MEDINFO 2025 — Healthcare Smart × Medicine Deep
M.S. Househ et al. (Eds.)

doi:10.3233/SHTI251294

Development and Validation of a Machine Learning-Based Model for Methimazole Dosage Adjustment in Children and Adolescents with Hyperthyroidism

Kanghyuck Lee[a,b], Joon Young Kim[c], Taehoon KO[a,b,1], Kyungchul Song[c,1]
[a] *Department of Biomedicine & Health Sciences, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea;*
[b] *Department of Medical Informatics, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea*
[c] *Department of Pediatrics, Yonsei University College of Medicine,211 Eonju-ro,Gangnam-Gu, Seoul, Republic of Korea*
ORCiD ID: Kanghyck Lee https://orcid.org/0009-0006-4257-4776
Joon Young Kim https://orcid.org/0000-0001-9595-3610
Taehoon Ko https://orcid.org/0000-0002-4045-0036
Kyungchul Song https://orcid.org/0000-0002-8497-5934

Abstract. This study developed a machine learning model using data from 142 children and adolescents with hyperthyroidism, with external validation conducted on 63 patients from another institution. Input variables included age, sex, height SDS, weight SDS, BMI SDS, T3, free T4, TSH, follow-up interval, and previous methimazole dose. SHAP analysis identified T3, TSH, and free T4 as the most influential factors. The model demonstrated robust performance with an RMSE of 5.15 mg (internal validation) and 3.54 mg (external validation), highlighting the potential of machine learning to optimize methimazole dose adjustment in pediatric hyperthyroidism.

Keywords. Hyperthyroidism, Methimazole, Machine learning

1. Introduction

Assessment of pediatric hyperthyroidism requires precise dosage adjustment of methimazole to avoid adverse effects and achieve optimal therapeutic outcomes. Traditional approaches often rely on clinician experience, which can be subjective. This study aimed to develop and validate a machine learning-based model to predict methimazole dosage in children and adolescents.

[1] Corresponding Author: Taehoon Ko, E-mail: thko@catholic.ac.kr.
Kungchul Song, E-mail: ENDOSONG@yuhs.ac

2. Methods

This retrospective study included 142 children and adolescents treated with methimazole for hyperthyroidism at a pediatric clinic between July 2005 and February 2024 as the training set. The test set consisted of 63 children and adolescents treated with methimazole for hyperthyroidism at another pediatric clinic between January 2020 and February 2024. Input variables included age, sex, height standard deviation score (SDS), Weight SDS, body mass index (BMI) SDS, T3, free T4, thyroid stimulating hormone (TSH), follow-up interval, and the previous dose of methimazole. An eXtremely Gradient Boosting(XGBoost) algorithm was employed to develop the prediction model. Internal validation was performed using 5-fold cross-validation and external validation was performed using the test set. SHapley Additive exPlanations (SHAP) analysis was conducted to determine the contribution of each variable to the model's predictions.

3. Results

The mean methimazole dosage was 10.39±11.92 mg in the training set and 9.28±8.3 mg in the test set, respectively. Root mean square error (RMSE) was 5.15±0.45 mg and mean absolute error (MAE) was 2.48±0.32 mg in internal validation, repectively. In external validation, RMSE was 3.54±0.73 mg and MAE was 2.11±0.73 mg. In SHAP analysis, T3, TSH, and free T4 were the most important contributor for the model, which were followed by BMI SDS, age, and weight SDS.

4. Conclusion

We developed a machine learning-based model for dosage adjustment of methimazole in children and adolescents with hyperthyroidism and performed external validation using an independent data set. This result suggests potential role of machine learning in the treatment of hyperthyroidism in children and adolescents.

References

[1] Chen, S. S., Zaborek, N. A., Doubleday, A. R., Schaefer, S. C., Long, K. L., Pitt, S. C., ... & Schneider, D. F. (2019). Optimizing levothyroxine dose adjustment after thyroidectomy with a decision tree. *Journal of Surgical Research*, *244*, 102-106.

[2] Kizilgul, M., Yigit, T., Ata, N., Celik, A., Ucan, B., Sahin, M., ... & Birinci, S. (2024). Machine learning prediction of pancreatitis risk with antithyroid drugs: a nationwide retrospective observational study. *The Journal of Clinical Endocrinology & Metabolism*, dgae628.

[3] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

[4] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

[5] Bleeker, S. E., Moll, H. A., Steyerberg, E. A., Donders, A. R. T., Derksen-Lubsen, G., Grobbee, D. E., & Moons, K. G. M. (2003). External validation is necessary in prediction research:: A clinical example. Journal of clinical epidemiology, 56(9), 826-832.