CLINICAL STUDY

Mass screening of newborns for congenital hypothyroidism of central origin by free thyroxine measurement of blood samples on filter paper

Masanori Adachi1,2, Akiko Soneda1, Yumi Asakura1, Koji Muroya1, Yuji Yamagami2,3 and Fumiki Hirahara2,4

1Department of Endocrinology and Metabolism, Kanagawa Children’s Medical Center, Mutsukawa 2-138-4, Minami-ku, Yokohama 232-8555, Japan, 2Neonatal Mass-screening Committee, Kanagawa Prefecture Medical Association, Yokohama, Japan, 3Kanagawa Health Service Association, Yokohama, Japan and 4Department of Obstetrics and Gynecology, Yokohama City University School of Medicine, Yokohama, Japan

(Correspondence should be addressed to M Adachi at Department of Endocrinology and Metabolism, Kanagawa Children’s Medical Center; Email: madachi@mars.sanet.ne.jp)

Abstract

Objective: To evaluate the effectiveness of mass screening of newborns for congenital hypothyroidism of central origin (CH-C) by measurement of free thyroxine (FT4) and thyroid-stimulating hormone (TSH).


Methods: TSH and FT4 levels in dried blood spots on filter paper were measured using ELISA kits, and CH-C was diagnosed at FT4 levels below a cutoff of 0.7 ng/dl (9.0 pmol/l). Survey results were collated with the database created by the screening organizer.

Results: Twenty-four CH-C patients (18 males) were identified, 14 of whom had multiple pituitary hormone deficiencies (group M), eight had isolated CH-C (group I), and two had undetermined pituitary involvement (group U). In groups M, I, and U, the number of patients with FT4 levels below the cutoff value at screening was five (36%), seven (88%), and one (50%) respectively; other patients had been diagnosed clinically. Thus, 13 patients were true positives, while nine were false negatives, yielding screening sensitivity of 59.1% and positive predictive value of 11.5%. The calculated sensitivity was 81.8% at a higher cutoff value of 0.9 ng/dl (11.6 pmol/l). The overall incidence of CH-C was estimated at 1 in 30 833 live births, while that of CH of thyroidal origin (CH-T) is 1 in 3472 live births in Kanagawa prefecture (CH-T/CH-C, 8.9).

Conclusions: Newborn screening with combined FT4 and TSH measurements can identify a significant number of CH-C patients before manifestation of clinical symptoms, but a more appropriate FT4 cutoff value should be considered.

European Journal of Endocrinology 166 829–838

Introduction

Screening of newborns for congenital hypothyroidism (CH) is now routinely used in most of the developed world and in an increasing number of developing countries, which has prevented serious intellectual sequelae in a considerable number of patients with CH (1, 2). While most CH cases are due to CH of thyroidal origin (CH-T) manifesting as thyroid dysgenesis or thyroid hormone synthesis defects, a significant number of CH cases are due to inadequate thyroid-stimulating hormone (TSH) secretion from the anterior pituitary (3, 4, 5, 6, 7, 8, 9). The latter category of CH cases is termed as CH of central origin (CH-C). The incidence of CH-C is estimated to be ~1 in 20 000–30 000 live births (3, 5, 6, 7, 10), which is much higher than previously thought. Nevertheless, CH screening in Japan is mainly based on the detection of elevated TSH levels in dried blood samples on filter paper (primary TSH strategy). This assay has demonstrated high sensitivity in detecting CH-T (11, 12) but failed to identify newborns with CH-C. On the other hand, screening based on the detection of low T4 levels (primary T4 strategy) can identify CH-C newborns only inefficiently, as false-positive cases are inevitable due to both thyroxine-binding globulin (TBG) deficiency and transient low T4 levels in critically ill newborns.

To overcome this situation, The Netherlands has implemented a system of assaying TSH, T4, and TBG, which can eliminate false-positive results caused by TBG deficiency (5, 6). Assaying free T4 (FT4) may be an alternative solution because FT4 is less influenced by TBG than T4. Moreover, determination of FT4 seems to be superior to that of T4 because this reduces false-positive cases in premature newborns, according to the report of a smaller difference between full-term and...
preterm newborns in FT₄ levels than in T₄ levels measured in dried blood samples on filter paper (13). Therefore, in Kanagawa prefecture, we have adopted a strategy of simultaneously measuring TSH and FT₄ in all newborns using a filter paper assay (9). Sapporo city has also adopted the same screening system. The report of a 5-year audit in Sapporo city was released in 2004, in which six CH-C cases were identified through this screening (7). However, the study in Sapporo included only patients showing positive screening results, which preclude evaluation of the sensitivity of screening in detecting CH-C. In addition, the annual birth rate in Sapporo is approximately one-fourth of Kanagawa prefecture.

To evaluate the effectiveness of our CH-C screening system, we have conducted a detailed, comprehensive survey of CH-C patients from Kanagawa region, Japan. In this study, all CH-C cases detected via screening and diagnosed clinically were included and used to estimate the sensitivity and positive predictive value (PPV) of the screening method.

Subjects and methods

Outline of newborn screening system

Kanagawa prefecture, in which Yokohama is the main city, is located in the central region of the Japanese islands, neighboring the Tokyo metropolitan area. The annual number of births in Kanagawa prefecture has been ~70,000 in recent years. The incidence of CH-T in Kanagawa prefecture is estimated to be 1 in 3472 births. Neonatal screening is exclusively conducted by the Neonatal Mass-screening Committee (NMC) of the Kanagawa Prefecture Medical Association (KPMA), which comprises executive officers, technical experts, gynecologists, general pediatricians, and pediatric endocrinologists. The screening procedure adopted by the NMC-KPMA is based on the determination of TSH and FT₄ in dried blood spots on filter paper obtained 4 to 7 days after birth (median sampling day was the fifth day). According to the standard practice followed, newborns with high TSH levels (≥30 μIU/ml) or low FT₄ levels (<0.7 ng/dl of serum (9.0 pmol/l)) are immediately sent to one of the several pediatric endocrine units within the prefecture. A second filter paper sampling is requested for those with borderline TSH levels (15–30 μIU/ml serum) or low FT₄ levels (<0.7 ng/dl of serum (9.0 pmol/l)). If the results again indicate borderline TSH or low FT₄, the baby is sent for a thorough evaluation. Thus, CH-C is suspected if FT₄ levels are low in two consecutive samples. To eliminate cases with transient low FT₄ due to prematurity, samples taken from the newborns with birth weight <2000 g are considered to be preliminary, and the results are sent to each attending physician as an unofficial report. Once the baby attains a weight of 2500 g or reaches 30 days of age, the first sample is requested.

TSH levels in filter paper samples were determined by ELISA using mouse monoclonal antihuman TSH antibodies (Eiken Chemical Co. Ltd., Tochigi, Japan). To determine FT₄ levels in filter paper samples, ENZAPLATE N-FT₄ was used (Siemens Healthcare Diagnostics K.K., Tokyo, Japan), which is an ELISA kit based on a competitive reaction between sample FT₄ and peroxidase-tagged human T₄ to bind to rabbit polyclonal antihuman T₄ antibody (first antibody). A 3 mm disc is punched out from the filter paper and is incubated with peroxidase-tagged T₄ and the first antibody in a reaction mixture of 150 μl for 4 h at 18–25 °C in a micro-well plate with immobilized caprine antirabbit IgG antibodies (second antibody). After removal of the filter paper disc and washing five times, O-phenylenediamine is added, and the absorbance is then measured at 492 nm. A calibration curve is established using standard filter paper samples of known FT₄ concentrations, which are provided by the manufacturer. FT₄ level in the sample is then determined by comparison with the calibration curve.

The performance of this kit, of which there is only one study, reported in a Japanese journal (14), is as follows. The FT₄ determination range is 0.5–5.0 ng/dl, which is based on a precision level lower than 15% of the coefficient of variation (CV). Intra-assay CV is 7.6–15.0%, whereas inter-assay CV is 9.4–18.5%. The correlation between the FT₄ levels measured by this kit and the electrochemiluminescence immunoassay (ECLIA) kit (Elecsys FT₄; Roche Diagnostics) is shown in Fig. 1.

Preliminary survey

A preliminary survey was conducted in December 2008. Questionnaires were sent to all 139 hospitals with a pediatric section in Kanagawa prefecture. The questionnaire included questions about the number of CH-C patients born in Kanagawa prefecture between January 1999 and December 2008 and treated continuously with levothyroxine (L-T₄). CH-C was defined as CH considered to be of hypothalamic or pituitary origin, excluding acquired sequelae of head trauma, brain tumor, etc., and irrespective of involvement of other pituitary functions. Cases of hypothyroxinemia due to prematurity were excluded.

Secondary survey

In April 2009, we requested the corresponding doctors caring for the probable CH-C patients identified in the preliminary survey to provide detailed information, including patient profile, medical complications, data on newborn screening, and results of thyroid function, thyroid imaging studies, and pituitary function tests with imaging information.
infants younger than 2 months. FT4 levels of blood samples on filter paper were measured by an ELISA kit, whereas serum FT4 levels measured by ECLIA was measured by ECLIA.

**Collation study**

After completion of the secondary survey, we collated the list of CH-C patients identified through the above surveys with the NMC-KPMA database, in which information from the first-line investigation at the pediatric endocrine unit and the screening results for all patients with positive screening results had been compiled.

**Patient categorization**

CH-C patients identified through the secondary survey and collation study were categorized into three groups according to the involvement of other pituitary hormones. Group M comprised CH-C patients with at least one pituitary hormone deficiency other than insufficient TSH secretion. These patients were considered to have congenital hypopituitarism with multiple pituitary hormone deficiencies. The diagnosis of each pituitary hormone deficiency was based on the attending physician’s evaluation, except for GH deficiency, which was verified by at least one pharmacological stimulation test. Group I comprised isolated CH-C patients without pituitary involvement other than TSH insufficiency. Group U consisted of CH-C patients for whom pituitary involvement was undetermined.

**Statistical analysis**

Statistical analysis was carried out using Microsoft Office Excel 2007 (Microsoft Corporation). Correlation between the assay results of FT4 (ELISA) in filter paper and serum FT4 by ECLIA was evaluated by linear regression analysis. Mann–Whitney U-test was used to compare FT4 values between groups M and I. Fisher’s exact probability test was used to compare the incidence of screening positive patients according to the etiological categories (groups M and I). P values of < 0.05 were considered to be significant.

The Ethics Committee of Kanagawa Children’s Medical Center reviewed and approved the study procedures.

**Results**

Out of the 139 hospitals from Kanagawa prefecture to which the preliminary survey questionnaire was sent, responses were obtained from 94 hospitals, including 14 hospitals stating that they currently had no pediatric section. Accordingly, the actual response rate was calculated to be 64.0% (80/125 hospitals with pediatric sections). Through this primary survey, 42 patients with probable CH-C (2–11 years old) were identified at 14 out of the 80 hospitals.

Figure 2 shows the number of CH-C patients, both probable and confirmed, identified through the surveys. The preliminary survey identified 42 probable patients, of which 20 patients were considered to represent true CH-C cases. After collation with the NMC-KPMA database, 24 CH-C patients (of which 18 were male) were finally identified. Details of each patient are summarized in Tables 1 and 2. As the total number of newborns screened during the study period was 740 003, we calculated the minimal incidence of samples and serum FT4 (ECLIA) was evaluated by linear regression analysis. Mann–Whitney U-test was used to compare FT4 values between groups M and I. Fisher’s exact probability test was used to compare the incidence of screening positive patients according to the etiological categories (groups M and I). P values of < 0.05 were considered to be significant.
Table 1 Characteristics of 14 patients with CH-C with multiple pituitary hormone deficiencies (group M). Patient 13 was already on L-thyroxine treatment at the time of screening.

<table>
<thead>
<tr>
<th>Pt. no., sex</th>
<th>Year</th>
<th>Wt (g)</th>
<th>Diagnostic symptom (age)</th>
<th>FT₄ values at Sc (ng/dl)</th>
<th>At first presentation</th>
<th>Deficient pituitary hormones</th>
<th>MR imaging of the CN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1st sample (1st sample (day)</td>
<td>Serum TSH (mIU/ml)</td>
<td>Serum FT₄ (ng/dl)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2nd sample (day)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1, F</td>
<td>1999</td>
<td>2872</td>
<td>Low vision (4M)</td>
<td>0.90 (5)</td>
<td>1.59</td>
<td>0.86</td>
<td>TSH, GH, AVP</td>
</tr>
<tr>
<td>2, F</td>
<td>2000</td>
<td>3352</td>
<td>SS (1Y)</td>
<td>1.42 (4)</td>
<td>4.34</td>
<td>0.75</td>
<td>TSH, GH, LH/FSH</td>
</tr>
<tr>
<td>3, F</td>
<td>2000</td>
<td>2994</td>
<td>SS (4Y)</td>
<td>1.80 (5)</td>
<td>0.79</td>
<td>0.84</td>
<td>TSH, GH, AVP</td>
</tr>
<tr>
<td>4, M</td>
<td>2000</td>
<td>4420</td>
<td>Icterus (2M)</td>
<td>0.81 (26)</td>
<td>2.90</td>
<td>1.00</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td>5, M</td>
<td>2001</td>
<td>3240</td>
<td>Shock (1D)</td>
<td>0.83 (7)</td>
<td>7.40</td>
<td>0.70</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td>6, F</td>
<td>2002</td>
<td>3150</td>
<td>Shock (1D)</td>
<td>0.58 (5)</td>
<td>10.28</td>
<td>0.65</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td>7, M</td>
<td>2002</td>
<td>2342</td>
<td>Seizure (1Y)</td>
<td>0.81 (7)</td>
<td>1.71</td>
<td>0.42</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td>8, M</td>
<td>2004</td>
<td>2275</td>
<td>Sc (23D)</td>
<td>0.48 (5)</td>
<td>3.77</td>
<td>0.76</td>
<td>TSH, ACTH</td>
</tr>
<tr>
<td>9, M</td>
<td>2005</td>
<td>3135</td>
<td>Sc (22D)</td>
<td>0.55 (5)</td>
<td>6.58</td>
<td>0.66</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td>10, M</td>
<td>2005</td>
<td>2972</td>
<td>SS, microgenis (1Y)</td>
<td>2.02 (5)</td>
<td>3.85</td>
<td>0.99</td>
<td>TSH, LH/FSH</td>
</tr>
<tr>
<td>11, M</td>
<td>2005</td>
<td>3168</td>
<td>Sc (31D)</td>
<td>0.37 (14)</td>
<td>3.29</td>
<td>0.53</td>
<td>TSH, ACTH</td>
</tr>
<tr>
<td>12, M</td>
<td>2007</td>
<td>1786</td>
<td>Follow up of HP (4M)</td>
<td>1.10 (6)</td>
<td>0.19</td>
<td>0.96</td>
<td>TSH, GH, ACTH, AVP</td>
</tr>
<tr>
<td>13, M</td>
<td>2007</td>
<td>3122</td>
<td>Hypoglycemia (2D)</td>
<td>Not tested</td>
<td>3.34</td>
<td>0.88</td>
<td>TSH, GH, ACTH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14, M</td>
<td>2007</td>
<td>3445</td>
<td>Sc (31D)</td>
<td>0.43 (6)</td>
<td>2.35</td>
<td>0.60</td>
<td>TSH, GH, ACTH</td>
</tr>
</tbody>
</table>

Pt. no., patient number; Wt, weight; Sc, screening; D, days old; M, months old; Y, years old; AVP, arginine vasopressin; PRL, prolactin; TSH, thyroid-stimulating hormone. APS, absent pituitary stalk; EPP, ectopic posterior pituitary; APP, absent posterior pituitary; ONH, optic nerve hypoplasia; ASP, absent septum pellucidum; PH, pituitary hypoplasia; HP, holoprosencephaly; SS, short stature; MR, magnetic resonance.

*aThis patient was born with low birth weight and hence this value was treated as unofficial.*
<table>
<thead>
<tr>
<th>Pt. no., sex</th>
<th>Year</th>
<th>Weight (g)</th>
<th>Diagnostic symptom (age)</th>
<th>FT₄ values at Sc (ng/dl)</th>
<th>At first presentation</th>
<th>Basis for diagnosis of hypothyroidism</th>
<th>MR imaging of the CN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1st sample</td>
<td>2nd sample</td>
<td>Serum TSH (μIU/ml)</td>
<td>Serum FT₄ (ng/dl)</td>
</tr>
<tr>
<td>15, M</td>
<td>2003</td>
<td>3370</td>
<td>Sc (14D)</td>
<td>0.14 (5)</td>
<td>0.48 (14)</td>
<td>2.86</td>
<td>0.45</td>
</tr>
<tr>
<td>16, M</td>
<td>2004</td>
<td>2770</td>
<td>SS (2Y)</td>
<td>1.79 (5)</td>
<td>2.20</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>17, M</td>
<td>2006</td>
<td>3450</td>
<td>Sc (15D)</td>
<td>0.60 (4)</td>
<td>0.47 (15)</td>
<td>2.79</td>
<td>1.01</td>
</tr>
<tr>
<td>18, M</td>
<td>2008</td>
<td>3060</td>
<td>Sc (24D)</td>
<td>0.68 (13)</td>
<td>0.68 (24)</td>
<td>1.86</td>
<td>0.94</td>
</tr>
<tr>
<td>19, M</td>
<td>2008</td>
<td>3868</td>
<td>Sc (12D)</td>
<td>0.43 (5)</td>
<td>0.66 (12)</td>
<td>3.02</td>
<td>0.72</td>
</tr>
<tr>
<td>20, F</td>
<td>2007</td>
<td>3262</td>
<td>Sc (13D)</td>
<td>0.50 (5)</td>
<td>0.60 (13)</td>
<td>2.28</td>
<td>0.73</td>
</tr>
<tr>
<td>21, M</td>
<td>2008</td>
<td>3440</td>
<td>Sc (13D)</td>
<td>0.69 (4)</td>
<td>0.53 (13)</td>
<td>2.13</td>
<td>0.70</td>
</tr>
<tr>
<td>22, M</td>
<td>2008</td>
<td>3145</td>
<td>Sc (20D)</td>
<td>0.21 (5)</td>
<td>0.50 (20)</td>
<td>2.34</td>
<td>0.43</td>
</tr>
<tr>
<td>23, F</td>
<td>2007</td>
<td>668</td>
<td>Follow-up of low birth weight (1M)</td>
<td>0.27* (4)</td>
<td>Unknown</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>24, M</td>
<td>2008</td>
<td>2542</td>
<td>Sc (15D)</td>
<td>0.57 (4)</td>
<td>0.57 (15)</td>
<td>1.24</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Pt. no., patient number; D, days old; M, months old; Y, years old; L-T₄, levothyroxine; PH, pituitary hypoplasia; SS, short stature; Sc, Screening; TSH-R, TSH response; NFR, normal FT₄ range; MR, magnetic resonance; ND, not done.

*This patient had low birth weight, and the data obtained at 4 days of age were treated as unofficial. L-Thyroxine therapy was initiated before her first official sample was obtained.
CH-C in Kanagawa prefecture as 1 in 30 833 births (24/740 003).

Among the 24 patients, 14 patients (58%, ten males) were categorized into group M (Fig. 3). Group M (n=14) consisted of five patients with septo-optic dysplasia, five patients with pituitary hypoplasia, one with holoprosencephaly, and three with normal pituitary morphology. Eight other patients out of the 24 (33%) were considered to have isolated CH-C, without pituitary involvement, and they were hence categorized as group I (Fig. 3). Pituitary function in the remaining two patients could not be fully evaluated because of their younger age, and they were therefore categorized as group U (Fig. 3).

Twelve patients (50%) were identified as having CH-C solely via the newborn screening system in Kanagawa prefecture (Fig. 3). Of these, four patients belonged to group M, seven patients to group I, and one patient to group U. In addition, patient 6 in group M was clinically diagnosed with CH-C because this patient exhibited shock; however, the screening result was actually positive (low FT4 levels), and hence, this was considered as a true-positive case of CH-C. Therefore, the total number of true-positive CH-C cases was 13. In contrast, nine other patients out of 24 (38%, eight patients in group M and one in group I) had normal screening results and were revealed to have CH-C through the evaluation of clinical symptoms such as shock and/or hypoglycemia during the neonatal period (n=2), short stature (n=4), and other features (n=3). These nine patients were considered to be false negatives. The remaining two patients (one in group M and one in group U, depicted as ‘?’ in Fig. 3) were already on l-T4 treatment before screening, and hence, they were excluded from the judgment as to whether the screening results were positive or negative as they had already been diagnosed with CH-C. Patients in group I were significantly identified more frequently through the screening program than those in group M: 88% (7/8) vs 29% (4/14), P<0.01.

Out of the 24 CH-C patients, for 22 patients the filter paper assay for FT4 showed clear positive or negative results during screening (Fig. 4). The remaining two patients had been started on l-T4 therapy before screening. Because no blood samples were collected from any patient between 8 and 11 days of age, FT4 measurements were arbitrarily divided into those obtained on or before 10 days of age (FT4 before 10D, n=18, collected from 18 patients) and those obtained on or after 11 days of age (FT4 after 11D, n=16, collected from 14 patients). Overall, the FT4 level before 10D was 0.82±0.56 ng/dl (median, 5 days of age; range, 4–7 days), whereas the FT4 level after 11D was 0.57±0.13 ng/dl (median, 17.5 days; range, 12–31 days; Fig. 4). In addition, when we analyzed the data exclusively obtained from patients whose FT4 levels had been determined twice (n=10), no significant difference was observed between FT4 before 10D (0.46±0.05 ng/dl) and FT4 after 11D (0.52±0.02 ng/dl). Thus, FT4 values in CH-C patients appeared to be stable during the neonatal period.

A comparison of FT4 levels in group M (n=17) with those in group I (n=15) also did not show a statistically significant difference (group M, 0.81±0.49 ng/dl; group I, 0.60±0.37 ng/dl), indicating that the severity of hypothyroidism did not differ significantly between these two groups, differentiated by pituitary involvement.

Evaluation of the performance of the screening system is depicted in Table 3. Our screening system yielded 13 true positives and nine false negatives, so that the sensitivity of detection of a true positive was calculated to be 59.1%. Specificity and PPVs were calculated to be 99.99 and 11.5% respectively. A total of 740 003 newborns were screened during the study period and 113 newborns were sent for thorough evaluation based on two consecutive FT4 measurements. The cutoff level used was 0.7 ng/dl serum (9.0 pmol/l). In the next step, we simulated the performance of the screening system with higher cutoff values. As depicted in Fig. 4, FT4 levels for nine patients who were not identified in the screening ranged from 0.81 to 2.02 ng/dl (median, 0.9 ng/dl), which was substantially lower than the reference range of

<table>
<thead>
<tr>
<th>Multiple pituitary hormone deficiencies (group M)</th>
<th>Isolated hypothyroidism (group I)</th>
<th>Undetermined (group U)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="n=10" alt="Symptom-based diagnosis" /></td>
<td><img src="n=12" alt="Screening-based diagnosis" /></td>
<td><img src="n=2" alt="High-risk follow up" /></td>
</tr>
</tbody>
</table>

**Figure 3** Summary of the 24 patients with CH-C, categorized by presence/absence of other pituitary hormone deficiencies and diagnostic symptoms. The red and blue figures indicate female and male patients respectively. Figures within a single-line box indicate CH-C patients who could have been identified as having CH-C by screening if the FT4 cutoff values were 0.9 ng/dl according to the results of the filter paper assay. lFT4 data with the filter paper assay were not available for two patients; l-thyroxine treatment was initiated in one male patient at 2 days of age. One female patient had low birth weight, and the data obtained at 4 days of age were treated as unofficial. l-Thyroxine therapy was initiated before the first official sample was obtained from this patient.
1.64–2.80 ng/dl (21.1–36.0 pmol/l; data obtained from the 67,933 normal newborns). If the cutoff value is raised to 0.9 ng/dl serum (11.6 pmol/l), then an additional five patients would have been found to be positive by the screening, and the estimated sensitivity would be increased by 81.8%.

Discussion

In Japan, two types of ELISA-based kits are available for measuring FT₄ levels in dried blood samples on filter paper: one developed by Siemens Healthcare Diagnostics K.K and another by Eiken Chemical Co. Ltd. Because TSH and FT₄ can be measured with a common detection module, additional costs for FT₄ measurements are only those incurred for reagents: 465 yen for TSH alone vs 705 yen for TSH and FT₄ determination per newborn examined. Most of the screening centers adopt a primary TSH and backup FT₄ system; the filter paper method is used for measuring TSH in all newborns, while it is used for measuring FT₄ only in those with high TSH values for confirmation of possible hypothyroidism (11, 12). To detect CH-C, certain areas, including Kanagawa prefecture and Sapporo city, have adopted a combined primary TSH–FT₄ screening system (7, 9).

After the report from Sapporo city (7), this report is the second audit of this CH-C newborn screening system, conducted on a larger population and for a longer study period. We also tried to trace CH-C patients not identified by neonatal screening (false-negative cases).

The ELISA-based filter paper FT₄ kits are almost exclusively used in Japan. One may argue against its accuracy in determining FT₄ levels, considering that some TBG-deficient patients were falsely detected to have low FT₄ levels and that the equilibrium dialysis method is the gold standard (15, 16). However, it has been difficult to introduce the equilibrium dialysis method in newborn screening because it requires a larger volume serum sample and longer measurement times. On the other hand, to use the FT₄ index instead of FT₄, tri-iodothyronine (T₃) uptake must also be measured, which increases cost. FT₄ determined by ELISA on filter paper blood samples seems to correctly reflect FT₄ status in newborns because most (88%) FT₄ values in CH-C patients were more than 2 S.D. below the mean of normal newborns and because FT₄ levels in CH-T were distributed in a substantially low range (0.04–2.32 ng/dl; Fig. 4). Moreover, Fig. 4 shows that the FT₄ levels measured using the filter paper method may be consistent even at lower concentrations of FT₄.

Thus, we believe that although FT₄ levels determined using the filter paper samples may not be identical to those measured by the equilibrium dialysis method, the assay is a promising, practical alternative for use in CH-C screening. Because combined TSH–T₄ is recommended as the ideal strategy for detecting both CH-C and CH-T by the American Thyroid Association and Pediatric Endocrine Societies in the US and Europe (2), we think it is justified to continue implementation of our combined TSH–FT₄ system as a new version of the TSH–T₄ system.

From our survey, the incidence of CH-C was calculated as 1 in 30,833 live births, while that of CH-T was 1 in 3472 live births. Thus, the CH-T/CH-C ratio in this study was 8.9, which is close to the ratio 8.4 reported from
The Netherlands (6). Although we previously reported a much lower CH-C incidence (1 in 160,516 births) (9), that survey was based on only the cases detected through screening. The incidence rate of 1 in 30,833 reported here is likely to be underestimated because this study was based on a questionnaire survey and false-negative cases may not have been recorded. Indeed, we could not obtain follow-up data on 11 cases identified in the preliminary survey, as well as on eight patients with positive screening results. In addition, because correct diagnosis of CH-C is difficult (17, 18), especially in those with isolated hypothyroidism, some cases may have been overlooked. Moreover, as shown in Fig. 4, the mean values and range of FT4 in CH-C patients were lower than those in CH-T patients, suggesting that milder forms of CH-C may escape detection.

A remarkable finding in this study is that isolated hypothyroidism (group I) was detected in one-third of the total CH-C population. Previous studies have found that 78% (5) to 98% (8) of CH-C patients had multiple pituitary hormone defects such as septo-optic dysplasia. There are some explanations for this discrepancy. First, isolated CH-C patients present less prominent symptoms (19, 20, 21, 22) and hence may be missed in the absence of screening. Indeed, all but one patient in group I was identified through the newborn screening. A Dutch screening system with TSH, T4, and TBG determination (5) reported a prevalence rate of 22% of isolated CH-C, which is closer to our findings. Secondly, ethnic differences may be a factor: in Sapporo city, two of six CH-C patients were reported to demonstrate isolated hypothyroidism (7). Thirdly, some patients may not have been correctly diagnosed: a patient (patient 21 in Table 2) with pituitary hypoplasia is likely to have other hormone deficiencies. Finally, transient hypothyroidism may not be definitively ruled out, especially in younger patients. However, the authors are aware of a patient in group I (patient 15) who demonstrated severe hypothyroidism when L-T4 therapy was tentatively interrupted. Reevaluation of all other patients in group I will determine the true incidence of isolated hypothyroidism.

Our current system yielded a sensitivity of 59.1% and PPV of 11.0% in detecting CH-C. In fact, 12 patients were diagnosed with CH-C entirely on the basis of low FT4 levels at newborn screening. Above all, the presence of four patients in group M, who were overlooked clinically but in whom low FT4 levels were detected at screening, underscores the usefulness of our combined primary TSH–FT4 system. The sensitivity of 59.1% seems superior to the reported sensitivity of 19.0% in the state of Indiana, USA, where T4 measurement was used (8). On the other hand, a study from The Netherlands reported the sensitivity to be 71.4% (6). Because our study relied on responses to a questionnaire, the actual sensitivity of our screening system may be lower: physicians who did not respond may have
CH-C patients who were missed in the screening, some unrecognized cases with isolated hypothyroidism may be present, and early death of patients with multiple pituitary hormone deficiencies may have been ignored. Thus, we cannot directly compare the performance of our system to that used in The Netherlands.

Setting a higher cutoff for FT4 has both advantages and disadvantages. At a cutoff of 0.9 ng/dl instead of 0.7 ng/dl, the estimated sensitivity rises to 81.8%; it may increase three times more considering the requests for a second filter paper test (retesting ratio, 0.50%). In Sapporo city, the FT4 cutoff has been set to 1.0 ng/dl (7); six CH-C patients were identified through the screening over 4 years, and the prevalence of CH-C was reported to be 1 in 13,872 live births (Table 3). If a cutoff of 0.7 ng/dl were to be applied to their cohort, only one of six CH-C patients would have been detected by screening. Thus, resetting the cutoff value to 0.9 ng/dl (or higher) may be necessary. This level is in accordance with the FT4 cutoff 0.9 ng/dl used in The Netherlands for the diagnosis of CH-C (5) and is ~2 S.D. of both the reported cord blood values (23) and our FT4 values (Fig. 3) for normal newborns.

Adequacy of the retesting ratio depends on many factors including population size, system performance parameters such as sensitivity and PPV, and local economic conditions. The retesting ratio was as high as 0.76% in Sapporo city, due to a higher cutoff value and inclusion of low-birth weight newborns. This ratio may be acceptable in a smaller city but may not be suitable for Kanagawa prefecture. The estimated retesting ratio of 0.50% (3735 samples during 10 years) resulting from a cutoff of 0.9 ng/dl may be more acceptable than a ratio of 0.76%. A comparative figure for retesting for congenital adrenal hyperplasia in Kanagawa prefecture is 0.3%. Because FT4 determinants did not change significantly according to collection dates (Fig. 4), as shown also in normal newborns (13), differential cutoff values according to the sampling dates are not expected to reduce the retesting ratio.

Another problem is the introduction of an immediate evaluation system to facilitate early treatment of CH-C patients, especially for those with multiple pituitary hormone deficiencies. As stated in the Subjects and methods section, unless two consecutive tests reveal low FT4 values, newborns will not be subject to a thorough evaluation in our system. The aim of performing a second sampling is to exclude false positives. Indeed, during the study period, second samples were requested for 1,220 newborns, but only 113 of these newborns were sent for thorough evaluation and 1,107 false-positive cases were eliminated (Table 3). Even with a cutoff of 1.0 ng/dl, the number of newborns sent for evaluation will increase minimally (79 additional cases across 10 years). We retrospectively analyzed the impact of introduction of an immediate evaluation system in which newborns with FT4 lower than 0.5 ng/dl (6.4 pmol/l) will be immediately evaluated.

Seven CH-C patients, including three patients in group M, would have been diagnosed without delay. However, according to our simulation, this strategy will create a false-positive number of more than 200 over 10 years, with a PPV of 2.8%. The question of whether this figure is reasonable is beyond the scope of our study. Nevertheless, we plan to improve our screening strategy by considering scientific, economical, ethical, and political issues.

In conclusion, measurement of FT4 in dried blood spots on filter paper is suitable for newborn screening for CH-C; moreover, the combined primary TSH–FT4 system applied in Kanagawa prefecture identified a significant number of CH-C patients before they manifested clinical symptoms. The survey identified 24 CH-C patients, 14 of whom had multiple pituitary hormone deficiencies, yielding an incidence rate of CH-C of 1 in 30,833 live births. Screening sensitivity was calculated to be 59.1%, based on 13 true-positive cases and nine false-negative cases, with a cutoff of 0.7 ng/dl of FT4. A more appropriate (higher) FT4 cutoff value and proper implementation of the screening would facilitate early detection of CH-C cases.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Author contribution statement
M Adachi, Y Yamagami, and F Hirahara conceptualized and designed the study. M Adachi and A Soneda contributed to the data collection, analysis, and writing of the manuscript. Y Asakura and K Muroya contributed to preparation of the manuscript by critically analyzing it.

Acknowledgements
The authors wish to thank all the pediatric doctors who participated in the survey. In addition, the authors are grateful to Ms. Kamiko Kagiya for her assistance with data handling. Finally, they wish to thank Mr. Masaru Fukushi (Sapporo Immuno Diagnostic Laboratory, Sapporo, Japan) for his invaluable advice in preparing the manuscript.

References


