HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 4 (2024)

Analysis of Health Factors Influencing Well-Differentiated Thyroid Cancer Recurrence

Eashan Shetty[1] and Deepak Shetty[#]

[1]Lawrence E Elkins High School, USA
[#]Advisor

ABSTRACT

This study analyzed specific health factors that influenced the outcome of recurrence in well-differentiated thyroid cancer. The dataset we used to analyze the health factors were from 383 patients with reported clinicopathologic features that either experienced recurrence or not. Before performing any statistical technique on dataset, categorical variables were encoded to numerical values. A logistic regression model was then performed on each independent variable to demonstrate its direct impact on the chance of thyroid cancer recurrence. Key indicators that suggested positive influence of health factor on cancer recurrence was p value below .05, positive beta coefficient, and odds ratio above 1. The results from the logistic regression shown that age, male, smoking history, adenopathy, papillary, multifocal, risk, tumor size, nodal involvement, stage, and radiotherapy all increased odds of recurrence. On the other hand, being a female and a unifocal tumor decreased chances of recurrence.

Introductory

Thyroid cancer is a disease where tissues of the thyroid gland form malignant (cancerous) cells (National Cancer Institute, n.d). There are various types of thyroid cancer's such as well-differentiated thyroid cancer and medullary thyroid cancer (National Cancer Institute, n.d). Well-differentiated thyroid cancer includes tumors such as papillary thyroid cancer or follicular thyroid cancer (National Cancer Institute, n.d). When a cancer is well-differentiated, this means cancer cells acts like normal cells and regularly grows slowly. Well-differentiated thyroid cancer in addition account for majority of thyroid cancer types as papillary and follicular thyroid carcinoma make up 80-84% and 6-10% of all thyroid carcinomas (Paschke, 2015). Hurthle cell cancer, which is also a type of thyroid well-differentiated cancer, is a rare type of cancer that is known to be more aggressive and makes up less than 5% of well differentiated cancer (Fariduddin, 2023). Symptoms of well-differentiated thyroid cancer include difficulty swallowing, enlarged lymph nodes, and pain (John Hopkins Medicine, n.d). Treatment of this disease involves surgery to remove the thyroid gland and, in some cases, using radiotherapy or chemotherapy (John Hopkins Medicine, n.d). Even though survival rates are extremely high when thyroid cancer is caught in its early stages, there is still a risk of recurrence after remission.

When a cancer returns after remission, it is known as cancer recurrence (Mayo Clinic, n.d). The cause of recurrence in most cases is leftover cancerous cells that remained after best efforts to get rid of the cancer (Mayo Clinic, n.d). Recurrent cancer can be diagnosed through a blood test or imaging test such as a CT scan. In thyroid cancer, people who survive the disease will have a 20% risk of thyroid cancer recurrence (Farnsworth, 2023). Too add on, males and being over 45 years of age at diagnosis can have a higher risk of recurrence (Fansworth, 2023). In a study that analyzed the predictive factors for recurrence of papillary thyroid carcinoma of 4,085 patients, it showed that tumor size > 10mm, multifocality, and lymph node metastasis were independent risk factors of recurrence (Kohler, 2021). The gap in this analysis however is that it only took in account of PTC (papillary thyroid carcinoma) and no other carcinomas like follicular thyroid that make up 10% of all well-

differentiated thyroid cancer. In addition, it did not analyze treatment or lifestyle factors being a possible risk of recurrence such as radiotherapy treatment or smoking history.

The primary objective of this study is to investigate significance of association between health factors and recurrence of well-differentiated thyroid cancer. Through analyzing a dataset using statistical methods, we aim to identify possible factors influencing recurrence.

Methods

Dataset Details

This study is based on 383 patients that had differentiated thyroid cancer (Borzooei, 2023). The dataset was collection in a period of 15 years with each patient followed for 10 years aiming to predict recurrence (Borzooei, 2023). It contains clinical attributes, 22 which will be analyzed to find significant association between the attribute and occurrence of recurrence. The attributes analyzed and included in the dataset are age, gender, smoking, radiotherapy, Adenopathy, Pathology, Focality, Risk, tumor classification, nodal classification, metastasis classification, stage of cancer. Most attributes listed are categorical or binary variables, as age is the only discrete variable. Variables like risk is divided into 3 options, low, intermediate, and high. On the other hand, tumor classification, nodal classification, and metastasis classification is based on the TNM staging of thyroid tumors. This dataset was approved by UC Irvine Machine Learning Repository and was publicly available.

Statistical Encoding and Technique

Too ensure a successful analysis on the dataset, it had to be encoded to replace categorical values with numerical values. Depending on the categorical variables, we had to use two different techniques on encoding it. One-Hot Encoding was used on nominal variables such as gender which has two or more options (female or male). Ordinal Encoding was used on ordinal variables that had a clear ordering like stages (I, II, III). These methods assigned numerical values to whole dataset allowing to perform independent logistic regression analyses on each variable. Logistic regression is a statistical method to demonstrate the relation between a binary dependent variable, in this case the recurrence event, and one or more independent variables. In this study we used, we used a separate logistic regression model for each predictor variable and was modeled in the following form:

$$\textit{Recurrence}\ (p) = \beta_0 + \beta_1 X_1 \quad (1)$$

In this single variable logistic regression model, *Recurrence* (p) is the probability of recurrence, β_0 is the intercept, β_1 is the coefficient of the independent variable, and X_1 is the independent variable being analyzed. Too determine significance of the independent variable on our dependent variable, I used three indicators for each independent logistic regression model. The B coefficient in the regression model is the change in log-odds of recurrence and is the slope. It represents how much it affects the dependent variable for a one-unit increase. A positive coefficient indicates an increased chance of recurrence, while negative coefficient can demonstrate a decreased chance. The p-value will test the null hypothesis that the coefficient of the variable has no effect on the dependent variable. A p value less than .05 indicates the independent variable is statistically significant on the dependent variable. Finally, Odds Ratio is calculated by exponentiating the beta coefficient. It measures how odds of recurrence changes with a unit increase and a odds ratio greater than 1 means increased odds.

Results

Table 1. Independent Logistic Regression Results

Variable	Beta Coefficient	P-value	Odds Ratio
Age	0.03734805203	0.0000009818118959	1.038054255
Male	1.686050582	0.0000000013943776	5.398119122
Smoking	1.963258787	0.000000003228849	7.1225
Euthyroid	0.540633337	0.1468407231	1.717094017
Hyperthyroidism	-0.8356018206	0.1894795841	0.4336134454
Hypothyroidism	-0.6931471806	0.3760704138	0.5
Subclinical Hyper	-0.09982799809	0.8516565296	0.9049930652
Adenopathy	1.703192882	0.0000001595334732	5.491452991
micropapillary	-21.73290024	0.9973921389	0.0000000003643522
Papillary	0.6829430104	0.01884841005	1.979695431
Follicular	0.7047977978	0.07817349904	2.0234375
Hurthle Cell	0.09224973365	0.854120004	1.096638655
Unifocal	-1.763588592	0	0.1714285714
Multifocal	1.763588592	0	5.833333333
Risk	3.590401623	0	36.24863129
T	1.409731932	0	4.094857555
N	1.687833786	0	5.407753654
M	22.84131457	0.9975562283	8314897083
Stage	2.719665009	0.0000000006189246	15.17523782
Radiotherapy	2.77991476233185	0.0104977179827166	16.1176470588234
Female	-1.6860505822028	0.00000000139437761337291	0.185249709639952

The single variable logistic regression results are shown in Table 1 with key indicators such as beta coefficient, p value, and odds ratio. Each variable was independently modeled to analyze its influence on recurrence of well-differentiated thyroid recurrence.

Statistically significant values according to their p values are age, male, smoking adenopathy, papillary, unifocal, multifocal, risk, T (Tumor size), N (Node), stage, radiotherapy, and female. This is because their p values fall below .05. On the other hand, variables that are not statistically significant were euthyroid, hyperthyroidism, hypothyroidism, subclinical hyperthyroidism, micropapillary, follicular, hurthle cell, and M (Metastaisis). These variables had no significant value due to their p values being above .05.

Discussion and Conclusion

Out of statistically significant variables, variables such as age, male, smoking history, adenopathy, papillary, multifocal, risk, tumor size, nodal involvement, stage, radiotherapy all increased the chance of recurrence in well-differentiated thyroid cancer. Age had a beta coefficient of .037 meaning that with a unit increase in age, chances of recurrence increase by 3.8%. Being a male in comparison of being a female increased your odds of recurrence at a high extent with a odds ratio of 5.398 meaning a 430% increase in chances. Being a smoker was one of the health factors that influenced thyroid cancer recurrence the most by increasing yours odd by 612% in comparison if you did not smoke. Adenopathy or enlarged lymph nodes had a positive beta coefficient of 1.703 and a high odds ratio of 5.491 indicating an influence of thyroid recurrence. The tumor type such as papillary carcinoma had significant increase on odds of recurrence with a odds ratio of 1.980. If multiple tumors originated from a primary tumor or multifocality, then your odds increase by 400% compared to being unifocal with a high odds ratio of 5.833. Other tumor characteristic such as size also influenced recurrence with a positive beta coefficient of 1.410 and nodal involvement of the cancer will increase the likelihood of recurrence by nearly 400%. Finally, radiotherapy history significantly increases your odds as it has a odds ratio of 16.118, one of the highest out of all variables analyzed. Variables like risk and stage influenced your chances of recurrence the most with odd ratios over 10. Although, there were some variables that were significant but decreased odds of recurrence. Being a female has a negative beta coefficient of -1.686 and decreased your odds of 81.5%. Additionally, a unifocal tumor had a similar negative beta coefficient of -1.764 and decreased odds of recurrence by 83% in comparison of a multifocal tumor. In the variables analyzed, there were more that had odd ratio's above 1 like euthyroid or follicular carcinoma, but they were not statistically significant and could not show association to recurrence.

In the end, the logistic regression model analyzed each variable independently and provided multiple health factors that impacted the outcome of well-differentiated thyroid cancer recurrence. Variable that increased odds of thyroid recurrence were age, male, smoking history, adenopathy, papillary, multifocal, risk, tumor size, nodal involvement, stage, radiotherapy. Additionally, variables that decreased odds of thyroid recurrence were unifocality and female. Finally, variable's that had no statistical significance were euthyroid, hyperthyroidism, hypothyroidism, subclinical hyperthyroidism, micropapillary, follicular, hurthle cell, and M (Metastaisis). From this conclusion, workers in the medical field can better calculate risk of recurrence and know what to target to prevent thyroid recurrence in well-differentiated cancer. Through this study, professionals can gain a better understanding of health factors influencing thyroid recurrence.

Limitations

The limitation in this study is the use of independent logistic regressions for each variable to see a direct analysis on its effect of recurrence. Even though it shows its independent impact on recurrence, it does not account for possible interactions between variables that cause recurrence chances to increase. Research that concerns on this matter should perform a multivariable logistic regression with other cofounding factor to explore the interaction of variables on recurrence of well-differentiated thyroid cancer.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Thyroid cancer treatment. (2023, July 21). Cancer.gov. https://www.cancer.gov/types/thyroid/patient/thyroid-treatment-pdq#:~:text=Thyroid%20cancer%20is%20a%20disease,the%20risk%20of%20thyroid%20cancer.

Farnsworth, C. (2023, November 13). What to know about thyroid cancer recurrence. Medical News. https://www.medicalnewstoday.com/articles/thyroid-cancer-recurrence

When cancer returns: How to cope with cancer recurrence. (2023, November 30). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/cancer/in-depth/cancer/art-20044575

De Carvalho, A. Y., Kohler, H. F., Gomes, C. C., Vartanian, J. G., & Kowalski, L. P. (2021). Predictive factors for recurrence of papillary thyroid carcinoma: analysis of 4,085 patients. Acta Otorhinolaryngologica Italica, 41(3), 236–242. https://doi.org/10.14639/0392-100x-n1412

Thyroid cancer. (n.d.). Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/conditions-and-diseases/thyroidcancer#:~:text=Most%20thyroid%20cancers%20are%20asymptomatic,ultrasound%20or%20another%20imaging%20study.

Fariduddin, M. M., Mathias, P. M., & Syed, W. (2023, February 13). Hurthle cell thyroid carcinoma. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK568736/

Byrne, I., Kanaoka, Y., Pollack, N. E., Rhee, H. J., & Sommers, P. M. (2019). An Analysis of Airport Delays Across the United States, 2012-2018. Journal of Student Research, 8(2). https://doi.org/10.47611/jsr.v8i2.775

Paschke, R., Lincke, T., Müller, S. P., Kreissl, M. C., Dralle, H., & Fassnacht, M. (2015). The Treatment of Well-Differentiated Thyroid Carcinoma. Deutsches Ärzteblatt International. https://doi.org/10.3238/arztebl.2015.0452

Borzooei,Shiva and Tarokhian,Aidin. (2023). Differentiated Thyroid Cancer Recurrence. UCI Machine Learning Repository. https://doi.org/10.24432/C5632J.