Objective: To develop statistical prediction tools to select patients for short-stay thyroidectomy based on dynamic quantification of individual risk for postoperative hypocalcemia.

Design: Clinical and biochemical factors that could influence postoperative calcium levels were analyzed. A multivariable logistic regression model was used to study the predictive ability of each variable for hypocalcemia. A step-down model reduction selection method was used to rank the predictors according to their predictive accuracy.

Setting: Memorial Sloan Kettering Cancer Center.

Patients: A test population of 393 patients who met our inclusion criteria and who underwent total thyroidectomy at Memorial Sloan Kettering Cancer Center in the year 2008 made up the modeling data set, 116 of whom developed biochemical hypocalcemia postoperatively (29.5%). The nomograms were validated on an independent data set consisting of 296 selected patients who underwent total thyroidectomy during the year 2005, using the same selection criteria for inclusion as those for the modeling data set.

Main Outcome Measures: The 8 predictors with the highest predictive accuracy were selected to generate a nomogram, which was validated both internally and externally using an independent data set. A second nomogram was developed for assessing the probability of a patient stay of 24 hours or shorter, based on preoperative and intraoperative factors.

Results: The 8 variables of highest predictive value were age, sex, medications, history of cancer, preoperative serum calcium level, creatinine concentration, central neck dissection, and alkaline phosphatase levels. A nomogram was created based on the final parsimonious model. The nomogram had excellent accuracy (concordance index of 74.6%) and scored high on internal validation tests. The concordance index of the second nomogram for predicting the likelihood of discharge from the hospital within 24 hours was 70%.

Conclusion: We have produced a set of nomograms that can dynamically quantify the risk of postthyroidectomy hypocalcemia and prolonged hospital stay based on preoperative clinical and biochemical variables and intraoperative surgical variables.


HYPOCALCEMIA FOLLOWING total thyroidectomy can be a cause of significant distress and anxiety for caregivers and patients alike. In addition, postoperative hypocalcemia and its management is generally the most significant contributor to length of hospital stay after thyroid surgery.1-3 With the increasing trend toward shorter hospitalization in the United States, the ability to triage patients with thyroidectomy and counsel them about their predicted length of hospital stay based on preoperative and intraoperative parameters has significant appeal for caregivers, patients, and hospital administrator alike.

The risk of postoperative hypocalcemia is obviously higher when more extensive thyroid surgery is carried out (eg, total thyroidectomy with central compartment neck dissection). These patients are generally not selected for the 24-hour hospital stay setting. However, in most other cases for which short-stay thyroidectomy is indicated, the risk of postoperative hypocalcemia is difficult to quantify because its pathogenesis is multifactorial. The present report is limited to the development of nomograms that could accurately quantify the risk of postopera-
The variables used for nomogram generation were selected based on a panel discussion (including clinicians and nurses) and a review of the literature. We identified clinical and biochemical preoperative and postoperative variables that are currently accepted as factors that affect the level of postoperative serum calcium levels (Table 1). A comprehensive list of the drugs that affect calcium level was taken from the United States Pharmacopeia and National Formulary (www.usp.org/USPNF/) and compared against the patient medication list recorded in the electronic medical record (EMR).

**METHODS**

The validation data set consisted of 268 patients of 326 who underwent total thyroidectomy during the year 2005, using the same selection criteria for inclusion as for those in the modeling data set. Of the 268 patients, 198 were female (73.9%), and 100 male (36.1%). Median age was 47.5 years (age range, 16-88 years). Of the 393 patients, 293 were female (75%), and 100 male (25%). Median age was 49 years (age range, 3-89 years). Central compartment lymph node dissection was carried out in 13%. A preoperative diagnosis of malignant disease by fine-needle aspiration biopsy was present in 88.8% patients. Of the 393 patients, 116 were recorded to have hypocalcemia (calcium level, <8 mg/dL) at first postoperative assessment (29.5%). (To convert calcium to millimoles per liter, multiply by 0.25.)

**VALIDATION DATA SET**

From a cohort of 544 patients who had thyroid surgery at Memorial Sloan-Kettering Cancer Center during the year 2008, 363 patients were selected as the modeling data set for the creation of all the nomograms. This selection was based on those patients who had total thyroidectomy and in whom biochemical parameters had been recorded both preoperatively and postoperatively, that could be used in the generation of our nomograms, regardless of the type of surgery performed. Patient demographics, blood parameters, and clinical and surgical data were reviewed and collected from the hospital EMR. Of the 393 patients, 293 were female (75%), and 100 male (25%). Median age was 49 years (age range, 3-89 years). Central compartment lymph node dissection was carried out in 13%. A preoperative diagnosis of malignant disease by fine-needle aspiration biopsy was present in 88.8% patients. Of the 393 patients, 116 were recorded to have hypocalcemia (calcium level, <8 mg/dL) at first postoperative assessment (29.5%). (To convert calcium to millimoles per liter, multiply by 0.25.)

**MODELING DATA SET**

Of the 393 patients, 293 were female (75%), and 100 male (25%). Median age was 49 years (age range, 3-89 years). Central compartment lymph node dissection was carried out in 13%. A preoperative diagnosis of malignant disease by fine-needle aspiration biopsy was present in 88.8% patients. Of the 393 patients, 116 were recorded to have hypocalcemia (calcium level, <8 mg/dL) at first postoperative assessment (29.5%). (To convert calcium to millimoles per liter, multiply by 0.25.)
The aim of nomogram 1 was to quantify the risk of postoperative hypocalcemia in an individual patient undergoing total thyroidectomy based on their unique preoperative variables that were readily available to the surgeon (Table 1). We hypothesized that this nomogram would be able to triage patients scheduled for total thyroidectomy into a 24-hour hospital stay rather than a prolonged hospital stay based on the anticipated risk of developing postoperative hypocalcemia.

End Point of Interest
Postoperative hypocalcemia was defined as a serum calcium level lower than 8 mg/dL (reference range at our institution, 8.5-10.5 mg/dL). The end point of interest for the first nomogram was hypocalcemia at the first postoperative assessment of serum calcium level (usually at 6 hours following completion of surgery). This time point of 6 hours was selected because of its practical utility as the first biochemical parameter that becomes available in the postoperative setting, which clinicians can use to guide their decision regarding length of hospital stay for a particular patient. Based on a review of the literature it is also the most widely used parameter to predict and treat hypocalcemia after thyroid surgery. We also performed statistical analysis to examine the correlation between a patient’s first postoperative calcium value and (1) the last postoperative calcium measurement during the hospital stay; (2) whether they were discharged with calcium supplementation; and (3) length of stay in the hospital.

Statistical Methods
Multivariable logistic regression was conducted to quantify the predictive value of each of the predictor variables in predicting the risk of hypocalcemia. To allow nonlinear relationships between numeric or ordinal variables and the outcome, restricted cubic spline functions with 3 knots were applied to these variables. Restricted cubic spline forces the piecewise polynomial functions to be linear in the tails and avoids poor performance in tails of regular cubic spline functions. All the predictor variables were ranked according to their predictive ability by the step-down model reduction method described previously. The step-down method was performed by first building a multivariable logistic model with all the 14 preoperative variables, which subsequently generated a predicted risk score (ie, the linear predictor) for each patient. After that, the predicted risk score served as the outcome for a multivariable linear model that used the same data and variables as used in the original logistic model but with the newly generated risk score as its outcome. This “full” linear regression model with all the 1 variables by definition has an \( R^2 \) value of 1. Next, each variable was removed from this model to evaluate the reduction in \( R^2 \) value. The variable with the least decrement in \( R^2 \) value was removed to reach a reduced linear model with 13 variables. In the same fashion, another variable was removed to attain a reduced model with 12 variables.

This process was repeated until there was only 1 variable left in the last reduced model. The predictive ability of each variable was therefore ranked in ascending order that the variable was eliminated from the full model (ie, the last eliminated variable had the highest predictive ability). Of the multiple variables recorded, the 8 variables with the highest predictive value were selected for parsimoniously building the final nomogram. The nomogram was internally validated by assessing discrimination and calibration. Discrimination was measured with the concordance index, similar to the area under the receiver operating characteristic curve: values range from 0.5 (no discrimination) to 1.0 (perfect discrimination). Calibration was assessed graphically by plotting the observed proportions against the predicted probabilities of hypocalcemia over a series of equally spaced values within the range of the predicted probabilities. Bootstrap was used to correct for overfitting bias. This nomogram was also validated with another validation data set that was not used for building the nomogram.

NOMOGRAM 2
The aim of nomogram 2 was to quantify the probability of hospital discharge within 24 hours based on all the variables available to the surgeon in the early postoperative period after thyroid surgery using both preoperative and intraoperative variables (Table 1). We hypothesized that this nomogram would allow the clinician to recalibrate the risk for prolonged hospital stay for an individual patient based on intraoperative information and the first postoperative serum calcium level in addition to the already available preoperative information.

End Point of Interest
The end point for this nomogram was hospital stay within 24 hours.

Statistical Methods
As with nomogram 1, a multivariable logistic regression model with restricted cubic splines for continuous variables was built to quantify the risk of hospital stay longer than 24 hours. The step-down method was used to reduce the modeling complexity by including only 8 predictor variables that had the highest predictive value among of all 21 variables. Discrimination and calibration were internally validated using bootstrap that eliminates the overfitting bias. \( P \) values less than .05 were used to determine the statistical significance. All statistical analyses and graphics were performed using open-source statistical software R, version 2.11.1 (http://www.R-project.org) with R packages of Design and Hmisc.

RESULTS

NOMOGRAM 1
The purpose of nomogram 1 was to predict the risk of hypocalcemia at first serum calcium assessment based on preoperative variables. All preoperative predictors were ranked by their predictive ability using the step-down method (Table 2). The 8 predictors with the highest predictive accuracy—age, sex, medications, history of cancer, preoperative calcium level, creatinine concentration, central neck dissection, and alkaline phosphatase level—were selected to generate a parsimonious model, which was the basis of the nomogram (Figure 2A). The nomogram had excellent accuracy for predicting hypocalcemia at the first postoperative serum calcium assessment, with a concordance index of 74.6%, and calibrated well.
In addition to validating this nomogram internally with bootstrapping, we also validated it using independent patients from our institution from the year 2005. It achieved a predictive accuracy of 67.7% and calibrated well for low-risk patients but overestimated the hypocalcemia risk for higher-risk patients.

Discrimination was measured with the concordance index, similar to the area under the receiver operating characteristic curve as area under curve: values range from 0.5 (no discrimination) to 1.0 (perfect discrimination) (Figure 2B). Calibration was expressed graphically by plotting the observed proportions against the predicted probabilities of hypocalcemia over a series of equally spaced values within the range of the predicted probabilities. Bootstrapping was used to correct for overfitting bias in the assessment of model discrimination and calibration.

The patients’ first measured postoperative serum calcium level correlated strongly to their last measured inpatient serum calcium level. Patients with lower first postoperative calcium levels tended to have lower calcium levels at their discharge ($P < .001$). Furthermore, patients with low first postoperative calcium readings (calcium level, <8 mg/dL) were more likely to have a longer hospital stay than patients with a higher value ($P = .01$). Whether or not a patient was discharged on calcium supplementation was also related to their first postoperative serum calcium level. A patient who was hypocalcemic at first postoperative assessment was 4.5 times more likely to be discharged with calcium supplementation than a patient who was not hypocalcemic ($P < .001$).

**NOMOGRAM 2**

The duration of hospital stay following surgical procedures has undergone a significant shift in recent years toward stays shorter than 24 hours. The advent of modern surgical instrumentation and anesthetic techniques has reduced surgical morbidity and decreased recovery times such that patients are more likely to fulfill day-case criteria for same-day discharge. Therefore, day surgery is now more feasible for operations that were traditionally deemed to require a longer hospital stay.

Over the last decade, there has been a shift toward same-day discharge after thyroid surgery because this has been deemed safe and is associated with a low complication rate. However, the development of postoperative hypocalcemia still causes problems in discharge planning. Hypocalcemia is the most frequent metabolic complication after thyroid surgery with a reported incidence from 0.5% to 50%. To preempt this and allow prompt discharge of patients on the same day of surgery, some institutions empirically treat patients with calcium supplementation. This practice, although widely prevalent, does not allow tailored treatment based on risk estimation. It is also not without risks and complications associated with overadministration of medications, leading to hypercalcemia, inhibition of intact parathyroid hormone (iPTH), an increased risk of myocardial infarction, and even status epilepticus.

A triage policy based on anticipated risk of hypocalcemia would therefore be useful to select patients for short hospital stay following thyroid surgery. Such a risk prediction tool would allow appropriate planning for a patient’s potential hospital length of stay and need for calcium supplementation at discharge. As yet, there is no reliable method to quantify the risk of postoperative hypocalcemia in individual patients based on preoperative parameters. Moreover, this risk is not static: it is dependent on factors that change during and/or after surgery. Our aim was to establish a nomogram that could be used to triage patients into a 24-hour stay category based on their preoperative parameters, using hypocalcemia on first blood draw after surgery as the initial end point. An ancillary nomogram was developed to allow dynamic recalibration of the risk as additional information becomes available during the patient’s hospital stay.

Currently, most institutions follow a standard protocol of a 6-hour serum calcium assessment after thyroid surgery. The total calcium test is the test most frequently ordered to evaluate calcium status. In most cases, it is a good reflection of the amount of free calcium present in the blood, since the balance between free and bound calcium is usually stable and predictable. Serum calcium levels measured 6 hours after surgery and day 1 postoperatively can be useful in predicting if the patient will develop hypocalcemia and whether the hypocalcemia...
mia will be temporary or permanent. The total calcium level reference range at our institution is 8.5 to 10.5 mg/dL. Other institutions report their ranges to be 8.8 to 10.3 mg/dL, or 2.20 to 2.57 mmol/L. There has been a call for more uniformity and standardization in the definitions used for reporting hypocalcemia rates because this will allow for more meaningful interinstitution comparison to be made.

In recent years, there has also been a move toward serum PTH assay solely or in combination with calcium as a predictor of hypocalcemia. Together, these tests can help determine whether the parathyroid glands are functioning normally after thyroid surgery with better predictability. A single early postoperative iPTH measurement may be the most cost-effective screening tool for hypocalcemia, but even greater specificity can be achieved by combining these findings with a serum calcium measurement taken 6 hours postoperatively.

Figure 2. Nomogram (A) and calibration curve (B). A, Nomogram for predicting hypocalcemia at first postoperative calcium assessment after thyroid surgery. C/M indicates a drug affecting calcium metabolism. B, Calibration curve validated on the independent group of patients. AUC refers to the C statistic, area under the curve for the receiver operating curve. Bars above the x axis indicate relative frequency of models for predicted hypocalcemia probability.

Figure 3. Nomogram for predicting the likelihood of a hospital stay shorter than 24 hours and its calibration plot. C/M indicates a drug affecting calcium metabolism.
tact parathormone measurement as early as 10 minutes after completion of surgery is helpful to predict early postoperative hypocalcemia. An iPTH level decrease of 30% or more at this time point suggests the risk of postoperative hypocalcemia.24 Other intraoperative assays of PTH are also available, but the PTH measurement on the first postoperative day is considered more useful and deemed more reliable and less expensive than intraoperative quick PTH assays.25

Although serum PTH measurement on its own may be a satisfactory tool to predict hypocalcemia, it has been shown that greater specificity can be achieved by combining it with a serum calcium measurement taken 6 hours postoperatively.2,26 An obvious disadvantage of using postoperative PTH as a predictor of hypocalcemia is that this information is not available to the clinician or the patient preoperatively and cannot be used to plan the patient’s anticipated hospital stay.

Certain thyroid disorders and surgical procedures place patients at higher risk of developing postoperative hypocalcemia. Patients undergoing thyroidectomy for Graves disease or Hashimoto thyroiditis are more likely to experience postoperative hypocalcemia owing to the difficulty in surgical dissection and increased risk for devascularization of the parathyroid glands.27,28 Patients with Graves disease are also more likely to require increased doses of calcium postoperatively.29 Total thyroidectomy, central compartment neck dissection, autotransplantation, and nonvisualization of parathyroid glands are also associated with a significantly higher risk of hypocalcemia.5,18,19,20

Serum calcium level can also be affected by many metabolic, endocrine, and dietary conditions. The corrected calcium level is a helpful assay if the albumin reading is abnormal and gives an estimate of what the calcium level would be if the albumin were within normal limits. In the present study, we used only serum total calcium values and did not correct for albumin. However, as a proxy for any affect this would have, we subdivided the comorbidity variable into the following categories: cardiac, respiratory, hypertension, endocrine/metabolic, bone conditions, and cancer history. Of note, we coded “bone condition” to include all those patients who had any metabolic bone disease or any arthritis because patients may have been taking alternative medications that were not listed in their EMR that could be of significance in their risk of hypocalcemia.

The 8 predictors with the highest predictive accuracy based on multiple logistic regression were age, sex, medications, history of cancer, preoperative serum calcium level, creatinine concentration, central neck dissection, and alkaline phosphatase level. This combination of patient clinical and biochemical parameters was the most useful in predicting the likelihood of a patient having low calcium readings at their first postoperative assessment. We also found that the first postoperative serum calcium value was a good predictor of length of stay (P = .01; Spearman correlation coefficient, −.19). Therefore, patients deemed to be at risk for hypocalcemia using this nomogram are more likely to require a longer hospital stay and can be appropriately triaged for admission.

Since the risk for postoperative hypocalcemia is dynamic, we generated a second nomogram incorporating intraoperative patient variables that would become available to the clinician in the immediate postoperative period. This nomogram (Figure 3) showed the 8 variables of highest predictive value to be age, sex, central compartment neck dissection, number of parathyroid glands excised and reimplanted, intraoperative fluids, procedure time, medication, and first postoperative serum calcium level. These findings reaffirm the commonly accepted surgical risk factors associated with hypocalcemia in the literature.5,18,19,20 The use of this nomogram in the immediate postoperative period has practical utility in allowing the clinician to counsel a patient about anticipated discharge from hospital, allowing patients and health care providers to make appropriate provisions.

The development of these nomograms may introduce a change in our ability to predict an individual patient’s risk of postoperative hypocalcemia based on a mathematical probability. Although the risk factors increasing the likelihood of postthyroidectomy hypocalcemia are well known, there has never before been a model where risk can be calculated for individual patients. This is an exciting new prospect in categorizing patients into high- or low-risk groups and should allow for tailored provision of inpatient care. Large series have shown day-case thyroid surgery to be both a safe and cost-effective health model.2,11,15,30 In addition, the routine implementation of postoperative calcium and vitamin D supplements has been suggested as a preventive measure for symptomatic hypocalcemia.17 However, our nomogram based on individual risk of developing hypocalcemia and probability of actual need of calcium supplementation based on patient variables shows promise and may offer an alternative for reliable early-discharge planning.

The obvious limitation of this study is that the nomograms were generated based on retrospective data. Therefore, patients were selected on the basis of availability of biochemical values, thus creating potential sampling bias. Although an independent internal set of patients treated during a different time period was used for validation, these data were also retrospective and involve the same limitations as the modeling set. To test the relevance of these nomograms in other clinical practices, we are in the process of validating them at other institutions.

We are also currently evaluating these nomograms prospectively and are exploring the relationship of several other patient variables in quantifying individual risk for postoperative hypocalcemia. For example, it has been shown that postoperative PTH levels better predict long-term hypocalcemia requiring vitamin D supplementation than do serum calcium levels alone.31 However, the precise interplay of the multitude of factors that might influence the development of hypocalcemia and the need for supplementation is difficult to define for an individual patient without using some form of mathematical prediction. Therefore, we are currently conducting a prospective study incorporating postoperative PTH assay into our nomogram and are prospectively assessing the utility of such a nomogram in risk prediction for patients undergoing thyroid surgery at our institution.
In conclusion, we have produced a set of nomograms that can dynamically quantify the risk of postthyroidectomy hypocalcemia based on preoperative clinical and biochemical and intraoperative surgical variables. These nomograms may be helpful in planning the length of a patient’s hospital stay and facilitate discharge planning after thyroid surgery.

Submitted for Publication: March 10, 2011; final revision received September 7, 2011; accepted September 18, 2011.

Correspondence: Snehal G. Patel, MD, Department of Head and Neck Surgery, Memorial Sloan Kettering Cancer Center, 1275 York Ave, New York, NY 10065 (patels@mskcc.org).

Author Contributions: Dr Patel had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Ali, Palmer, Ganly, Shaha, Shah, and Patel. Acquisition of data: Ali and Palmer. Analysis and interpretation of data: Ali, Yu, Palmer, Shaha, Shah, Kattan, and Patel. Drafting of the manuscript: Ali, Yu, Palmer, Ganly, and Patel. Critical revision of the manuscript for important intellectual content: Ali, Yu, Palmer, Ganly, Shaha, Shah, Kattan, and Patel. Statistical analysis: Yu and Kattan. Administrative, technical, and material support: Palmer and Shaha. Study supervision: Ganly, Shaha, Shah, and Patel.

Financial Disclosure: None reported.

Previous Presentation: This study was presented at the American Head and Neck Society 2011 Annual Meeting; April 27, 2011; Chicago, Illinois.

REFERENCES