British Journal of Nutrition (2017), **117**, 441–449
doi:10.1017/S0007114517000162

Reliability of thyroglobulin in serum compared with urinary iodine when assessing individual and population iodine nutrition status

Stig Andersen[1,2,3]*, Paneeraq Noahsen[1,4], Louise Westergaard[2] and Peter Laurberg[1,4]

[1]*Arctic Health Research Centre, Institute of Clinical Medicine, Aalborg University Hospital, 9000 Aalborg, Denmark*
[2]*Department of Geriatric and Internal Medicine, Aalborg University Hospital, 9000 Aalborg, Denmark*
[3]*Ilisimatusarfik, University of Greenland, 3400 Nuuk, Greenland*
[4]*Department of Endocrinology, Aalborg University Hospital, 9000 Aalborg, Denmark*

(Submitted 10 September 2016 – Final revision received 13 December 2016 – Accepted 3 January 2017 – First published online 22 February 2017)

Abstract

The occurrence of thyroid disorders relies on I nutrition and monitoring of all populations is recommended. Measuring I in urine is standard but thyroglobulin in serum is an alternative. This led us to assess the reliability of studies using serum thyroglobulin compared with urinary I to assess the I nutrition level and calculate the number of participants needed in a study with repeated data sampling in the same individuals for 1 year. Diet, supplement use and life style factors were assessed by questionnaires. We measured thyroglobulin and thyroglobulin antibodies in serum and I in urine. Participants were thirty-three Caucasians and sixty-four Inuit living in Greenland aged 30–49 years. Serum thyroglobulin decreased with rising I excretion (Kendall's τ −0·29, $P=0{\cdot}005$) and did not differ with ethnicity. Variation in individuals was lower for serum-thyroglobulin than for urinary I (mean individual CV: 15·1 *v.* 46·1 %; $P<0{\cdot}01$). It required 245 urine samples to be 95 % certain of having a urinary I excretion within 10 % of the true mean of the population. For serum-thyroglobulin the same precision required 206 samples. In an individual ten times more samples were needed to depict I deficiency when using urinary I excretion compared with serum-thyroglobulin. In conclusion, more participants are need to portray I deficiency in a population when using urinary I compared with serum-thyroglobulin, and about ten times more samples are needed in an individual. Adding serum-thyroglobulin to urinary I may inform surveys of I nutrition by allowing subgroup analysis with similar reliability.

Key words: Thyroglobulin: Urinary iodine excretion: Number of samples needed: Reliability of iodine nutrition surveillance: Iodine nutrition study design

I nutrition is important and the World Health Organization[1] recommends monitoring of all populations. Urinary iodine concentration (UIC) is a recommended method to assess I status in a population but it has limitations. First, it portrays the I intake over the preceding hours only[2–5]. Second, it holds uncertainties due to a marked variation in urinary I excretion[5,6]. Hence, a considerable number of individuals or samples are required for a reliable estimate of I nutrition level[7,8]. Alternative approaches to assess I nutrition include measurement of thyroglobulin in serum (s-TG)[5,9–11] and thyroid volume by ultrasound[12,13]. They differ from UIC in that thyroid volume reflects the I status over the preceding years and it correlates positively to s-TG[14].

TG is a protein produced exclusively by the thyroid gland. It plays an important role in the synthesis of thyroid hormones and an increased amount of TG is released into the blood in I deficiency[5,9–11]. Thus, s-TG is a sensitive marker of I deficiency in a population.

S-TG is used to monitor patients treated for differentiated thyroid cancer and the effectiveness is well established in this group of patients[15,16]. However, even though it is well recognised that s-TG is elevated with I deficiency, data are lacking on the number of samples needed and on the reliability of s-TG as a measure of I deficiency. These measures are well established for urinary I excretion[7,8,17] and similar estimates are relevant for s-TG.

This led us to conduct a survey with repeated collection of serum and urine for measurement of s-TG and I in urine. Data were used to establish (a) the number of participants needed in surveys of I nutrition and (b) to calculate the reliability of studies of I nutrition by both s-TG and UIC with a given number of participants. Finally, we assessed the potential benefit of the combined use of both methods.

Methods

Participants

We enrolled ninety-seven healthy subjects living in Ilulissat or Saqqaq in the Disco Bay area in North Greenland. A random

Abbreviations: s-TG, serum thyroglobulin; TGAb, thyroglobulin antibodies; UIC, urinary iodine concentration.

* **Corresponding author:** Professor S. Andersen, email stiga@dadlnet.dk

sample of individuals was drawn from the National Civil Registration System in which every person living in Denmark, the Faeroe Islands and Greenland is registered. Subjects invited to participate showed an interest in supporting the points raised and agreed to contribute. In all, eighty-five out of ninety-seven (88 %) participated in three or four of the data collections. Participants in Ilulissat were stratified by age, sex and place of birth to approximate equal participation in the age groups 30–39 and 40–49 years, men and women, and the three groups consisting of (a) subjects not born in Greenland, (b) subjects born in Greenland living in town and (c) in a settlement. Subjects with both parents born in Greenland are hereafter named Inuit, and subjects whose parents are both born outside Greenland are named Caucasians.

Procedures

Data were collected four times during a full year to include also seasonal differences in the estimate of individual variance (30 March through 10 April, 25 June through 5 July, 25 September through 5 October and 7 January through 19 January).

Participants were invited by letter to participate in the investigation. The investigation took place at the local hospital or nursing station or, at request, as home visits.

A physical examination was performed by one of the doctors collecting data (S. A., P. L.). We measured height and weight in indoor clothing, calculated BMI(weight in kilograms divided by height in metres squared) and recorded if any disability was present. Information on smoking habits (present/past/never), alcohol intake (units/week), use of I containing supplements and medication was obtained by a questionnaire. None took medication containing I or known to influence the thyroid. Information regarding sex and age was obtained from the National Civil Registration System.

Information on dietary habits was obtained by a FFQ as described previously[18]. In brief, the frequency of intake of seven traditional Inuit food items and seven imported food items were given in six categories ranging from never to daily. Inuit food items scored positively and imported food items scored negatively. The sum of food frequency score for all food items consumed by each participant was calculated based on this recording and participants were categorised into groups of intake of <40, 40–60 and >60 % traditional Inuit food items scores on a scale where 100 % was purely Inuit foods and 0 % was purely imported food. Moreover, participants were asked how many days of the week the main meal was of Greenlandic food items and the number of days it was imported foods for cross-validation.

A non-fasting blood sample was drawn from the antecubital vein using minimal tourniquet. Blood was allowed to clot and serum was separated and kept at −20°C until analysis. A spot urine sample was collected in I-free polyethylene containers and stored at −20°C until analysis.

Ethical approval was obtained from the Commission for Scientific Research in Greenland before the commencement of this study (505-63), and all subjects gave informed written consent in Danish or Greenlandic by participant choice.

Assays

TG was measured in serum by the LUMItest (BRAHMS) that had a working range of 1–500 μg/l. All samples from an individual were included in the same assay run. Median values about 9, 10 and 15 μg/l are seen in I replete, mildly deficient and moderately deficient Caucasians, respectively[9,10]. Thyroglobulin antibodies (TGAb) were measured using Dynotest RIA (BRAHMS Diagnostica) with a functional sensitivity of 20 kU/l. TGAb in serum did not influence measurement of s-TG when <100 U/ml and all individuals with TGAb <100 U/ml were included in the analysis. Thus, eighty-one participants were included in the analysis including s-TG.

Urine samples were analysed for I by using the Sandell–Kolthoff reaction modified after Wilson & van Zyl[19] as described in detail previously[18,20].

Statistics and calculations

Population characteristics were compared using Mann–Whitney U test for comparison of two groups. Kendall's τ was used to describe associations between groups. Frequencies were compared using Fisher's exact test or χ^2 test with groups of less than five participants included in the adjacent group when appropriate.

The number of samples needed to assess the I status of a population or an individual was calculated from the equation developed to estimate the precision of a set-point, D, in biochemical variables[21]: $n = (Z \times \mathrm{CV}\%/D)^2$. Similar calculations are recommended for use when estimating the number of specimens required in biochemical measures. This provides an estimate of the certainty or reliability of the results of sample collections. It is described in detail by Fraser & Harris[21] and we followed these recommendations in keeping with our previous reports on urinary I in different populations[7,8]. The percentiles of standard normal distribution (Z) used were 2·58 for 99 %, 2·33 for 98 %, 1·96 for 95 %, 1·64 for 90 %, 1·28 for 80 %, 1·04 for 70 %, 0·84 for 60 %, 0·67 for 50 %. The precision range (D), a measure of reliability of the assessment of I nutrition level, used in the calculations was set to vary from ±50 to ±1 %. Using the z-statistics may underestimate the sample size for small n by up to 30 % compared with using t-statistics but this was chosen in order to comply with the recommendations as noted above[7]. Mean within-individual variances were similar whether assessed as the mean variance among individual or using ANOVA techniques. Mean, highest and lowest within-individual CV was used for calculation of number of samples needed to assess the I excretion in an individual. The CV% was the percentage of variance square root divided by the mean. Variances were compared by Levene's test for homogeneity of variances.

The statistical program for the social sciences version 13.0, Corel Quattro Pro X3 and a Texas Instruments TI-30X IIS calculator were used to process data and perform the calculations.

Results

Totally, ninety-seven participants were enrolled with thirty-three Caucasians, thirty-nine Inuit in the town Ilulissat and

Table 1. Subjects enrolled in the study of reliability of measures of population iodine deficiency and of number of subjects needed† (Numbers and percentages; medians and interquartile ranges (IQR); mean values and standard deviations)

	All participants		Caucasians		Inuit		
	n	%	*n*	%	*n*	%	*P**
	97	100	33	34·0	64	66·0	
Age (years)							NS
30–39	55	56·7	21	63·6	34	53·1	
40–49	42	43·3	12	36·4	30	46·9	
Sex							NS
Men	48	49·5	18	54·5	30	46·9	
Women	49	50·5	15	45·5	34	53·1	
BMI (kg/m^2)							NS
<18·5	1	1·1	0	0·0	1	1·6	
18·5–25	39	41·1	15	45·5	24	38·7	
25–30	34	35·8	11	33·3	23	37·1	
30+	21	22·1	7	21·2	14	22·6	
Inuit diet‡							0·004
<40 %	67	69·1	29	87·8	38	59·4	
40–60 %	29	29·9	4	12·1	25	39·1	
60+%	1	1·0	0	0·0	1	1·6	
Smoker§							
Present	50	52·6	12	37·5	38	60·3	0·054
Past	15	15·8	5	15·6	10	15·9	
Never	30	31·6	15	46·9	15	23·8	
Alcohol‖¶							0·029
Abstainers	17	17·5	1	3·1	16	25·8	
<7 units/week	69	73·4	27	84·4	25	40·3	
7+ units/week	25	26·6	4	12·5	21	33·9	
Data collections**							NS
<3	12	12·4	3	9·1	9	14·1	
3 or 4	85	97·6	30	90·9	55	85·9	
4	73	75·3	27	81·8	46	71·9	
TGAb (U/ml)							
>20††	29	32·2	10	30·3	19	33·3	NS
>100‡‡	5	5·2	1	3·0	4	6·3	
Urinary I (μg/l)							
Median	87		78		88		0·036
IQR	72		66		82		
Mean	116		103		128		
SD	96		84		136		

TGAb, thyroglobulin antibodies.
* *P* value for ethnic differences: χ^2 test, except alcohol that was tested using Fisher's exact test and urinary I using Mann–Whitney test; NS designates $P > 0{\cdot}1$. Groups with less than five participants were merged with adjacent groups (BMI <25 kg/m^2; Inuit diet scores <40 %; alcohol <7 units/week; less than four data collections).
† Of the 97 subjects enrolled, eighty-one were included in the calculations as they participated in three or four data collections and did not have antibodies against thyroglobulin that could influence the measurement of thyroglobulin.
‡ Calculated from a FFQ on the intake of seven Greenlandic and seven imported food items.
§ 2 missing.
‖ 4 missing.
¶ One unit equals 2 g of alcohol.
** Number of participations in data collection with questionnaire and collection of specimens.
†† Level for positive thyroglobulin antibody as given by the manufacturer.
‡‡ Serum TG was influenced only when TGAb was above 100.

twenty-five Inuit in the settlement Saqqaq (Table 1). None of the participants had disease affecting the thyroid, were pregnant or took I-containing supplements. The number of subjects participating in four, three, two and one data collection was seventy-three, twelve, nine and three. Participants who attended less than three times (*n* 12) and participants who had TGAb that could affect the measurement of s-TG (*n* 5) were excluded from the calculations. The calculations thus included eighty-one subjects.

Caucasians were bigger than Inuit (men/women, height $P < 0{\cdot}001/P < 0{\cdot}001$; weight $P = 0{\cdot}01/P = 0{\cdot}026$). Still, BMI was similar (Table 1). Height, weight and BMI did not differ with residence or age (residence, all, $P > 0{\cdot}1$; age in Caucasians, all, $P > 0{\cdot}1$; age in Inuit, women's height, $P = 0{\cdot}054$, all other $P > 0{\cdot}1$). Dietary habits differed with ethnicity, and more Inuit than Caucasians were smokers and alcohol abstainers. Inuit had slightly higher urinary I excretion (Table 1).

The frequency of TGAb positive individuals did not differ with ethnicity (Table 1). Median s-TG did not differ with ethnicity ($P = 0{\cdot}70$). S-TG was higher when UIC was <50 μg/l, and the frequency of s-TG >20 μg/l decreased with rising UIC (Kendall's τ −0·29, $P = 0{\cdot}005$) (Fig. 1).

Variation

Table 2 shows mean, variance and CV% for s-TG and UIC among Inuit and Caucasians. Overall, between individual, mean individual and median individual variation was higher in Inuit compared with Caucasians. Overall variation was higher in UIC compared with s-TG ($P<0{\cdot}01$) and the two measures of individual variation were markedly higher for UIC compared with s-TG (Table 2).

Number of samples needed

Table 3 lists the number of samples need for a chosen precision range of 1–50 % with a confidence of 95 %. This is shown both for an individual (right columns) and for populations (left column). It can be seen that 982 subjects need to donate one sample each to be 95 % certain of having a UIC within 5 % of the true mean of that population. For s-TG the same precision requires 823 subjects. As can be read from Table 3, 20 % more participants are needed to portray I deficiency in a population using UIC compared with s-TG.

For an individual, the results of two samples gives a precision range of 20 % when using s-TG, whereas the same precision range for UIC requires twenty-one urine samples from that individual. In addition, a single blood sample for s-TG measurement in an individual gives a precision of 30 % CI, whereas nine urine samples are need from that individual to match this precision with 95 % CI. The difference between individuals in variation for s-TG is lower than that found for urinary I.

Table 4 lists the number of samples needed in an individual (right column) and the number of participants needed in population surveys (left column) to detect I deficiency when the CI is chosen to vary in parallel with the precision range. In a population, a 10 % precision with 90 % CI requires 144 participants for s-TG and 172 participants for urinary I excretion. A precision range of 2 % with 98 % CI requires 7268 participants for s-TG and 8669 participants for urinary I analysis. Thus, 20 % more participants are needed to settle I deficiency when using UIC compared with s-TG.

For an individual, a precision range of 20 % with 80 % CI requires a single sample for s-TG and nine samples for UIC. A precision range of 10 % with 90 % CI requires six blood samples for s-TG and fifty-eight urine samples for UIC. Thus, just under ten times more urine samples are needed in an individual to depict I deficiency when using urinary I excretion compared with measurement of s-TG.

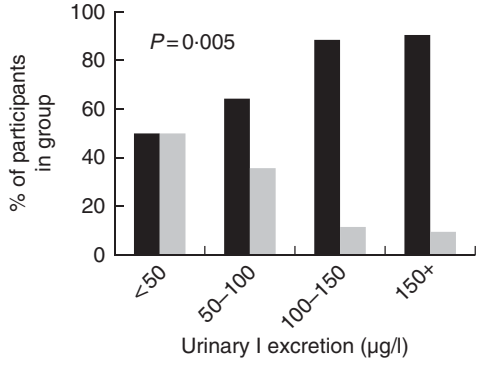

Fig. 1. The frequency of thyroglobulin (TG) in serum above or below 20 μg/l is illustrated for different levels of urinary iodine excretion. The frequency of s-TG > 20 μg/l decreased with rising urinary iodine excretion (Kendall's τ, $P=0{\cdot}005$). ■, TG < 20 μg/l; □, 20 > μg/l.

Reliability of surveys

The reliability of studies of I nutrition can be read from the *x*-axis in Fig. 2 matching a specified number of participants on the *y*-axis. The upper panel illustrates that about ten times more urine samples are needed compared with serum to obtain a

Table 2. Participants' descriptive and variation in thyroglobulin in serum and iodine in urine in eighty-one participants who participated in at least three data collections and had thyroglobulin antibodies < 100 U/ml

	Age	BMI	Thyroglobulin in serum (μg/l)			I content in urine (μg/l)		
	Mean (years)	Mean (kg/m^2)	Mean*	Variance*	CV%*†	Mean*	Variance*	CV%*†
Inuit	39·6	26·8						
Overall			14·4	114·9	74·2	128·3	18 446	128·4
Between individual			14·4	114·5	74·1‡	128·4	17 520	95·6‡
Mean individual			14·5	17·0	22·7	129·0	13 308	59·4
Median individual			11·5	2·2	14·5	118·5	1774	51·1
Caucasians	39·3	26·8						
Overall			11·6	60·8	66·9	102·7	6990	81·4
Between individual			11·6	59·5	65·0‡	102·9	6590	76·3‡
Mean individual			11·6	6·0	16·6	101·7	4044	48·7
Median individual			10·4	1·3	15·5	91·1	1321	44·4
All participants	39·5	26·8						
Overall			13·5	98·7	73·5	115·6	9293	83·4
Between individual			13·5	98·1	73·2‡	115·7	8954	79·9‡
Mean individual			13·6	13·4	20·9	116·1	5859	53·7
Median individual			11·4	1·9	15·1	106·9	1496	46·5

* Based on three or four samples in each of eighty-one participants.

† Calculated as $(\Sigma \mathrm{cv}\,\%^{2}{}_{1-n})^{1/2}$. Calculations using ANOVA techniques gave similar results.

‡ Square root of sum of squared CV for each data collection: $(\Sigma \mathrm{cv}\,\%^{2}{}_{1-n})^{1/2}$.

Table 3. Number of participants needed to be 95 % confident of being within a specified range for serum thyroglobulin (TG) to describe the I nutrition status of a population
(Numbers; median, lowest and highest variations)

	Number of samples needed for estimation of I deficiency with a specified precision range*							
	Serum thyroglobulin (μg/l)				Urinary I concentration (μg/l)			
	In a population†	In an individual‡			In a population†	In an individual‡		
Precision range (%)§	*n*	Median variation	Lowest	Highest variation	*n*	Median variation	Lowest	Highest variation
±1	20572	877	13	40736	24593	8295	86	58724
±2	5143	219	3	10184	6135	2074	21	14681
±5	823	35	1	1629	982	332	3	2349
±10	206	9	1	407	245	83	1	587
±20	51	2	1	102	61	21	1	147
±30	23	1	1	45	27	9	1	65
±40	13	1	1	25	15	5	1	37
±50	8	1	1	16	10	3	1	23

* Calculated from $n=(Z\times CV/D)^2$, where $Z=1{\cdot}96$ for 95 % CI; D = precision range.
† Number of individuals needed was calculated based on the average CV in the population.
‡ Variation differs between individuals. Number of samples needed to sample in an individual are given for individuals with median, lowest and highest variation.
§ Calculated with a 95 % CI ($Z=1{\cdot}96$).

similar reliability in an individual. The lower panel shows that 20 % more urine samples are needed in a population. The logarithmic scale on the *y*-axis illustrates the steep rise in the number of participants required when the reliability of I nutrition studies is increased.

Discussion

I deficiency may cause a spectrum of disorders that can be prevented by cheap and simple I supplementation[22]. Identification of I-deficient populations and monitoring of I fortification programmes are done by population surveys[18,23–25], and collection of urine samples for measuring I excretion is the standard method. However, this method has two major limitations. First, I is excreted within hours after ingestion and urinary I thus represents I intake only over the preceding hours[2–5] and spot urine samples may underestimate the true I excretion in populations with I-rich main meals[26]. The thyroid has the capacity to store large amounts of I and the urinary I excretion may thus not represent the true I intake level in I-deficient populations[6,7]. Second, I excretion varies considerably due to both differences in diet and due to dilution depending on the fluid intake, perspiration, ambient temperature and other factors[5,6]. This marked variation causes a considerable number of samples to be required for a reliable estimate of the true I nutrition level of that population or individual[7,8,17]. Both of these issues are addressed by the use of TG in serum.

S-TG is used to monitor patients treated for differentiated thyroid cancer by detection of low levels of s-TG. Elevated s-TG is seen with sustained I deficiency and it has been suggested and used in surveys of I nutrition[9–11,27]. The present study provides the first data to describe the reliability of surveys using s-TG to assess and monitor I nutrition. In addition, we calculated the number of samples needed for a certain level of confidence in the I nutrition level estimated. Similar data have been published for urinary I excretion[7,8,17] and we included data on UIC in the present study to allow for direct comparisons.

Reliability of iodine nutrition studies

We found a slightly lower between-individual variation for s-TG compared with UIC providing an advantage to s-TG over UIC in population surveys. Thus, 20 % more samples are needed for the same precision when using UIC as compared with s-TG. For example, if a precision of 5 % is aimed for, a survey requires one sample from each of 823 individuals when measuring s-TG and one sample from each of 982 individuals when measuring UIC. Conversely, if 1000 subjects are surveyed then the result is within ±4·5 % of the true value for s-TG and ±5·0 % for UIC. If 100 participants are included then this precision range is 14 % for s-TG and 16 % for UIC with 95 % CI. In other words, a mean value of 100 suggests that the true population mean is between 86 and 114 for s-TG and between 84 and 116 for UIC in the example.

Within-individual variation is much lower for s-TG than for UIC. Consequently, the number of samples needed in an individual is much lower and about ten times more samples are needed for the same precision when using UIC as compared with s-TG. It takes twenty-one urine samples to obtain a precision of ± 20 %, whereas this is seen with just two measurements of s-TG. Thus, ten times more samples are needed in an individual for the same reliability of the results when comparing UIC with s-TG.

The risk of error is lower for s-TG compared with UIC. The importance of number of samples for the risk of error can be estimated by comparing Tables 3 and 4. A 90 % confidence of being within 10 % from the true I nutrition level requires 6 (33) less samples for s-TG (UIC) than a 95 % CI of being within 10 % from the actual level.

British Journal of Nutrition

Table 4. Number of participants necessary to describe the iodine deficiency level with a defined precision range and with parallel confidence interval calculated from the variation in serum thyroglobulin and iodine excretion among healthy men and women in Greenland
(Numbers; median, lowest and highest variations)

		Number of samples needed for estimation of I deficiency with a defined precision range and CI*								
		Serum thyroglobulin (μg/l)					Urinary I concentration (μg/l)			
		In a population†	In an individual‡				In a population†	In an individual‡		
CI (%)§	Precision range§	*n*	Median variation	Lowest	Highest variation		*n*	Median variation	Lowest	Highest variation
99	±1 %	35 645	1520	23	70 583		42 519	14 373	149	10 1753
98	±2 %	7268	310	5	14 392		8669	2931	30	20 747
95	±5 %	823	35	1	1629		982	332	3	2349
90	±10 %	144	6	1	285		172	58	1	411
80	±20 %	22	1	1	43		26	9	1	63
70	±30 %	6	1	1	13		8	3	1	18
60	±40 %	2	1	1	5		3	1	1	7
50	±50 %	1	1	1	2		1	1	1	3

* Calculated from $n=(Z\times CV/D)^2$, where Z=CI (Z=2·58 for 99 %, 2·33 for 98 %, 1·96 for 95 %, 1·64 for 90 %, 1·28 for 80 %, 1·04 for 70 %, 0·84 for 60 %, 0·67 for 50 %); D=precision range.
† Number of individuals needed was calculated based on the average CV in the population.
‡ Variation differs between individuals. Number of samples needed to sample in an individual are given for individuals with median, lowest and highest variation.
§ CI set to vary in parallel with the precision range.

Fig. 2. The relation between number of participants and the precision of the estimate of iodine nutrition by both thyroglobulin (TG, – – – –) and urinary iodine excretion (UIC, ——). The precision of studies of iodine nutrition (*x*-axis) and the corresponding number of participants needed (*y*-axis) for that specific precision. About ten times more urine samples are needed compared with serum to obtain a similar precision in an individual (a) while 20 % more samples are needed in a population (b). The logarithmic scale on the y-axis illustrates the marked decrease in the number of participants required when lowering the demand for precision in iodine nutrition studies.

The variance found for urinary I was high, similar to previous studies of variation in urinary I excretion[7,8,17]. This is comparable across populations[7,17], I excretion levels[7,8] and after initiation of an I fortification programme[8]. This reinforces the reliability of the results of these studies of variation and reliability that support design of future I nutrition studies.

Design of iodine nutrition studies

Urinary I portrays the I intake during the hours before sampling while s-TG is rather a measure of long-term I nutrition[10,28]. The two measures of I nutrition thus provide information on different aspects of the I nutrition level of the population surveyed. The two measures combined could provide a more detailed description of the true I nutrition level of a population. Moreover, the use of both measures could be speculated to be a more accurate predictor of I deficiency disorders than the single measure of UIC. Thus, the combined UIC and s-TG could provide a two-dimensional insight into I nutrition level of a population. A description of the importance of variation of both of these measures is available from this study and may guide an intelligent study design.

The intelligent study design would seek to benefit from the advantages of each of the two measures. We suggest first to

assess the nutrition level from UIC and subsequently to add details and subgroup analysis within that population by the use of s-TG.

For example, we aimed for a precision of 5 % in an I nutrition survey. First, we need to survey 982 participants to be 95 % sure of this precision for UIC. The confirmation and subgroup analysis by the use of s-TG requires 823, 20 % less participants, and the precision would still be 5 %. Thus, the combined use of UIC and s-TG allows for further analysis with unaltered validity of the estimates.

Populations may be heterogeneous. They may cover different geographical areas with different subsurface geology[28,29], differences in water supply, differences in dietary habits[18,30] or different ethnic groups that call for subgroup analysis. The I nutrition level should be assessed for each of such groups and our data guide the analysis if s-TG is used.

Level of serum thyroglobulin

A cut-off level for s-TG has to be decided upon. S-TG concentrations were between 5 and 14 μg/l in healthy adults[9,10,14,31,32] and between 94 and 208 μg/l in adults in an area with endemic goitre[33,34]. A s-TG cut-off of 13 μg/l was suggested for children but it should not be concluded that this is a suitable level for I status in adults[35].

The occurrence of s-TG above a cut-off level of 40 μg/l was used to delineate I deficiency in children[11] and adults[10]. This contrasted the cut-off level of 13 μg/l reported to be the median value that delineated I-deficient adults[35]. The authors of the latter review emphasised the need for further investigation to settle a cut-off level[35]. A recent randomised trial reported a median s-TG of 16·6 μg/l in adults with mild I deficiency and a decrease in s-TG with I supplementation[36]. In our population the overall mean level was 13·5 μg/l and the individual median was 11·5 μg/l. The cut-off for I deficiency may be set to detect those with raised values rather than just above the median value. Thus, the 75th-percentile in our population of 19·1 μg/l guided the level of 20 μg/l set to delineate individuals with I deficiency in our investigation. In addition, this was in between previous suggestions[9,10,14,29,31,35] and provided a distinct separation of I deficiency groups in our data. It may thus be suggested to use this level to delineate I deficiency in an adult population.

Limitations

Serum concentration of TG has limitations. It is not specific for I deficiency but increases also with excess I intake, increasing thyroid mass, inflammation of the thyroid, cold, if the TSH receptor is stimulated and in pregnancy[23,37–41]. These groups should be identified and excluded or taken into consideration and corrected for in surveys of I deficiency using s-TG.

The use of s-TG should also consider inter-assay differences and possibly detection differences in s-TG between I nutrition levels. Attempts have been made to reduce the consequences thereof by standardisation[42] but differences between assays remain and should be taken into consideration[10,42]. Hence, we suggest a two dimensional approach that includes I in urine for the overall assessment of I nutrition and s-TG for validation and individual I nutrition assessment. The results will thus be reliable based on UIC and supported by using s-TG.

The spot urine I excretion of our study population suggested borderline to mild I deficiency. Different I intake levels associate with different s-TG levels. Whether variance of s-TG also differs with levels of s-TG remains to be settled. Also, the variance for other groups such as school children needs to be clarified.

Conclusion

Variation was lower for s-TG than for UIC. Thus, more participants are needed for similar reliability of the results for UIC compared with s-TG. Consequently, 80 % of samples are redundant in an individual if s-TG is used to assess I nutrition level, and 20 % less individuals are needed in a population when using s-TG rather than UIC. Consequently, s-TG provides the opportunity for either fewer participants or a higher reliability.

This difference may be used to gain more from I nutrition surveys. The reliability of the study is upheld when UIC is used to assess the overall I nutrition level and s-TG is added to assess I nutrition in subgroups. This is important when planning and evaluating I nutrition surveys in populations and in individuals. We thus suggest using both measures to assess I nutrition in a smart I nutrition survey design.

Acknowledgements

The authors gratefully acknowledge Carla Hame and Ruth Møller Jensen for their support during the data collection in Saqqaq and Eskild Boeskov for support in Ilulissat.

This work was supported by grants from Greenland Government's Research Council and by Karen Elise Jensen Fond. They had no role in the design, analysis or writing of this article.

S. A. contributed to study design, raising of funds, data collection, analysis and interpretation of data and writing of the manuscript. P. N. contributed substantially to interpretation of data and critical review of the manuscript. L. W. contributed substantially to interpretation of data and critical review of the manuscript. P. L. contributed to study design, raising of funds, data collection, analysis and interpretation of data and reviewing of the manuscript. P. L. tragically passed away but did approve the final version of the manuscript before the initial review.

The authors declare that there are no conflicts of interest.

References

1. World Health Organization (2007) *Assessment of Iodine Deficiency Disorders and Monitoring their Elimination: A Guide for Programme Managers*, 3rd ed. WHO/UNICEF/ICCIDD. Geneva: WHO.
2. Keating FR & Albert A (1949) The metabolism of iodine in man as disclosed with the use of radioiodine. *Rec Prog Horm Res* **4**, 429–481.
3. Vought RL, London WT, Lutwak L, *et al.* (1963) Reliability of estimates of serum inorganic iodine and daily fecal and urinary iodine excretion from single casual specimens. *J Clin Endocrinol Metab* **23**, 1218–1228.

4. Als C, Helbling A, Peter K, *et al.* (2000) Urinary iodine concentration follows a circadian rhythm: a study with 3023 spot urine samples in adults and children. *J Clin Endocrinol Metab* **85**, 1367–1369.
5. Pearce EN & Caldwell KL (2016) Urinary iodine, thyroid function, and thyroglobulin as biomarkers of iodine status. *Am J Clin Nutr* **104**, 898S–901S.
6. Andersen S, Pedersen KM, Pedersen IB, *et al.* (2001) Variations in urinary iodine excretion and thyroid function. A 1-year study in healthy men. *Eur J Endocrinol* **144**, 461–465.
7. Andersen S, Karmisholt J, Pedersen KM, *et al.* (2008) Reliability of studies of iodine intake and recommendations for number of samples in groups and in individuals. *Br J Nutr* **99**, 813–818.
8. Karmisholt J, Laurberg P & Andersen S (2014) Recommended number of participants in iodine nutrition studies is similar before and after an iodine fortification programme. *Eur J Nutr* **53**, 487–492.
9. Knudsen N, Bülow I, Jørgensen T, *et al.* (2001) Serum Tg – a sensitive marker of thyroid abnormalities and iodine deficiency in epidemiological studies. *J Clin Endocrinol Metab* **86**, 3599–3603.
10. Vejbjerg P, Knudsen N, Perrild H, *et al.* (2009) Thyroglobulin as a marker of iodine nutrition status in the general population. *Eur J Endocrinol* **161**, 475–481.
11. Zimmermann MB, Aeberli I, Andersson M, *et al.* (2013) 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 μg/L: a UNICEF/ICCIDD study group report. *J Clin Endocrinol Metab* **98**, 1271–1280.
12. Hegedus L, Perrild H, Poulsen LR, *et al.* (1983) The determination of thyroid volume by ultrasound and its relationship to body weight, age, and sex in normal subjects. *J Clin Endocrinol Metab* **56**, 260–263.
13. Knudsen N, Bols B, Bulow I, *et al.* (1999) Validation of ultrasonography of the thyroid gland for epidemiological purposes. *Thyroid* **9**, 1069–1074.
14. Rasmussen LB, Ovesen L, Bülow I, *et al.* (2002) Relations between various measures of iodine intake and thyroid volume, thyroid nodularity, and serum thyroglobulin. *Am J Clin Nutr* **76**, 1069–1076.
15. Schlumberger M, Fragu P, Gardet P, *et al.* (1991) A new immunoradiometric assay (IRMA) system for thyroglobulin measurement in the follow-up of thyroid cancer patients. *Eur J Nuclear Med* **18**, 153–157.
16. Haugen BR, Alexander EK, Bible KC, *et al.* (2016) 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. The American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* **26**, 1–133.
17. König F, Andersson M, Hotz K, *et al.* (2011) Ten repeated collections for urinary iodine from spot samples or 24 hour samples are needed to reliably estimate individual iodine status in women. *J Nutr* **141**, 2049–2054.
18. Andersen S, Hvingel B, Kleinschmidt K, *et al.* (2005) Changes in iodine excretion in 50-69-y-old Denizens of an Arctic society in transition and iodine as a biomarker of the frequency of consumption of traditional Inuit foods. *Am J Clin Nutr* **81**, 656–663.
19. Wilson B & van Zyl A (1967) The estimation of iodine in thyroidal amino acids by alkaline ashing. *S Afr J Med Sci* **32**, 70–82.
20. Laurberg P (1987) Thyroxine and 3,5,3'-triiodothyronine content of thyroglobulin in thyroid needle aspirates in hyperthyroidism and hypothyroidism. *J Clin Endocrinol Metab* **64**, 969–974.
21. Fraser CG & Harris EK (1989) Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* **27**, 409–437.
22. Laurberg P, Bülow Pedersen I, Knudsen N, *et al.* (2001) Environmental iodine intake affects the type of nonmalignant thyroid disease. *Thyroid* **11**, 457–469.
23. Teng W, Shan Z, Teng X, *et al.* (2006) Effects of iodine intake on thyroid function in China. *N Engl J Med* **354**, 2783–2793.
24. Hollowell JG, Staehling NW, Hannon WH, *et al.* (1998) Iodine nutrition in the United States. Trends and public health implications: iodine excretion data from National Health and Nutrition Examination Surveys I and III (1971–1974 and 1988–1994). *J Clin Endocrinol Metab* **83**, 3401–3408.
25. Laurberg P, Jørgensen T, Perrild H, *et al.* (2006) The Danish investigation on iodine intake and thyroid disease, DanThyr: status and perspectives. *Eur J Endocrinol* **155**, 219–228.
26. Andersen S, Waagepetersen R & Laurberg P (2015) Misclassification of iodine intake level from morning spot urine samples with high iodine excretion among Inuit and non-Inuit in Greenland. *Br J Nutr* **113**, 1433–1440.
27. Krejbjerg A, Bjergved L, Bulow Pedersen I, *et al.* (2016) Serum thyroglobulin as a biomarker of iodine deficiency in adult populations. *Clin Endocrinol* **85**, 475–482.
28. Andersen S, Petersen SB & Laurberg P (2002) Iodine in drinking water in Denmark is bound in humic substances. *Eur J Endocrinol* **147**, 663–670.
29. Andersen S, Guan H, Teng W, *et al.* (2009) Speciation of iodine in high iodine groundwater in China associated with goitre and hypothyroidism. *Biol Trace Elem Res* **128**, 95–103.
30. Rasmussen LB, Ovesen L, Bülow I, *et al.* (2002) Dietary iodine intake and urinary iodine excretion in a Danish population: effect of geography, supplements and food choice. *Br J Nutr* **87**, 61–69.
31. Nakamura S, Sakata S, Minamori Y, *et al.* (1984) Serum thyroglobulin (Tg) concentration in healthy subjects: absence of age- and sex-related differences. *Endocrinol Jpn* **31**, 93–98.
32. Pacini F, Pinchera A, Giani C, *et al.* (1980) Serum thyroglobulin in thyroid carcinoma and other thyroid disorders. *J Endocrinol Invest* **3**, 283–292.
33. van Herle AJ, Hershman JM, Homabrook RW, *et al.* (1976) Serum thyroglobulin in inhabitants of an endemic goiter region of New Guinea. *J Clin Endocrinol Metab* **43**, 512–516.
34. Bayram F, Beyazyildiz A, Gökce C, *et al.* (2009) The prevalence of iodine deficiency, serum thyroglobulin, antithyroglobulin and thyroid peroxidase antibody levels in the urban areas of Kayseri, Central Anatolia. *Exp Clin Endocrinol Diabetes* **117**, 64–68.
35. Ma ZF & Skeaff SA (2014) Thyroglobulin as a biomarker of iodine deficiency: a review. *Thyroid* **24**, 1195–1209.
36. Ma ZF, Venn BJ, Manning PJ, *et al.* (2016) Iodine supplementation of mildly iodine-deficient adults lows thyroglobulin: a randomized controlled trial. *J Clin Endocrinol Metab* **101**, 1737–1744.
37. Feldt-Rasmussen U, Bech K & Date J (1979) Serum thyroglobulin in patients with toxic and non-toxic goitres compared to sex- and age-matched control subjects. *Acta Endocrinol* **91**, 264–270.
38. Andersen S, Kleinschmidt K, Hvingel B, *et al.* (2012) Thyroid hyperactivity with high thyroglobulin in serum despite

sufficient iodine intake in chronic cold adaptation in an Arctic Inuit hunter population. *Eur J Endocrinol* **166**, 433–440.
39. Braverman LE (1996) Evaluation of thyroid status in patients with thyrotoxicosis. *Clin Chem* **42**, 174–178.
40. Gardner DF, Rothman J & Utiger RD (1979) Serum thyroglobulin in normal subjects and patients with hyperthyroidism due to Graves' disease: effects of T3, iodide, 131I and antithyroid drugs. *Clin Endocrinol* **11**, 585–594.
41. Laurberg P, Andersen S, Bjarnadottir RI, *et al.* (2007) Evaluating iodine deficiency in pregnant women and young infants – complex physiology with a risk of misinterpretation. *Public Health Nutr* **10**, 1547–1552.
42. Feldt-Rasmussen U, Profilis C, Colinet E, *et al.* (1996) Human thyroglobulin reference material (CRM 457). 1st Part: assessment of homogeneity, stability and immunoreactivity. *Ann Biol Clin (Paris)* **54**, 337–342.