Longitudinal study of thyroid function in Down’s syndrome in the first two decades

P A Gibson, R W Newton, K Selby, D A Price, K Leyland, G M Addison

Aims and Methods: Thyroid function tests were initially carried out on 122 children with Down’s syndrome aged 6–14 years and then repeated four to six years later in 103 adolescents (85% of the group of 122) when they were aged 10–20 years (median 14.4 years). At the second test two were hypothyroid and two with isolated raised thyroid stimulating hormone (IR-TSH) were receiving thyroxine.

Results: At the first test there were 98 (80%) euthyroid children: 83 were retested and four (5%) had IR-TSH. At the first test 24 had IR-TSH: 20 were retested and 14 (70%) had become normal. Seventeen with IR-TSH on initial testing had a thyrotrophin releasing hormone test within three months; TSH had become normal in eight (47%) of these children. There was no association between reported clinical symptoms and IR-TSH, but there were clear symptoms in one of the two with definite hypothyroidism.

Conclusions: The likelihood ratio for a positive result on second testing when raised TSH and positive antibody status on first testing are combined is 20. This suggests initial testing results could be used as a basis to select a subgroup for further testing at say five yearly intervals unless new symptoms emerge in the interim. It also suggests that yearly screening (as recommended by the American Academy of Pediatrics, 2001) is probably not justified in the first 20 years of life.

The association between thyroid dysfunction and Down’s syndrome is well recognised, but the natural history of this dysfunction is not known. To date published results are presented as cumulative incidence or prevalence rather than incidence data. This serves to overemphasise the frequency of abnormality and has led many paediatricians to check thyroid function in those with Down’s syndrome on an annual basis in childhood and adolescence.

We report a longitudinal cohort study. The design has allowed us to define the properties of thyroid function tests carried out late in the first decade and to what extent they predict subsequent abnormality in the second decade. This enables a rational approach in decision making analysis to define target groups for subsequent testing.

Prasher1 cites 30 case reports of hypothyroidism and 19 of hyperthyroidism. Smith2 was probably the first to report the use of thyroid extract in a person with Down’s syndrome. A number of cross-sectional studies of thyroid function in individuals with Down’s syndrome have shown an increased prevalence of both congenital hypothyroidism and acquired thyroid dysfunction. More recently, larger studies have defined the pattern of abnormality in more detail, though study design has varied. The studies of Toledo and colleagues,5 Prasher,1 Gruneiro de Papendieck and colleagues,6 and Noble and colleagues7 were cross-sectional studies, while those of Rubello and colleagues,8 Ivarsson and colleagues,9 and Tuysuz and Beker12 were all longitudinal. The research of Toledo and colleagues8 and Rubello and colleagues9 was hospital based, whereas the work of Prasher,1 Rooney and Walsh,10 Karlsson and colleagues,11 and Noble and colleagues7 was community based. Prasher studied adults,1 Rooney and Walsh10 and Rubello and colleagues9 studied individuals from childhood to the sixth decade, and the others mainly studied children and adolescents.

All reported a significantly raised frequency of thyroid abnormality, which Rubello and colleagues10 and Rooney and Walsh10 showed rises with age, particularly after the age of

Figure 1 Study population.

<table>
<thead>
<tr>
<th>Hester Adrian Research (Social Science)</th>
<th>referrals (1973–1980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>180 children recruited</td>
<td></td>
</tr>
</tbody>
</table>

161

Lost contact

17

144

Refused study

22

122

Receiving thyroxine

1

Medical Study

(1986–1987)

121

Medical Study Second Assessment

(1990–1993)

103

Died

19

Lost contact or refused reassessment

18
Euthyroidism. The study of van Trotsenburg and colleagues observed the absolute thresholds between hypothyroidism, IR-TSH, and euthyroidism. Variability in definitions as well as laboratory techniques and population identification in part explain the wide range in quoted rates of thyroid dysfunction in Down's syndrome. Between 1973 and 1980, Cunningham, working at the Hester Adrian Research Centre, University of Manchester, recruited a population based cohort of 180 children with Down's syndrome in Manchester, England. These children and their families were studied prospectively in detail. Study population details are described by Selby and colleagues and summarised in fig 1.

Blood was first collected from 122 of this cohort when they were aged 6–14 years (median 9.8 years). Blood samples were analysed for thyroid stimulating hormone (thyrotrophin, TSH), thyroid binding globulin, total thyroxine (T4) (quoted according to thyroid binding globulin), and the two autoantibodies thyroglobulin and thyroid microsomal antibody. The more specific anti-thyroperoxidase assay was not available at first testing and was not pursued at second testing.

The investigating team usually performed blood sampling and analysis. Occasionally blood had already been taken by the local paediatric team; in these cases results were collated from the participating hospital.

At the time of the initial sampling, clinical data were collected on hair, appetite, bowel function, height, weight, and family history of autoimmune disease.

For the purposes of this study the following definitions are used:
- **Hypothyroidism**: low thyroxine and TSH of 6 µu/ml or more
- **Isolated raised TSH (IR-TSH)**: normal thyroxine and TSH of 6 µu/ml or more
- **Euthyroid**: normal thyroxine and TSH less than 6 µu/ml
- **Positive autoantibodies**: titre greater than 1:64.

We assessed whether testing aspects of thyroid function in the first decade (“first testing” in table 1) was predictive of hypothyroidism in the second decade. We express results in terms of the following functions to help guide clinical decision making.

The properties of tests are defined as follows:

- **Sensitivity** is the proportion of children with hypothyroidism aged 10–20 positive at first testing, a/(a+c)
- **Specificity** is the proportion of children with normal thyroid function aged 10–20 negative at first testing, d/(b+d)
- **Positive predictive value** is the proportion of children with hypothyroidism aged 10–20, of all children positive at first testing, a/(a+b)
- **Negative predictive value** is the proportion who are normal aged 10–20, of all children negative at first testing d/(c+d)
- **Likelihood ratio for a positive test result** is the likelihood that a positive first test predicts hypothyroidism aged 10–20 compared to the likelihood that a positive first test predicts normal thyroid function aged 10–20, [a/(a+c)]/[b/(b+d)]
- **Likelihood ratio for a negative test result** is the likelihood that a negative first test predicts hypothyroidism aged 10–20 compared to the likelihood that a negative first test predicts normal thyroid function aged 10–20, [c/(c+d)]/[d/(d+b)]

### Table 1: Testing of thyroid function in the first decade

<table>
<thead>
<tr>
<th></th>
<th>Hypothyroidism, 2nd test (aged 10–20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>First testing</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>a</td>
</tr>
<tr>
<td>Negative</td>
<td>c</td>
</tr>
</tbody>
</table>

Figure 2: Fagan’s likelihood ratio nomogram. Known population pre-test probabilities set against likelihood ratios derived for a given test set against the nomogram allow the prediction an individual’s probability of a future abnormal result. Adapted from Fagan TJ (N Engl J Med 1975;293:257). Copyright 1975, New England Journal of Medicine. All rights reserved.
Likelihood ratios can be used in clinical practice to influence management decisions. Referring to Fagan's nomogram (fig 2) it can be seen that a likelihood ratio of unity makes the probability of a positive or negative result aged 10–20 the same as the pretest probability. Likelihood ratios greater than unity magnify the probability of a like result on testing aged 10–20. Likelihood ratios less than unity diminish the probability of a like result on testing aged 10–20.

RESULTS

In 1986–87, of 122 children, one child had been commenced on thyroxine by a local paediatrician prior to the first test. We reviewed this 13 year old's data: TSH 9.2 mU/l, T4 93 nmol/l, thyroid binding globulin 26.3 mg/l. We identified this as a case of IR-TSH.

Ninety-eight of the remaining 121 children had normal thyroid function and 23 had IR-TSH. Between first sampling and resampling in 1990–93, 18 cases were lost to follow up, and the individual on thyroxine was excluded from further analysis because she remained on thyroxine (fig 1). Of the group of 103 individuals resampled, 91 had normal thyroid function, 10 had IR-TSH (one of whom had been commenced, in our view unnecessarily on thyroxine by her local paediatrician), and two cases of definite hypothyroidism were identified. Treatment was commenced on the two newly identified cases (one from the 1986–87 normal thyroid function group and one from the IR-TSH group).

Table 2 compares thyroid function on second testing with thyroid function on first testing. Fourteen who remained untreated with initial IR-TSH showed normal results on second testing. Table 3 combines the euthyroid and IR-TSH groups aged 10–20 as “normal”.

Table 4 shows the autoantibody status at first and second testing. Data are available on 101 at initial testing (data missing on two of the 103, one of whom developed hypothyroidism) and eight were positive (all eight for antimicrosomal antibody and two of these additionally for antithyroglobulin antibody). Between first and second testing three became normal, and two initially negative became positive. For the eight with positive autoantibodies on first testing, five were still positive on second testing, three retained IR-TSH, and one developed hypothyroidism. The association between positive autoantibodies on first testing and a second abnormal test (combining hypothyroidism, IR-TSH, and positive antibodies) was significant (Fisher’s exact p < 0.05), but not with hypothyroidism on its own (Fisher’s exact p < 0.5). There was no significant sex difference in autoantibody status.

Table 5 shows the results of 101 children with available results at 10–20 years and their combined thyroid function and antibody status at initial testing.

DISCUSSION

The study confirms the high prevalence of thyroid dysfunction in Down’s syndrome with 12 of 103 showing abnormal results on second testing. The cause of the IR-TSH has not been elucidated. The results may indicate a predisposition in young people with Down’s syndrome to a self-limiting autoimmune process resulting in IR-TSH without clinical symptoms. However, autoantibodies were only seen in nine at first testing and in seven at second testing. On both occasions a positive association was seen with IR-TSH but not with hypothyroidism. There are other possible mechanisms for self-limiting subclinical thyroid dysfunction. Possible explanations include an inappropriate release of TSH related to a central disorder, the production of a less active form of TSH, or some form of TSH insensitivity in the thyroid gland. It is not clear how this would involve a self-limiting course but slow maturation of negative feedback control systems in the hypothalmo-pituitary axis in Down’s syndrome, is possible. Van Trotsenburg and colleagues reported similar findings in neonates with Down’s syndrome and recommended more functional and molecular studies where autoimmunity was not involved.

To justify the introduction of a screening programme the condition to be identified should be relatively common, cause a significant health risk if not identified, be distinguished by a test that is relatively specific, reliable, acceptable, and...
cheap, and have a satisfactory treatment. Thyroid function testing in Down’s syndrome meets these.

There are however, other considerations. Is there a latent or early stage and is the natural history understood? Our study highlights how results may change during a short period of time. Eight of 17 children with a raised TSH on initial testing had a normal result on retesting within three months and 14 of the 20 children with IR-TSH were normal on second testing. This indicates that the natural history of the condition for many young people is one of recovery. While the debate continues over the best screening method and frequent retesting of IR-TSH is not indicated. Our study shows three potentially important points for clinical practice:

- Where the development of age specific normal ranges for TSH and T4 in Down’s syndrome is one possible approach for the future, we note that there are only small variations where this has been done for the general population. We recommend the simpler approach of recognising that IR-TSH in Down’s syndrome is frequently self-limiting without the need for treatment.
- There is no association between reported clinical symptoms and IR-TSH. There were clear symptoms in one of the two with definite hypothyroidism. Clinicians should pay particular attention to symptoms explainable on the basis of hypothyroidism when parents or the young people with Down’s syndrome themselves report these, remembering some of these symptoms can at times be features of Down’s syndrome.
- We emphasise that the positive likelihood ratio for the combined abnormal autoantibody status and IR-TSH result on first testing is 20 (and 14 for autoantibody alone). The clinical relevance of this can be seen from the use of Fagan’s nomogram. Before a child with Down’s syndrome has a first decade test, the probability of needing thyroxine in the second decade is, say, 2%. After testing positive for autoantibodies alone the probability of hypothyroidism in the second decade is 28%, which rises to 34% when IR-TSH is also identified and included. It is equally important to note that testing negative for autoantibodies with a normal TSH causes the probability of hypothyroidism in the second decade to become exceptionally small.

We therefore propose that a decision on retesting is based on the initial test results. In our first decade group of 122, regular retesting could have been confined to 25. This would have been cost effective and minimised inconvenience and distress for the majority. Our results challenge the validity of population based screening programmes for thyroid dysfunction in Down’s syndrome. The vast majority of those young people with thyroid dysfunction were shown to have a self-limiting condition and those with definite hypothyroidism in our series had definite symptomatology.

Conclusions
Hypothyroidism in Down’s syndrome should not be over-diagnosed. Knowledge of IR-TSH and its frequent self-limiting natural history needs dissemination. Treatment and frequent retesting of IR-TSH is not indicated. Our data suggest that early positive results for autoantibodies or

<p>| Table 3 | Results of 103 children with available results for outcome aged 10–20 years |
|-----------------------------------------------|
| Hypothyroidism, 2nd test                      |</p>
<table>
<thead>
<tr>
<th>Hypothyroidism</th>
<th>“Normal”</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroid function, 1st test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolated raised TSH</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Euthyroid</td>
<td>2</td>
<td>81</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>100</td>
</tr>
</tbody>
</table>

The children are defined at age 10–20 years as having hypothyroidism or “normal” (IR-TSH is included as “normal”, meaning not having hypothyroidism, and not needing therapy). Sensitivity 50% (95% CI 40–60%); specificity 82% (95% CI 74–90%). Positive predictive value 3%; negative predictive value 99%. Likelihood ratio for positive test result 2.7; likelihood ratio for negative test result 0.62.

<p>| Table 4 | Results of 101 children with available results at 10–20 years and their antibody status at first testing |
|-----------------------------------------------|
| Hypothyroidism, 2nd test                      |</p>
<table>
<thead>
<tr>
<th>Hypothyroidism</th>
<th>“Normal”</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroid antibody, 1st test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive antibody</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Negative antibody</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>100</td>
</tr>
</tbody>
</table>

The children are defined at age 10–20 years as having hypothyroidism or “normal” (IR-TSH is included as “normal”, meaning not having hypothyroidism, and not needing therapy). Sensitivity 100%; specificity 93% (95% CI 82–99%). Positive predictive value 13%; negative predictive value 100%. Likelihood ratio for positive test result 14; likelihood ratio for negative test result very small (≤0.001).
IR-TSH can be used as a basis to select a subgroup for further testing at, say, five yearly intervals unless new symptoms emerge in the interim. They also suggest yearly screening as recommended by the American Academy of Pediatrics,24 is probably not justified in the first 20 years of life.

Authors’ affiliations
P A Gibson, Department of Paediatrics, Royal Lancaster Infirmary, Lancaster, UK
R W Newton, D A Price, G M Addison, Departments of Neurology, Endocrinology, and Clinical Biochemistry, Royal Manchester Children’s Hospital, Pendlebury, UK
K Selby, Department of Paediatric Neurology, Vancouver Children’s Hospital, Vancouver, Canada
K Leyland, Department of Community Child Health, South Bank Child Centre, Glasgow, UK

Competing interests: none declared

REFERENCES