**Title: Development and internal-external validation of a prognostic model for long-term hypoparathyroidism after total or completion thyroidectomy.**

Sam P.J. van Dijk1, BSc; Carolien C.H.M. Maas2,3, Ali Alshangi1, BSc; Caroline M.J. van Kinschot4, MD; Ivona Lončar1, BSc; MSc; Charlotte van Noord5, MD, PhD; Linetta B. Koppert1, MD, PhD, MSc; David van Klaveren2, MD, PhD; Cornelis Verhoef 1, MD, PhD; Robin P. Peeters4, MD, PhD; Tessa M. van Ginhoven1, MD, PhD.

1 Department of Surgical Oncology and Gastrointestinal Surgery, Erasmus MC Cancer Institute P.O. 2040, 3000, CA, Rotterdam, the Netherlands. 2 Department of Public Health, Erasmus University Medical Center, P.O. 2040, 3000, CA, Rotterdam, The Netherlands.  
3 Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands.  
4 Department of Internal medicine, Erasmus Medical Center Rotterdam, P.O. 2040, 3000, CA, Rotterdam, the Netherlands.  
5 Department of Internal Medicine, Maasstad Hospital Rotterdam, P.O. 3079 DZ, Rotterdam, the Netherlands.

**Corresponding author:**

Sam P.J. van Dijk

Erasmus MC, University Medical Center Rotterdam, Department Surgical Oncology and Gastrointestinal Surgery

Doctor Molewaterplein 40, 3015 GD Rotterdam, The Netherlands

Email: [s.p.j.vandijk@erasmusmc.nl](mailto:s.p.j.vandijk@erasmusmc.nl)

**Keywords** Long-term hypoparathyroidism; Prediction Model; Thyroid

**Word count:** 2318

**Abstract**

**Objective**

To develop and internal-externally validate a clinical prediction model to predict the occurrence of long-term hypoparathyroidism after total or completion thyroidectomy.

**Design**

Multicenter retrospective cohort study.

**Setting**

Eleven hospitals Dutch hospitals.

**Participants**

Patients were included if they underwent a total or completion thyroidectomy between January 2010 and June 2021.

**Main outcome**

The occurrence of long-term hypoparathyroidism. Long-term hypoparathyroidism was defined as the need for active vitamin D supplementation 1 year after surgery, and inability to be tapered of vitamin D supplementation. The prediction model was developed using multivariable logistic regression analysis and was internal-externally validated using leave-one-center-out cross-validation.

**Results**

Of the 366 included patients, 44 (12.0%) patients developed hypoparathyroidism in the first year follow-up after surgery. No patients had long-term hypoparathyroidism when PTH levels decreased less than 70 percent 24 hours after surgery. Multivariable logistic regression analysis showed that delta PTH 24 hours after surgery (OR 1.08; 95% CI: 1.05-1.12), corrected calcium 24 hours after surgery (OR 1.43; 95% CI: 1.12-1.82) and not identifying at least one parathyroid during surgery (OR 3.78; 95% CI: 1.58-9.04) were predictors for developing long-term hypoparathyroidism. The discriminative ability of the model was excellent (optimism-corrected C-index 0.89, 95% CI: 0.85-0.92), but calibration assessed through internal-external leave-one-center-out cross-validation was poor due to low number of events.

**Conclusion**

The model proposed in this study showed good performance and could be used to perform an individual assessment of patients at risk for long-term hypoparathyroidism after total thyroidectomy.

**Background**

Hypoparathyroidism is a prevalent complication following total or completion thyroidectomy that can lead to hypocalcemia as a result of unintended damage to the parathyroid glands. While iatrogenic hypoparathyroidism is often temporary and resolves within a few weeks 1, in some patients hypoparathyroidism becomes a chronic condition known as long-term hypoparathyroidism. The incidence of long-term hypoparathyroidism varies in the literature, partly due to the absence of consensus on its definition, ranging from less than 5% to 15% 2-4. Patients with long-term hypoparathyroidism experience lower health-related quality of life and recent studies suggest an increased mortality risk compared to individuals without this condition 5-7.

Several risk factors for developing long-term hypoparathyroidism are described in literature including extensive surgical procedures (e.g., central lymph node dissection) and low post-operative parathyroid hormone (PTH) levels 8, 9. Post-operative PTH levels play a crucial role in predicting long-term hypoparathyroidism after total or completion thyroidectomy 5, 10, 11. A recent meta-analysis demonstrated the existence of multiple PTH-based methods for predicting hypoparathyroidism, varying in terms of the timing of PTH measurements and the threshold levels employed 12. Although most studies report sensitivity and specificity outcomes of postoperative PTH levels, a comprehensive prognostic model that assesses individual patient risks of developing long-term hypoparathyroidism is currently lacking in the literature.

Identifying patients at high risk of developing long-term hypoparathyroidism can play a crucial role in preventing post-discharge adverse events, including readmissions caused by hypocalcemia. Conversely, the recognition of patients with a significantly low risk can de-escalate the intensity of follow-up care. In this study, we aimed to address this critical gap by developing and validating a comprehensive clinical prognostic model predicting the occurrence of long-term hypoparathyroidism following total or completion thyroidectomy.

**Methods**

**Study design**

This was a multicenter retrospective cohort study.

**Patients and data collection**

Patients were included in this study if 1) they underwent a total or completion thyroidectomy, 2) PTH has been determined preoperatively (up to one year preoperatively) or postoperatively on day 1 and if 3) follow-up data up to 1 year is available. Exclusion criteria were: surgery in the central neck compartment or external beam radiation therapy of the neck within 1 year of the index procedure. Model development and internal-external validation was performed in a cohort of adult patients who underwent surgery between January 2010 and June 2021 in one of the ten hospitals in the South-Western region of the Netherlands participating in a collaborative organization called the Thyroid Network and one academic hospital outside the region (Supplement X).

Baseline characteristics such as age, sex, preoperative diagnosis, medical history, blood values (calcium, albumin, PTH and vitamin D serum levels), surgical procedure, pathological specimen and follow-up data were obtained. Serum calcium levels were corrected for albumin levels according to the formula: corrected calcium (mmol/L) = measured calcium (mmol/L) + 0.016 x (34-albumin (g/L)). The reference value for calcium is 2.20-2.65 mmol/L. The change in PTH levels (delta PTH) was calculated by (PTH at baseline - postoperative PTH after 24 hours) / (PTH at baseline) ⋅ 100%. The change in corrected calcium levels (delta corrected calcium) was calculated by (corrected calcium at baseline - postoperative corrected calcium after 24 hours) / (corrected calcium at baseline) ⋅ 100%. The TRIPOD (transparent reporting of a multivariable model for individual prognosis or diagnosis) guidance for development and reporting of multivariable prediction models was followed (Supplement X) 13.

**Outcome measures**

The primary outcome of this study was long-term hypoparathyroidism. Long-term hypoparathyroidism was defined as the need for active vitamin D supplementation 1 year after surgery, and inability to be tapered of vitamin D supplementation 14. Secondary outcome was the incidence of readmissions after initial discharge.

**Statistical analysis for model development**

Missing values were imputed using Multivariate Imputations by Chained Equations (MICE) 15. Variables used in the multiple imputation model included the candidate predictors, auxiliary variables and the outcome 16. Estimates of the imputed datasets were pooled using Rubin’s Rule 21.

Multivariable logistic regression analysis was used to predict the probability of long-term hypoparathyroidism based on certain patient characteristics. Identification of candidate predictors of the outcome was based on clinical expertise and the existing literature. Selection of variables for the final model was conducted using backward selection with p<0.05 17, 18 (ref). We calculated the importance of each predictor in the model using Wald tests. To compare the fit of nested models, we used likelihood ratio tests. Predictions of the final model were made easily accessible through a web application.

To prevent overfitting of the model, we utilized a bootstrap approach with 1000 resamples to derive a uniform shrinkage factor19, 20. This shrinkage factor was then multiplied by the final regression coefficient of each selected predictor and the intercept was adjusted appropriately.

The performance of the model was assessed using the concepts of calibration and discrimination. Calibration of the model was assessed by calibration plots, i.e., plotting observed frequencies versus predicted probabilities in groups. Perfect predictions should lie on the 45-degree line for agreement with the outcome in the calibration plot. We used leave-one-center-out cross-validation to assess the validity of the model, i.e., models were fit using data leaving one center out of the development set and making predictions for the left out center. Discrimination of the model was assessed by providing the C-index. A C-index of 1.0 indicates perfect discrimination, i.e., the model perfectly distinguishes between high and low risk patients, whereas an C-index of 0.5 indicates that the model is no better than chance.

Lastly, to assess the relationship between readmission due to symptoms related to hypocalcemia and the risk of long-term hypoparathyroidism was assessed by dividing patients into risk groups based on predicted probabilities of the model (%); low risk (0-10%), intermediate risk (10%-40%) and high risk (>40%) patients. All statistical analysis was performed using the R version 4.1.0.

**Results**

The flow diagram of patient inclusion is shown in Supplemental Figure 1, and the patients characteristics are summarized in Table 1. The cohort included 366 patients after exclusion of 16 patients due to surgery in the central neck compartment 1 patient due to external beam radiation therapy in the head and neck region within one year after initial surgery. The median age of the derivation cohort was 56.0 [IQR, 42.0-69.0] and 32% was male (n=117). Total thyroidectomy was performed in 292 (80%) patients and 74 (20%) patients underwent a completion thyroidectomy. Long-term hypoparathyroidism occurred in 44 (12.0 %) patients one year after surgery. We observed low percentages of missing values for PTH measurements in the derivation cohort (n=19, 5.2% baseline PTH; n=28, 7.7% 24-hour PTH). All eligible patients were included in the model development after imputing missing values.

After backwards selection methods, multivariable logistic regression analysis of all 366 patients showed that delta PTH 24 hours after surgery (OR 1.08; 95% CI 1.05-1.12 after applying uniform shrinkage of 0.868), corrected calcium 24 hours after surgery (OR 1.44; 95% CI 1.11-1.86), and not identifying at least one parathyroid during surgery (OR 3.90; 95% CI 1.62-9.37) were all significantly associated with long-term hypoparathyroidism (Table 2). The model demonstrated a strong discriminatory capacity with an optimism-corrected C-index of 0.88 (CI 0.84-0.92) (Table 2). Internal-external validation of the model showed an overall high c-index, but poor calibration (Figure 2). Although delta PTH is the most important predictive factor, 24-hour corrected calcium and the identification of one parathyroid gland during surgery were found improve the model significantly (LR-test see Supplemental Figure 1; C-index 0.85 versus 0.88 see Table 2).

No patients with a PTH decrease of less than 70% in this cohort developed long-term hypoparathyroidism. In the derivation cohort, 228 (62.6%) were classified as low risk, 75 patients (20.2%) were classified as intermediate risk and 63 patients (17.2%) were located in the high-risk group. Two (0.9%) low-risk patients were readmitted, whereas seven (9.3%) patients in the intermediate-risk group and 18 (28.6%) patients in the high risk group were readmitted (Table 4).

**Discussion**

This study identified predictors for long-term hypoparathyroidism in patients who underwent total or completion thyroidectomy and developed a prognostic model for this population. The prediction model proposed in this study showed high discriminating power (C-index 0.88, 95% CI: 0.84-0.92) and is the first model for prediction long-term hypoparathyroidism that has been internally or externally validated. Larger multicenter studies are necessary to enhance accuracy and validate the the current model.

There has been one other study that proposed a prediction tool for developing long-term hypoparathyroidism after total thyroidectomy 22. That prediction tool was not validated and contained two parameters which could only be measured after one month: PTH and calcium levels at 1 month after surgery. The current study considered the selection of variables who are available just one day after surgery. This facilitates a quick individual assessment of patients at risk for long-term hypoparathyroidism after total thyroidectomy and could aid in personalized discharge instructions and supplementation regimens. By identifying patients at higher risk, healthcare providers can further allocate appropriate follow-up care, calcium and vitamin D supplementation, and potentially reduce the burden on healthcare resources by targeting interventions to those who need them most.

The incidence of long-term hypoparathyroidism in our cohort was 12.0%, which seems to be in line with incidences reported in other nation-wide multicenter studies 4, 23. Furthermore, we substantiated findings of Loncar *et al*., who found that patients with a delta PTH of less than 70% 24-hour after surgery had no risk of long-term hypoparathyroidism 24. In the current study, no patient developed long-term hypoparathyroidism who had a delta PTH of less than 70%. This observation suggests that evaluating the 24-hour PTH decrease can serve as an initial assessment tool to determine if a patient is at risk for long-term hypoparathyroidism. Moreover, our study highlights the significance of incorporating additional variables other than PTH into the prognostic model. Although a decrease in PTH levels emerged as the most significant predictor accounting for X% of the models’ predictive ability, we observed that parathyroid gland identification and 24-hour calcium measurements significantly enhanced the discriminatory ability of the model. Consequently, this comprehensive approach provides clinicians with a more robust tool for prognostication and aids in optimizing patient care strategies following thyroidectomy procedures.

During the early postoperative phase following thyroidectomy, the main objectives are to ensure that patients remain free of symptoms and can be discharged without any risks. However, unnecessary calcium supplementation comes with important health risks such as cardiovascular events and kidney stones 25 and should be prescribed purposefully, accounting for individual patients' risks and benefits. While the primary focus of the model developed in this study was to predict the occurrence of long-term hypoparathyroidism, we also observed a notable correlation between a higher risk of long-term hypoparathyroidism and an increased likelihood of readmissions due to symptoms related to hypocalcemia. We observed no readmissions in patients with a PTH decrease <70% and patients with a low predicted probability of developing long-term hypoparathyroidism (<10%). This suggests that providing clear information regarding the symptoms of hypoparathyroidism and hypocalcaemia at discharge may be sufficient for patients in low-risk patients. Patients with an intermediate to high risk of long-term hypoparathyroidism should receive early interdisciplinary care and close follow-up in collaboration of general practitioner and endocrinologist 14. Future research endeavors should establish distinct prediction models that can accurately assess the risk of symptoms and readmissions during the initial postoperative period. These efforts should aim to enhance and tailor the management strategies implemented after surgery to each individual patient.

Future studies should aim to overcome certain limitations observed in this study. We had a relatively small sample of patients with long-term hypoparathyroidism (44 cases) which makes the model more prone to incidental findings. Therefore, we used uniform shrinkage obtained by bootstrapping and validated the model using leave-one-center-out cross-validation26. Bootstrapping facilitates obtaining optimal estimates of internal validity of logistic regression models developed in smaller samples (e.g., events per variable ≤10) 27. Larger-scale studies are warranted to update the model and perform external validation, ensuring its reliability and generalizability. Furthermore, it is a retrospective cohort study. We tried to carefully extract the data but possibly some information bias remains. Furthermore, following Loncar et al., but a different threshold could lead to different conclusions Lastly, we assumed that there was no substantial difference in in-laboratory measurement variation between the different hospitals.

**Conclusion**

The model proposed in this study showed good performance measures and could be used to perform an individual assessment of patients at risk for long-term hypoparathyroidism after total thyroidectomy. External validation of the model proposed in this study is required to determine its usefulness in other patient populations.

**Ethical approval**

The Medical Ethics committee of the Erasmus Medical Center approved this study ( (MEC-2018-1195, MEC-2013-233, MEC-2017-1041).

**References**

1. Ritter K, Elfenbein D, Schneider DF, et al. Hypoparathyroidism after total thyroidectomy: incidence and resolution. *J Surg Res* 2015; 197(2):348-53.

2. Asari R, Passler C, Kaczirek K, et al. Hypoparathyroidism after total thyroidectomy: a prospective study. *Arch Surg* 2008; 143(2):132-7; discussion 138.

3. Kim SM, Kim HK, Kim KJ, et al. Recovery from Permanent Hypoparathyroidism After Total Thyroidectomy. *Thyroid* 2015; 25(7):830-3.

4. Díez JJ, Anda E, Sastre J, et al. Prevalence and risk factors for hypoparathyroidism following total thyroidectomy in Spain: a multicentric and nation-wide retrospective analysis. *Endocrine* 2019; 66(2):405-415.

5. Almquist M, Hallgrimsson P, Nordenström E, et al. Prediction of permanent hypoparathyroidism after total thyroidectomy. *World J Surg* 2014; 38(10):2613-20.

6. Siggelkow H, Clarke BL, Germak J, et al. Burden of illness in not adequately controlled chronic hypoparathyroidism: Findings from a 13-country patient and caregiver survey. *Clin Endocrinol (Oxf)* 2020; 92(2):159-168.

7. Jørgensen CU, Homøe P, Dahl M, et al. Postoperative Chronic Hypoparathyroidism and Quality of Life After Total Thyroidectomy. *JBMR Plus* 2021; 5(4):e10479.

8. Godlewska P, Benke M, Stachlewska-Nasfeter E, et al. Risk factors of permanent hypoparathyroidism after total thyroidectomy and central neck dissection for papillary thyroid cancer: a prospective study. *Endokrynol Pol* 2020; 71(2):126-133.

9. Cho JN, Park WS, Min SY. Predictors and risk factors of hypoparathyroidism after total thyroidectomy. *International Journal of Surgery* 2016; 34:47-52.

10. Loncar I, Noltes ME, Dickhoff C, et al. Persistent Postthyroidectomy Hypoparathyroidism in the Netherlands. *JAMA Otolaryngol Head Neck Surg* 2021; 147(11):959-965.

11. Wang W, Xia F, Meng C, et al. Prediction of permanent hypoparathyroidism by parathyroid hormone and serum calcium 24 h after thyroidectomy. *Am J Otolaryngol* 2018; 39(6):746-750.

12. Nagel K, Hendricks A, Lenschow C, et al. Definition and diagnosis of postsurgical hypoparathyroidism after thyroid surgery: meta-analysis. *BJS Open* 2022; 6(5).

13. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Annals of Internal Medicine* 2015; 162(1):W1-W73.

14. Bollerslev J, Rejnmark L, Zahn A, et al. European expert consensus on practical management of specific aspects of parathyroid disorders in adults and in pregnancy: recommendations of the ESE Educational Program of Parathyroid Disorders (PARAT 2021). *European Journal of Endocrinology* 2022; 186(2):R33-R63.

15. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; 45(3):1 - 67.

16. Moons KGM, Donders RART, Stijnen T, et al. Using the outcome for imputation of missing predictor values was preferred. *Journal of Clinical Epidemiology* 2006; 59(10):1092-1101.

17. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. *Fam Med Community Health* 2020; 8(1):e000262.

18. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression. Vol. 398: John Wiley & Sons, 2013.

19. Smith GC, Seaman SR, Wood AM, et al. Correcting for optimistic prediction in small data sets. *Am J Epidemiol* 2014; 180(3):318-24.

20. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.: New York: Springer, 2009.

21. Rubin DB. Multiple imputation after 18+ years. *Journal of the American statistical Association* 1996; 91(434):473-489.

22. Sitges-Serra A, Gómez J, Barczynski M, et al. A nomogram to predict the likelihood of permanent hypoparathyroidism after total thyroidectomy based on delayed serum calcium and iPTH measurements. *Gland Surgery* 2017:S11-S19.

23. Jørgensen CU, Homøe P, Dahl M, et al. High incidence of chronic hypoparathyroidism secondary to total thyroidectomy. *Dan Med J* 2020; 67(5).

24. Lončar I, Dulfer RR, Massolt ET, et al. Postoperative parathyroid hormone levels as a predictor for persistent hypoparathyroidism. *Eur J Endocrinol* 2020; 183(2):149-159.

25. Li K, Wang XF, Li DY, et al. The good, the bad, and the ugly of calcium supplementation: a review of calcium intake on human health. *Clin Interv Aging* 2018; 13:2443-2452.

26. Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Statistics in Medicine* 2019; 38(7):1276-1296.

27. Steyerberg EW, Harrell FE, Borsboom GJJM, et al. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology* 2001; 54(8):774-781.