Thyroid Function in Pregnancy: What is Normal?  
Marco Medici,1,2† Tim I.M. Korevaar,1,2† W. Edward Visser,1,2 Theo J. Visser,1,2 and Robin P. Peeters1,2

BACKGROUND: Gestational thyroid dysfunction is common and associated with maternal and child morbidity and mortality. During pregnancy, profound changes in thyroid physiology occur, resulting in different thyroid-stimulating hormone (TSH) and free thyroxine (FT4) reference intervals compared to the nonpregnant state. Therefore, international guidelines recommend calculating trimester- and assay-specific reference intervals per center. If these reference intervals are unavailable, TSH reference intervals of 0.1–2.5 mU/L for the first trimester and 0.2–3.0 mU/L for the second trimester are recommended. In daily practice, most institutions do not calculate institution-specific reference intervals but rely on these fixed reference intervals for the diagnosis and treatment of thyroid disorders during pregnancy. However, the calculated reference intervals for several additional pregnancy cohorts have been published in the last few years and show substantial variation.

CONTENT: We provide a detailed overview of the available studies on thyroid function reference intervals during pregnancy, different factors that contribute to these reference intervals, and the maternal and child complications associated with only minor variations in thyroid function.

SUMMARY: There are large differences in thyroid function reference intervals between different populations of pregnant women. These differences can be explained by variations in assays—as well as population-specific factors, such as ethnicity and body mass index. The importance of using correct reference intervals is underlined by the fact that even small subclinical variations in thyroid function have been associated with detrimental pregnancy outcomes, including low birth weight and pregnancy loss. It is therefore crucial that institutions do not rely on a fixed universal cutoff concentrations, but calculate their own pregnancy-specific reference intervals.

Thyroid dysfunction during pregnancy is common, with a prevalence of 2%–4% (1, 2). Maternal thyroid dysfunction is associated with an increased risk of various adverse maternal and child outcomes, including miscarriage, intrauterine growth retardation, hypertensive disorders, preterm delivery, and a decreased child IQ (2–4). During pregnancy, profound changes in thyroid physiology occur to provide sufficient thyroid hormone (TH)³ to both the mother and fetus. This is particularly important during early pregnancy because the fetal thyroid starts to produce considerable amounts of TH only from approximately 20 weeks of gestation, until which time the fetus heavily depends on the maternal supply of TH. This supply of TH to the fetus, as well as increased concentrations of TH binding proteins (thyroxine-binding globulin) and degradation of TH by placental type 3 iodothyronine deiodinase necessitate an increased production of maternal TH (1, 2). This requires an intact thyroid gland and adequate availability of dietary iodine and is in part mediated by the pregnancy hormone human chorionic gonadotropin, which is a weak agonist of the thyroid-stimulating hormone (TSH) receptor (5). As a consequence, serum free thyroxine (FT₄) concentrations increase and TSH concentrations decrease from approximately the eighth week throughout the first half of pregnancy, resulting in different reference intervals for TSH and FT₄ compared to the nonpregnant state.

Given these pregnancy-related changes in thyroid physiology and the complications associated with thyroid dysfunction, it is important to determine reference intervals for normal thyroid function during pregnancy. This is crucial to identify women who would potentially benefit from treatment. For this reason, the guidelines of the Endocrine Society, American Thyroid Association, and European Thyroid Association recommend that trimester-specific reference intervals be calculated for each center (6–8). If these calculated intervals are not

© 2015 American Association for Clinical Chemistry

1 Department of Internal Medicine, 2 Rotterdam Thyroid Center, Erasmus Medical Center, Rotterdam, the Netherlands.
† Address correspondence to this author at: Department of Internal Medicine, Erasmus Medical Center, Dr. Molewaterplein 50, 3015 GE, Rotterdam, the Netherlands. Fax +31-10-7035430, e-mail m.medici@erasusmc.nl.
† Marco Medici and Tim I.M. Korevaar contributed equally to the work, and both should be considered as first authors.
Received December 1, 2014; accepted March 3, 2015.
Previously published online at DOI: 10.1373/clinchem.2014.236646

3 Nonstandard abbreviations: TH, thyroid hormone; TSH, thyroid-stimulating hormone; FT₄, free thyroxine; TPOAb, thyroid peroxidase antibody; TRAb, TSH receptor antibody; MoM, multiple of medians; UIC, urinary iodine concentration; BMI, body mass index; SGA, small size for gestational age; OR, odds ratio.
available in the laboratory. TSH reference intervals of 0.1–2.5 mU/L for the first trimester and of 0.2–3.0 mU/L for the second trimester are recommended (6–8). These reference interval estimations were predominantly based on the published reference intervals of 6 pregnancy cohorts (9–14). Although the center-specific reference intervals for many additional pregnancy cohorts have been published following the publication of these guidelines and show substantial differences in cutoffs for TSH, most institutions still rely on these fixed reference intervals. This is particularly relevant because even subclinical thyroid dysfunction, mostly defined according to population-based cutoffs, is associated with an increased risk of adverse maternal and child outcomes. Therefore, this review provides an overview of studies of thyroid function reference intervals during pregnancy and different factors that contribute to these intervals, as well as the clinical complications associated with minor variations in thyroid function.

Studies of Thyroid Function Reference Intervals During Pregnancy

In accordance with recommendations by the International Federation of Clinical Chemistry (15), international thyroid guidelines advise that reference intervals should be based on the 2.5th and 97.5th percentile of the respective population with an optimal iodine intake (6–8). In addition, each study on a specific endpoint should also incorporate a sensitivity analysis on neighboring cutoff percentiles to explore the optimal cutoff point. Analyses using a nonparametric cutoff should be performed in a sufficiently sized, nonselected population that consist of “healthy” reference individuals. Because of the high interindividual variability and skewness for TSH but also to some extent FT4, a minimum of approximately 400 individual measurements per partition is required as opposed to the minimum of 120 measurements recommended for standard parametric 90% coverage interval calculations (16–19). Although the term “healthy” can be interpreted in many ways for TSH and FT4 reference interval determinations, this at least means a population free of major known factors inhibiting or stimulating thyroid function. Preferably, this population would consist of thyroid antibody [i.e., thyroid peroxidase antibody (TPOAb)-negative women without preexisting thyroid disease or other thyroid-interfering factors such as medication use or twin pregnancies]. Exclusion of TSH receptor antibody (TRAb)-positive subjects could further improve reference interval estimations, although most of the TRAb-positive subjects are also TPOAb-positive and TRAbs are far less common in the general population than TPOAbs (20).

Table 1 shows reference intervals for TSH and FT4 during early pregnancy, calculated according to the international guidelines, in sufficiently sized population-based cohorts among TPOAb-negative women (9, 14, 21–32). For both hormones, a wide range of reference interval values has been reported, with the upper limit of TSH varying between 2.15 and 4.68 mU/L between different cohorts. Importantly, 90% of all upper limits of TSH are higher than the recommended fixed TSH cutoff concentrations of 2.5 and 3.0 mU/L for the first and second trimesters, respectively. The clinical relevance of this finding is that the use of these fixed upper limits of 2.5 mU/L and 3.0 mU/L results in significant overtreatment in euthyroid women, which may have negative effects on maternal and/or fetal outcomes. This is illustrated in Fig. 1 for a large iodine-sufficient population-based cohort in the Netherlands, where 8.6% and 4.9% of the TPOAb-negative women with TSH levels within the reference interval had a TSH concentration above 2.5 mU/L and 3.0 mU/L in the first and second trimesters, respectively (28). These data underline the importance of calculating population-based pregnancy-specific thyroid parameter reference intervals, instead of using fixed upper limits of 2.5 and 3.0 mU/L.

Factors Influencing Thyroid Function Reference Intervals during Pregnancy

As illustrated in Table 1, various commercial TSH and FT4 assays have been used to evaluate thyroid function during pregnancy. Although previous studies have shown that the interassay differences for TSH are relatively small (r = 0.91–0.98), FT4 measurements seem much more prone to interference and have larger interassay differences (r = 0.68–0.89) (33, 34). Pregnancy results in a shift of potentially interfering factors such as thyroxine-binding globulin and albumin. Not only does the extent of this shift vary per individual, it also affects measurements by each immunoassay differently. Therefore, the population differences in FT4 concentrations can be at least partly attributed to assay-related factors. As recently suggested by Bestwick et al., TSH and FT4 values can be expressed as multiple of medians (MoM) to interpret and compare the upper and lower limits obtained via different assays (21). An MoM value is calculated by dividing each individual’s value by the population median, which creates a value that is standardized for the assay median. These values are independent of interassay differences, and therefore cutoff points for different assays can be generalized more easily. Table 2 shows the calculated lower and upper limits expressed as MoM values for the same studies as Table 1, which resulted in more uniform reference intervals. This is especially the case for FT4, suggesting that TSH is more subject to change by nonanalytical factors.

It has long been known that iodine is an essential component of TH that is subject to physiologic changes
<table>
<thead>
<tr>
<th>Reference and country (analysis method)</th>
<th>No.</th>
<th>Gestation (weeks)</th>
<th>TSH in mU/L</th>
<th>FT4 in pmol/L (ng/dl)</th>
<th>Population characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bestwick et al. (21), Italy (AutoDELFIA)</td>
<td>5505</td>
<td>&lt;16</td>
<td>1.07</td>
<td>0.04-3.19</td>
<td>NR</td>
</tr>
<tr>
<td>Bestwick et al. (21), UK (Advia Centaur)</td>
<td>16334</td>
<td>&lt;16</td>
<td>1.11</td>
<td>0.06-3.50</td>
<td>Moderate-Mild</td>
</tr>
<tr>
<td>Bocos-Terraz et al. (9), Spain (Architect)</td>
<td>481</td>
<td>&lt;14</td>
<td>0.94</td>
<td>0.41-2.63</td>
<td>Mild</td>
</tr>
<tr>
<td>Gilbert et al. (22), Australia (Architect)</td>
<td>1817</td>
<td>9-13</td>
<td>0.74</td>
<td>0.02-2.15</td>
<td>Borderline</td>
</tr>
<tr>
<td>Lamberts-Messerlian et al. (23), USA (Immulate 2000)</td>
<td>8351</td>
<td>T1</td>
<td>1.00</td>
<td>0.12-3.37</td>
<td>Mild</td>
</tr>
<tr>
<td></td>
<td>8415</td>
<td>T2</td>
<td>1.19</td>
<td>0.35-3.35</td>
<td>Mild</td>
</tr>
<tr>
<td>La'ulu and Roberts (24, 25), USA</td>
<td>2172</td>
<td>10-13</td>
<td>0.94</td>
<td>0.02-2.69</td>
<td>Mild</td>
</tr>
<tr>
<td></td>
<td>2683</td>
<td>14-20</td>
<td>1.14</td>
<td>0.15-3.11</td>
<td>Mild</td>
</tr>
<tr>
<td>Li et al. (26), China (Cobas Elesys 601)</td>
<td>640</td>
<td>7-12</td>
<td>1.47</td>
<td>0.10-4.34</td>
<td>Proven sufficient</td>
</tr>
<tr>
<td>Mannisto et al. (27), Finland, (Architect 2000)</td>
<td>4333</td>
<td>T1</td>
<td>1.11</td>
<td>0.08-3.54</td>
<td>Sufficient</td>
</tr>
<tr>
<td></td>
<td>747</td>
<td>T2</td>
<td>1.37</td>
<td>0.11-4.24</td>
<td>Sufficient</td>
</tr>
<tr>
<td>Medici et al. (28), the Netherlands (Vitros ECI)</td>
<td>5186</td>
<td>8-18</td>
<td>1.30</td>
<td>0.03-4.04</td>
<td>Proven sufficient</td>
</tr>
<tr>
<td>Pearce et al. (29), USA (Advia Centaur)</td>
<td>585</td>
<td>&lt;14</td>
<td>1.1</td>
<td>0.04-3.60</td>
<td>Borderline</td>
</tr>
<tr>
<td>Quinn et al. (30), Russia (Abbott AxSYM)</td>
<td>380</td>
<td>T1</td>
<td>1.66</td>
<td>0.09-4.67</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>549</td>
<td>T2</td>
<td>2.00</td>
<td>0.20-4.68</td>
<td>Moderate</td>
</tr>
<tr>
<td>Springer et al. (31), Czech Republic (ADIVA Centaur)</td>
<td>4337</td>
<td>9-11</td>
<td>1.21</td>
<td>0.06-3.67</td>
<td>Mild</td>
</tr>
<tr>
<td>Stricker et al. (14), Switzerland (Architect 2000SR)</td>
<td>575</td>
<td>6-12</td>
<td>0.95</td>
<td>0.07-2.82</td>
<td>Sufficient</td>
</tr>
<tr>
<td></td>
<td>528</td>
<td>T2</td>
<td>1.02</td>
<td>0.20-2.79</td>
<td>Sufficient</td>
</tr>
<tr>
<td>Vaidya et al. (32), UK (Modular E 170)</td>
<td>1089</td>
<td>&lt;12</td>
<td>1.08</td>
<td>0.14-3.19</td>
<td>Mild-moderate</td>
</tr>
</tbody>
</table>

*Studies were selected according to the following criteria: n ≥ 500, exclusion of PPOAb-positive women, and availability of data from the manuscript or via personal communication. Iodine status was estimated on the basis of references from the respective article, WHO iodine status reports, or the Vitamin and Mineral Nutrition Information System.

*Weight reported (median weight 59 kg in Italian and 67 kg in UK population (21)).

*NR, Not reported; T1, first trimester; T2, second trimester.

*Reported FT4 level is a mean.

*Limits are 5th and 98th percentiles for TSH and 2nd and 95th percentiles for FT4.

*Based on reports of the total PASTER (First and Second Trimester Evaluation of Risk) trial population.

*FT4 determined only for TSH within the reference interval.

*Based on iodine measurements in study population.

*High human chorionic gonadotrophin values excluded.

*Free T4 index (reference interval 1.0–4.0).
during pregnancy, including an increased turnover and renal excretion, necessitating increased intake during pregnancy. It is therefore expected that populations with an abnormal iodine status have a higher prevalence of thyroid dysfunction, which would lead to unreliable reference interval estimations. For this reason, the international guidelines recommend calculating reference intervals in populations with an optimal iodine intake (6–8). Despite this, few data are available about the exact effects of iodine status on thyroid function reference intervals during pregnancy. A Chinese study recently measured first-trimester serum thyroid function and urinary iodine concentrations (UIC) in 7190 pregnant women from an iodine-sufficient population (35). No effects of low UIC on mean serum TSH or FT4 concentrations were observed. However, compared to women with adequate iodine intake (UIC 150–249 μg/L), women with excessive iodine intake (UIC >500 μg/L) had higher mean TSH (2.32 vs 1.86 mU/L) and lower FT4 (15.27 vs 16.12 pmol/L) concentrations (all P < 0.001). Calculated serum TSH and FT4 reference intervals were 0.24–5.63 mU/L and 12.23–21.01 pmol/L in women with adequate iodine intake, and 0.36–6.12 mU/L and 12.14–20.64 pmol/L in women with excessive iodine intake. More studies in various trimesters of pregnancy and different ethnicities are needed to extrapolate the exact extent of these effects.

The extent to which other population characteristics, such as ethnicity, body mass index (BMI), and smoking, influence TSH or FT4 measurements is more quantifiable. These characteristics have all been associated with differences in serum thyroid parameters as well (21, 24, 25, 27, 29, 36–43). With regard to ethnicity, for a wide range of serum thyroid function tests both upper and lower limits differ according to ethnic background in the first and second trimester. La’ulu et al. showed substantial differences in TSH upper limits, ranging from 2.73 in blacks (MoM 2.81) to 3.64 mU/L in Asians (MoM 3.17), reaching borderline statistical significance (24, 25). Recently, we have shown significant differences in TSH reference intervals between various ethnic groups in a population-based pregnancy cohort of European origin (see Fig. 1 in the Data Supplement that accompanies the online version of this article at http://wwwclinchemorg/content/vol61/issue5) and additionally demonstrated that these ethnic differences in thyroid parameter reference intervals may lead to considerable misclassification of thyroid disease in up to 18% of cases (39).

BMI has also been associated with both TSH and FT4 concentrations during pregnancy (21, 27, 37, 41). Mannisto et al. found that among women with a BMI >30 kg/m², the upper limit (95th percentile) for TSH increased from 2.86 mU/L in women with a BMI <20 kg/m² to 3.50 mU/L. For the same groups, they also showed that the lower limit for FT4 (fifth percentile) decreased from 12.3 to 11.6 pmol/L, respectively (27). Bestwick et al. expressed these values in MoMs and found an increase in TSH of 0.025 MoM and a decrease in FT4 of 0.009 MoMs per 10-kg increase in body weight (21).

In this context it is noteworthy that the prevalence of overt hypothyroidism in morbidly obese subjects (BMI >40 kg/m²) was found to be 11.8% (44). The guidelines of the American Thyroid Association therefore recommend TSH screening in morbidly obese pregnant women (7).
In line with the above, Table 1 additionally shows data on BMI, iodine status, and specific ethnic backgrounds for the various studies on thyroid function reference intervals. However, it is hard to comment on these associations from this table because these characteristics were incompletely reported in many studies.

Finally, various studies have shown that smoking has only limited effects on mean TSH and FT₄ concentrations during pregnancy (21, 29, 40, 42, 45). This is illustrated by a study of 4317 Finnish pregnant women, which found that smokers had TSH concentrations identical to those of nonsmokers (1.02 mU/L), whereas there was a small difference in FT₄ concentrations (15.02 vs 15.24 pmol/L; \( P = 0.006 \) (40)). Because effect sizes were small, it seems unlikely that population differences in smoking prevalence have any noteworthy effect on TSH and FT₄ reference intervals.

**Minor Variations in Thyroid Function and the Risk of Maternal and Child Complications**

The previous paragraphs showed that there are substantial differences in thyroid parameter reference intervals between populations. However, to illustrate the clinical relevance of using these population-based pregnancy-specific intervals instead of fixed or nonpregnancy reference intervals, we calculated these effects in the Generation R study (see online Supplemental Table 1). Women with TSH concentrations above the population-based reference interval had an increased risk of premature delivery and children with intrauterine growth retardation (small size for gestational age (SGA)), whereas women with TSH concentrations below the lower limit of this reference interval had an increased risk of hypertensive disorders (46, 47). However, the use of fixed TSH cutoffs did not identify women with an increased risk of premature delivery or SGA, and the association between suppressed TSH and hypertensive disorders remained similar. This suggests that the use of fixed instead of population-based reference intervals would lead to overtreatment, particularly in women with high-normal TSH concentrations. In recent years, other studies have investigated the effects of minor subclinical variations in thyroid function on the risk of adverse maternal and child outcomes. These studies are important in the clinical context of this review because they provide insight into...
Pregnancy loss is a difficult study end point because early fetal loss naturally occurs in approximately 30% of pregnancies, of which the majority occurs even before pregnancy is clinically recognized (66). Negro et al. studied the relationship between thyroid function and the combined endpoint of miscarriage and stillbirth in TPOAb-negative pregnant women and concluded that women with serum TSH concentrations of 2.5–5.0 mU/L had a 6.1% risk of pregnancy loss, compared to 3.6% in women with a TSH concentration below 2.5 mU/L (49). However, the fact that no population-based reference intervals were calculated or sensitivity analyses done makes the 2.5 mU/L cutoff somewhat arbitrary and hard to interpret in relation to other studies. Further analyses showed a positive linear association between TSH concentrations and pregnancy loss. This is in line with the results of a Dutch cohort of 2497 pregnant women, in which it was shown that the incidence of miscarriage and fetal and neonatal death (combined into child loss) increased by 80% by every doubling of the maternal TSH concentration (67). However, given the limited number of 27 cases and the heterogeneity of cases included in this group, these results should be interpreted with caution.

Ashoor et al. retrospectively measured thyroid parameters in early-pregnancy samples taken from 202 pregnancies that would subsequently end in miscarriage or fetal loss and 3592 normal pregnancies (68). Although the associations with subclinical thyroid dysfunction were not formally tested (i.e., abnormal TSH with still normal FT₄), the pregnancies complicated by child loss had higher mean TSH and lower FT₄ concentrations, and a higher prevalence of TSH concentrations >97.5th percentile and FT₄ concentrations <2.5th percentile. Finally, early-pregnancy TSH concentrations >95th percentile were associated with an increased risk of miscarriages [odds ratio (OR) 3.66, P = 0.002] in an Australian pregnancy cohort, although subclinical and overt hypothyroid cases were pooled (69). Taken together, these studies do suggest an increased risk of pregnancy loss in pregnancies with subclinical hypothyroidism, but large prospective studies from conception onwards are needed to determine the exact magnitude of effects.

PREMATURE DELIVERY
Premature delivery is the leading direct cause of child death in almost all high- and middle-income countries and is associated with substantial morbidity later in life (70–72). Subclinical hypothyroidism has been described as a risk factor for premature deliveries, although the pathophysiological mechanism remains poorly understood. The largest study on this association has been performed by Casey et al. on a cohort of 17 298 pregnant women presenting for prenatal care (53). Subclinical hypothyroidism (TSH >97.5th percentile and FT₄ within the reference interval) was associated with a slightly increased risk of prematurity <34 weeks (4% vs 2.5%, P = 0.01), borderline significantly associated with prematurity <32 weeks (2.5 vs 1%, P = 0.07), and not associated with prematurity <36 weeks (7 vs 6%, P = 0.39). This is in line with a later study by Cleary-Goldman et al. showing that subclinical hypothyroidism (TSH >97.5th percentile and FT₄ within the reference interval) was not associated with prematurity <37 weeks, whereas the effects on earlier premature deliveries were not

### Table 3. Subclinical thyroid dysfunction during pregnancy and the risk of maternal and child adverse outcomes.

<table>
<thead>
<tr>
<th>Thyroid (dys)function group</th>
<th>Pregnancy loss</th>
<th>Prematurity</th>
<th>Hypertensive disorders</th>
<th>Low birth weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subclinical hyperthyroidism</td>
<td>? (48, 50)</td>
<td>? (46, 48, 50, 56)</td>
<td>(&lt;50, 57, 58)</td>
<td>(&lt;48, 50, 56)</td>
</tr>
<tr>
<td>FT₄ within reference intervals</td>
<td>? (51)</td>
<td>(&lt;51) ^c</td>
<td>(?47, 51)</td>
<td>(?51, 64, 65)^d</td>
</tr>
</tbody>
</table>

^ Reference numbers are shown for studies on respective thyroid (dys)function group and adverse outcome as follows: Korevaar et al. (46), Medici et al. (47), Mannisto et al. (48), Negro et al. (49), Casey et al. (50), Haddow et al. (51), Stagnaro-Green et al. (52), Casey et al. (53), Cleary-Goldman et al. (54), Karakosta et al. (55), Su et al. (56), Mannisto et al. (57), Allan et al. (58), Wilson et al. (59), Ashoor et al. (60), Sahu et al. (61), Karagiannis et al. (62), Wang et al. (63), Medici et al. (64), Shields et al. (65).

^ a Tested only gestational age at birth <37 weeks.

^ b FT₄ concentrations at the upper limit of the reference interval (center-specific reference ranges) associated with lower birth weight and higher risk of SGA (small size for gestational age) newborns.

^ c FT₄ concentrations at the upper limit of the reference interval (center-specific reference ranges) associated with lower birth weight and higher risk of SGA (small size for gestational age) newborns.

^ d FT₄ concentrations at the upper limit of the reference interval (center-specific reference ranges) associated with lower birth weight and higher risk of SGA (small size for gestational age) newborns.
investigated (54). Various other studies have also investigated these relations, with conflicting results (48, 49, 52, 55, 56, 58, 69). This can be partly explained by the fact that some studies pooled overt and subclinical hypothyroid cases (58, 69) and some included a limited number of premature deliveries (55, 56), whereas others used different TSH cutoff values (49, 52). We therefore studied the association between increased TSH concentrations and the risk of premature deliveries using a population-based 97.5th percentile (4.0 mU/L) and a fixed 2.5 mU/L cutoff (46). Although no associations were seen with a TSH >2.5 mU/L, a 1.9- and 2.5-times increased risk of prematurity <37 and <34 weeks was seen among women with a TSH >4.0 mU/L. However, this association no longer persisted after exclusion of TPOAb-positive women or women with comorbidities. This shows that these factors confound the observed associations and underlines the importance of performing indepth analyses in a detailed cohort, taking the interfering role of various confounders into account.

Far fewer data are available on the effects of subclinical hyperthyroidism on prematurity. In a study in women presenting for prenatal care, subclinical hyperthyroidism (n = 224 cases) was not associated with prematurity <36, ≤34, and ≤32 weeks (50). This is in line with a population-based cohort study by Mannisto et al. in which subclinical hyperthyroidism (n = 224 cases) was not associated with prematurity <37 and <34 weeks either (48). Although 2 other population-based studies also did not find any associations, it should be noted that their analyses were limited by a small number of subclinical hyperthyroid cases (n = 77 and 31) (46, 56).

**HYPERTENSIVE DISORDERS**

Hypertensive disorders, including gestational hypertension and (pre)clampsia, are common during pregnancy and are an important cause of maternal and fetal morbidity and mortality (73, 74). Both hypothyroidism have vascular effects, including endothelial cell dysfunction (75, 76), and are associated with an increased risk of hypertensive disorders during pregnancy. Therefore, many studies have also investigated the effects of subclinical thyroid dysfunction on the risk of hypertensive disorders. Although some studies were limited by the small number of subclinical hypothyroid or hypertensive cases (47, 55, 60, 61), a few of these studies were carried out in large pregnancy cohorts (53, 54, 57–59).

In a prospective cohort study in nearly 25,000 pregnancies by Wilson et al., subclinical hypothyroidism (TSH >97.5th percentile and FT₄ 2.5–97.5th percentile) was associated with a 1.6-fold increased risk of severe pre-eclampsia (59). However, the fact that this association disappeared when only women screened before 20 weeks of gestation were included is suggestive of reverse causality (53). This could be due to, for example, placental factors that are increased in preeclampsia and affect thyroid function (77). Indeed, the other large studies did not find a relation between subclinical hypothyroidism in early pregnancy and the risk of subsequent hypertensive disorders (54, 57, 58). The previously mentioned study by Wilson et al. also studied individuals with subclinical hyperthyroid and did not find any effects either (59), as replicated in Finnish and Dutch population-based cohorts (47, 57). Whereas the latter cohort was limited by a small number of subclinical hyperthyroid cases (n = 62), it also was used to investigate the effects of variation in thyroid function within the population-based calculated 2.5–97.5th percentile intervals and revealed an increased risk of preeclampsia in pregnancies with high-normal FT₄ concentrations (47). In contrast, a decreased risk of preeclampsia in pregnancies with high-normal FT₄ concentrations was detected in a recent study by Haddow et al., although these effects were borderline significant and P values were not corrected for multiple testing (51). Therefore, future studies will have to clarify if even variation in FT₄ concentrations within population-specific reference intervals affects the risk of hypertensive disorders during pregnancy.

**LOW BIRTH WEIGHT**

A low birth weight can be due to either SGA or prematurity and has been associated with an increased risk of perinatal morbidity and mortality (78, 79). The previously mentioned study by Cleary-Goldman et al. was the first large study to investigate the relationship between subclinical hypothyroidism and birth weight and showed no effect on the risk of newborns with very low (<2500 g) or high (>4000 g) birth weights (54). A subsequent study by Mannisto et al. investigated these relations with both subclinical hypo- and hyperthyroidism in more detail and did not find effects on the risk of SGA or large size for gestational age newborns either, and it also showed no differences in mean birth weights between these groups (48). A few other studies have investigated these relations with conflicting results, likely to be due to their substantially smaller sample sizes (55, 56, 61–63).

As opposed to studying subclinical thyroid dysfunction groups, Shields et al. were the first to study the relation between continuous FT₄ concentrations and birth weight in a population-based cohort after excluding women with overt thyroid dysfunction, and these authors found a statistically significant negative relation between FT₄ and birth weight (65). These relations have been subsequently studied in Dutch pregnant women with FT₄ concentrations within their center-specific reference intervals, showing that high-normal FT₄ concentrations are associated not only with lower mean birth weights, but also with more SGA and <2500-g newborns (64). These results have recently been convincingly
replicated in the recent study by Haddow et al., which additionally showed that these children do not suffer from more labor/delivery complications (51). Because a low birth weight is a risk factor for cardiovascular and psychiatric diseases in later life (78, 80), it would be interesting to follow these children up for the occurrence of these complications.

Conclusions

In the last decade a large number of studies have been published on thyroid function reference intervals during pregnancy. In the current review we show that there are large differences in TSH and FT4 reference intervals between these populations, with 90% of these studies having higher upper limits of TSH than the fixed TSH cutoff concentrations of 2.5 and 3.0 mU/L that are currently advocated in the guidelines (6–8). Nevertheless, most institutions still rely on the fixed TSH cutoff concentrations of 2.5 and 3.0 mU/L for the first and second trimesters, respectively.

The use of MoMs has illustrated that part of the differences in these intervals between populations can be explained by the use of different assays, and a number of population-specific characteristics such as ethnicity and BMI have also been identified as determinants of reference intervals. Provided that institutions determine their own population-based intervals, there is no direct need for using MoMs in clinical practice. However, the universal use of MoMs in clinical studies on the effects of thyroid dysfunction during pregnancy would certainly be useful, since it will facilitate comparison and metaanalysis of results.

We therefore conclude that institutions should not rely on a fixed universal cutoff concentration worldwide, but should calculate their own pregnancy-specific population-based reference intervals. If such reference intervals are not available, adopting population-based reference intervals from a population with similar characteristics is the best option.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

References