Primary Hyperparathyroidism and Intraoperative Parathyroid Hormone Monitoring: Application of a Modified Interpretation in Patients With “Parathyroid Hormone Spikes”

Julian Heidtmann, MD, Daniela Dunkler, PhD,1 Lindsay Hargitai, MD, Christian Scheuba, MD, Bruno Niederle, MD, Philipp Riss, MD,* and Andreas Selberherr, MD, PhD

Section of Endocrine Surgery, Division of Visceral Surgery, Department of General Surgery, Medical University of Vienna, Vienna, Austria

ARTICLE INFO

Article history:
Received 20 December 2021
Received in revised form 2 August 2022
Accepted 20 August 2022
Available online xxx

Keywords:
Intraoperative parathyroid hormone monitoring
Parathyroid surgery
Primary hyperparathyroidism
PTH spikes
Visualization value

ABSTRACT

Introduction: Intraoperative parathyroid hormone (PTH) spikes occur in up to 30% of patients during surgery for primary hyperparathyroidism. This can lead to a prolonged PTH decline and cause difficulties in using current interpretation criteria of intraoperative PTH monitoring. The aim of this study was to evaluate an alternative interpretation model in patients with PTH spikes during exploration.

Methods: 1035 consecutive patients underwent surgery for primary hyperparathyroidism in a single center. A subgroup of patients with intraoperative PTH spikes of >50 pg/mL were selected (n = 277; 27.0%). The prediction of cure applying the Miami and Vienna criteria was compared with a decay of ≥50% 10 min after excision of the enlarged parathyroid gland using the “visualization value” (VV; = PTH level immediately after visualization of the gland) as basal value. Sensitivity, specificity, accuracy, positive predictive value, and negative predictive value were calculated.

Results: Using the VV, sensitivity was 99.2% (Vienna 71.0%; Miami 97.7%), specificity was 18.2 (Vienna 63.6%; Miami 36.4%), and accuracy was 92.8 (Vienna 70.4%; Miami 92.8%). Of 255 single-gland disease patients, 72 were identified correctly as cured (P < 0.001), yet 10 of 22 patients with multiple-gland disease were missed compared with the Vienna Criterion (P = 0.002). The comparison with the Miami Criterion showed that six more patients were correctly identified as cured (P = 0.219), whereas four patients with multiple-gland disease were missed (P = 0.125).

Conclusions: Using the VV as a baseline in patients with intraoperative PTH spikes may prove to be an alternative and therefore can be recommended. However, if the VV is higher than the preexcision value, it should not be applied.

© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Introduction

Surgery with open “targeted” (focused, limited) exploration by means of open minimally invasive parathyroidectomy (OMIP; i.e., a short skin incision) is the standard treatment in biochemically documented primary hyperparathyroidism (PHPT) and localized single-gland disease (SGD) and has replaced bilateral neck exploration in most patients.5 Concordant results of high-resolution ultrasound (US) and 99mTc-sestamibi scintigraphy (MIBI) or F-Choline-PET-CT are invaluable for the success of OMIP.2 However, the finding of a localized SGD does not definitely rule out multiple-gland disease (MGD).3 Therefore, the application of intraoperative parathyroid hormone (iOPTH) monitoring and a standardized interpretation of PTH decay are important requisites to predict cure with high certainty after removing the presumed hyperfunctioning tissue.6

The suitable criteria of PTH decay ensuring cure are still under discussion. Various interpretation models of PTH decrease have been described in the literature.5-10 The “ideal” criterion in predicting cure should confirm SGD with high sensitivity, thus avoiding more extended exploration and should identify patients with MGD who in turn require more extended exploration.

Although a “strict” criterion may inevitably lead to a higher rate of more extended and prolonged operations, more “generous” criteria may miss patients with MGD,11 leading to a higher rate of persistence and recurrence.

Interpretation of PTH decrease applying various criteria

To interpret intraoperative PTH decrease, two well-defined criteria were used—the “Vienna Criterion” (Criterion 2) and the “Miami Criterion” (Criterion 3),13 and the results were compared with the “Modified Vienna Criterion” (Criterion 1) using the VV value:8

Criterion 1—Modified Vienna Criterion: The VV was applied as basal level: If a PTH increase of >50 pg/mL from BL to excision was documented, the VV was used as an “alternative BL value.” A decline of >50% from the VV 10 min after excision was considered to reflect a normalization of elevated parathyroid metabolism.

Criterion 2—Vienna Criterion: Prediction of complete resection was defined as a subsequent decrease of >50% from BL within 10 min after excision of the suspected hyperfunctioning parathyroid tumor.9

Criterion 3—Miami Criterion: Surgical success was defined as a >50% decrease from the highest (preincision or preexcision) value within 10 min after resecting the presumed hyperfunctioning parathyroid tissue.15

Statistical analysis

The metric parameters were described by median and quartiles, and categorical variables were described by absolute and relative frequencies. To obtain accuracy, specificity, sensitivity, and positive predictive value (PPV) and negative predictive value (NPV) were defined: True positives as correct predictors of permanent normocalcemia; true negatives as correct predictors of incomplete excision (resection of an additional gland necessary or operative failure |persistent
hypercalcemia); false positives as incorrect predictors of complete resection (resulting in persistent hypercalcemia); and false negatives as incorrect predictors of incomplete excision (postoperative normocalcemia).

For each of the three criteria, accuracy, sensitivity, and specificity were calculated as well as exact 95% confidence intervals.

McNemar’s test (with exact P values) was applied to statistically compare sensitivity and specificity between the three criteria. Differences were considered statistically significant if a two-sided P value was < 0.05. The SAS system V 9.4 (SAS Institute Inc, Cary, NC) was used for statistical calculations.

The study was approved by the Ethics Committee of the Medical University of Vienna. All patients gave their written informed consent for all diagnostic and therapeutic procedures.

Results

The study included 1035 consecutive patients (916 females [88.8%] and 119 males [11.2%]) with biochemically proven sporadic PHPT, who had undergone initial surgery at the Medical University of Vienna between 1999 and 2015.

The mean preoperative serum calcium level (albumin adjusted) was 2.78 mmol/L (2.20-2.55 mmol/L), mean intact PTH was 178 pg/mL (15-65 pg/mL), whereas the mean postoperative albumin-adjusted serum calcium level was 2.39 mmol/L (2.20-2.55 mmol/L), and the mean postoperative intact PTH was 52 pg/mL (15-65 pg/mL; Table 1). The median age was 62 y (range 15-93).

In 992 patients (96.6%), a single parathyroid adenoma (SGD) was the cause for sporadic PHPT, whereas MGD was present in 35 patients (3.4%). SGD was defined by normal serum calcium levels >6 mo postsurgery, and no contradictory histologic findings.

Surgical procedure

OMIP was performed in 349 patients (34.0%), whereas 115 (11.2%) received unilateral and 339 (33.0%) initial bilateral neck exploration. A total of 224 (21.8%) operations were intraoperatively converted from OMIP to uni- or bi-lateral explorations.

Because of various thyroid nodules, 227 (22.1%) patients received concomitant thyroidectomy, and 199 (19.4%) patients received hemithyroidectomy.

Subcategory—patients with PTH spikes

Patients with intraoperative PTH spikes of >50 pg/mL versus BL (see below) were analyzed (n = 277; 27.0%).

The distribution regarding gender (female:male 241:36 [87.0%:13.0%]), age (median 63; minimum 25, maximum 93; 25th percentile: 53, 75th percentile: 71), surgical method, concomitant thyroid surgery, and histological findings (SGD 268 [96.8%], MGD 9 [3.2%]) was comparable with the overall cohort.

Comparison of the criteria— all patients

Comparing the three criteria when applied to the whole collective of all operated patients, the sensitivity was 93.3% (Criterion 1), 89.1% (Criterion 2), and 97.9% (Criterion 3), respectively. In terms of specificity, Criteria 1 and 3 were at 55%, whereas Criterion 2 showed 70%. Accuracy was 91.1% for Criterion 1, 88% for Criterion 2, and 95.5% for Criterion 3 (Tables 2 and 3). The PPV was nearly equal in all criteria (Criterion 1: 97.1, Criterion 2: 97.9, Criterion 3: 97.3), whereas the NPV varied widely between 33.7 for Criterion 1, 28.4 for Criterion 2 and 62.3 for Criterion 3 (Table 2).

The decisions made by every criterion are shown in Table 3.

Comparison of the criteria—patients with PTH spikes >50 pg/mL

With respect to the selected patients with intraoperative PTH spikes of 50 pg/mL or more, Criterion 1 reached the highest sensitivity of 99.2%, followed by Criterion 3 (97.7%) and Criterion 2 (71%) (Tables 4 and 5). Regarding specificity, Criterion 2 showed 63.6%, followed by Criterion 3 with 36.4% and Criterion 1 with 18.2%. This results in an accuracy of 92.8% in Criteria 1 and 3 and of 70.4% in Criterion 2. The PPV ranged from 95.8 in Criterion 2 to 94.7 in Criterion 3 and 93.4 in Criterion 1. The NPV was highest in Criterion 1 (66.7), followed by 57.1 in Criterion 3, and the lowest in Criterion 2 (15.9; Table 4).

The decisions made by every criterion for patients with PTH spikes >50 pg/mL are shown in Table 5.

Discussion

High cure rates are achieved in parathyroid surgery when applying careful preoperative workup, as well as critical interpretation of IOPTH monitoring and the surgeons’ experience.5,9

### Table 1 – Overview of patients’ characteristics.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Actual value</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium preoperative</td>
<td>2.78 mmol/L</td>
<td>2.20-2.55 mmol/L</td>
</tr>
<tr>
<td>Calcium postoperative</td>
<td>2.39 mmol/L</td>
<td></td>
</tr>
<tr>
<td>PTH preoperative</td>
<td>178 pg/mL</td>
<td>15-65 pg/mL</td>
</tr>
<tr>
<td>PTH postoperative</td>
<td>52 pg/mL</td>
<td></td>
</tr>
</tbody>
</table>
In 1991, Irvin et al. introduced intraoperative PTH monitoring, and several interpretation models have been published since then.\(^{13-16}\) Commonly, OMIP is the surgical strategy of choice in localized primary hyperparathyroidism. This approach was made possible by using detailed preoperative localization studies, such as US and functional imaging (MIBI) in combination with intraoperative PTH monitoring. Various criteria have been published to facilitate the sensitive and specific interpretation of intraoperative PTH monitoring values regarding the cure of patients. In most of these criteria, a ≥50% decline from a defined BL or (highest) preexcision value within 10 min after excision of the enlarged gland has been considered to indicate long-term “cure.”\(^{9,15,16}\) Due to the exploration of just the single (expected) hyperfunctioning gland detected by imaging, PTH decline is crucial for the surgical concept to minimize the risk of missing additional hyperfunctioning parathyroid glands.\(^{16}\) However, the criterion guiding the interpretation of intraoperative PTH decline is a matter of ongoing discussion.

The need for IOPTH monitoring in patients with localized SGD and without signs of MGD is presently also under debate, although higher rates of persistent disease after parathyroid surgery have been reported in series without IOPTH monitoring.\(^{5,19}\) Although US and MIBI lack to detect MGD with high certainty, IOPTH monitoring is indicated. Although great efforts to improve the results of imaging studies, that is, F-Choline-PET/CT, have recently been made, the experience with these modalities is currently still low.\(^{20,21}\)

IOPTH spikes are documented in up to 30% of the patients during exploration.\(^{7}\) PTH spikes may be attributed to intraoperative manipulation of the enlarged gland\(^{16}\) or the manual stimulation (respectively, the rupture) of a parathyroid cyst.\(^{22}\) The emergence of PTH spikes is unrelated to the surgical procedure applied (unilateral versus bilateral exploration), the preoperative PTH level or serum creatinine.\(^{18}\) Patients with SGD and large adenoma are more likely to develop higher PTH spikes, whereas PTH spikes detected in patients with MGD are lower.\(^{5}\) Attempts to evaluate shorter interpretation protocols for the evaluation of intraoperative PTH decline with 5-min postexcision criteria have shown to be effective in patients without PTH spikes. However, a shortened criterion may not be used safely in patients experiencing intraoperative PTH increase.\(^{23}\)

Previous studies have indicated that patients with an IOPTH rise face a higher risk of unnecessarily extended surgical explorations and that interpretation using Criterion 2 (Vienna) may show a high rate of interpretational errors in patients with PTH spikes.\(^{9,11,14,24}\) Previous data therefore recommended using Criterion 3 (Miami) when PTH spikes were detected.\(^{13}\) The present data used to compare Criteria 2 and 3 are in favor of this finding.

Because of its stringent limits, Criterion 2 (Vienna) misses less MGD, which results in higher specificity but may also lead to a higher rate of intraoperative conversions to unilateral or bilateral neck explorations because of its lower sensitivity compared with Criterion 3 (Miami). By using the VV, our new approach is an attempt to improve surgical decision-making in this defined subgroup of patients and achieve higher sensitivity compared with strict BL criteria. When applying our modified interpretation model, higher sensitivity is achieved at the cost of lower specificity, resulting in the same level of accuracy when comparing it to the Miami criterion. Such low specificity may be ascribed to the fact that in all patients with PTH spikes, the VVs were higher than the preexcision values. Therefore, MGD may be missed. Furthermore, the PPV was documented to be just slightly in favor of the third criterion (Miami versus VV: 94.7 versus 93.4). In terms of the NPV, Criterion 1 (VV) was clearly ahead of other interpretation models (Miami versus VV: 57.1 versus 66.7).

Compared with Criterion 2, the use of Criterion 1 changed the surgical strategy in 127 of 975 cured patients (\(P = 0.001\)) and in 11 of 60 noncured patients (\(P = 0.012\)) in the entire collective. With the use of the VV (Criterion 1) instead of Criterion 2, 127 more patients would have been correctly detected as cured, whereas 11 MGD would have been missed. Compared with Criterion 3, the surgical strategy would have been altered in 65 of 975 cured patients (\(P < 0.001\)) and in eight of 60 patients who were not cured (\(P = 1.000\)). While 65 additional patients would have been identified correctly as cured, eight patients with persisting disease would have been missed.

### Table 2 – Sensitivity, specificity, and accuracy analyzing all patients (n = 1035).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>Accuracy, % (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Vienna Criterion (Criterion 1)</td>
<td>93.3 (91.6, 94.8)</td>
<td>55.0 (41.6, 67.9)</td>
<td>91.1 (89.2, 92.8)</td>
<td>97.1 (95.8, 98.1)</td>
<td>33.7 (24.4, 43.9)</td>
</tr>
<tr>
<td>Vienna Criterion (Criterion 2)</td>
<td>89.1 (87.0, 91.0)</td>
<td>70.0 (56.8, 81.2)</td>
<td>88.0 (85.9, 89.9)</td>
<td>97.9 (96.8, 98.8)</td>
<td>28.4 (21.3, 36.4)</td>
</tr>
<tr>
<td>Miami Criterion (Criterion 3)</td>
<td>97.9 (96.6, 98.7)</td>
<td>55.0 (41.6, 67.9)</td>
<td>95.5 (94.0, 96.6)</td>
<td>97.3 (96.0, 98.2)</td>
<td>62.3 (47.9, 75.2)</td>
</tr>
</tbody>
</table>

CI = confidence interval.

### Table 3 – Comparison of the decisions of the three criteria in all patients (n = 1035).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Not cured (n = 60)</th>
<th>Cured (n = 975)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Vienna Criterion (Criterion 1)</td>
<td>33</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Correct</td>
<td>27</td>
</tr>
<tr>
<td>Vienna Criterion (Criterion 2)</td>
<td>42</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Correct</td>
<td>18</td>
</tr>
<tr>
<td>Miami Criterion (Criterion 3)</td>
<td>33</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Correct</td>
<td>27</td>
</tr>
</tbody>
</table>
Comparing the different criteria in patients with PTH spikes $>50$ pg/mL, we identified a significant number of different decisions in favor of Criterion 1 compared with Criterion 2, but no significant difference between Criterion 1 and Criterion 3. In detail, the comparison between Criteria 1 and 2 showed different surgical decisions in 72 of 255 patients who were identified correctly as cured by the VV ($P \leq 0.001$). On the other hand, 10 of 22 noncured patients would have been overseen by the VV but identified correctly by Criterion 2 ($P = 0.002$). When comparing Criteria 1 and 3, the VV would have changed the surgical strategy in six of 255 patients who had been cured ($P = 0.219$) and four of 22 patients who had not ($P = 0.125$). Therefore, six correct decisions in cured patients match against four wrong decisions in patients who had not been cured.

Overall, the use of the VV as a modified BL seems to be reliable and should be favored in patients with IOPTH spikes. However, the Modified Vienna Criterion is not recommended in general because of its minority in patients without IOPTH spikes. A “perfect” criterion for the interpretation of obligatory IOPTH monitoring, which may universally be used in all possible situations, is still to be found. The surgeon’s personal expertise remains the key feature to success under difficult circumstances. Applying the Modified Vienna Criterion as a modified BL in patients with IOPTH spikes would thus appear to be a valuable alternative and can be recommended in patients with spikes, especially for clinicians using the Vienna Criterion as the standard interpretation model. This modified interpretation model seems superior to the Vienna Criterion and more efficient than the Miami criterion in this subgroup of patients. When the VV measured is higher than the pre-excision value, the VV should be used with caution and cannot generally be recommended. Regardless of which criterion is used to interpret IOPTH decline, the VV is a valuable addition that provides more detailed information about PTH decay and may therefore prove helpful in surgical decision-making in questionable cases.

**Limitations**

As all data were gathered retrospectively, the impact on surgical decision-making can only be assumed.

The analysis is based on single-center experience, and only a handful of surgeons were thus included in the decisional process, all with the same educational background. As, in reality, all patients were interpreted with the Vienna Criterion as the standard regime, the consequences for patients interpreted with the Miami Criterion or the Modified Vienna Criterion are hypothetical.

Another limitation concerns IOPTH monitoring and arterial line placement. Although its effectiveness and advantages have been extensively published, these issues remain an issue discussed among endocrine surgeons. In addition, the arterial line needed for multiple PTH level testing can cause complications such as thrombosis, temporary vascular occlusion, hematoma and catheter-related infections.

**Conclusions**

In conclusion, the use of the VV as a modified BL is reliable and should be given preference in patients with IOPTH spikes. Using the Modified Vienna Criterion as a modified BL in patients with IOPTH spikes would appear to be a valuable alternative and can be recommended in patients with spikes, especially for clinicians using the Vienna Criterion as the standard interpretation model.

**Author Contributions**

J.H. contributed to study conception and design, acquisition of data, and drafting of article. D.D. contributed to analysis and interpretation of data. L.H. contributed to acquisition of data and drafting of article. C.S. and B.N. contributed to acquisition of data and critical revision of article. P.R. contributed to study conception and design, acquisition of data, and critical revision of article.
Disclosure

None of the authors declare a conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES