Stratifying the Risk of Developing Clinical Hypocalcemia after Thyroidectomy with Parathyroid Hormone

Alejandro Castro, MD1, Laura del Rio, MD1, and Javier Gavilan, MD1

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. To identify the risk of clinical hypocalcemia in the first hours after thyroidectomy.

Study Design. Prospective observational study.


Subjects and Methods. A total of 123 patients who underwent total or completion thyroidectomy between June 2010 and March 2012 were included. Pre- and postoperative intact parathyroid hormone (iPTH) levels were obtained. Patients remain hospitalized a minimum of 48 hours until blood calcium stabilized. Calcium and/or vitamin D supplements were prescribed only when signs or symptoms of hypocalcemia developed. Receiver operating characteristic curve analysis was employed to evaluate the postoperative iPTH level and the pre-to postoperative decrease in iPTH levels. Two cutoff values were determined to stratify the risk of developing clinical hypocalcemia into 3 groups.

Results. The areas under the receiver operating characteristic curve were 0.991 for the postoperative iPTH and 0.998 for the decrease in iPTH. An iPTH decrease of 80% had 100% sensitivity to detect patients who developed clinical hypocalcemia, while a postoperative iPTH \(<3\) pg/mL had 100% specificity. Thus, patients with an iPTH decrease \(<80\)% are at a very low risk of clinical hypocalcemia, and patients with a postoperative iPTH \(<3\) pg/mL are at a very high risk. Patients with an iPTH decrease \(>80\)% and a postoperative iPTH \(\geq3\) pg/mL are at intermediate risk. No significant correlation was found between the time when the sample was obtained and iPTH values.

Conclusion. This study establishes a very accurate test to stratify the risk of clinical postthyroidectomy hypocalcemia based on pre- and postoperative iPTH levels.

Keywords

thyroidectomy, hypocalcemia, parathyroid hormone

Received February 6, 2017; revised July 27, 2017; accepted August 17, 2017.
hypercalcemia and hypercalciuria. Moreover, this strategy does not completely suppress the risk of developing severe hypocalcemia at home.\(^{12,17}\)

An alternative strategy consists of selective supplementation to patients at risk of needing treatment. Different parameters have been employed to select these patients, including the slope of calcium levels during the first postoperative hours,\(^{18-22}\) although PTH levels are nowadays considered the most reliable predictor.\(^{19,23-26}\) Unfortunately, there is wide variability in the literature regarding the timing to obtain the PTH sample, the variable to employ (only postoperative PTH levels vs preoperative-postoperative difference), and the cutoff levels to discharge/start treatment.

In an effort to contribute to the standardization of a test to predict the development of postthyroidectomy hypocalcemia, we performed a study to evaluate different variables based on PTH levels.\(^{27}\) The current study was designed to validate our previous results and generate a reliable protocol to stratify the risk of needing treatment for clinical postthyroidectomy hypocalcemia. The main objective is to identify, in the first hours after surgery, patients at no risk of developing clinical hypocalcemia to proceed to an early discharge. The secondary objective is to identify patients at a very high risk of clinical hypocalcemia to start treatment before the onset of symptoms.

**Methods**

The study was designed as a prospective observational study. Every consecutive adult patient who underwent total or completion thyroidectomy at our department during the study period was considered a candidate. Exclusion criteria consist of (1) clinical conditions that could interfere with calcium homeostasis (previous diagnosis of hyper-/hypoparathyroidism, previous treatment with calcium and/or vitamin D supplements, creatinine clearance <40 mL/min, previous treatment with diuretics, hematologic malignant tumor with hypercalcemia), (2) loss of PTH samples and/or relevant data, and (3) refusal to participate.

Intact PTH (iPTH) was determined via the Immulite 2000 system (Siemens AG, Munich, Germany). The analytic sensitivity is 3 pg/mL, and values <3 pg/mL are reported as <3. A preoperative blood sample for PTH measurement was obtained at the ward anytime between admission and moving to the operating room. A postoperative sample was taken in the ward when the patient came back from the recovery room.

Symptoms of hypocalcemia and Chvostek and Trousseau signs were monitored during hospitalization every 8 to 12 hours or more frequently if necessary. Serum calcium levels were determined at 8 PM the day of surgery and 8 AM on the following days; additional serum calcium levels were obtained with suspicious symptoms.

Oral or intravenous calcium was started in the presence of clinical hypocalcemia (signs or symptoms), iPTH levels were blinded to the physician who prescribed the treatment. If symptoms of hypocalcemia were mild and signs were doubtful, the treatment was delayed until they evolved and/or a low serum calcium level was confirmed. Asymptomatic patients with low serum calcium levels detected on routine blood test were closely monitored but did not receive treatment until signs or symptoms of hypocalcemia developed.

Patients remained hospitalized a minimum of 48 hours and were discharged when serum calcium increased according to 2 consecutive levels drawn at least 8 hours apart. If a patient was started on supplements, he or she remained hospitalized to determine if calcium dosage should be increased and/or vitamin D added, until 2 consecutive levels showed an increment in serum calcium. In these cases, the same regimen was prescribed after discharge. If the patient did not receive treatment during hospitalization, no supplements were prescribed after discharge. All the patients came back to the clinic 2 or 3 weeks after surgery and were questioned about symptoms of hypocalcemia. No patient discharged without supplements reported symptoms.

The study protocol was approved by our institutional Ethics Committee for Clinical Research, and an informed consent was signed by every included patient.

**Statistical Analysis**

Sample size was calculated by fixing the lowest acceptable sensitivity to 95% (lowest range of confidence interval) and an alpha error of 0.05. The number of hypocalcemic patients included (true positive + false negative) was calculated to be at least 24.\(^{28,29}\) Considering the results of our previous study,\(^{27}\) the incidence of hypocalcemia was set to 20% and the proportion of excluded patients to 30%. Thus, at least 156 patients should undergo total or completion thyroidectomy during the study period.

The receiver operating characteristic (ROC) curve analysis was employed to determine the performance of 2 classifiers: the postoperative iPTH level (iPTH\(_{\text{post}}\)) and the preoperative-postoperative iPTH decrease (iPTH\(_{\text{decr}}\)) calculated by the following equation: \([\text{iPTH}_{\text{post}} - \text{iPTH}_{\text{pre}}] / \text{iPTH}_{\text{pre}} \times 100.\) To determine the gold standard to construct the ROC curve, a patient was considered hypocalcemic only if he or she needed treatment during hospitalization due to signs or symptoms related to hypocalcemia according to the study protocol described before. The area under the curve (AUC) and its significance and confidence interval were calculated with SPSS 20 software (IBM, Chicago, Illinois). To compare AUCs and to calculate partial AUCs, the pROC package for the R Project for Statistical Computing software was employed.\(^{30}\)

To maximize sensitivity and specificity, 2 cutoff values were calculated for each classifier: one with 100% sensitivity and the higher specificity and the other with 100% specificity and the higher sensitivity. The values were rounded to the nearest whole number, to simplify the definitive diagnostic test protocol.

Student’s \(t\) test was employed to compare continuous variables between treated and untreated patient groups.

**Results**

A total of 162 patients underwent total or completion thyroidectomy at our institution between June 2010 and March
2012. Thirty-nine patients were excluded: 28 for clinical conditions, 10 from sample loss, and 1 for refusal to participate. Thus, 123 patients were included in the study. Demographic and clinical data are summarized in Table 1.

A total of 29 patients (23.6%) received treatment due to signs or symptoms of hypocalcemia. Preoperative iPTH was 47 ± 21 pg/mL (mean ± SD) for untreated patients and 52 ± 34 pg/mL for treated patients; no significant difference was found (P = .287). Postoperative iPTH was significantly lower for treated patients: 3 ± 3 versus 31 ± 18 pg/mL, P = .001.

The distribution of serum calcium levels in the treated and untreated patient groups are shown in Figure 1. Statistically significant differences between both groups were found regarding total and ionized calcium levels at the 3 sample times (P < .001 for all comparisons).

Hospital stay was 3 ± 2 days for untreated patients (mode = 2 days) and 5 ± 2 days for treated patients (mode = 4 days). Box plots are shown in Figure 2.

Regarding clinical hypocalcemia, the AUC for the classifier iPTH_{post} was 0.991, with a 95% CI of 0.979 to 1.000, whereas the AUC for iPTH_{decr} was 0.998 (95% CI, 0.995-1.000). No statistically significant difference was found between AUCs—not between total AUCs (P = .23), partial AUCs bounded by 100%-95% sensitivity (0.043 vs 0.049, P = .09), or partial AUCs bounded by 100%-95% specificity (0.045 vs 0.048, P = .14).

Table 2 shows the AUCs calculated for both classifiers according to different criteria to define hypocalcemia: clinical, any calcium <7 mg/dL, any calcium <7.5 mg/dL, and any calcium <8 mg/dL. Both iPTH variables predicted hypocalcemia determined by a clinical criterion significantly more accurately than when determining hypocalcemia by any biochemical criterion.

Two cutoff values were considered for each classifier to detect clinical hypocalcemia, one with 100% sensitivity and the other with 100% specificity: 3 and 24 pg/mL for iPTH_{post} and 98% and 80% for iPTH_{decr}. Sensitivity, specificity, and accuracy for each cutoff value, as well as their

Table 1. Demographic and Clinical Data of Patients Included in the Study.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>%a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, N</td>
<td>123</td>
</tr>
<tr>
<td>Age, y, mean ± SD</td>
<td>50 ± 15</td>
</tr>
<tr>
<td>Male:female</td>
<td>18:82</td>
</tr>
<tr>
<td>Thyroidectomy, total:completion</td>
<td>91.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multinodular hyperplasia</td>
<td>30.9</td>
</tr>
<tr>
<td>Diffuse hyperplasia</td>
<td>10.6</td>
</tr>
<tr>
<td>Thyroiditidis</td>
<td>5.7</td>
</tr>
<tr>
<td>Follicular adenoma</td>
<td>3.3</td>
</tr>
<tr>
<td>Papillary carcinoma</td>
<td>45.5</td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>0.8</td>
</tr>
<tr>
<td>Medullary carcinoma</td>
<td>0.8</td>
</tr>
<tr>
<td>C cell hyperplasia</td>
<td>2.4</td>
</tr>
</tbody>
</table>

* Values are presented as percentages unless noted otherwise.
usually, being ill or not. Accuracy is measured by sensitivity and specificity. To evaluate the accuracy of a diagnostic test, a gold standard is needed that defines whether a patient is ill or not. Using a 2-category classifier and comparing its results with the gold standard, we can calculate true and false positives and negatives and determine sensitivity and specificity. If our classifier is a continuous variable, a cutoff value should be settled to transform it into a 2-category one.

Thus, sensitivity and specificity both depend on the cutoff value settled to determine if the test is positive or negative. As multiple cutoff values can usually be chosen, a statistical method has been designed to compare the accuracy of a classifier irrespective of them. The ROC curves analysis and the AUC estimate the global accuracy of a test, considering at the same time all possible cutoff values (and subsequent sensitivity-specificity pairs). If we are interested in a test with a high sensitivity or a high specificity, partial AUCs restricted to the desired range of sensitivity or specificity can be also calculated. Statistical nonparametric tests have also been designed to compare total and/or partial AUCs from different classifiers. They compare the global accuracy of the classifiers, independent of the cutoff values that will be settled later.

Thus, the ROC curve analysis is the ideal method to determine the performance of a continuous variable to classify patients into 2 categories. The ability of PTH to predict postthyroidectomy hypocalcemia has been noted in the literature, although many investigators employed statistical methods different to the ROC curve analysis. In the present study, we evaluated the ability of 2 variables (iPTH post and iPTH decr) to classify patients according to their probability of developing clinical postthyroidectomy hypocalcemia. The AUCs of both variables were extremely high, close to 1, demonstrating the strong correlation that exists between PTH levels and the onset of clinical hypocalcemia after total or completion thyroidectomy.

Different publications have made use of the ROC curve analysis in a similar setting to ours. The AUCs obtained by many of these investigators were lower than ours, in the range of 0.65 to 0.88. We believe that this difference can be explained by the definition of postthyroidectomy hypocalcemia adopted as gold standard in each study. This idea is supported by our own AUCs when based on biochemical criteria. Defining postthyroidectomy hypocalcemia is crucial to obtain a representative ROC curve. The problem is that there is no uniform definition of postthyroidectomy hypocalcemia, mostly due to the lack of a uniform serum calcium threshold for the development of symptoms. Moreover, asymptomatic hypocalcemia has been described after nonthyroid surgery, probably due to hemodilution. This factor may confound calcium measurements in the early postoperative period. We believe that a disease should be defined by its signs and symptoms, not only by a biochemical parameter. For these reasons, we decided to define postthyroidectomy hypocalcemia as the development of related signs or symptoms. Our data show that iPTH levels predict clinical hypocalcemia more accurately than biochemical hypocalcemia.

Discussion

The accuracy of a diagnostic test corresponds to its ability to correctly classify patients into different categories—

#### Table 2. Areas under the Curve Calculated for Both Classifiers per Different Criteria to Define Hypocalcemia.

<table>
<thead>
<tr>
<th>Hypocalcemia Criteria</th>
<th>Classifier</th>
<th>Clinical</th>
<th>Ca &lt;7 mg/dL</th>
<th>Ca &lt;7.5 mg/dL</th>
<th>Ca &lt;8 mg/dL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iPTH post</td>
<td>0.991</td>
<td>0.858&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.896&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.783&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>iPTH decr</td>
<td>0.998</td>
<td>0.852&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.905&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.785&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Abbreviations: iPTH decr, intact parathyroid hormone decrease; iPTH post, postoperative intact parathyroid hormone.

<sup>a</sup>Each biochemical criterion was compared with the corresponding clinical criterion.

<sup>b</sup>P = .01.

<sup>c</sup>P < .01.

<sup>d</sup>P = .02.

95% CIs, are shown in Table 3. iPTH decr of 80% was selected as the most accurate 100% sensitivity cutoff value and iPTH post of 3 pg/mL as the most accurate 100% specificity cutoff value. By combining both cutoff values, a test was established to stratify the risk of needing treatment for hypocalcemia into 3 groups. Table 4 shows the distribution of the population according to this test.

Time since preoperative blood sample to beginning of surgery was 7 ± 6.75 hours. Time since the end of surgery to extraction of postoperative sample was 10.5 ± 8.5 hours. No significant correlation was found between the time when the sample was obtained and the values of the classifiers: r = −0.08 (P = .37) for iPTH decr and hours since preoperative sample to surgery, r = 0.06 (P = .5) for iPTH post and hours since surgery to postoperative sample, and r = 0.09 (P = .36) for iPTH decr and hours since surgery to postoperative sample.

![Figure 2. Box plots showing length of hospital stay for treated and untreated patient groups.](image)
Table 3. Sensitivity, Specificity, and Accuracy (and Their 95% CIs) for Each Considered Cutoff Value.\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>iPTH(_{\text{post}})</th>
<th>iPTH(_{\text{decr}})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;3 / ≥3 pg/mL</td>
<td>≥8% / &lt;98%</td>
</tr>
<tr>
<td></td>
<td>≥24 / ≥24 pg/mL</td>
<td>&gt;80% / ≤80%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89.7 (73.6-96.4)</td>
<td>24.1 (12.2-42.1)</td>
</tr>
<tr>
<td>Specificity</td>
<td>100 (96.1-100)</td>
<td>100 (96.1-100)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>97.6 (93.1-99.2)</td>
<td>82.1 (74.4-87.9)</td>
</tr>
</tbody>
</table>

Abbreviations: iPTH\(_{\text{decr}}\), intact parathyroid hormone decrease; iPTH\(_{\text{post}}\), postoperative intact parathyroid hormone.

\(^a\)Values are presented as percentages.

Table 4. Distribution of the Population according to the Risk Groups Defined by the Test and the Gold Standard.\(^a\)

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>Clinical Hypocalcemia, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high (iPTH(_{\text{post}}) &lt;3 pg/mL)</td>
<td>Yes 26 (21.1) No 0 (0) Total 26 (21.1)</td>
</tr>
<tr>
<td>Intermediate (iPTH(<em>{\text{post}}) ≥3 pg/mL + iPTH(</em>{\text{decr}}) &gt;80%)</td>
<td>3 (2.5) 4 (3.2) 7 (5.7)</td>
</tr>
<tr>
<td>Very low (iPTH(_{\text{decr}}) ≤80%)</td>
<td>0 (0) 90 (73.2) 90 (73.2)</td>
</tr>
<tr>
<td>Total</td>
<td>29 (23.6) 94 (76.4) 123 (100)</td>
</tr>
</tbody>
</table>

Abbreviations: iPTH\(_{\text{decr}}\), intact parathyroid hormone decrease; iPTH\(_{\text{post}}\), postoperative intact parathyroid hormone.

\(^a\)If they developed clinical hypocalcemia during hospitalization.

Other authors have utilized a clinical criterion to construct the ROC curve and found AUCs >0.9.\(^{43,47-50}\) However, they also treat asymptomatic patients with variable degrees of biochemical hypocalcemia, which may bias the development of symptoms. In our study, to ensure the natural history of disease, patients remained hospitalized without calcium or vitamin D supplements until signs or symptoms occurred, or until serum calcium levels stabilized. Under these conditions, we found that iPTH levels strongly predict the development of clinical hypocalcemia.

It is obvious that patients who develop symptoms related to hypocalcemia have lower serum calcium levels than those who do not. Our patients with clinical hypocalcemia had significantly lower total and ionized calcium levels than patients without clinical hypocalcemia in the evening of surgery and on the following 2 mornings. However, it is known that no uniform inter- or intraindividual serum calcium level is correlated with clinical manifestations. There should be other factors associated with biochemical hypocalcemia that determine the onset of symptoms. We advocate that the direct effect of surgery is parathyroid gland damage, and it produces symptomatology mediated by an indirect biochemical hypocalcemia. Thus, the term postthyroidectomy hypocalcemia should be replaced by postthyroidectomy hypoparathyroidism.

The timing to extract the blood sample has been another matter of debate.\(^{19,23,26,51,52}\) In a previous study,\(^{27}\) we found no correlation between the iPTH levels and the number of hours from surgery to when the sample was extracted. The results of the present study confirm this lack of correlation, and the outstanding AUCs calculated suggest that the timing of the sample is probably irrelevant. We believe that the blood sample should be obtained under ideal conditions, as the iPTH is a very labile molecule and deteriorates rapidly. For this reason, we decide to make more flexible the timing of the sample, allowing our nurses to decide in their daily practice when the conditions were optimal to manipulate the sample correctly.

After the global test accuracy is determined and compared with the AUC, cutoff values should be settled to calculate sensitivity and specificity. Increasing the sensitivity usually implies a decrease in specificity, and vice versa. By determining 2 cutoff values instead of only 1, we could maximize both sensitivity and specificity. Two cutoff values classify the population into 3 groups. Patients in 2 of these 3 groups will or will not develop clinical hypocalcemia with substantial certainty. The test is not accurate for patients in the third, intermediate, group. As shown in Table 3, patients who undergo total or completion thyroidectomy can be classified into 3 risk groups based on our 2 cutoff values. Those with an iPTH\(_{\text{decr}}\) ≤80% can be safely discharged home without treatment, as the risk of developing clinical hypocalcemia is essentially null. Calcium supplements can be prescribed to those with an iPTH\(_{\text{post}}\) <3 pg/mL, as virtually all of them will otherwise develop signs or symptoms. The small proportion of patients in the intermediate risk group (iPTH\(_{\text{decr}}\) >80% and iPTH\(_{\text{post}}\) ≥3 pg/mL) would have an approximate 50% chance of developing clinical hypocalcemia; they could be discharged home with treatment or maintained hospitalized and monitored, depending on the philosophy of the institution.

Conclusions

iPTH levels strongly predict the development of signs or symptoms of postthyroidectomy hypocalcemia. This study establishes a very accurate test to stratify the risk of clinical postthyroidectomy hypocalcemia based on pre- and postoperative iPTH levels.

Acknowledgments

We thank the members of the Department of Otorhinolaryngology at Hospital Universitario La Paz (Madrid, Spain) for their contribution in managing patients during their hospital stay, especially Ricardo Bernaldez and Antonio del Palacio, that operated on some
of the patients included in the study. We also thank the members of endocrinology department for their contribution in managing patients with hypocalcemia, especially Cristina Alvarez-Escola, Beatriz Lecumberri, and Garcilaso Riesco-Eizaguirre.

Author Contributions
Alejandro Castro, conception and design of the study, collected data, analysis and interpretation of data, write article and approval of final version; Laura del Rio, collected data, contribute to analysis of data, revision and approval of final version; Javier Gavilan, contribute to design of the study and interpretation of data, revision and approval of final version.

Disclosures
Competing interests: None.
Sponsorships: None.
Funding source: FIBHULP grant, no role in the study.

References