An Alternative Analysis of Intraoperative Parathyroid Hormone Data May Improve the Ability to Detect Multiglandular Disease

Paul G. Gauger, MD; Michelle H. Mullan, MD; Norman W. Thompson, MD; Gerard M. Doherty, MD; Keith A. Matz, BS; Barry G. England, PhD

Hypothesis: A nomogram based on regression analysis of intraoperative parathyroid hormone level decay discriminates single gland disease from multiglandular (MG) disease more accurately than the currently used 50% rule.

Design: Retrospective case series.

Setting: Academic health center.

Patients: Two hundred thirty-five patients (222 patients with single gland disease and 13 patients with MG disease) who underwent parathyroidectomy.

Interventions: Intraoperative parathyroid hormone level analysis at baseline, time 1 (about 5 minutes), and time 2 (about 10 minutes) after excision of the first gland.

Main Outcome Measures: The mean slope was calculated at time 1 and time 2 and analyzed using one-way analysis of variance and the Fisher least significance difference post hoc tests using data normalized to baseline intraoperative parathyroid hormone levels to compare patients with single gland disease with patients with MG disease. A regression-based nomogram was created to analyze individual kinetic decay data.

Results: The mean (SEM) single gland disease slope was significantly steeper than the MG disease slope at both time 1 (−0.91 [0.02] vs −0.66 [0.05]; P < .01) and time 2 (−0.77 [0.01] vs −0.56 [0.05]; P < .01). When the standard threshold rule of a 50% decrease from baseline was used, only 23% of the patients with MG disease were correctly predicted by intraoperative parathyroid hormone values (77% false-positive result rate) at time 1. However, the nomogram correctly predicted 54% of the patients with MG disease at time 1 (46% false-positive result rate). At time 2, the standard threshold 50%-rule method correctly predicted 38% of the patients with MG disease (62% false-positive result rate), while the nomogram still correctly classified 54% of the patients with MG disease (46% false-positive result rate).

Conclusions: A regression-based nomogram incrementally improves prediction of MG disease compared with the standard 50%-rule method and accounts for variability in the exact timing of samples. Slope analysis suggests that the earliest time point best isolates the kinetics of the excised gland. The nomogram will need to be validated prospectively.

Arch Surg. 2004;139:164-169

Intraoperative parathyroid hormone (IOPTH) monitoring as an adjunct to parathyroidectomy has an important role in minimally invasive parathyroidectomy, yet it serves a less certain role in an open parathyroid gland exploratory procedure. Optimal interpretation of IOPTH level kinetic data in the discrimination of single gland (SG) disease and multiglandular (MG) disease has not been entirely defined and is under scrutiny. The inability of IOPTH monitoring to detect remaining enlarged parathyroid glands after removal of a single hypersecreting gland in up to 55% of these cases is certainly not an actual deficiency of the measurement technique itself.1 Two plausible alternative explanations exist. The first possibility is that these data may merely reflect an aspect of MG primary hyperparathyroidism that requires a shift in the accepted pathophysiologic paradigm. This is to say that not all enlarged parathyroid glands are hypersecretory (ie, size is not immutably linked to hyperfunction). The second possibility is that our current method of data analysis may not be optimal in terms of sensitivity and specificity. Many different variations in data interpretation criteria exist, but one has become most common.2 Although standard analysis of IOPTH data has been based on extensive operative experience, the arithmetic considerations are admittedly somewhat arbitrary. Analysis includes normalization of all values to the individual baseline value (usually the highest of ei-
ther postincision or preexcision values). A fall to below 50% of the baseline value by 10 minutes after adenoma excision is considered to reflect biochemical cure and this criterion (also known as the 50% rule) is widely used.\textsuperscript{3,4} This 50% rule is easy to use, but it is not universally accepted. Libutti et al\textsuperscript{7} have suggested that a kinetic analysis based on first-order decay characteristics may be more reflective of the true disappearance of circulating parathyroid hormone (PTH) levels after excision of a hypersecreting parathyroid gland and that such an approach can account for individual variations in PTH half-life levels and the variability in timing of the postexcision blood samples. This study attempted to define the influence of the background or residual PTH levels that may contribute in a variable manner to the PTH level disappearance rates after excision of a single abnormal parathyroid gland.

For this study, we chose not to ask whether IOPTH monitoring is absolutely reliable for detecting MG disease since there is good evidence that it is not. Instead, we questioned whether the specific method of data analysis could be improved to enhance the intended purpose of IOPTH monitoring. Our aim was to define a simple, yet sensitive, method of analysis to discriminate SG from MG disease using early decay kinetics, while minimizing the effect of background PTH concentrations. An additional requirement was the “portability” of the method for real-time use in the operating room during parathyroidectomy. The hypothesis was that the initial slope of the kinetic decay curve could be used to discriminate SG from MG disease more accurately than the currently accepted 50%-rule threshold criterion. An additional hypothesis was that a nomogram could be constructed that allowed the plotting of individual PTH values at a given time point to predict the presence of SG or MG disease.

A total of 363 parathyroidectomies were performed during a 2-year period at a single institution. A subset of 233 patients was retrospectively reviewed after being selected for the following criteria: (1) the patients had a clearly documented biochemical diagnosis of primary hyperparathyroidism and were undergoing initial parathyroidectomy (either standard or targeted); (2) IOPTH monitoring was performed with at least 2 postexcision values available; (3) the patients had standard biochemical evidence of cure (IOPTH value of <50% of baseline value by 10 minutes after the initial gland excision or by the conclusion of the procedure if multiple glands were excised); and (4) the patients were clinically cured (normocalcemia documented during follow-up at 2 weeks, 3 months, and up to 36 months postoperatively). Of the 233 patients analyzed, all were undergoing their first operation with the findings of SG disease in 222 patients and MG disease in 13 patients. In the 222 patients with SG disease, 40 parathyroidectomies were performed in a minimally invasive (targeted) fashion, while the remainder were performed as a standard 4-gland exploratory procedure.

Intraoperative parathyroid hormone monitoring was performed using venous blood with the baseline level drawn after anesthetic induction (termed “time 0” or T\textsubscript{0}). During the study period, it was not routine to draw additional preexcision baseline venous blood samples. Levels were also measured at a minimum of 2 time points after excision of the first enlarged parathyroid gland. These times were termed “time 1” (T\textsubscript{1}) (mode 5; range, 5-9 minutes) and “time 2” (T\textsubscript{2}) (mode 10; range, 10-15 minutes). Analysis was performed using the Immulite Turbo system (DPC Inc, Los Angeles, Calif). Intraoperatively, the PTH values were analyzed to determine biochemical cure by using previously published criteria (fall to <50% of the baseline value by 10 minutes after excision).\textsuperscript{3,4}

An alternative method of analyzing kinetic data was developed as follows: the PTH concentration disappearance rate was analyzed based on the first 2 values subsequent to excision of the first parathyroid gland. The IOPTH values were normalized as percentages of the baseline value. The slope of the kinetic decay rate was calculated for each patient from the regression equation for the elapsed time of T\textsubscript{0} to T\textsubscript{1} and T\textsubscript{1} to T\textsubscript{2}. This dual analysis was performed for each of the 2 treatment groups (those with SG and those with MG diseases). Data from all 4 groups were subjected to 1-way analysis of variance using the Fisher least significant difference post hoc test comparisons. Results are expressed as mean (SEM). For each group (at both T\textsubscript{1} and T\textsubscript{2}), the geometric mean of the slopes was used to calculate mean half-life (t\textsubscript{1/2}) as t\textsubscript{1/2}=−0.5/mean slope. As slope is essentially an expression of a change in value (PTH level) per change(Δ) in defining unit (in this case, time), the residual PTH level (percentage of baseline value) was plotted as a function of Δ time. As the exact timing of T\textsubscript{1} and T\textsubscript{2} varied, Δ time was converted to a logarithmic scale to allow a linear expression of the regression equation. One minute was added to all T\textsubscript{0}, T\textsubscript{1}, and T\textsubscript{2} time points to permit transforming data to the logarithm and for plotting data on a logarithmic scale (it is not mathematically possible to calculate the log of 0). Regression lines were determined for the mean of each treatment group (those with SG disease and those with MG disease). These thresholds were used to categorize individual data from the MG disease group to determine a biochemical cure. Specifically, if a value was found to the right of the MG disease mean, the patient was classified as having MG disease (true-negative [TN] result). Similarly, individual points to the left of the MG regression line were considered to have false-positive (FP) results (see the following paragraph). To be certain that sensitivity for MG disease was not gained at the cost of decreasing specificity for SG disease, a similar analysis was done for patients with SG disease to determine the false-negative (FN) and true-positive (TP) result rates.

In interpreting the following data, it is important to understand some standard definitions regarding IOPTH level test performance. If one considers a TP test result to be one in which the IOPTH values fall to curative levels after excision of a single adenoma and, indeed, no additional abnormal parathyroid glands remain, then an FP test result is one in which the IOPTH values fall to curative levels despite the fact that abnormal parathyroid glands remain in situ. Similarly, a TN test result is one in which IOPTH values fail to fall to levels indicating cure, thus, correctly indicating the presence of additional abnormal parathyroid glands. An FN test result is one in which the IOPTH values fail to fall to curative levels, yet there is no additional abnormal parathyroid gland(s) remaining. Thus, if the IOPTH level were perfect in detecting the presence of MG disease, the TN result rate would be 100% and the FP result rate would be 0%.

Initial analysis of the SG disease group suggests that not only the normalized PTH value but also the rate of PTH level decay begin to change immediately after the excision of the hypersecreting parathyroid gland. The mean slope is −0.91 (0.02) at T\textsubscript{1} that corresponds to a half-life of 2.54 minutes. The mean slope is −0.77 (0.01) at T\textsubscript{2} that yields a half-life of 3.55 minutes (Table 1). The T\textsubscript{1} and
Table 1. Mean Slope of Parathyroid Hormone Level Decay and Calculated Half-life Values

<table>
<thead>
<tr>
<th>Clinical Group</th>
<th>Time Point</th>
<th>Mean (SEM) Slope</th>
<th>Half-life, min</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG disease</td>
<td>T1</td>
<td>-0.91 (0.02)</td>
<td>2.54</td>
</tr>
<tr>
<td>SG disease</td>
<td>T2</td>
<td>-0.77 (0.01)</td>
<td>3.55</td>
</tr>
<tr>
<td>MG disease</td>
<td>T1</td>
<td>-0.66 (0.05)</td>
<td>4.72</td>
</tr>
<tr>
<td>MG disease</td>
<td>T2</td>
<td>-0.56 (0.05)</td>
<td>6.81</td>
</tr>
</tbody>
</table>

Abbreviations: MG, multiglandular; SG, single gland; T1, time 1 (mode 5; range, 5-9 minutes after the first parathyroid gland is excised); T2, time 2 (mode 10; range, 10-15 minutes after the first parathyroid gland is excised).

Figure 1. Box-and-whisker plot of slope values (mean) in Table 1 expressed for patients with single gland (SG) disease and patients with multiglandular (MG) disease. The 2 plots on the left are data from time 1 (T1) (mode 5; range, 5-9 minutes) after excision of the first adenoma, while the data on the right are from time 2 (T2) (mode 10; range, 10-15 minutes). Statistically significant relationships (all P < .01) are as follows: SG disease T1 vs MG disease T1, -0.91 vs -0.77, respectively; and SG disease T2 vs MG disease T2, -0.66 vs -0.56, respectively. 

T1 slopes are significantly different for the patients with SG disease (P < .01) (Figure 1). The mean slope of the SG disease group (-0.91 [0.02]) is significantly steeper (representing faster PTH level disappearance) than the mean slope of the MG disease group (-0.66 [0.05]) at T1 (P < .01). As one would expect in the presence of residual hypersecretory parathyroid tissue, the calculated half-life is also longer in the MG disease group (4.72 minutes) vs the SG disease group (2.54 minutes) at T1. This same relationship holds true at T2 in which the SG disease slope of -0.77 (0.02) is significantly different than the mean MG disease slope of -0.56 (0.05) (P < .01). Calculated half-lives are 3.55 minutes for patients with SG disease compared with 6.81 minutes for patients with MG disease (Table 1).

Further analysis was performed on both T1 and T2 values. With an understanding of these differences, 2 nomograms (one each for T1 and T2 data) were created based on the mean disappearance rates of PTH level in the 2 clinical treatment groups. Time was converted to a logarithmic scale to allow linear expression of the regression relationship. The x-axis had to be expressed as time + 1 (in minutes) as a mathematical artifact. This is because to be clinically useful and relevant, the y-intercept of the regression lines should be at 100% of baseline PTH value and T0. It is not mathematically possible to take the log of 0 minutes and, thus, the nomogram must reflect this. Therefore, each point on the graph actually occurred 1 minute earlier than expressed on the x-axis. A line indicating the presently used threshold value (50% of the baseline value at 10 minutes) was also superimposed. With these 3 key lines, individual patient data were plotted to determine sensitivity for MG disease for both the standard 50% rule and IOPTH monitoring.

Neither method perfectly predicted the presence of a second abnormal parathyroid gland after excision of the first parathyroid gland in the entire group of 13 patients with MG disease. However, the newly described method allowed an increase in correct identification of patients with MG disease. To be classified correctly, a plot for a patient with MG disease needed to reside at a point to the right of the defining regression line. At T1, the standard arithmetic threshold method correctly predicted only 3 of 13 patients with MG disease—giving a TN result rate of 23% and a high FP result rate of 77%. However, the nomogram method detected an additional 4 patients with MG disease—giving a TN result rate of 54% and decreasing the FP result rate to 46% (Figure 2). At T2, the standard arithmetic threshold method correctly predicted only 5 of 13 patients with MG disease—giving a TN result rate of 38% and a high FP result rate of 62%. The nomogram method correctly classified an additional 2 patients with MG disease—giving a TN result rate of 54% and decreasing the FP result rate to 46% (unchanged from performance at T1) (Figure 3).
To be certain that an increase in TN results did not occur at the expense of an unacceptable increase in FN results, the same nomograms constructed from data at both data points were used to analyze the individual data plots for the patients with SG disease (Figure 4). At T1, the FN result rate was 5% using the standard 50%-rule method based on the misclassification of 12 of 222 patients with SG disease (TP result rate, 95%). Using the nomogram method, the FN result rate increased to 8% as it misclassified 18 of 222 patients with SG disease (TP result rate, 92%) (Table 2).

The same nomogram used in Figure 3 was used to analyze the individual data plots for the patients with SG disease at T2 (Figure 5). At T2, the FN result rate was 4% using the standard 50%-rule method based on the misclassification of 8 of 222 patients with SG disease (TP result rate, 96%). Using the nomogram method, the FN result rate increased to 5% as it misclassified 11 of 222 patients with SG disease (TP result rate, 95%) (Table 2).

**COMMENT**

Intraoperative parathyroid hormone monitoring is becoming more common in the practice of parathyroid surgery. This is appropriate and important as minimally invasive or targeted parathyroid gland exploratory procedures are used. However, there is appropriate skepticism about the true potential of IOPTH monitoring to enhance the success rate of standard parathyroid gland surgery.6,7 Another concern is the cost-effectiveness of the technique.8 Multiple research groups, including ours, have noted that the ability of IOPTH monitoring to detect remaining enlarged parathyroid glands is imperfect.1,9-11 As most primary hyperparathyroidism is caused by SG disease, it may be said that IOPTH monitoring performs best in the very situation in which it is least valuable (ie, SG disease).

Despite initial zealous enthusiasm, many endocrine surgeons are beginning to accept that IOPTH monitoring does not act as a foolproof stopgap to detect all MG disease.1,2 There are 2 likely explanations and perhaps both are simultaneously operative. First is that the old paradigm of MG primary hyperparathyroidism is changing. It is clear that an enlarged parathyroid gland is not necessarily a hypersecretory gland at the time of parathyroidectomy. In fact, if the IOPTH level is used to define the existence of MG disease instead of gross morphologic criteria, the incidence is only 5%, which is less than the incidence reported in almost all series as defined by gland size alone.12 It has been suggested that each abnormal parathyroid gland in MG disease may have its own set point and that excision of the dominant gland may be met with a profound fall in IOPTH levels. However, if the additional enlarged gland(s) is not hypersecretory at the time of operation, its presence may not be obvious. If the calcium level does not fall below the set point of a remaining abnormal parathyroid gland, that gland will not secrete excess levels of PTH...
Table 2. Comparison of Standard 50%-Rule Threshold Method and Nomogram Method

<table>
<thead>
<tr>
<th>Time Point</th>
<th>Method</th>
<th>Result Rate, % (No. of Patients/Total No. of Patients)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50%-rule threshold</td>
<td>TN: 23 (3/13) FP: 77 (10/13) FN: 5 (12/222) TP: 95 (204/222)</td>
</tr>
<tr>
<td>T1</td>
<td>Nomogram</td>
<td>TN: 54 (7/13) FP: 46 (6/13) FN: 8 (18/222) TP: 92 (204/222)</td>
</tr>
<tr>
<td>T2</td>
<td>Nomogram</td>
<td>TN: 54 (7/13) FP: 46 (6/13) FN: 5 (11/222) TP: 95 (211/222)</td>
</tr>
</tbody>
</table>

Abbreviations: FN, false negative; FP, false positive; TN, true negative; TP, true positive; T1, time 1 (mode 5; range, 5-9 minutes after the first parathyroid gland is excised); T2, time 2 (mode 10; range, 10-15 minutes after the first parathyroid gland is excised).

*There were 13 patients with single glandular disease and 222 patients with multiglandular disease.

Figure 5. The same nomogram in Figure 3 was used to classify patients with single gland (SG) disease at time 2 (T2) (mode 10; range, 10-15 minutes). Time is expressed as x + 1 as a mathematical necessity. The dashed line indicates the current threshold method of 50% fall from baseline 10 minutes after excision. Note the individual plots for patients with SG at T2. Using the multiglandular (MG) disease regression line, the nomogram incorrectly classified 11 (5%) of 222 patients with SG disease as being patients with MG disease, while the standard 50%-rule threshold method incorrectly classified 8 (4%) of 222 patients with SG disease as being patients with MG disease. All circles indicate the data points that account for the difference in the classification by either method. The plus signs show the points plotted.

and, thus, will not be detected by IOPTH monitoring. This is not to say, however, that the gland will not someday begin to overproduce levels of PTH itself, and it is this scenario on which the concerns of decreased long-term durable cure rates are founded. This explanation is supported by the earlier work of Brown et al who showed in dispersed parathyroid gland cells from patients with primary hyperparathyroidism that adenomatous tissue had an elevated and more variable set point than healthy tissue or most hyperplastic tissue (set point defined as the calcium concentration half-maximally inhibiting the levels of PTH release).

Another possibility for this shortcoming is that the accepted method of analyzing kinetic data is suboptimal for the physiologic features of MG disease. The half-life of PTH is variable among different patients and the standard arithmetic analysis may be too generic to encompass all patients. Libutti et al suggested a kinetic analysis based on first-order decay of the PTH level. This article is especially important, as it was the first to address the contribution of background or residual PTH secretion by other healthy remaining parathyroid glands. This was done by a sophisticated iterative curve-fit function to estimate the unknown postexcision baseline function. However, the method requires the use of calculus and is slightly cumbersome to implement in real time during parathyroidectomy. Therefore, it has not been easy to supplant the conceptually appealing and user-friendly 50%-rule threshold method.

The fact that statistically significant differences exist between the mean slopes at T1 and T2 in the patients with SG disease has 2 likely explanations—both of which may be contributory. The most obvious is that first-order decay kinetics theoretically demand that the rate of disappearance decreases as a function of time while the curve asymptotically approaches the x-axis. However, as Libutti et al has noted, one must consider the fact that the new background contribution of the other existing parathyroid glands will further influence this curve. This is possibly already evident at T1 in our series. When one considers the clinically observed wide variability in apparent parathyroid gland functional suppression (as caused by the excised adenoma), the variation among normal decay curves should increase as a function of increasing time. In clinically relevant terms, this would suggest that analysis most proximate to adenoma excision might have the most reliable discriminatory ability. The fact that the calculated mean half-life of each study group began to increase between T1 and T2 may relate to previous observations of PTH half-life variability.

While this method serves to improve detection of MG disease by increasing the TN rate, there is a slight increase in the FN rate. In clinical terms, this would translate to an increased ability to detect MG disease at the cost of some patients who would have their targeted parathyroidectomy aborted to pursue a conventional 4-gland exploratory procedure. Although individual surgeon and patient values and expectations may influence this equation, it is our belief that this is a worthwhile tradeoff at this point when the conventional parathyroidectomy is still considered the gold standard. This principle will aid in preserving the excellent cure rate that parathyroidectomy enjoys.
One explanation for the FN rate in this study period, regardless of method of interpretation used, is the possibility of undetected baseline value elevation from tumor manipulation. For example, a postanesthetic baseline value may be misleading if tumor manipulation during dissection or retraction causes the physical release of PTH. If this occurs, the actual value from which decay begins after tumor excision may be increased substantially. If this occurs unobserved, the T1 and T2 values will be artifactually elevated as well. This factor was not accounted for during most of the study period but has since been described by another group. We routinely measure immediate preexcision baseline IOPTH levels and use that value in the decay rate calculations if it has risen above the level of the postanesthetic baseline value. We expect this to decrease the FN result rate as we continue to use the nomogram. One could argue that the same factor could have influenced the data for patients with MG disease. Aside from being statistically less likely to occur in a smaller patient group, that hypothetical influence would have served to displace individual patient data points rightward, thus, potentially increasing equally the TN result rate for both the standard and nomogram methods.

While this study finds increased clinical discriminatory ability in the assay by changing the standard analysis of the derived data, it remains to be seen how this will assist evaluation of the individual patient during parathyroidectomy. We have taken individual patient data and related it to the mean data of large groups of patients in clinically defined disease subsets. To provide a reliable and reproducible process, this method will need to be used and validated by other groups using their own patients. As this method appears to hold some promise, we will next validate it prospectively during parathyroidectomy using these nomograms (similar to that in Figures 2 and 3) at the time points in the T1 and T2 ranges to predict MG disease based on a single individual time-indexed IOPTH value. The nomogram based on T1 data will be used for the first postexcision value. If classification is unclear at that point, the nomogram based on T2 data will be used to categorize the second postexcision value. If the point lies to the right of the MG regression line, one would predict the presence of another abnormal parathyroid gland with a high degree of certainty. If the point lies to the left of the SG regression line, one would be confident in the classification as SG disease. If the point is to the right of the SG regression line but to the left of the MG regression line, even this method will not reliably predict the presence of another abnormal parathyroid gland—and one cannot overemphasize the importance of sound clinical judgment in this situation.

If this nomogram were to be computer-based, one could continually recalculate the regression lines as the patients accrue. This would allow some flexibility and adaptability for individual practice variations in patient characteristics—essentially allowing the nomogram to “learn” as it is used. Of course, these patients will need long-term follow-up to ensure that the prediction of cure based on the nomogram is borne out. This is, of course, true of all patients undergoing targeted or minimally invasive parathyroidectomy procedures, which rely on IOPTH level as an intraoperative measure of cure regardless of the specific method of data interpretation used.

### CONCLUSIONS

A nomogram based on time-indexed IOPTH data better predicts MG disease compared with standard arithmetic (50%-rule) analysis. It serves to increase the TN result rate and decrease the FP result rate allowing better detection of MG disease with an acceptable decrease in the TP result rate classification of SG disease. It also considers variability in the exact timing of the postexcisional samples. Early slope analysis suggests that the earliest time point best isolates the kinetics of the excised parathyroid gland. With the data established by this large series, further prospective use and validation of this method will be possible to demonstrate its true ability to predict the presence of MG disease during parathyroidectomy.

Accepted for publication July 12, 2003.

Corresponding author and reprints: Paul G. Gauger, MD, Division of Endocrine Surgery, Department of Surgery, University of Michigan, 1500 E Medical Center Dr, Taubman Center, Room TC2920H, Ann Arbor, MI 48109-0331 (e-mail: pgauger@umich.edu).

### REFERENCES


(Reprinted) Arch Surg/Vol. 139, Feb 2004

www.archsurg.com

©2004 American Medical Association. All rights reserved.