

Published in final edited form as:

Clin Chim Acta. 2004 September ; 347(0): 61–69. doi:10.1016/j.cccn.2004.03.033.

Differences between measurements of T4 and T3 in pregnant and nonpregnant women using isotope dilution tandem mass spectrometry and immunoassays: are there clinical implications?

Offie P. Soldin^{a,b,*}, Rochelle E. Tractenberg^c, and Steven J. Soldin^{b,d,e}

^aDivision of Cancer Genetics and Epidemiology, Lombardi Cancer Center, Georgetown University School of Medicine, Room S-165A, 3800 Reservoir Road, NW, Washington, DC 20007-2197, USA

^bDivision of Endocrinology and Metabolism, Georgetown University School of Medicine, Washington, DC, USA

^cDepartment of Biomathematics and Biostatistics, Washington, DC, USA

^dBioanalytical Core Laboratory, Georgetown Clinical Research Center, Georgetown University Medical Center, Washington, DC, USA

^eDepartment of Laboratory Medicine, Children's National Medical Center and Departments of Pediatrics and Pathology, The George Washington University School of Medicine, Washington, DC, USA

Abstract

Background—It has been established that triiodothyronine (T3) and thyroxine (T4) measurements by tandem mass spectrometry (MS/MS) are more specific and are significantly different from immunoassay (IA) measurements (all $p < 0.05$) throughout pregnancy. In this study, we examined the clinical implications of these discrepancies.

Methods—Kappa statistics were used to determine the degree to which IA and MS/MS agreed in their identification of out-of-reference interval observations of circulating T4 and T3 from 52 normal, iodine-sufficient women during each trimester of pregnancy.

Results—After taking chance agreement into account, the two methods had poor agreement on classification of T3 values at the first ($\kappa = 0.185$) and second ($\kappa = 0.183$) trimesters, with extremely poor agreement for the third trimester and 1-year postpartum. Agreement on T4 was poor only for the third trimester ($\kappa = 0.183$). The two methods agreed on out-of-range values in only 0–25% of T3 cases and 25–66.7% of T4 cases.

Conclusions—The areas of disagreement suggest that women at risk (i.e., with analyte values outside of the 5–95% range) will not be detected using IA. Based on this cohort, our preliminary estimates are that 25–100% of such women would be missed if IA were used to assay the analyte.

© 2004 Elsevier B.V. All rights reserved.

*Corresponding author. Endocrinology and Metabolism and Division of Cancer Genetics and Epidemiology, Lombardi Cancer Center, Georgetown University Medical Center, Room S-165A, 3800 Reservoir Road, NW, Washington, DC 20007-2197, USA. Tel.: +1-202-687-4717; fax: +1-301-229-5285. os35@georgetown.edu (O.P. Soldin).

Keywords

Thyroxine T4; Triiodothyronine T3; Thyroid hormones; Pregnancy; Tandem mass spectrometry MS/MS; Immunoassay

1. Introduction

While the thyroid gland synthesizes the bulk of thyroxine (T4) and approximately 20% of triiodothyronine (T3), most of circulating T3 is produced enzymatically via monodeiodination of T4 by specific deiodinases present in the cells of target tissues [1]. In the circulation, 99.98% of T4 and 99.7% of T3 are bound to plasma-binding proteins. Binding to plasma proteins increases their biological half-life and enables their transport. The plasma-binding proteins include thyroxine-binding globulin (TBG), transthyretin, and albumin [2,3]. In serum drawn from healthy adults, total serum T4 (TT4; free + protein-bound hormone) is present at about 60-fold higher concentration than total T3 (TT3).

Pregnancy induces complex changes in circulating maternal steroid hormones and in TBG concentrations, leading to changes in thyroid hormone concentrations. In addition to the stimulatory effects of estrogen on TBG synthesis, a major contribution to the increased TBG concentration during pregnancy is their reduced plasma clearance due to changes in TBG glycosylation induced by estrogen. The concentrations of TBG double by the 16th to 20th week of gestation [4]. In addition to the 2- to 3-fold increase in serum TBG, there are also modest decreases in both serum transthyretin and albumin [5]. It is commonly thought that TT4 and TT3 increase in the setting of pregnancy-induced increases in serum TBG concentrations. After delivery, there is a rapid reversal of this process—serum TBG concentrations return to normal within 4–6 weeks and serum T4 and T3 return to pregestational serum levels.

The lack of specificity, bias, and imprecision are major analytical problems associated with some assays for free and total thyroid hormones. The presence of circulating iodothyronine-binding autoantibodies that interfere with total T4 and T3 immunoassays is a known phenomenon [6-8]. These autoantibodies may give falsely high or falsely low values of thyroid hormone measurements depending on the assay separation method used and are often in discordance with the clinical features [6-8]. Direct measurements of the unbound serum free T4 (FT4) and free T3 (FT3) are method-dependent [9,10] and are technically difficult to determine because they are measured in the picomole range, and to be valid, they must be free from interference by the much higher total hormone concentrations. Total thyroid hormone concentrations are measured at nanomolar levels and are therefore easier to measure. Currently, serum T4 and T3 concentrations are most commonly measured by immunoassay methods.

The College of American Pathologists Proficiency Testing (CAP PT) Program reported major differences for thyroid hormone immunoassays (Table 1). The mean results on proficiency testing samples for different immunoassay (IA) methods reported by CAP PT can vary by a factor of 2. The table shows the lowest and highest means obtained (for each of the samples), tested by the various IA methods. Table 1 illustrates the difference in specificity of the various antibodies used in the IAs. Certain conditions such as pregnancy, estrogen therapy, or genetic abnormalities in protein binding have also reportedly made IA methods for T4 and T3 diagnostically unreliable [11,12].

Recent work has shown that tandem mass spectrometry (MS/MS) of TT3 and TT4 provides excellent sensitivity and specificity [13]. The specific and sensitive determination of thyroid

hormones, as is possible using MS/MS, is especially important during pregnancy due to the potential for detrimental effects on fetal neurodevelopment. It has already been established that T3 measurements by MS/MS were significantly higher at every time point than IA measurements (all $p < 0.05$) and that thyroxine (T4) measurements by MS/MS were significantly lower at every time point (all $p < 0.005$; Soldin et al. in review); the present study focused on identifying and quantifying clinical implications of these discrepancies.

2. Materials and methods

Circulating T4 and T3 from 52 healthy, iodine-sufficient women during week 12, wk22 and wk 32 of pregnancy and approximately 1-year postpartum. All women were thyroid antibody negative throughout pregnancy. All women were of Caucasian inheritance/non-Hispanic, with the mean age of 30 years (25–38 years). Serum samples were measured simultaneously in serum, using an isotope dilution tandem mass-spectrometer (MS/MS; API-3000) and Dade Behring (Newark, DE) immunoassay (T4) and DPC-Immulite (T3). The research was conducted at the Georgetown University Medical Center's General Clinical Research Center.

2.1. Chemicals and reagents

Standards of thyroxine (T4) and 3,5,3'-triiodothyronine (T3) were from Sigma (St. Louis, MO). A stable deuterium-labeled internal standard, L-thyroxine-d₂ was synthesized according to procedures described in the literature [14,15] by Dr. Tomas Class from the Chemistry Department of Georgetown University. HPLC grade methanol was from VWR Scientific. All other chemicals were of analytical grade and were from Sigma.

2.2. Sample preparation

Informed consents were obtained and the study was IRB-approved. The samples were obtained from healthy, iodine-sufficient women on normal diets undergoing their first pregnancy. All of the women were of Caucasian inheritance/non-Hispanic with the mean age of 30 years (25–38 years) attending a pregnancy clinic at the Karolinska Institute, Sweden. The samples were obtained during gestation week 12 (first trimester), gestation week 22 (second trimester), gestation week 32 (third trimester), and approximately 1-year postpartum. Gestational dates were confirmed by ultrasound. All pregnancies were viable, nonmultiple pregnancies.

2.3. LC/MS/MS and immunoassay procedures

An API 3000 tandem mass-spectrometer (SCIEX, Toronto, Canada) equipped with TurboIonSpray and Shimadzu HPLC system was used to perform the analysis. Negative ion multiple reaction-monitoring (MRM) mode was used. LC/MS/MS for T4 and T3 was performed as previously described [13]. During pregnancy, the reference intervals for T4 using MS/MS were between 81.9 and 176.6 nmol/l (6.4–13.8 µg/dl; Ref [16]). Serum T3 reference intervals for pregnant women measured by MS/MS range between 1.71 and 4.50 nmol/l (111–292 ng/dl; Ref [16]).

Serum T4 and TSH concentrations were measured by immunoassay using the Dade Behring RxL Dimension Clinical Chemistry Analyzer, while concentrations of serum T3 were measured using the DPC-Immulite Automated Analyzer (Diagnostics Product Corporation, California) [17]. Serum T4 immunoassay values are method-dependent; the typical non-pregnancy reference intervals approximate 58–160 nmol/l (4.5–12.6 µg/dl) [18], while during pregnancy, the reference intervals for T4 range between 93.4 and 193.3 nmol/l (7.3–15.1 µg/dl) [16]. Serum T3 non-pregnant reference intervals approximate 1.2–2.7 nmol/l (80–180 ng/dl) [18], while during pregnancy, the reference intervals for T3 range between

1.52 and 3.96 nmol/l (99–257 ng/dl) [16]. The manufacturer's reference range for TSH was 0.34–4.82 μ IU/ml.

2.4. Statistical analysis

SPSS version 11.5 (SPSS, Chicago, IL) was used to carry out these analyses. Descriptive statistics were calculated for each analyte separately by trimester. Using the 5th and 95th percentile values for this cohort, together with TSH values (cutoff: 4.5 mIU/l), we identified those women whose IA or MS/MS measurements suggested a clinically important out-of-range observation at each trimester. Kappa (κ) is a measure of agreement in classification by IA and MS/MS that takes into account the fact that assay results were paired [19]. We also calculated the percents of the cohort at each trimester on which the assay criteria (a) agreed overall, (b) agreed on within-range classifications, and (c) agreed on out-of-range classifications.

3. Results

Table 2 presents the descriptive statistics for the analytes at each trimester and 1-year postpartum for each method. The mean values, together with their 95% confidence intervals (CIs), are presented in Fig. 1a (T3) and b (T4). All TSH concentrations were within normal range; the geometric means and SE(GM) for TSH concentrations were 0.98 ± 0.096 mIU/l in the first trimester, 1.29 ± 0.10 mIU/l in the second trimester, 1.25 ± 0.09 mIU/l in the third trimester, and 1.17 ± 0.09 mIU/l 1-year postpartum.

Paired sample *t*-tests demonstrated that measurements of T3 by IA were significantly lower than those by MS/MS at every time point, including postpartum (all $p < 0.05$). The mean values for T4 by IA are significantly higher than those derived by MS/MS at every time point, including postpartum (all $p < 0.05$). Using IA to measure T3 gives a different idea of the range of T3 values compared to MS/MS; although MS/MS measurements of T4 are significantly lower than measurements of T4 by IA, the distributions are not as divergent as they are for T3.

Agreement between classifications of values as “within-” or “out-of-” range was determined using kappa statistics, which can range in absolute value from 0 to 1 and can quantify the degree of association present after accounting for the probabilities that methods will agree simply by virtue of the fact that the assays are applied to the same sample (i.e., chance agreement; see Ref. [19]). *P*-values for kappa statistics can be calculated, but the more important characterization of the kappa value is in terms of the level of agreement it reflects. Kappa values between 0.0 and 0.4 reflect poor agreement, suggesting that beyond the chance agreement arising when classification of the same person or material is carried out under two raters or methods, the two classifications are independent. Values between 0.4 and 0.7 reflect fair to good agreement and values greater than 0.7 reflect excellent agreement [19]. The agreement matrices for the characterization of T3 and T4 values for each woman at each time point by IA and MS/MS are shown in Table 3.

Between 33.3% and 67.7% of observations on which IA- and MS/MS-based decisions resulted in a “disagreement” were either false positives (MS/MS classified the observation as within-range and IA classified it as out-of-range) or false negatives (MS/MS classified the observation as out-of-range and IA classified it as within-range); that is, the tendencies for false positive or false negative results using IA were similar. The two methods had poor agreement on classification of T3 values at the first ($\kappa = 0.185$) and second ($\kappa = 0.183$) trimesters, with extremely poor agreement (essentially 0) for the third trimester and 1-year postpartum. Agreement on T4 was fair for the first ($\kappa = 0.54$) and second ($\kappa = 0.54$)

trimesters and 1-year postpartum ($\kappa = 0.455$; all $p < 0.05$); agreement was poor for T4 at the third trimester ($\kappa = 0.183$).

The two methods agreed on out-of-range T3 values in only 0–25% of cases; agreement on out-of-range T4 values was 25–66.7%. Agreement and the criteria on which the classifications were based are shown in Fig. 2. The overall agreement levels presented in Table 3 are reflected in the center ‘block’ of each scatter plot in Fig. 2, which represents the values that both methods call “within-range”. It is the out-of-range values that are of greatest clinical concern; these are the small numbers of observations that fall outside the center block in each scatter plot in Fig. 2.

4. Discussion

The use of tandem mass spectrometry has been shown to be accurate, specific, reliable, simple, and fast [13]. Using this technique, the proteins are precipitated, and both T3 and T4 are measured in the samples simultaneously. In this study, MS/MS results were compared with those obtained by the traditional IA.

Kappa values that were calculated for each pair of observed analyte values (by IA and MS/MS) showed poor agreement on classification of T3 values at the first and second trimesters, with extremely poor agreement (essentially 0) for the third trimester and 1-year postpartum. In addition, our mass spectrometric method for T3 was used on CAP PT samples. Most immunoassay methods for T3 reported lower values than those obtained by the isotope dilution LC/MS/MS. However, while the agreement of the two methods on T4 at the third trimester was poor, the agreement for T4 was fair to good for the first and second trimesters and 1-year postpartum. We estimate that agreement on within-range T3 values was 90.9–93.3%, but the two methods agreed on out-of-range values in only 0–25% of cases. Similarly, we estimate that agreement on within-range T4 values was 93.3–97.8%, but agreement on out-of-range values was 25–66.7%.

Large variations exist in serum concentrations of thyroid hormones. The sources of variation include biological variation comprised of within-person variation of circadian and seasonal sources as well as the variation between individuals. In addition, despite major technological advances and methodologic ingenuity, most T4 and T3 measurements are subject to error, either preanalytical or analytical method-dependent artifacts [9,10,20].

Population-based reference intervals are useful for clinical interpretation of serum thyroid hormone concentrations; however, it is important to recognize their limitations for use in individuals. A consistently abnormal TSH probably indicates that T4 and T3 concentrations are not normal for the individual even when these fall inside the laboratory reference interval. This dissociation underlines the importance of TSH in the diagnosis and monitoring of thyroid dysfunctions. It also implies that subclinical thyroid disease may be defined in purely biochemical terms. Under critical circumstances, such as pregnancy where normal thyroid function is of importance for maintaining the pregnancy and for fetal brain development and growth, subclinical thyroid disease should be treated promptly. Even TSH within the reference range may be associated with slightly abnormal thyroid function of the individual. Although the clinical importance of slight abnormalities in thyroid function in pregnant women remains to be elucidated, the determination of thyroid hormones by MS/MS provides a more precise and specific result than is currently available with IA. Both IA and MS/MS reflect the increase in T3 and T4 during pregnancy and the drop-off in these values in the women postpartum; the problem with IA arises at the extremes of the analyte value distributions.

In a comparison of the concentrations of T3 and T4 of the samples using IA and MS/MS, we observed high levels of agreement overall, and in particular, on women with T3 and T4 values that fall in the normal (5th–95th percentile) ranges. However, these individuals do not represent the pregnancies at greatest risk for the potential detrimental effects on fetal neurodevelopment and growth [21,22]. Our results are preliminary because such a small proportion of the overall cohort was misclassified. That is, by definition, only 10% of the sample at each time point could fall outside the normal range of values. Because of these small samples (i.e., three to four disagreements on five to six out-of-range values), these data must be considered to represent only preliminary estimates. Our findings that 25–100% of women with out-of-range T3 or T4 values will be missed if IA is used to assay the analyte suggest that women at risk are not detected using IA. The confidence intervals associated with these values are far too large to draw firm conclusions (these are healthy, normal pregnant and nonpregnant women and a small number of women fall out of range, 10% [$n = 5$] by definition); therefore, further studies of the clinical implications of using IA to detect at-risk pregnancy are warranted.

Acknowledgments

This study was supported in part by grant 5-M01 RR-13297-S1 from the National Center for Research Resources, National Institutes of Health. We are grateful to Dr. L. Hilakivi-Clarke, Georgetown University, and Dr. E. Weiderpass of the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, and the International Agency for Research on Cancer, Lyon, France, for supplying the serum specimens, and to Dr. N. Soukhova of the Bioanalytical Core Laboratory, Georgetown Clinical Research Center, for technical assistance.

References

- [1]. Lum SM, Nicoloff JT, Spencer CA, Kaptein E. Peripheral tissue mechanism for maintenance of serum triiodothyronine values in a thyroxine-deficient state in man. *J Clin Invest.* 1984; 73(2): 570–5. [PubMed: 6699177]
- [2]. Robbins, J. Thyroid hormone transport proteins and physiology of hormone binding. In: Braverman, LE.; Utiger, RD., editors. *Werner and Ingbar's "The Thyroid". A Fundamental and Clinical Text.* Lippincott-Raven; Philadelphia: 2000. p. 103-20.
- [3]. Schussler GC. The thyroxine-binding proteins. *Thyroid.* 1999; 10(2):141–9. [PubMed: 10718550]
- [4]. Glinoe D. What happens to the normal thyroid during pregnancy? *Thyroid.* 1999; 9(7):631–5. [PubMed: 10447005]
- [5]. Bartalena L. Recent achievements in studies on thyroid hormone-binding proteins. *Endocr Rev.* 1990; 11(1):47–64. [PubMed: 2108013]
- [6]. Beck-Pecoz P, Romelli PB, Cattaneo MG, Faglia G, White EL, Barlow JW, et al. Evaluation of free thyroxine methods in the presence of iodothyronine-binding autoantibodies. *J Clin Endocrinol Metab.* 1984; 58:736–9. [PubMed: 6546576]
- [7]. Sakata S, Nakamura S, Miura K. Autoantibodies against thyroid hormones or iodothyronine. Implications in diagnosis, thyroid function, treatment, and pathogenesis. *Ann Intern Med.* 1985; 103(4):579–89. [PubMed: 3898955]
- [8]. Klee GG. Human anti-mouse antibodies. *Arch Pathol Lab Med.* 2000; 124(6):921–3. [PubMed: 10835540]
- [9]. Stockigt JR. Free thyroid hormone measurement. A critical appraisal. *Endocrinol Metab Clin North Am.* 2001; 30(2):265–89. [PubMed: 11444163]
- [10]. d'Herbomez M, Forzy G, Gasser F, Massart C, Beaudonnet A, Sapin R. Clinical evaluation of free thyroxine assays: persistent problems in particular populations. *Clin Chem Lab Med.* 2003; 41:942–7. [PubMed: 12940522]
- [11]. Despres N, Grant AM. Antibody interference in thyroid assays: a potential for clinical misinformation. *Clin Chem.* 1998; 44:440–54. [PubMed: 9510847]

- [12]. Sarne DH, Refetoff S, Nelson JC, Linarelli LG. A new inherited abnormality of thyroxine-binding globulin (TBG-San Diego) with decreased affinity for thyroxine and triiodothyronine. *J Clin Endocrinol Metab.* 1989; 68:114–9. [PubMed: 2491856]
- [13]. Soukhova N, Soldin OP, Soldin SJ. Isotope dilution tandem mass spectrometric method for T4/T3. *Clin Chem Acta.* 2004; 343:185–90.
- [14]. Burman KD, Bongiovanni R, Garis RK, Wartofsky L, Boehm TM. Measurement of serum T4 concentration by high performance liquid chromatography. *J Clin Endocrinol Metab.* 1981; 53:909–12. [PubMed: 7287878]
- [15]. Tai SS, Sniegowski LT, Welch MJ. Candidate reference method for total thyroxine in human serum: use of isotope-dilution liquid chromatography-mass spectrometry with electrospray ionization. *Clin Chem.* 2002; 48:637–42. [PubMed: 11901062]
- [16]. Soldin, OP.; Hilakivi-Clarke, L.; Weiderpass, E.; Soldin, SJ. Trimester-specific thyroid hormone reference intervals by isotope dilution tandem mass spectrometry and immunoassays- a longitudinal study; The 86th Annual Meeting of the Endocrine Society; New Orleans, Louisiana. 2004;
- [17]. Babson AL, Olson DR, Palmieri T, Ross AF, Becker DM, Mulqueen PJ. The IMMULITE assay tube: a new approach to heterogeneous ligand assay. *Clin Chem.* 1991; 37:1521–2. [PubMed: 1893580]
- [18]. Demers LM, Spencer CA. Laboratory medicine practice guidelines: laboratory support for the diagnosis and monitoring of thyroid disease. *Clin Endocrinol (Oxf).* 2003; 58:138–40. [PubMed: 12580927]
- [19]. Fleiss, JL.; Levin, B.; Paik, MC. Statistical methods for rates and proportions. third ed.. Wiley; Hoboken (NJ): 2004. p. 598-626.
- [20]. Christofides ND, Wilkinson E, Stoddart M, Ray DC, Beckett GJ. Assessment of serum thyroxine binding capacity-dependent biases in free thyroxine assays. *Clin Chem.* 1999; 45:520–5. [PubMed: 10102912]
- [21]. McDermott MT, Ridgway EC. Subclinical hypothyroidism is mild thyroid failure and should be treated. *J Clin Endocrinol Metab.* 2001; 86:4585–90. [PubMed: 11600507]
- [22]. Kempers MJ, van Tijn DA, van Trotsenburg AS, de Vijlder JJ, Wiedijk BM, Vulsma T. Central congenital hypothyroidism due to gestational hyperthyroidism: detection where prevention failed. *J Clin Endocrinol Metab.* 2003; 88:5851–7. [PubMed: 14671180]

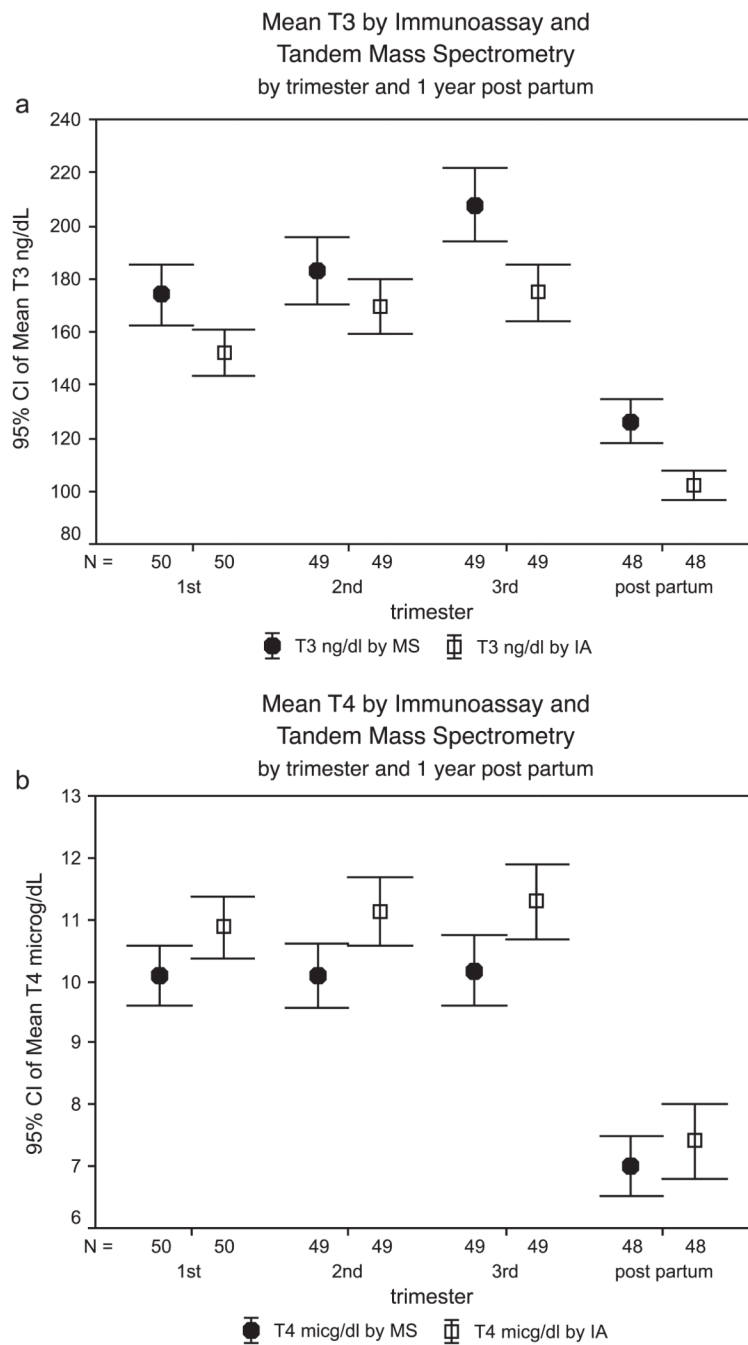


Fig. 1. (a)–(b) 95% confidence intervals for mean serum T3 and T4 concentrations in pregnancy by tandem mass spectrometry and by immunoassay. Note to Fig. 1a and b: For conversion, T3 ng/dl \times 0.0154 = nmol/l; T4 μ g/dl \times 12.8 = nmol/l.

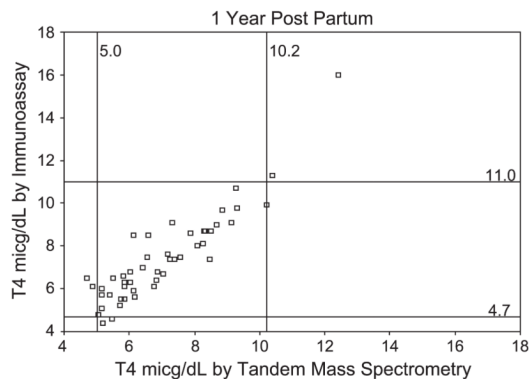
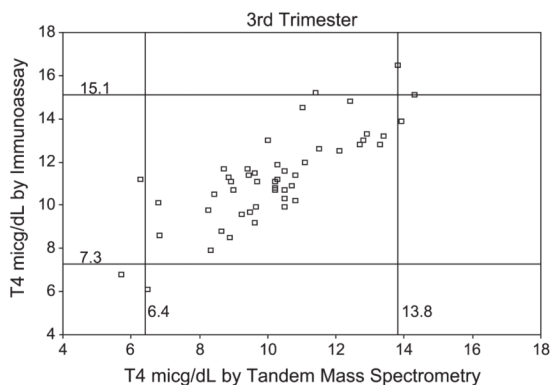
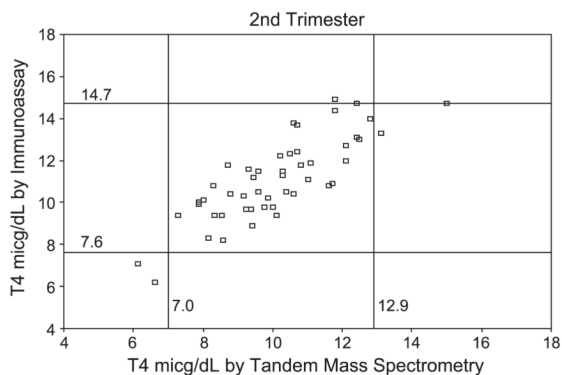
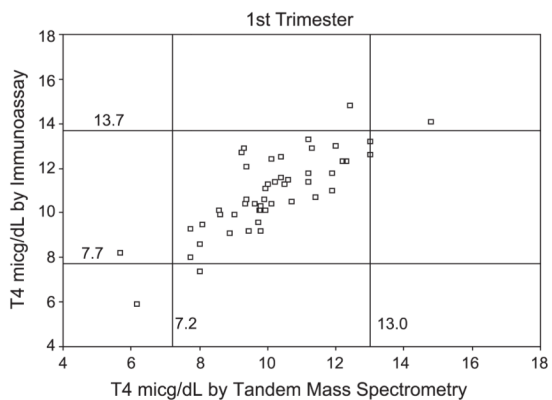
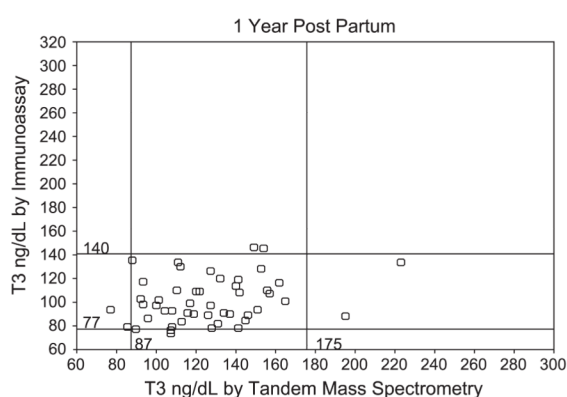
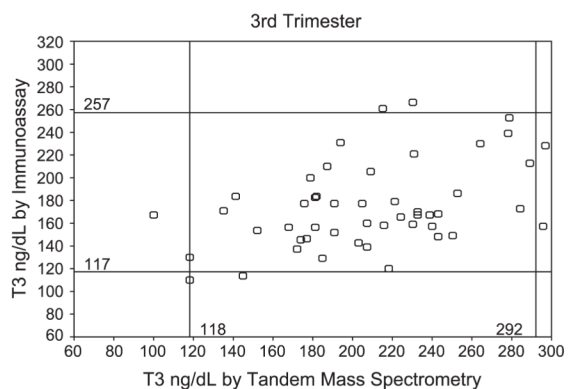
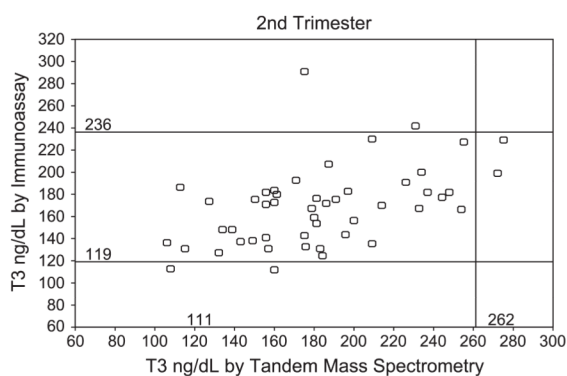
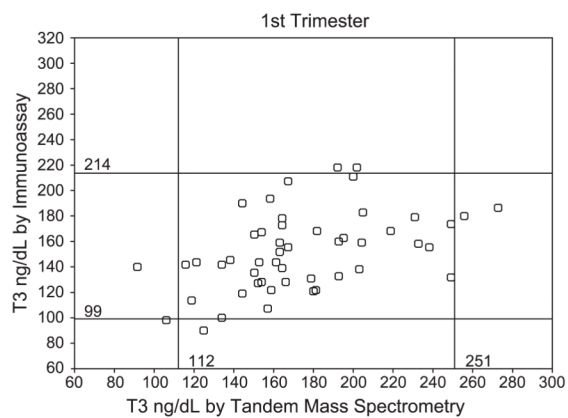


Fig. 2.
Serum T3 and T4 concentrations in pregnancy outside the reference intervals defined by tandem mass spectrometry and by immunoassay.

Table 1CAP PT Program 2003 - Immunoassays of T3 and T4^a

Thyroid hormone immunoassays, data acquired from CAP PT Program 2003		
Triiodothyronine T3	Mean-Low result ng/dl (nmol/l)	Mean-High result (Mean) ng/dl (nmol/l)
Sample 1	108.5 (1.67)	190.2 (2.93)
Sample 2	364.8 (5.62)	610.1 (9.40)

Thyroxine T4	Mean-Low result µg/dl (nmol/l)	Mean-High result µg/dl (nmol/l)
Sample 1	5.64 (72.2)	10.09 (129.1)
Sample 2	1.64 (21.0)	3.65 (46.7)
Sample 3	8.73 (111.7)	13.12 (167.9)

^aComparison of IA test results obtained for five test samples. The lowest and highest means for each of the samples are shown. Mean were calculated for each method.

Table 2

Mean (\pm S.D.), Medians and *t*-tests for T4 and T3 using isotope dilution tandem mass spectrometry and immunoassays

	IA mean ng/dl \pm (S.D.); (nmol/l \pm S.D.)	IA median ng/dl (nmol/l)	MS/MS mean ng/dl \pm (S.D.); (nmol/l \pm S.D.)	MS/MS median ng/dl (nmol/l)	<i>t</i> -test
<i>Triiodothyronine TT3</i> [†]					
1st trimester	<i>n</i> = 50 152.1 (30.6) (2.3 \pm 0.5)	148.5 (2.3)	<i>n</i> = 52 174.5 (40.2) (2.7 \pm 0.6)	164.0 (2.5)	4.1 ^{**}
2nd trimester	<i>n</i> = 49 169.2 (35.4) (2.6 \pm 0.5)	171.0 (2.6)	<i>n</i> = 51 183.2 (42.5) (2.8 \pm 0.7)	181.0 (2.8)	2.4 [*]
3rd trimester	<i>n</i> = 49 174.9 (37.4) (2.7 \pm 0.6)	167.0 (2.6)	<i>n</i> = 51 208.0 (46.6) (3.2 \pm 0.7)	207.0 (3.2)	5.1 ^{**}
1-year postpartum	<i>n</i> = 48 102.0 (19.2) (1.6 \pm 0.3)	97.5 (1.5)	<i>n</i> = 52 125.1 (28.4) (1.9 \pm 0.4)	121.0 (1.9)	5.6 ^{**}
<i>Thyroxine TT4</i> [‡]					
1st trimester	<i>n</i> = 50 10.9 (1.7) (139.2 \pm 22.3)	10.6 (136.3)	<i>n</i> = 52 10.1 (1.7) (129.2 \pm 29.4)	10.0 (127.7)	-5.4 ^{**}
2nd trimester	<i>n</i> = 49 11.1 (1.9) (142.4 \pm 25.0)	10.9 (139.5)	<i>n</i> = 51 10.1 (1.7) (129.3 \pm 22.3)	10.2 (130.6)	-6.4 ^{**}
3rd trimester	<i>n</i> = 49 11.3 (2.1) (144.5 \pm 26.8)	11.2 (143.4)	<i>n</i> = 51 10.2 (2.0) (130.5 \pm 25.5)	10.2 (130.6)	-6.0 ^{**}
1-year postpartum	<i>n</i> = 48 7.4 (2.1) (94.8 \pm 26.5)	6.9 (88.3)	<i>n</i> = 52 7.0 (1.6) (89.2 \pm 20.8)	6.7 (85.1)	-3.0 [*]

t-tests were paired, all MS/IA, and were based on measurements in ng/dl. Positive *t*-score means MS gave larger measurement; negative *t*-score means MS gave smaller measurement.

[†]For conversion: T3 ng/dl \times 0.0154 = nmol/l

[‡]T4 Ag/dl \times 12.8 = nmol/l.

^{*} *P* < 0.05.

^{**} *P* 0.001.

Table 3

Agreement matrix for T3 and T4 by IA (columns) and MS/MS (rows) by time point

		T3		T4	
		IA		IA	
MS/MS		IN	OUT	MS/MS	OUT
IN	43 ^a 3 ^b	1st Trimester		IN	45 2
OUT	3 ^c 1 ^d	T3: $\kappa = 0.185$		OUT	1 2
IN	42 3	2nd trimester		IN	44 1
OUT	3 1	T3: $\kappa = 0.185$		OUT	2 2
IN	42 4	3rd Trimester		IN	42 3
OUT	3 0	T3: $\kappa = -0.075$		OUT	3 1
IN	40 4	1-year Postpartum		IN	42 2
OUT	4 0	T3: $\kappa = -0.091$		OUT	2 2

IN = within-range, OUT = out-of-range; both based on 5th and 95th percentiles under respective assay method (IA or MS/MS).

^aBoth assays agree on normal patient.

^bFalse positive by IA.

^cFalse negative by IA.

^dBoth assays agree on at-risk patient.