ORIGINAL ARTICLE

Cost-effectiveness analysis of universal screening for thyroid disease in pregnant women in Spain

Sergio Donnay Candil, José Antonio Balsa Barro, Julia Álvarez Hernández, Carlos Crespo Palomo, Ferrán Pérez-Alcántara, Carlos Polanco Sánchez

Key words:
Cost effectiveness analysis; Hypothyroidism; Subclinical hypothyroidism; Pregnant women

Abstract
Objective: To assess the cost-effectiveness of universal screening for thyroid disease in pregnant women in Spain as compared to high risk screening and no screening.
Methodology: A decision-analytic model comparing the incremental cost per quality-adjusted life year (QALY) of universal screening versus high risk screening and versus no screening was used for the pregnancy and postpartum period. Probabilities from randomized controlled trials were considered for adverse obstetrical outcomes. A Markov model was used to assess the lifetime period after the first postpartum year and account for development of overt hypothyroidism. The main assumptions in the model and use of resources were assessed by local clinical experts. The analysis considered direct healthcare costs only.
Results: Universal screening gained 0.011 QALYs over high risk screening and 0.014 QALYS over no screening. Total direct costs per patient were €5786 for universal screening, €5791 for high risk screening, and €5781 for no screening. Universal screening was dominant compared to risk-based screening and a very cost-effective alternative as compared to no screening. Use of universal screening instead of high risk screening would result in €2,653,854 annual savings for the Spanish National Health System.
Conclusions: Universal screening for thyroid disease in pregnant women in the first trimester is dominant in Spain as compared to risk-based screening, and is cost-effective as compared to no screening (incremental cost-effectiveness ratio of €374 per QALY). Moreover, it allows diagnosing and treating cases of clinical and subclinical hypothyroidism that may not be detected when only high-risk women are screened.

© 2014 SEEN. Published by Elsevier España, S.L.U. All rights reserved.

Introduction

The prevalence of clinical hypothyroidism (CH) during pregnancy ranges from 0.3% to 0.5%, while the prevalence of subclinical hypothyroidism (SH) ranges from 2% to 3%. These subclinical thyroid function disorders are the most common functional abnormalities in pregnant women with thyroid diseases.

Hypothyroidism during pregnancy is usually asymptomatic. In the most severe cases, both SH and CH may cause some signs and symptoms such as inadequate weight gain, cold intolerance, asthenia, and dry skin. During pregnancy, SH and CH may cause adverse obstetric outcomes such as premature placental detachment and preterm delivery, as well as more frequent admission to neonatal intensive care units. On the other hand, pregnant women with positive thyroid peroxidase (TPO) antibodies have significantly higher thyroid-stimulating hormone (TSH) levels as compared to women with negative TPO antibodies and have been reported to have an increased fetal loss rate. The diagnosis of CH and SH during pregnancy requires understanding of the specific changes in thyroid function during each trimester. Since CH/SH was identified as a cause of maternal and fetal morbidity, there has been controversy about the convenience of the universal screening of thyroid function during pregnancy or selective screening based on risk factors (women with a family history of thyroid disease, type 1 diabetes mellitus, thyroid disorder, premature delivery, radiation therapy to the head or neck, amongst others). Selective screening for risk factors has been the preferred method because of its feasibility and because of the lack of studies showing the superiority of universal screening over risk-based screening. Recent studies have shown that screening women considered at high risk only means that 30–50% of women with CH or SH who could benefit from treatment are not identified.

In Spain, only women at high risk for CH are currently screened. There is thus no program for thyroid dysfunction screening in the population, but only recommendations from the Working Group on Iodine Deficiency Disorders and Thyroid Dysfunction of the Spanish Society of Endocrinology and Nutrition and the Spanish Society of Gynecology and Obstetrics.
A recent cost-effectiveness analysis in the United States has shown that universal screening for autoimmune thyroid disease of women in the first trimester of pregnancy is cost-effective both as compared to no screening during pregnancy and to selective screening of women at high risk. Our study was intended to assess the cost-effectiveness of universal screening for SH and CH during pregnancy as an alternative to risk-based or no screening in Spain.

**Subjects and methods**

A short-term (pregnancy time and one year after pregnancy) decision model and a long-term Markov model of annual cycles were developed to assess the clinical and financial benefits of CH and SH screening strategies during pregnancy (Fig. 1). The model assumed a hypothetical cohort of pregnant women aged 31.46 years (the mean age of pregnant women in Spain), 20% of whom belonged to some of the so-called high-risk groups.

The cost-effectiveness indicator used in the study was cost by quality-adjusted life year (QALY). Cost per miscarriage avoided and cost per preterm delivery avoided were also analyzed.

The following three screening strategies were examined:

1. Universal screening: this consists of the measurement of TPO antibodies and TSH in all pregnant women. If TSH is >5 mIU/L, T4 levels are measured to differentiate CH from SH before referral to an endocrinologist. Levothyroxine is administered to pregnant women with TSH levels >5 mIU/L or with TSH levels ranging from 2.5 to 5 mIU/L with positive TPO antibodies. Women undergo all measurements during the first trimester of pregnancy, but only the TSH and T4 measurements at each of the other trimesters.

2. Selective screening based on risk factors: this includes the same measurements as universal screening, but only in women classified as high risk. No additional measurement is performed in any other pregnant women. Levothyroxine is only administered to pregnant women with TSH levels >5 mIU/L.

3. No screening: measurements of TSH, T4, or TPO antibodies are not routinely performed.

The analysis was performed from the perspective of the healthcare system, discounting costs and effects at
3%, as recommended by the Spanish financial evaluation guidelines. Medical costs, probabilities and utilities related to treatment, and the potential lifetime sequelae for women were taken into account.

**Decision model**

The decision tree starts with the categorization of pregnant women based on TSH, TPO antibodies, and T4 levels. Based on the screening, the physician may categorize a pregnant woman during the first trimester as having CH, SH, or positive TPO antibodies, or will perform no laboratory tests until some of the characteristic symptoms of CH or SH appear, suggesting that screening takes place from the second trimester (Fig. 1). During the trimesters of pregnancy, the woman may continue with the course of her pregnancy and experience miscarriage during the first and second trimesters or preterm delivery in the third trimester. During the course of pregnancy, it has been planned that patients with CH and SH are treated with levotyroxine and that all other pregnant women may increase their TSH > 5 mIU/L during the second and third trimesters.

After term or preterm delivery, the chances of a woman experiencing post-partum thyroiditis (PPT) depending on whether or not she has positive TPO antibodies is assessed during the following year (Fig. 1). Screening has been taken into account in assessing the chances of diagnosing patients with symptomatic PPT.

After miscarriage or one year after delivery, the Markov model starts, all women with a non-CH status, except for those who evolve to permanent CH after experiencing PPT or those with CH detected during pregnancy (Fig. 1) being included. Over time, women may have CH (undiagnosed, diagnosed and complying with drug treatment, and diagnosed but not complying with drug treatment) or die from any cause. The probability of cardiovascular events and their associated mortality depending on patient type have been incorporated into the model.

**Probabilities**

Probabilities for the decision tree and transitions between Markov model states have mainly been obtained from Dosiou et al. An 11.16% probability of positive TPO antibodies
was associated in women (10% low risk and 15.8% high risk). TSH was differentiated based on whether women had positive or negative TPO antibodies, so that the probability of TSH < 2.5 mIU/L was 69.6% and 87.8% with positive and negative antibodies respectively; the probability of TSH ranging from 2.5 to 5 mIU/L was 16.9% versus 11%; and the probability of TSH > 5 mIU/L was 13.5% and 1.2% for those with positive and negative antibodies respectively. 

From the second trimester, women not treated with levothyroxine may progress to TSH levels >5 mIU/L, so that progression in the second trimester occurred in 14.3% of those with positive TPO antibodies and TSH ranging from 2.5 to 5 mIU/L; in 6.80% with negative TPO antibodies and TSH ranging from 2.5 to 5 mIU/L, in 5.70% with positive TPO antibodies and TSH < 2.5 mIU/L; and in 0.81% with negative TPO antibodies and TSH < 2.5 mIU/L. In the third trimester, progression occurred in 6.60% of women with positive TPO antibodies and TSH ranging from 2.5 to 5 mIU/L; in 2.60% with negative TPO antibodies and TSH ranging from 2.5 to 5 mIU/L; in 2.20% with positive TPO antibodies and TSH < 2.5 mIU/L; and in 0.31% with negative TPO antibodies and TSH < 2.5 mIU/L. The proportion of patients with increased TSH levels and low T4 levels has been assumed to be 23% for women with positive TPO antibodies and 8% for those with negative TPO antibodies. Thus, the model assumes that 100% of pregnant women with CH or SH receive treatment, except for those not screened or those from the subgroup of low-risk patients, so that only 28% of women with SH are diagnosed and treated. The model assumes full compliance with levothyroxine treatment in the pregnancy phase and 90% compliance with long-term treatment.

The probability of miscarriage or premature delivery comes from the Negro et al. study, where the treatment of thyroid dysfunction in 4562 women was analyzed based on whether or not universal screening had been performed (Table 1). In adjusting the data to the Spanish population, it has been taken into consideration that 85% of miscarriages occur in the first trimester.

The fact that, after pregnancy, the chances of PPT range from 40% to 60% in women with positive TPO antibodies, and 0.5% in women with negative TPO antibodies was taken into account. Such PPT will be symptomatic in 68% of cases and will be fully diagnosed in universal screening or in high-risk patients. By contrast, for women with no screening or from the low-risk subgroup without screening, it was assumed that only 25% of patients with symptomatic PPT will be diagnosed. 

Although it is known that approximately 10–20% of women who are euthyroid after an initial phase of hypothyroidism due to PPT will develop permanent hypothyroidism in the next 3–10 years, this effect was not included because it is indivisibly linked to the probability of transition to CH in the Markov model. The probabilities of transition between health states as a function of TSH and TPO antibodies come from the Dosiou et al. study. The probability of death from any cause comes from Spanish data, adjusted for age and sex. Based on the Rodondi et al. study, in which a higher risk of cardiovascular events as a function of TSH, sex, and age was identified, specific hazard ratios were estimated for patients with undiagnosed CH, diagnosed CH with no treatment compliance, and diagnosed CH with treatment compliance (Table 1). The application of these hazard ratios to the probability of experiencing cardiovascular events in Spain allowed us to estimate cardiovascular events in the age-adjusted cohort of women. Cardiovascular mortality was incorporated using the same procedure.

### Resource utilization and costs

During pregnancy, there is an association between the routine utilization of visits to the gynecologist and taking diagnostic tests. Diagnostic tests (TSH, T4, and TPO antibodies) during the three trimesters, as well as visits to endocrinologists and levothyroxine treatment (first trimester, 50 mcg; second and third trimester, 75 mcg) were taken into account as a function of patient type and screening.

For women with preterm delivery, the use of resources for newborns with weights less than 1500 g having specific needs (incubator, surgery, intensive care, etc.) was related. In this regard, resource utilization and the proportion of each type of case were based on diagnosis-related groups with codes 602 to 608 of the minimal basic data set of the Ministry of Health.

In the post-partum period, visits to the endocrinologist, diagnostic tests as a function of TPO antibodies and, if required, levothyroxine treatment at doses of 50 mcg/day from the first visit to the endocrinologist were identified. If PPT was identified, closer follow-up was performed (up to four times) and treatment was administered for approximately four months.

In addition, in the long-term follow-up period, women diagnosed with SH by universal screening were not treated after the first year after delivery, and only women diagnosed with HC incurred costs (two primary care visits, one visit to endocrinologist, 1.4 TSH measurements and 1.4 T4 measurements). Women with CH diagnosed, but with treatment noncompliance, used 80% of drug treatment (levothyroxine 100 mcg/day).

The costs of resource utilization were obtained from the Spanish cost database eSalud, and drug costs were based on the selling price of the company; specifically, the costs of levothyroxine treatment were obtained from the medicinal product database BotPLUS (Table 2).

The financial impact of incorporating the most cost-effective strategy into the national health system was also estimated based on the estimated number of pregnant women in Spain, multiplied by the difference in costs from the current risk-based screening approach. The number of pregnant women was estimated based on the fertility rate by age and 12-month period multiplied by the number of Spanish women.

### Utilities

To convert life years into QALY, utility values for each health state were assigned using the mean of the EQ-5D indicator. EQ-5D values represent numerically the value assigned by society to the current state of health. A value of 1 represents the best possible state of health, while 0 represents
A maximum utility (a value of 1) was assigned to euthyroid women and patients with CH, SH, or PPT, either asymptomatic or symptomatic but treated. It was estimated that utilities were 0.7085 for patients with untreated CH or who did not comply with treatment, 0.9 for those with symptomatic SH, and 0.81 for those with symptomatic, untreated PPT.11 Utilities of 0.9 and 0.91 were applied for miscarriage and premature delivery respectively.12 The utility associated with cardiovascular diseases was estimated at 0.379 based on a Spanish quality of life study in patients with heart failure.24

Sensitivity analysis

A battery of sensitivity analyses were performed modifying, mainly, the probability of positive TPO antibodies (10–30% high risk/5–15% low risk), the proportion of women at high risk (10–30%), the proportion of patients with PPT (40–60%), the associated utilities, the probability of diagnosing CH in women screened but with TSH levels < 5 (100%), the probability of diagnosing CH in women screened but with TSH levels > 5 (25%), and the age of pregnant women (18–45 years).

An additional probabilistic sensitivity analysis was performed using the second-order Monte Carlo simulation. This analysis makes it possible to assess the impact of uncertainty on the parameters and the robustness of the results. The results of the base case of the model were simulated on a hypothetical cohort to see the probability of the results remaining stable in the event of a multivariate and simultaneous change in the main parameters included. The results of the probabilistic sensitivity analysis were displayed using the cost-effectiveness plane, each simulation being represented as dots, on the abscissa axis the incremental QALYs, and on the ordinate axis the incremental costs of universal screening as compared to risk-based screening or no screening. The Spanish willingness-to-pay threshold normally used (30,000€/QALY gained) was included in the design of the cost-effective option.25 Variable variation was done by assigning a triangular distribution to probabilities and age, a log-normal distribution to hazard ratios, and a gamma distribution to utilities and costs. In addition, the probability that alternatives are cost-effective as a function of a variety of thresholds of willingness-to-pay for QALY increases is shown using the willingness-to-pay curve.

Results

The probability of spontaneous abortion was 4.12% for universal screening, 4.52% for risk-based screening, and 4.57% for no screening. Preterm delivery was also seen more commonly with the no screening (5.65%) and risk-based screening (5.63%) options. Universal screening therefore prevents 0.4–0.45% of miscarriages, and 0.22–0.24% of preterm deliveries.

Total direct costs per patient were €5786 for universal screening, €5791 for risk-based screening, and €5781 for no screening. In the group subject to risk-based screening, high-risk women showed the highest mean cost (€5923), while low-risk women had the lowest cost (€5758). Universal screening increased by 4.02 days (0.011 QALYs) the quality-adjusted life time as compared to risk-based screening, and is therefore a dominant alternative because it better shows effectiveness and lower cost. As compared to no screening, universal screening increased quality-adjusted life time by 4.99 days (0.014 QALY), with an incremental cost-effectiveness ratio of €374 per QALY. When risk-based screening was compared to no screening, quality-adjusted life time increased by 0.97 days (0.003 QALYs), with an incremental cost-effectiveness ratio of €4020 per QALY (Table 3).

Thus, based on the number of pregnancies in Spain (477,132) and the saving of €5.56 per patient with universal screening versus risk-based screening, the Spanish health system could save approximately €2,653,854.

Sensitivity analysis

The variables most influencing the result were the probability of positive TPO antibodies in low-risk and high-risk women, so that if such a probability substantially decreases (5% and 10% respectively), the incremental cost-effectiveness ratio of universal screening versus risk-based screening is a maximum of €3062 per QALY gained, and as compared to no screening is a maximum of €3735 per QALY gained, much lower than the efficiency threshold (€30,000/QALY gained). As regards the age of pregnancy, increased QALYs favoring universal screening were seen as age increased; by contrast, differential cost versus risk-based screening decreased from €5.92 to €4.90, but increased from €4.74 to €5.76 as compared to no screening. It should be noted that these changes in variables did not modify the conclusions in any case.

The probabilistic sensitivity analysis showed that universal screening is a more effective and less costly option than risk-based screening in 50% of simulations, and shows

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Unit costs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resource</strong></td>
<td><strong>Cost</strong></td>
</tr>
<tr>
<td>Visits</td>
<td></td>
</tr>
<tr>
<td>Gynecologist</td>
<td>€58</td>
</tr>
<tr>
<td>Endocrinologist</td>
<td>€64</td>
</tr>
<tr>
<td>Primary care physician</td>
<td>€43</td>
</tr>
<tr>
<td>Diagnostic tests</td>
<td></td>
</tr>
<tr>
<td>Positive thyroid antibodies</td>
<td>€16</td>
</tr>
<tr>
<td>Thyroid-stimulating hormone (TSH) levels</td>
<td>€25</td>
</tr>
<tr>
<td>T4</td>
<td>€14</td>
</tr>
<tr>
<td>Ultrasonography</td>
<td>€56</td>
</tr>
<tr>
<td>Complete blood count</td>
<td>€4</td>
</tr>
<tr>
<td>Events</td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td>€2226</td>
</tr>
<tr>
<td>Preterm delivery</td>
<td>€31,584</td>
</tr>
<tr>
<td>Miscarriage</td>
<td>€1470</td>
</tr>
<tr>
<td>Cardiovascular (hospitalization)</td>
<td>€5361</td>
</tr>
<tr>
<td>Cardiovascular (monitoring)</td>
<td>€300</td>
</tr>
</tbody>
</table>

* Mean costs in euros in 2012 taken from the eSalud database. 
Source: Gisbert and Broa.22
Table 3  Cost-effective analysis.

<table>
<thead>
<tr>
<th></th>
<th>Universal screening</th>
<th>Risk-based screening</th>
<th>No screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (€)</td>
<td>5786</td>
<td>5791</td>
<td>5781</td>
</tr>
<tr>
<td>Incremental cost (€)</td>
<td>5791</td>
<td>5791</td>
<td>5781</td>
</tr>
<tr>
<td>QALY</td>
<td>26.703</td>
<td>26.692</td>
<td>26.689</td>
</tr>
<tr>
<td>Differential QALYs</td>
<td>0.011</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Cost per QALY gained</td>
<td>Dominant</td>
<td>€374.28</td>
<td></td>
</tr>
<tr>
<td>Differential miscarriage/preterm delivery</td>
<td>9.54%</td>
<td>10.15%</td>
<td>10.23%</td>
</tr>
<tr>
<td>Cost per event prevented</td>
<td>Dominant</td>
<td>€737.19</td>
<td></td>
</tr>
</tbody>
</table>

QALY: quality-adjusted life years.

Discussion

This study has demonstrated, using the reported methods, that universal screening for hypothyroidism in pregnant women is cost-effective in Spain when compared to risk-based screening, and that no screening is not a valid alternative. This same effect was noted in two prior cost-effectiveness analyses conducted in the United States showing that universal screening of pregnant women is cost-effective as compared to no screening.26,27 Our results also agree with those found in another study conducted in the United States comparing universal and risk-based screening in pregnant women, which showed very similar costs and provided a higher number of QALYs gained.12 It should be noted that this is the first cost-effective study conducted in Spain assessing in a long-term decision model the suitability of universal screening based on its benefit-risk balance. Moreover, this study suggests that the use of universal screening instead of risk-based screening may represent savings of 2.6 million euros to the national health system.

One strength of this study was the use of data on obstetric adverse effects from randomized clinical trials and the inclusion of cardiovascular effects related to untreated CH. It should be noted that in a recent study which assessed the burden of SH for Spain, SH was seen to account for the loss of more than 30,000 disability-adjusted life years, which is equivalent to a cost of almost 67 million euros from the cardiovascular events it causes.28 Moreover, the assumptions made were very conservative, thus possibly underestimating the benefits of universal screening. It should be stressed that other greatly significant potential benefits of universal screening, such as an improvement in the intelligence quotient of the newborn, were excluded from the model.8

Our analysis has several limitations. It should first be noted that this is a complex theoretical model which attempts to represent in turn a simplified simulation of reality. Because of the complexity of reality, its theoretical simulation is not free from uncertainty. To minimize such uncertainty, various sensitivity analyses were performed.

Other limitations of our study include the fact that some probabilities were estimated when few data were available. An attempt was made to minimize the effects of the uncertainty introduced by taking into consideration the most conservative assumptions in the model and by analyzing it...
in depth in the sensitivity analysis, which allowed for more robust conclusions by supporting the results achieved, but for a much wider spectrum of women than the base case.

It should also be noted that the probabilities of adverse effects with levothyroxine were obtained from two clinical trials in south Italy, where iodine deficiency is mild. Any extrapolation of results to other geographical areas with different iodine nutrition levels or where early detection tests are performed at a more advanced gestational age should therefore be made with caution.

Conclusions

Universal screening for CH and SH in pregnant women is a more effective and less costly strategy in Spain as compared to risk-based screening. In addition, universal screening is highly cost-effective as compared to no screening. The results of our study therefore support the convenience of performing universal screening in the pregnant Spanish population.

Funding

This study was supported by Merck S.L.

Conflict of interest

The authors state their independence from the sponsoring and funding body in the analysis of the results and the preparation of the conclusions. Doctors S. Donnay, J. Álvarez, and J.A. Balsa received no honoraria for writing this manuscript. F. Pérez-Alcántara and C. Crespo work at an independent consulting company which has received funding from Merck S.L. C. Polanco is an employee of Merck S.L.

References


Cost-effectiveness analysis of screening for thyroid disease in pregnant women 329
