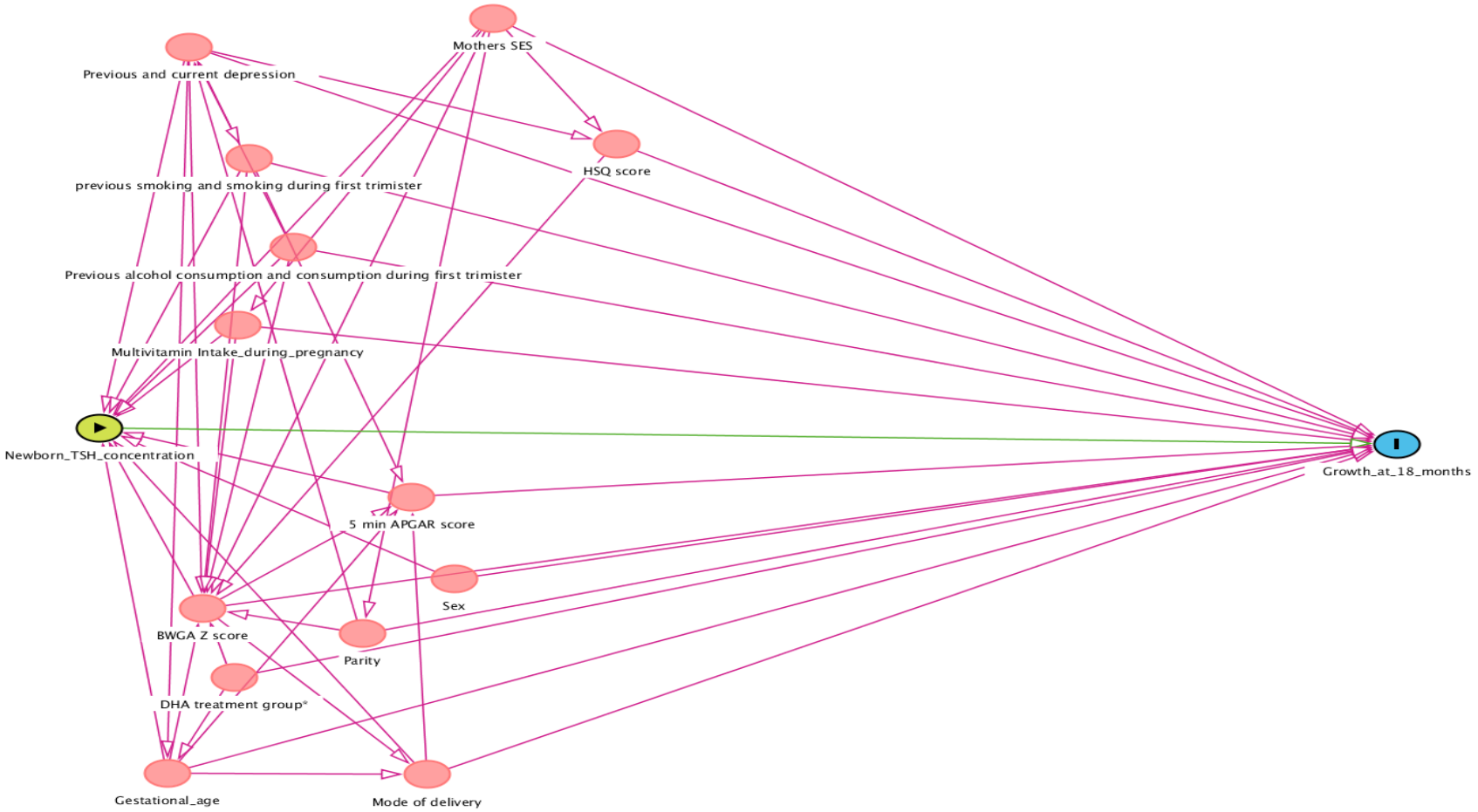
[6]TSH was categorised into quartiles in both DOMInO and PINK. DOMInO quartile 1 (Lowest): <1.1 mIU/L; quartile 2: 1.1-1.7 mIU/L; quartile 3: 1.8-2.9 mIU/L; quartile 4 (Highest): ≥3.0 mIU/L. PINK quartile 1 (Lowest): <1.5 mIU/L; quartile 2: 1.5-2.2 mIU/L; quartile 3: 2.3-3.3 mIU/L; quartile 4 (Highest): ≥3.4 mIU/L.

Appendix II Supplemental Figures



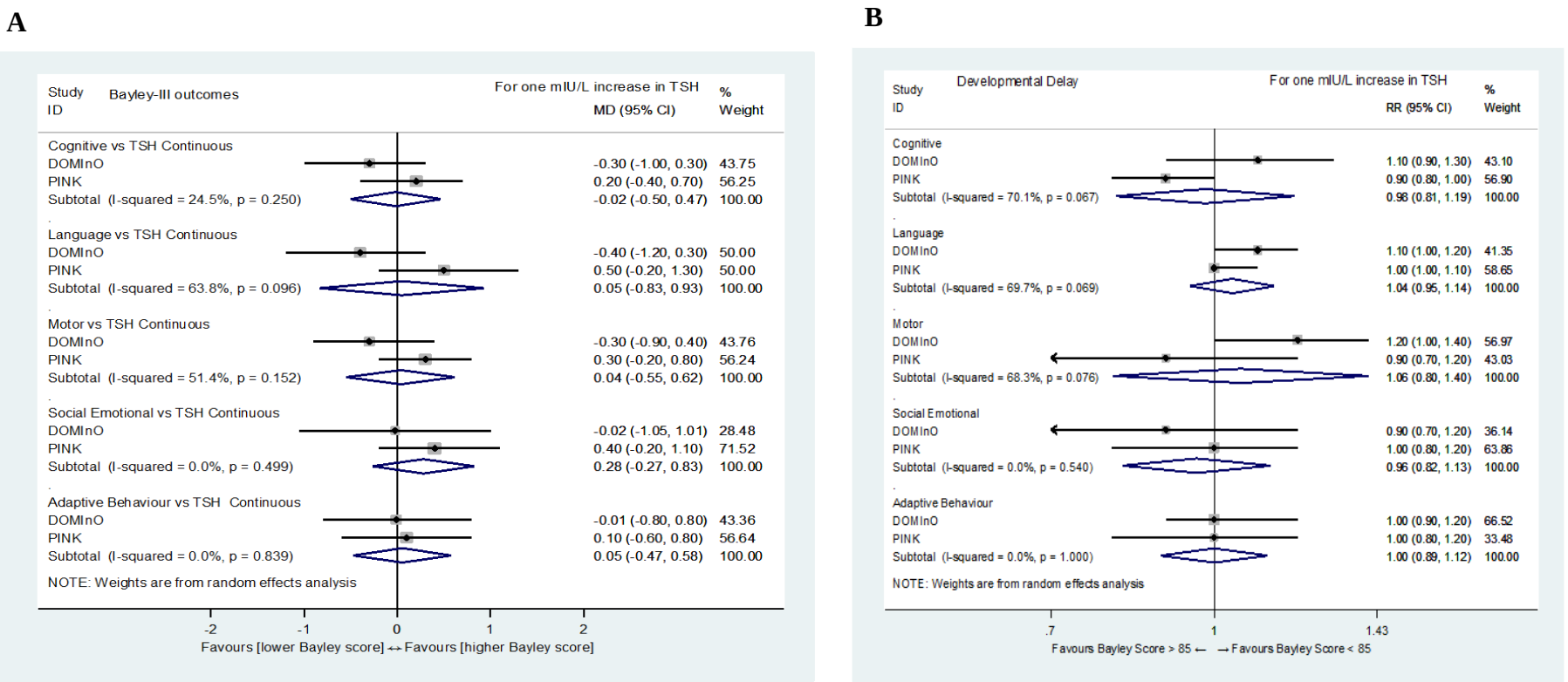
Supplemental Figure 7. 1 Direct Acyclic Graph for association between newborn TSH and child neurodevelopment.

BWGA Z score, birth-weight for gestational age z score; DHA, Docosahexaenoic acid; HSQ, home screening questionnaire; SES, socio-economic status; TSH, thyroid stimulating hormone. DHA treatment group was only for the DHA to Optimize Mother Infant Outcome (DOMInO) study [274].



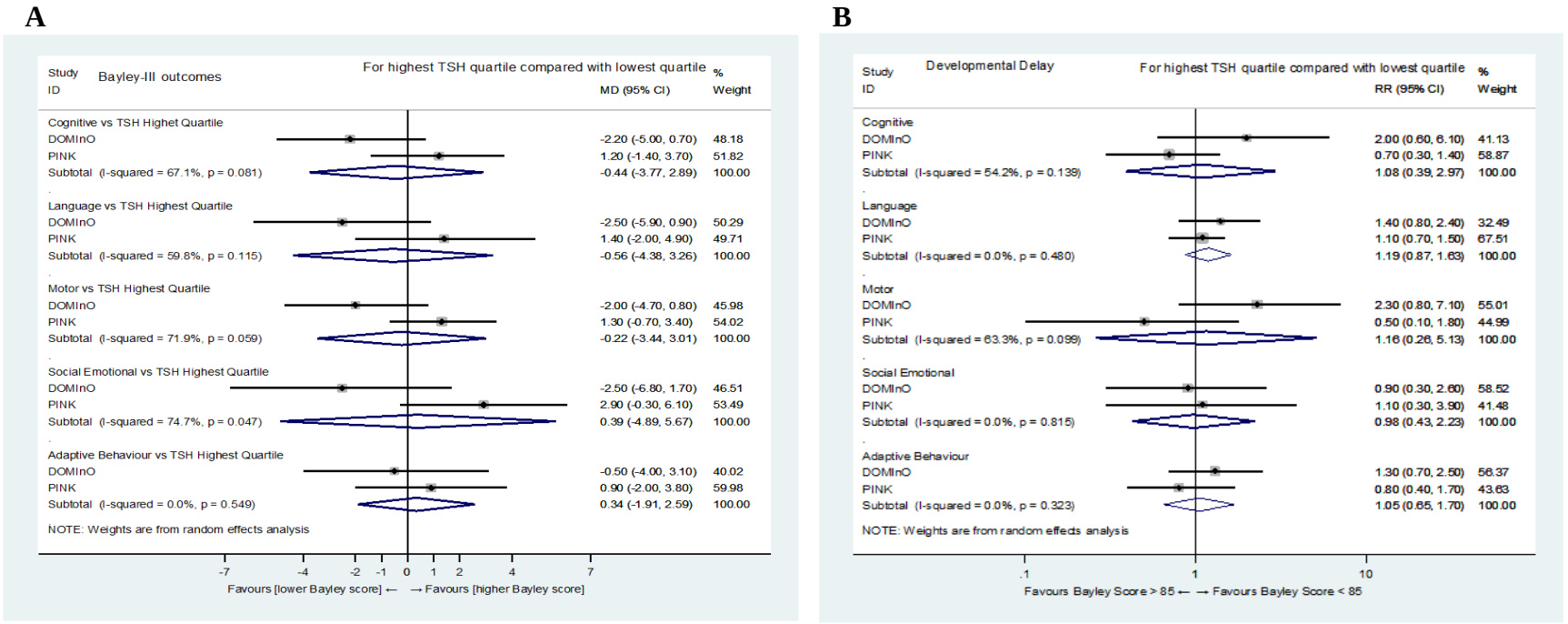
Supplemental Figure 7. 2 Direct Acyclic Graph for association between newborn TSH and child growth.

BWGA Z score, birth-weight for gestational age z score; DHA, Docosahexaenoic acid; HSQ, home screening questionnaire; SES, socio-economic status; TSH, thyroid stimulating hormone. DHA treatment group was only for the DHA to Optimize Mother Infant Outcome (DOMInO) study [274].



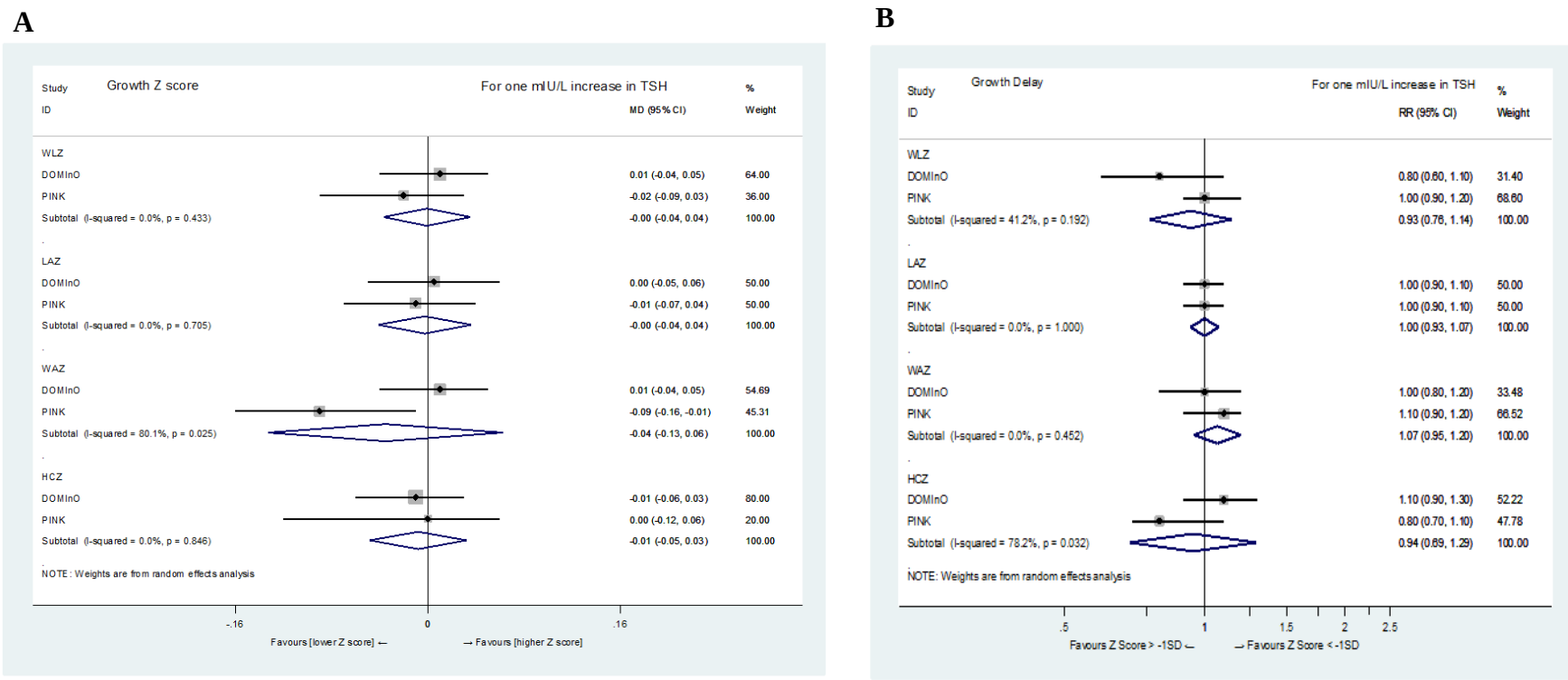
Supplemental Figure 7. 3 Associations between newborn TSH as continuous exposure and A) Bayley-III scores and B) neurodevelopmental delay at 18 months of age.

CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; PINK, Pregnancy Iodine and Neurodevelopment in Kids; RR, relative risk; TSH, thyroid-stimulating hormone.



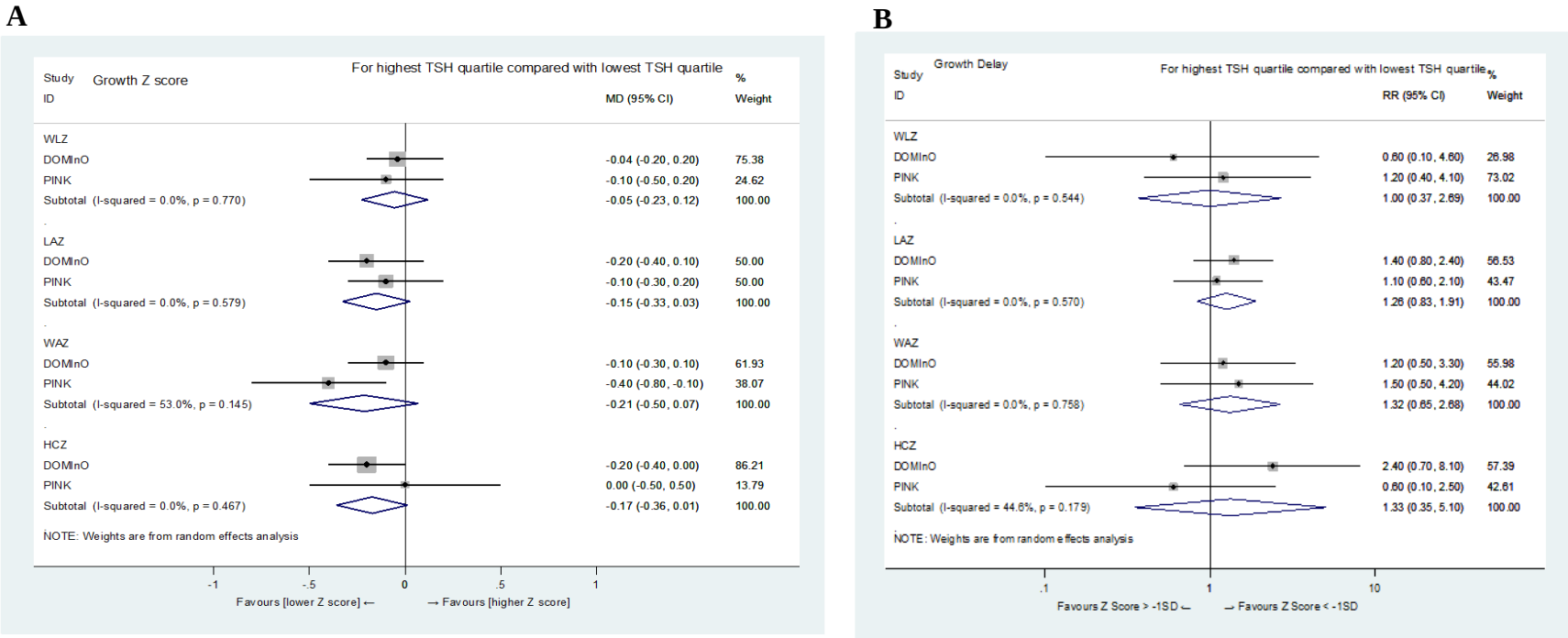
Supplemental Figure 7. 4 Associations between newborn TSH in quartiles and A) Bayley-III scores and B) neurodevelopmental delay at 18 months of age.

CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; MD, mean difference; PINK, Pregnancy Iodine and Neurodevelopment in Kids; RR, relative risk; TSH, thyroid-stimulating hormone.



Supplemental Figure 7. 5 Associations between newborn TSH as continuous exposure and A) anthropometric indices and B) growth delay at 18 months of age.

CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD, mean difference; PINK, Pregnancy Iodine and Neurodevelopment in Kids; RR, relative risk; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.



Supplemental Figure 7. 6 Associations between newborn TSH in quartiles and A) anthropometric indices and B) growth delay at 18 months of age.

CI, confidence interval; DOMInO, DHA to Optimize Mother Infant Outcome; HAZ, length-for-age z score; HCZ, head circumference-for-age z score; MD, mean difference; PINK, Pregnancy Iodine and Neurodevelopment in Kids; RR, relative risk; TSH, thyroid-stimulating hormone; WAZ, weight-for-age z score; WLZ, weight-for-length z score.