Surgery 165 (2019) 92–98

Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

The optimal dosing scheme for levothyroxine after thyroidectomy: A comprehensive comparison and evaluation ☆,☆☆

Nick A. Zaborek, MA [a,*], Andy Cheng, MD [b], Joseph R. Imbus, MD [a], Kristin L. Long, MD [a], Susan C. Pitt, MD, MPHS [a], Rebecca S. Sippel, MD [a], David F. Schneider, MD, MS [a]

[a] Department of Surgery, University of Wisconsin, Madison, WI
[b] St. Joseph Medical Center, Tacoma, WA

ARTICLE INFO

Article history:
Accepted 3 April 2018
Available online 6 November 2018

ABSTRACT

Background: Patients often struggle to attain euthyroidism after thyroidectomy, and multiple dosing schemes have been proposed to supplant the standard weight-based approach for initial levothyroxine dosing after thyroidectomy. The objectives of this study were to review the literature for existing levothyroxine dosing schemes and compare estimation accuracies with novel schemes developed with machine learning.

Methods: This study retrospectively analyzed 598 patients who attained euthyroidism after total or completion thyroidectomy for benign disease. A scoping review identified existing levothyroxine dosing schemes. Thirteen machine learning algorithms estimated euthyroid dose. Using 10-fold cross-validation, we compared schemes by the proportion of patients having a predicted dose within 12.5 μg/day of their euthyroid dose.

Results: Of 264 reviewed articles, 7 articles proposed retrospectively implementable dosing schemes. A novel Poisson regression model proved most accurate, correctly predicting 64.8% of doses. Incorporating 7 variables, Poisson regression was significantly more accurate than the best scheme in the literature (body mass index/weight based) that correctly predicted 60.9% of doses ($P = .031$). Standard weight-based dosing (1.6 μg/kg/day) correctly predicted 51.3% of doses, and the least effective scheme (age/sex/weight based) correctly predicted 27.4% of doses.

Conclusion: Using readily available variables, a novel Poisson regression dosing scheme outperforms other machine learning algorithms and all existing schemes in estimating levothyroxine dose.

Introduction

Levothyroxine (LT4) is the most widely prescribed drug in the United States.[1] LT4 is prescribed after thyroidectomy with the intent to restore normal thyroid hormone function, and for patients with benign thyroid disease, the typical starting dose of LT4 is 1.6 to 1.7 μg/kg.[2] Often, this weight-based initial dosing is inadequate, with about 70% of patients requiring dose adjustments at the first postoperative follow-up.[3-5] Overdosing of LT4 increases patient risk of accelerated bone loss, fractures, heat intolerance, diarrhea, and arrhythmias.[6-9] Underdosing of LT4 results in symptoms of hypothyroidism, including fatigue and weight gain.[10] Hence, incorrect thyroid hormone dosing after thyroidectomy greatly affects patient quality of life. Many authors have proposed alternative LT4 dosing schemes to improve initial LT4 dosing after thyroidectomy to decrease the time patients spend in either a hypo- or hyperthyroid state postoperatively.

The literature contains a range of approaches for LT4 dosing. Some authors propose modified weight-based schemes[3,11]; others incorporate body weight and 1 or 2 additional patient factors, including body mass index (BMI), age, sex, or lean body mass (LBM).[4,5,12-14] We previously developed and tested a BMI-based regression algorithm that was superior to simple, weight-based dosing[5]; however, prospective evaluation indicated that it still incorrectly dosed 61.1% of patients.[14] Overall, the existing dosing schemes lack accuracy and no consensus exists for the optimal dosing scheme.

One potential method to improve prediction of the correct LT4 dose is with machine learning (ML). ML approaches focus on the creation of a predictive model using labeled datasets, predicting a dependent variable from explanatory factors. In clinical medicine,

☆ Supported by NIH grants UL1TR000427 and KL2TR000428.
☆☆ Presented at the 39th Annual Meeting of the American Association of Endocrine Surgeons in Durham, NC, May 6–8, 2018.
* Reprint requests: Nick Zaborek, 600 Highland Ave, K6/100E Clinical Science Center, Madison, WI 53792-7375.
E-mail addresses: zaborek@surgery.wisc.edu, zaborek@wisc.edu (N.A. Zaborek).

https://doi.org/10.1016/j.surg.2018.04.097
0039-6060/

ML approaches are increasingly used to improve the accuracy of decision-making. For example, ML can successfully detect hyperparathyroidism[15] and identify risk of breast cancer from mammographic findings.[16] To date, ML has not been applied to LT4 dosing. Moreover, existing LT4 dosing schemes have not been compared to more sophisticated ML methods using all relevant patient factors.

The purpose of this study was 2-fold: (1) to evaluate the accuracy of ML approaches in predicting LT4 dose, and (2) to compare the accuracy of LT4 dosing schemes proposed in the literature. We hypothesized that a predictive model developed with supervised ML could improve initial LT4 dosing after thyroidectomy. We propose a predictive model that correctly doses significantly more patients than all other existing approaches.

Methods

This is a 2-part study. The first part comprises a scoping review to identify LT4 dosing schemes proposed in existing literature, and the second part includes developing predictive algorithms using a retrospective dataset and comparing their performance to previously proposed schemes identified in the scoping review. The Institutional Review Board of the University of Wisconsin School of Medicine and Public Health approved data collection for this Health Insurance Portability and Accountability Act-compliant study.

Scoping review

The scoping review identified existing LT4 dosing schemes after thyroidectomy for benign disease. Databases PubMed, Cochrane, Scopus, and Web of Science were queried using variations of the following search terms: levothyroxine, thyroidectomy, dosing, and algorithm. Additional articles were identified through citations of articles known to be relevant. To be included in our study, the article must have proposed a method to dose LT4 in adults ($\geq$18 years of age) after noncancerous total or completion thyroidectomy, or dose LT4 for adult patients needing full thyroid hormone replacement. In addition, the article had to be written in English with no time frame restrictions. Articles were excluded if they proposed schemes that could not be implemented retrospectively. Articles returned from the search were independently reviewed and agreed on by 3 authors for relevance.

Study sample and definitions

We performed a retrospective review of our institutional thyroid surgery database containing records for 598 patients who underwent total or completion thyroidectomy with pathology showing benign thyroid disease. Included were patients who achieved euthyroidism, defined as a serum thyroid-stimulating hormone (TSH) level of 0.45–4.50 mIU/mL, between 2008 and May 2017. We excluded patients who were under 18 years old, were taking liothyronine in addition to levothyroxine, had thyroid cancers greater than 1 cm in size (because these patients need adjusted doses for TSH suppression), or did not achieve euthyroidism by the time of data collection.

The recommended starting dose of LT4 was generally 1.6 μg/kg/day, although starting from October 2012, patients were dosed according to our BMI-based algorithm.[5] TSH values were checked 6 to 8 weeks after thyroidectomy, and levothyroxine doses were adjusted if necessary to meet the TSH goal. Subsequent dose titration occurred at 6- to 8-week intervals.

We collected the following 44 variables for each patient: initial and euthyroid dose; clinical demographics such as age, sex, race, weight, and BMI; pertinent medical history such as autoimmune thyroid disease, comorbidities, and medications; surgical pathology; and laboratory data such as preoperative TSH and postoperative parathyroid hormone levels.

ML model development and evaluation

We evaluated the ability of various supervised ML predictive models to predict daily LT4 requirements after thyroidectomy. A wide variety of predictive models were examined, including but not limited to support vector machines, Bayesian recurrent neural networks, decision trees, random forests, ordinary least squares (OLS) regression, Poisson regression, gamma regression, ridge regression, and the lasso.

A subset of explanatory variables was selected as predictors based on the following process. First, a random forest (built on the entire dataset) generated feature importance values for all predictor variables. Next, the least important feature was omitted from the building of another random forest built on the entire dataset, and the out-of-bag mean squared error was noted. This step was repeated successively for the 2 least important features, then the 3 least important features, and so on. The variables from the random forest with the lowest out-of-bag mean squared error were chosen as the model building features.

Before any comparison or evaluation, all predictions were rounded to the nearest 12.5 μg to represent a practical initial dose because 12.5 μg is the smallest increment between dosing strengths. We evaluated the predictive accuracy of each algorithm by calculating the percentage of patients whose predicted LT4 dose was within 12.5 μg of their actual euthyroid dose. In this case, we say the algorithm correctly predicted the LT4 dose. We determined the most accurate model to be the one with the greatest proportion of correct predictions.

Statistical analysis

All predictive algorithms were evaluated using repeated 10-fold cross-validation. We formally tested the difference in performance between the best ML algorithm and the best scheme proposed in the literature using the method outlined by Bouckaert[17] for cross-validated results. All models and descriptive statistics were generated with R (version 3.4.2; Vienna, Austria), a free and open source software environment for statistical computing.[18] The caret package for R was used extensively for building ML models.[19]

Results

Scoping review

Figure 1 details the results of the literature search. From an initial pool of 264 articles meeting our search criteria, a total of 9 articles proposed full replacement LT4 dosing schemes. Ultimately, 7 articles proposed schemes that could be implemented retrospectively and compared. The methods proposed by Sukumar/Agarwal et al[20] and Banovac et al[21] were excluded because they used a dual-absorption x-ray densitometry machine to measure LBM and hence could not be implemented retrospectively. Table 1 details the 7 proposed schemes identified in the literature and included in this comparison.

Cohort characteristics

A total of 598 patients were included in the retrospective cohort—504 patients (84.3%) were female, the median age was 51 years (range 18–84 years), and the median weight was 81 kg (range 38–206 kg) with a median BMI value of 29.3. Most patients (57.5%) underwent thyroidectomy for thyroid nodular disease or

Table 1
Existing schemes proposed in the literature for the estimation of LT4 requirements after thyroidectomy

Author/location	*Formula or scheme*
Cunningham et al. 1984/New Haven, CT	LT4 dose (μg/day) = 3.4 × LBM − 11; LBM (male) = (79.5 − 0.24 × weight − 0.15 × age) × weight/73.2* LBM (female) = (69.8 − 0.26 × weight − 0.12 × age) × weight/73.2
Olubowale et al. 2005/Chesterfield, UK	LT4 dose (μg/day) = 100 if weight <53 = 125 if 53 ≤ weight ≤ 86 = 150 if 86 < weight ≤ 108 = 175 if weight >108
Mistry et al. 2011/Hull, UK	LT4 dose (μg/day) = (0.943 × weight) + (−1.165 × age) + 125.8
Ojomo et al. 2013/Madison, WI	LT4 dose (μg/day) = (−0.018 × BMI + 2.13) × weight
Jin et al. 2013/Cleveland, OH	LT4 dose (μg/day) = 1.5 × weight
Di Donna et al. 2014/Rome, Italy	LT4 dose (μg/kg/day) =

BMI	≤23	23–28	>28
Age			
≤40	1.8	1.7	1.6
>40–55	1.7	1.6	1.5
>55	1.6	1.5	1.4

Elfenbein et al. 2015/Madison, Wisconsin — LT4 dose (μg/kg/day) =

BMI	Male	Female
<21	2.1	1.8
22–26	1.9	1.7
27–32	1.7	1.6
33–40	1.5	1.4
>40	1.3	1.2

* All weights are in kg.

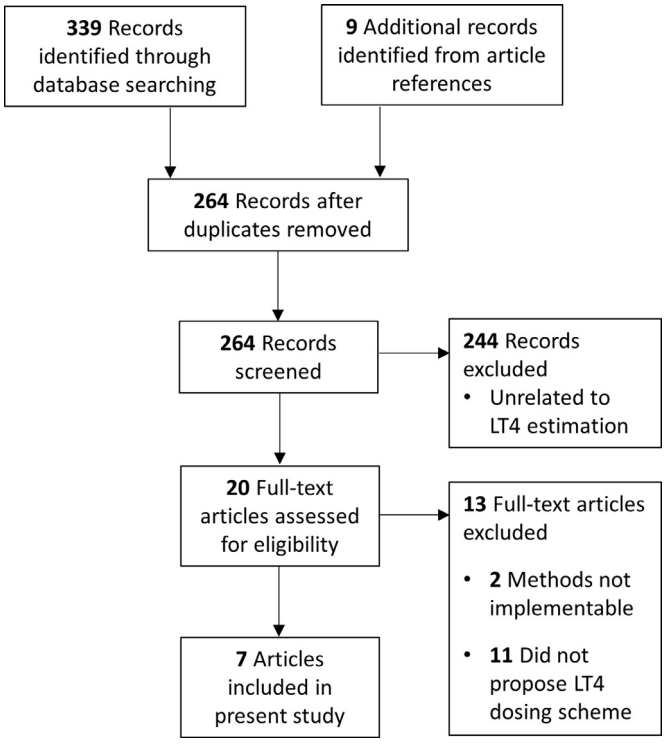

Fig 1. Scoping review search results.

Table 2
Baseline characteristics of the cohort

Cohort characteristics	
Sex, % (n)	
Male	15.7% (94)
Female	84.3% (504)
Weight, kg (median ± SD)	80.7 ± 22.6
Height, in (median ± SD)	65.0 ± 3.3
Age, y (median ± SD)	51.1 ± 14.1
BMI, kg/m^2 (median ± SD)	29.3 ± 8.1
Smoker, % (n)	
Yes	18.7% (112)
No	81.3% (486)
Preoperative TSH (median ± SD)	0.8 ± 2.4
Indication for surgery, % (n)	
Thyroid nodular disease/multinodular goiter	57.5% (344)
Hyperthyroidism	42.5% (254)
Graves' hyperthyroidism	66.5% (169)
Hashimoto's disease	18.4% (110)
Iron supplementation, % (n)	
Yes	2.5% (15)
No	97.5% (583)
Multivitamin-mineral use, % (n)	
Yes	13.2% (79)
No	86.8% (519)
Weeks to euthyroidism (median ± SD)	16.5 ± 17.9
Initial LT4 dose, μg/day (median ± SD)	125 ± 29.2
Euthyroid LT4 dose, μg/day (median ± SD)	125 ± 38.9
Preoperative LT4 use, % (n)	
Yes	17.2% (103)
No	82.8% (495)

SD, standard deviation.

multinodular goiter. The second most common indication for thyroidectomy was hyperthyroidism (42.5%). In this cohort, 10.5% of thyroidectomy specimens contained a micropapillary thyroid carcinoma, although this was not the initial indication for thyroidectomy. The median time to euthyroidism was 17 weeks (range 2–123 weeks). Nearly one-fifth of the cohort (18.7%) were smokers, 103 (17.2%) were taking LT4 preoperatively, and 79 (13.2%) were taking multivitamin-mineral supplements. Baseline characteristics of the cohort are summarized in Table 2.

ML results

Seven patient variables were identified by the random forest variable selection process to develop ML algorithms: weight, age, BMI, sex, preoperative TSH, iron supplementation use, and multivitamin/mineral use. After testing multiple predictive algorithms, we

Table 3
Accuracy of various machine learning algorithms in predicting LT4 requirements

Method	*% euthyroid doses*
Random forest	51.3%
SVM (Gaussian kernel)	55.6%
Ordinal regression	57.0%
K nearest neighbors	59.5%
Boosted decision trees	60.5%
SVM (polynomial kernel)	63.2%
Lasso	63.4%
Ridge regression	64.0%
Bayesian recurrent neural networks	64.0%
Boosted GLM	64.2%
Gamma regression	64.4%
OLS regression	64.4%
Poisson regression	64.8%

**GLM*, generalized linear model; *SVM*, support vector machine.

found Poisson regression to be most successful at predicting LT4 dose, correctly predicting 64.8% of LT4 doses. Poisson regression is a simple extension of the ubiquitous OLS linear regression, wherein the dependent variable is assumed to have a Poisson distribution. Other linear algorithms provided similar but less accurate results, such as OLS and gamma regression, each correctly predicting 64.4% of doses. Nonlinear methods, such as random forests and Gaussian kernel support vector machines, performed relatively poorly with 51.3% and 55.6% correctly predicted doses, respectively. Table 3 describes the performance for 13 different ML algorithms.

Comparison and evaluation of dosing schemes

Poisson regression LT4 dosing correctly dosed 64.8% of patients, proving more accurate than the best existing dosing scheme in the literature (proposed by Ojomo et al) that correctly predicted 60.9% of doses ($P = .031$). Standard weight-based dosing (1.6 μg/kg/day) correctly dosed 51.3% of patients, and the least effective dosing scheme proposed in the literature (proposed by Cunningham et al) correctly dosed 27.4% of patients.

Poisson dosing provided the least rate of extreme over- or under-replacement (errors greater than 25 μg) at 19.1%, and the least rate of any over-replacement at 19.5%. Of all schemes performing better than standard weight-based dosing, Poisson dosing resulted in the least rate of under-replacement at 15.7%. The comparison between Poisson dosing and existing methods is shown in Table 4.

Because we previously described the difficulties in achieving euthyroidism for patients at the extremes of BMI, we further evaluated dosing schemes by BMI subsets (tertiles). Table 5 describes the accuracy of each scheme within each BMI tertile. Poisson dosing outperforms all existing methods within each BMI group. Notably, standard weight-based dosing dosed 64.1% of the lesser BMI patients correctly, but only 38.0% of the patients with greater BMI; the 26.1% difference across BMI extremes is the greatest of any method.

Table 5
Accuracy of existing LT4 prediction schemes and a novel Poisson dosing algorithm by BMI tertile subsets

Scheme	*BMI tertile*		
	≤26	27–32	>32
1.6 μg/kg/day	64.1%	52.8%	38.0%
Cunningham et al	35.9%	19.3%	25.9%
Olubowale et al	47.1%	45.5%	38.4%
Mistry et al	41.7%	39.2%	39.4%
Ojomo et al	72.3%	55.1%	54.6%
Jin et al	57.3%	55.7%	47.2%
Di Donna et al	70.4%	60.2%	50.0%
Elfenbein et al	68.4%	53.4%	57.9%
Poisson dosing	73.3%	63.6%	59.7%

$$\begin{aligned} \text{Daily LT4 Dose } (\mu g) &= e^X, \\ X &= 2.02 + 0.01(W) - 0.0037(A) \\ &\quad - 0.098(F) - 0.01(B) \\ &\quad + 0.007(T) + 0.108(I) - 0.014(M) \end{aligned}$$

Fig 2. Poisson regression formula. *W*, patient weight (kg); *A*, patient age (years); *F*, patient sex (1 for female, 0 for male); *B*, patient BMI; *T*, preoperative TSH value; *I*, iron supplementation (1 for supplementation, 0 otherwise); *M*, multivitamin/mineral supplementation (1 for supplementation, 0 otherwise).

In addition, we observed how Poisson dosing performed in other patient subsets: Poisson dosing had an accuracy of 70.1% for patients with hyperthyroidism, 62.2% for patients with thyroid nodule disease, and 60.0% for patients with Hashimoto's disease. Figure 2 presents the Poisson regression formula for predicting daily LT4 requirements.

Discussion

Seven LT4 dosing schemes after thyroidectomy were identified in the literature, each using some combination of patient body weight, age, sex, or BMI. The Poisson regression LT4 dosing algorithm we developed correctly predicted significantly more doses compared to all other schemes proposed in the literature. In addition, the algorithm was more accurate across patient BMI levels and made fewer large dosing errors than existing schemes. This comparison is the second and most extensive review of LT4 dosing schemes.

Previously, Di Donna et al[13] compared several LT4 dosing schemes using a retrospective cohort of 92 patients. Our comparison extends on Di Donna's study by examining a comprehensive collection of LT4 schemes on a substantially greater retrospective cohort of 598 patients. Our comparison agrees with Di Donna's in finding the scheme proposed by Ojomo et al to be the most accurate of existing schemes, and in finding the scheme proposed by

Table 4
Efficacy of existing LT4 prediction schemes and a novel Poisson dosing algorithm

Scheme	*% Euthyroid doses*	*% dose errors >25 μg*	*% Hypothyroid*	*% Hyperthyroid*
1.6 μg/kg/day	51.3%	26.9%	16.4%	32.3%
Cunningham et al	27.4%	46.5%	7.0%	65.6%
Olubowale et al	43.5%	25.9%	15.7%	40.8%
Mistry et al	40.1%	40.0%	9.2%	50.7%
Ojomo et al	60.9%	23.1%	16.1%	23.1%
Jin et al	53.2%	25.1%	24.4%	22.4%
Di Donna et al	60.0%	21.6%	18.4%	21.6%
Elfenbein et al	60.2%	22.1%	17.4%	22.4%
Poisson dosing	64.8%	19.1%	15.7%	19.5%

Mistry et al to be less accurate than standard weight-based dosing. As detailed here, the Poisson regression method outperforms the scheme proposed by Ojomo et al.

The 7 identified schemes predicted LT4 requirements using 3 or fewer explanatory factors. Our ML approach resulted in a predictive LT4 algorithm using the 4 factors found in proposed schemes (body weight, age, sex, and BMI), and using 3 additional factors: preoperative TSH, iron supplementation, and vitamin-mineral supplementation. The influence of the supplemental factors is consistent with recent research demonstrating iron and various vitamins interfere with LT4 therapy.[22,23] Although each ML algorithm used these same 7 predictors, their accuracy varied substantially, reflecting the importance of their varying assumptions. Poisson regression assumes a model most aligned with LT4 estimation, and is correspondingly the most accurate ML algorithm.

Two of the most predictive schemes of those in the literature were proposed by Ojomo et al and Elfenbein et al. These 2 schemes represent prior iterations of progressive inquiry at our institution and were developed with an earlier version of the dataset used in this study. The Ojomo scheme was developed first and estimated LT4 needs using linear regression with BMI and weight as predictors.[5] After prospectively evaluating this scheme, differences in LT4 needs were observed between men and women, and a new scheme was developed by Elfenbein et al[14] also using linear regression to adjust for these differences according to sex. Yet, as noted here, a significant portion of incorrect dosing persisted, and we therefore hypothesized that more advanced ML techniques could improve the accuracy of dosing prediction. In addition, we collected all potential variables that could affect dosing accuracy, such as medication and supplement use, labs, and postoperative parathyroid function. Accordingly, this current study expanded these earlier schemes and used ML to develop an even more accurate scheme that adjusts for additional variables influencing LT4 requirements.

To promote provider use of LT4 dosing with Poisson regression, we developed an easy-to-use web application allowing users to input patient characteristics to estimate LT4 needs. A snapshot of this application is shown in Fig 3. Going further, our dosing algorithm could be implemented within electronic health systems as either an easy-to-use clinical decision-support tool or a means of an automatic calculation using variables within the electronic health record. These implementations would make it convenient for providers to prescribe LT4 according to our proposed algorithm, thereby decreasing time to euthyroidism after thyroidectomy.

Our study had important limitations. First, the proposed algorithm has not been tested prospectively. Although our algorithm was developed using cross-validation and a substantially larger dataset than other authors, we have yet to observe how our proposed algorithm performs in clinical practice. Second, the dataset came from a single institution, subjecting our results to the uncontrollable biases inherent in our patient population, thereby potentially limiting the generalizability of our algorithm. Our results, however, were similar to those of the comparison conducted by Di Donna et al,[13] a study using a cohort from another country; this observation suggests that LT4 dosing requirements are robust to variant populations and adds credibility to the generalizability of our results. Third, explanatory variables of our dataset were limited to those readily collected during clinic. Research has shown associations between LT4 needs and LBM measured by dual-energy x-ray absorptiometry, as well as genetic factors and TSH serum levels in patients undergoing LT4 therapy.[20,21,24] Absence of these predictors in our model development leaves space for future algorithm improvement, but these predictors require more invasive or expensive testing, and our goal was to develop a method with readily obtainable data. Our starting dataset included over 40 predictive variables to represent a robust set of potential factors that could affect thyroid hormone dosing. Because 36% of patients still required dose adjustment, other nonmeasured factors clearly affect LT4 dosing. Such factors might include patient compliance and individual metabolism of LT4. To address these limitations, we intend to pursue a prospective application of our algorithm in a multisite clinical trial to determine its predictive ability in practice.

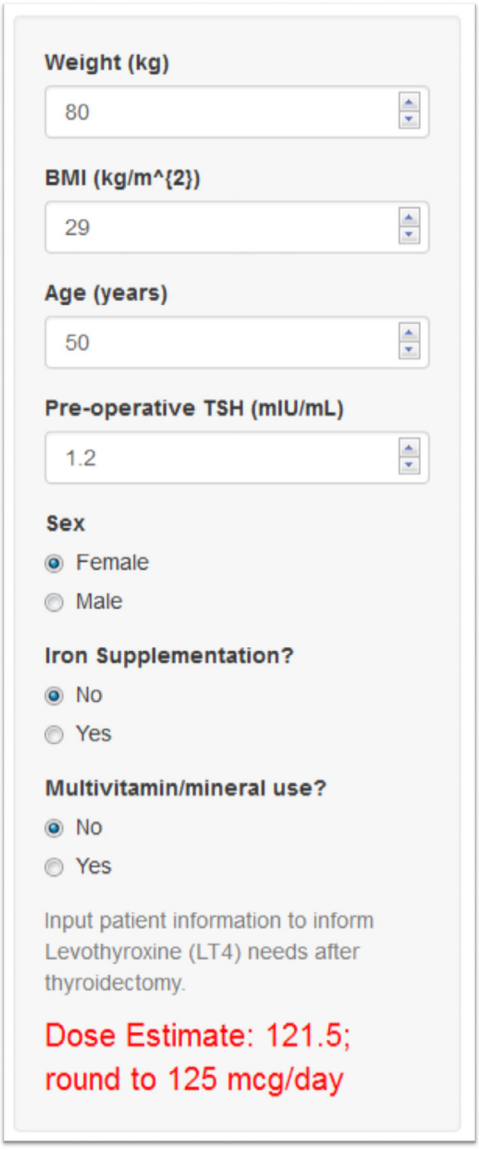

Fig 3. Web application to estimate daily LT4 requirements after thyroidectomy using Poisson regression. The URL for this application is https://zaborek-uwsurgery.shinyapps.io/ShinyDosing/.

In conclusion, using readily available variables, our novel Poisson regression dosing scheme outperforms other ML algorithms and all existing dosing schemes in calculating the appropriate dose of LT4 after thyroidectomy. Defining a correct prediction as a dose within 12.5 μg of the euthyroid dose, Poisson regression yielded significantly better dose estimates, including better estimation across patient BMI levels. Use of Poisson dosing by providers to calculate LT4 needs, either within electronic medical systems or with online calculators, could potentially decrease morbidity associated with LT4 replacement after thyroidectomy.

The authors thank Mary Hitchcock for her assistance with the scoping review.

References

1. Aitken M, Kleinrock M. *Medicines use and spending in the U.S.*. IQVIA Institute for Human Data Science; 2017.
2. Fish LH, Schwartz HL, Cavanaugh J, Steffes MW, Bantle JP, Oppenheimer JH. Replacement dose, metabolism, and bioavailability of levothyroxine in the treatment of hypothyroidism. Role of triiodothyronine in pituitary feedback in humans. *N Engl J Med*. 1987;316:764–770.
3. Olubowale O, Chadwick DR. Optimization of thyroxine replacement therapy after total or near-total thyroidectomy for benign thyroid disease. *Br J Surg*. 2006;93:57–60.
4. Mistry D, Atkin S, Atkinson H, Gunasekaran S, Sylvester D, Rigby AS, et al. Predicting thyroxine requirements following total thyroidectomy. *Clin Endocrinol (Oxf)*. 2011;74:384–387.
5. Ojomo KA, Schneider DF, Reiher AE, Lai N, Schaefer S, Chen H, et al. Using body mass index to predict optimal thyroid dosing after thyroidectomy. *J Am Coll Surg*. 2013;216:454–460.
6. Biondi B, Fazio S, Carella C, Amato G, Cittadini A, Lupoli G, et al. Cardiac effects of long term thyrotropin-suppressive therapy with levothyroxine. *J Clin Endocrinol Metab*. 1993;77:334–338.
7. Sawin CT, Geller A, Wolf PA, Belanger AJ, Baker E, Bacharach P, et al. Low serum thyrotropin concentrations as a risk factor for atrial fibrillation in older persons. *N Engl J Med*. 1994;331:1249–1252.
8. Bauer DC, Ettinger B, Nevitt MC, Stone KLStudy of Osteoporotic Fractures Research Group. Risk for fracture in women with low serum levels of thyroid-stimulating hormone. *Ann Intern Med*. 2001;134:561–568.
9. Uzzan B, Campos J, Cucherat M, Nony P, Boissel JP, Perret GY. Effects on bone mass of long term treatment with thyroid hormones: a meta-analysis. *J Clin Endocrinol Metab*. 1996;81:4278–4289.
10. Vigário PoS, Vaisman F, Coeli CM, Ward L, Graf H, Carvalho G, et al. Inadequate levothyroxine replacement for primary hypothyroidism is associated with poor health-related quality of life–a Brazilian multicentre study. *Endocrine*. 2013;44:434–440.
11. Jin J, Allemang MT, McHenry CR. Levothyroxine replacement dosage determination after thyroidectomy. *Am J Surg*. 2013;205:360–364.
12. Cunningham JJ, Barzel US. Lean body mass is a predictor of the daily requirement for thyroid hormone in older men and women. *J Am Geriatr Soc*. 1984;32:204–207.
13. Di Donna V, Santoro MG, de Waure C, Ricciato MP, Paragliola RM, Pontecorvi A, et al. A new strategy to estimate levothyroxine requirement after total thyroidectomy for benign thyroid disease. *Thyroid*. 2014;24:1759–1764.
14. Elfenbein DM, Ojomo KA, Schaefer S, Shumway C, Chen H, Sippel RS, et al. Prospective intervention of a novel levothyroxine dosing protocol based on body mass index after thyroidectomy. *J Am Coll Surg*. 2014;219:S125.
15. Somnay YR, Craven M, McCoy KL, Carty SE, Wang TS, Greenberg CC, et al. Improving diagnostic recognition of primary hyperparathyroidism with machine learning. *Surgery*. 2017;161:1113–1121.
16. Burnside ES, Rubin DL, Fine JP, Shachter RD, Sisney GA, Leung WK. Bayesian network to predict breast cancer risk of mammographic microcalcifications and reduce number of benign biopsy results: initial experience. *Radiology*. 2006;240:666–673.
17. Bouckaert RR. Choosing between two learning algorithms based on calibrated tests. *Proceedings of the Twentieth International Conference on International Conference on Machine Learning* AAAI Press; 2003:51–58.
18. . *R Foundation for Statistical Computing*. R: A language and environment for statistical computing 3.4.2 ed; 2017.
19. Kuhn M. *R Foundation for Statistical Computing*. caret: Classification and regression training R package version 6.0-77 ed.; 2017.
20. Sukumar R, Agarwal A, Gupta S, Mishra A, Agarwal G, Verma AK, et al. Prediction of LT4 replacement dose to achieve euthyroidism in subjects undergoing total thyroidectomy for benign thyroid disorders. *World J Surg*. 2010;34:527–531.
21. Banovac K, Carrington SAB, Levis S, Fill MD, Bilsker MS. Determination of replacement and suppressive doses of thyroxine. *J Int Med Res*. 1990;18:210–218.
22. Irving SA, Vadiveloo T, Leese GP. Drugs that interact with levothyroxine: an observational study from the Thyroid Epidemiology, Audit and Research Study (TEARS). *Clin Endocrinol (Oxf)*. 2015;82:136–141.
23. Jubiz W, Ramirez M. Effect of vitamin C on the absorption of levothyroxine in patients with hypothyroidism and gastritis. *J Clin Endocrinol Metab*. 2014;99:E1031–E1034.
24. Brigante G, Spaggiari G, Santi D, Cioni K, Gnarini V, Diazzi C, et al. The TRHR gene is associated with hypothalamo-pituitary sensitivity to levothyroxine. *Eur Thyroid J*. 2014;3:101–108.

Discussion

Dr Ashok R Shaha (New York, NY): Excellent study. Very well presented.

One of the most difficult problems with total thyroidectomy is handling the patients' thyroid medication and their symptoms. Many of these patients will keep calling almost every week. In young people, I add about 25 mcg more.

The question I have for you is at what point after total thyroidectomy we should do the TSH? Should it be based on time or based on the patient's symptoms?

And even though we all shoot for a normal level, what is the ideal TSH level that will correspond to the best symptom relief for the patients?

Nick A Zaborek: I know we reach out to our patients very consistently at 6 to 8 weeks to see how they are doing.

We define euthyroidism as TSH levels within the normal range, but even if TSH levels are within that range, a patient can still experience symptoms.

So I'm not sure what should be the exact definition of a euthyroid state, and I think that's better answered by Dr. Schneider, the clinician.

Dr David Schneider: For this study, we used just the normal TSH, and as you all know, the range is fairly wide. There's often further dose adjustments that occur even within the normal range until that patient feels normal. But for the purposes of having a clear outcome for this study, we just used normal TSH.

Dr Jacob Moalem (Rochester, NY): My question actually follows Dr. Shaha's.

Levothyroxine replacement has 2 different goals depending on what the indication for thyroidectomy was. For example, in thyroid cancer cases, the dose is also meant to be suppressive. So how can you have 1 formula or 1 way in which it's prescribed when you are actually trying to reach 2 separate targets?

Dr Nick A Zaborek: Actually, to be clear, this is specifically for the benign thyroidectomy indication.

Dr Barbra Miller (Ann Arbor, MI): If you are using TSH as 1 of the 7 inputs for your system, what happens if that patient is hypothyroid or hyperthyroid preop? How does that affect things?

Then in terms of your target for these patients, what about the patients with cardiac disease or who are elderly and you want to start them out lower and then gradually go up so they don't go into atrial fibrillation?

Dr Nick A Zaborek: To answer your second question first, I think when a patient has other comorbidities, that should be considered on an individual basis. This is meant for a generic patient, if you will.

To answer your first question, preoperative TSH levels are just informative in predicting LT4 requirements. Clinically speaking, I'm not sure if it works perfectly for hyperthyroid patients, but, nonetheless, it is still informative in predicting the dose.

Dr Akira Miyauchi (Kobe, Japan): Your study is very nice. My question for you is whether you think that more than the TSH level is important when patients are taking levothyroxine? How about

T4 or T3? We think normal TSH does not necessarily mean a euthyroid condition.

Dr Nick A Zaborek: That's a very good point that we have already discussed. I think measuring the euthyroid state is a complicated thing, and I am certainly not the best person to answer that. Perhaps Dr. Schneider might be able to address that.

Dr David Schneider: For the purposes of this study, we just consider euthyroidism as defined by TSH, but your point is a valid one. There are probably other factors that go into it, and maybe this is oversimplified. Regardless, there's a huge problem with people achieving a euthyroid state no matter how you define it. So we think this is a step forward that can help a lot of people get closer to euthyroidism a lot faster.

Dr Dina Elaraj (Chicago, IL): Thank you for that fascinating and timely study as machine learning and artificial intelligence seem to be in the news every day.

My question is regarding other factors that affect the time to achieve euthyroidism like patient compliance, timing of eating their first meal, timing of taking other medications in relation to taking their thyroid hormone, etc. How do you account for some of those factors when you are trying to choose a dose?

Dr Nick A Zaborek: That's a great point. These things are completely uncontrollable in the postoperative setting and we couldn't really measure those things at 6 to 8 weeks. We could ask the patient, but what they say might not be very accurate. So this was kind of meant as a catchall or a generic approach.

Surgery is abstracted and/or indexed in *Index Medicus, Science Citation Index, Current Contents/Clinical Medicine, Current Contents/Life Sciences,* and MEDLINE.