Research Article

Factors Affecting Thyroid Cancer in Patients with Thyroid Nodules Using Logistic Regression in Interval Censored Data

Jamshid Yazdani-Charati,[1] Ozra Akha,[2] and Farideh Khosravi[3,*]

[1]Department of Biostatistics, Health Sciences Research Center, Faculty of Health, Mazandaran University of Medical Sciences, Sari, IR Iran
[2]Diabetes Research Center, Mazandaran University of Medical Sciences, Sari, IR Iran
[3]Department of Biostatistics, Occupational and Environmental Health Research Center, Shahroud University of Medical Sciences, Shahroud, IR Iran

[*]*Corresponding author*: Farideh Khosravi, Department of Biostatistics, Occupational and Environmental Health Research Center, Shahroud University of Medical Sciences, Shahroud, IR Iran. E-mail: faridehkhosravi50@yahoo.com

Abstract

Background: Thyroid nodules in the majority of cases are symptom-free; but, their clinical significance is due to being differentiated from malignancy. The outbreak of thyroid cancer (TC) is gradually soaring because of lifestyle turning into westernization in developing nations such as Iran. Thus, the current research has been performed pursuing the goal to identify the most critical risk factors in the thyroid cancerous individuals with thyroid nodules in northern Iran during a decade.

Methods: This study is of historical cohort type. A total of 33,530 files of the patients referring to the medical clinic through 2003 to 2013 have been examined and 1,817 thyroid nodule afflicted patients have been diagnosed and the data about these patients includes 70 recorded variables The data have been analyzed by logistic regression in interval Censored data at significance level 0.05 applying SPSS and Strata software.

Results: A total of 1 594 (87.7%) of the patients consist of women in the mean age bracket 41.07 ± 13.81. Also, the mean age of those with TC is 38.15 ± 13.52. Among the most significant factors identified in thyroid malignancies in the patients with thyroid nodules, we can note severe obesity, bilateral nodules, abnormal thyroid stimulating hormone (TSH) and T4 hormones, phosphorus, calcium, and thyroglobulin.

Conclusions: Regarding the incidence raise of TC, it is required to identify the malignancy risk factors. It is recommended to carry out some studies on the role of abnormal T4 hormone, phosphorus, calcium, and thyroglobulin in thyroid malignancies in the patients suffering from thyroid nodules.

Keywords: Thyroid Nodules, TSH Hormone, Bilateral Nodule, Iran

1. Background

Any sort of swelling, observable, and palpable lump in thyroid refers to thyroid nodule (1). Thyroid nodule is a common disease and in physical examination, it is discovered among around 3% to 7% adults (2, 3). Plethora of studies indicate this issue that half of the people in society, in case of being examined via sonography, will come up with thyroid nodule and this level increases by aging (4). These nodules may be single or multiple and with performance or without performance. The prevalence of these nodules is high in the iodine deficient regions among women and the elderly. The diameter of the majority of palpable nodules is more than 1 cm. However, the palpability of a nodule depends on its location inside the gland (superficial or deep), the patient's neck anatomy, and the examiner's adroitness (5).

Cases such as nodule growth, voice hoarseness, vocal cords paralysis, palpable neck lymph nodes, and nodule adhesion to the surrounding tissues suggest malignancy. In addition, the patients with the record of their whole body exposed to radiation due to bone marrow transplantation or the head and neck radiation exposure are at higher risk for malignancy (6). As the reports suggest, in 2014, totally 1,665,540 new cases of cancer infection and 585,720 cancer induced deaths occurred in the US. Also, TC affliction and its induced death rate have been reported as 62,980 and 1,890 individuals, respectively (7, 8).

Based on a review conducted in Italy, the outbreak of TC is rising worldwide and it has been recognized as the most prevalent endocrine cancer (1% - 1.5%) annually in the US (9, 10).

One of the most important specifications of survival analysis is censored data. The censoring of the survival studies is of different types. Interval censoring occurs when an event is not observable directly and happens in a

random time range. In fact, it can be asserted that a person in the variable T is an interval censored if only we know that a < T < b. The interval censored data often occurs in longitudinal studies and clinical trials, which requires more follow-up (11). The prevalence of interval censored data is rising in medical studies given the biomarkers increase in a disease development. Interval censoring results from periodically monitoring a disease progress. Logistic regression is of applied techniques for classified data analysis. In binary logistic regression, the response variable has 2 modes. One of the methods to estimate the parameters is error least squares, in which the goal is to find the regression coefficients in order to minimize the error least squares. Another approach is maximum likelihood pursuing the goal behind the model likelihood maximization is to address the model parameters (12). The logistic regression model equals:

$$\pi\left(x\right) = \frac{exp\left(\alpha + \beta x\right)}{1 + exp\left(\alpha + \beta x\right)} \tag{1}$$

Because of such patients frequently referring, doing check-ups, sonography, recording the data in the files, the possibility of the thyroid nodules turning into cancer and concerning the thyroid gland performance being so vital in human body, in metabolism process and its impairment significant effect on other organs, we have decided to study this issue.

2. Objectives

The current research has been performed pursuing the goal to identify the most critical risk factors in the thyroid cancerous individuals with thyroid nodules in northern Iran during a decade.

3. Methods

The current research is of historical cohort type and targets to detect the most important factors influencing thyroid cancer in the patients with thyroid nodules, referring to Endocrinology and Metabolism diseases clinic from July 2003 to mid-March 2013. The statistical community consists of the patients with glands or symptoms like pain in nodule area or suffocation, hoarseness, swelling in the neck referring to the endocrinologist, introduced by another physician or referring due to other reasons such as diabetes. The question event in the research is thyroid cancer affliction in thyroid nodule patients. Out of 33,530 files surveyed, 1,817 files of the thyroid nodule sufferers have been recorded including 169 patients with different TCs. The data about the patients cover 70 variable types. Due to

the missing data being massive in the study, expectation-maximization (EM) algorithm has been employed to compute the missing data. This study was approved by forth meeting of ethics committee of Research at Mazandaran University of Medical Sciences (July 8th, 2014).

When all the data are interval censored, one of the methods used is logistic regression. In data analysis with logistic regression, the time interval has been taken as 6 months and in case of the individual referring during the considered interval, 1 and otherwise 0 has been recorded. Under the following conditions, this is useful:

- Having the sufficient number of events in each range;
- The distribution determination f (t) not being important.

For instance, the logistic model with 3 indicator variables (defined for the intervals) and an explanatory variable is as it follows:

Logit p $(Y = 1) = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 X$ \qquad (2)

P (y = 1) equals the probability that the event occurs in a certain range provided that it did not occur in the prior intervals (13).

Applying univariate logistic regression, some meaningful variables have entered the final logistic regression model along with the variables having P < 0.3 and the final analyses have been performed. The patients' files have been examined fully confidentially and considering the obtained results in general, the patients' privacy has been appreciated. The data have been analyzed with SPSS and Stata software at significance level 0.05.
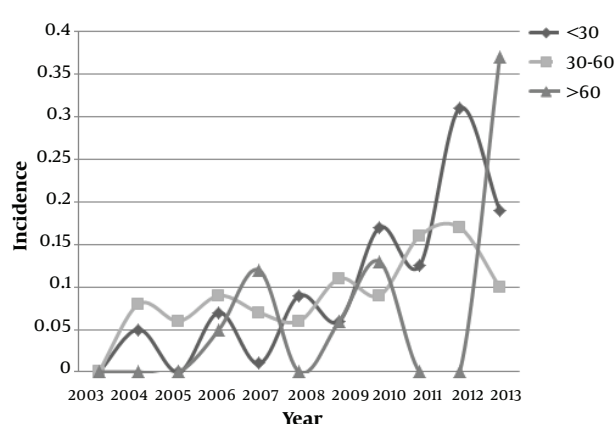
4. Results

The following tables include the descriptive data about some of the most principal variables existing in the study (Table 1).

Out of the cancer sufferers, 141 (83.4%) were women and 28% (16.6%) were men. The patients' age, when referring, has been recorded quantitatively. The patients' mean age has been yielded as 41.07 ± 13.81. Moreover, the TC suffering patients' mean age has been 38.15 ± 13.52. Their mean education years number has been 10.96 ± 4.25. The graph below illustrates annual TC incidence in thyroid nodule patients by age separation (Figure 1):

Twenty-two (22) 6-month time intervals have been defined in the research that in case of the patient referring at that time interval has been viewed as 1 and otherwise 0.These time intervals have been considered as independent variables along with other variables in the analysis. The results from the logistic regression model fitness have been given in the table below (merely the significant variables have been reported) (Table 2):

Table 1. Descriptive Statistics of Some Variables in Patients with Thyroid Nodules

Variable	No. (%)	Variable	No. (%)
Sex		**T3 hormone**	
Female	1594 (87.7)	Normal	464 (25.5)
Male	323 (12.3)	Abnormal	204 (11.2)
Age		**TSH hormone**	
< 30	448 (24.7)	Normal	1193 (65.7)
30 - 60	1202 (66.2)	Abnormal	492 (27.1)
> 60	160 (8.8)		
Marital status		**Hemoglobin**	
Single	1525 (84)	Normal	678 (37.3)
Married	153 (8.4)	Abnormal	454 (25)
Hypothyroidism		**Hb A1C**	
Yes	196 (10.8)	Normal	43 (2.4)
No	1621 (89.2)	Abnormal	71 (3.9)
Hyperthyroidism		**Vitamin D**	
Yes	93 (5.1)	Normal	193 (10.6)
No	1724 (94.9)	Abnormal	144 (7.9)
Diabetes		**Calcium**	
Yes	69 (3.8)	Normal	401 (22.1)
No	1748 (96.2)	Abnormal	34 (1.9)
Body Mass Iindex		**Phosphorus**	
< 18.5	26 (4.1)	Normal	366 (20.1)
18.5 - 25	117 (6.4)	Abnormal	18 (1)
25 - 30	174 (9.6)	**Thyroglobulin**	
30 - 35	136 (7.5)	Normal	49 (2.7)
> 35	10 (0.6)	Abnormal	34 (1.9)
Cholesterol		**Thyroglobulin antibodies**	
Normal	478 (26.3)	Normal	33 (1.8)
Abnormal	431 (23.7)	Abnormal	10 (0.6)
HDL		**Right lobe nodule**	
Normal	313 (17.2)	Yes	1130 (62.3)
Abnormal	150 (8.3)	No	687 (37.7)
LDL		**Left lobe nodule**	
Normal	320 (17.6)	Yes	1000 (55)
Abnormal	246 (13.5)	No	617 (45)
Triglyceride		**Bilateral nodules**	
Normal	664 (36.5)	Yes	618 (34)
Abnormal	194 (10.7)	No	1199 (66)
Fasting blood sugar		**Multi nodular goiter**	
Normal	983 (58.3)	Yes	1102 (60.6)
Abnormal	314 (17.3)	No	715 (39.4)
Blood pressure		**Sleeping disorder**	
Normal	748 (41.2)	Yes	86 (4.7)
Abnormal	279 (15.4)	No	1731 (95.3)
T4 hormone		**Sore throat and irritation**	
Normal	1486 (81.8)	Yes	105 (5.8)
Abnormal	164 (9)	No	1702 (94.2)

Figure 1. Annual TC Incidence in Thyroid Nodule Patients by Age Separation

According to the achieved results, the odds to get the cancer in the first half of 2005, if not happening in the prior intervals, equals 0.53. This odds is 0.45, 0.64, 1.4, 1.36, and 1 .45 in the second half of the years 2005, 2007, 2011, 2012, and 2013, respectively. Given the above tables, the odds ratio logarithm for the variables having bilateral nodules has been gained 2.38 so that it can be stated that the cancer infection odds in the bilateral thyroid nodules patients is 2.38 times than other sufferers. The cancer odds in the bilateral thyroid nodules with abnormal T4 hormone is 1.29 times than other sufferers .This odds has been obtained as 2.98, 4.96, 4.77, and 12.95 in the patients with abnormal TSH hormone, calcium, phosphorus, and thyroglobulin. Also, in the overweight individuals, based on the logistic regression model fitness, this odd has been gained as 1.34 times than the other patients. Other variables reported in the above table have odds ratio less than 1.

5. Discussion

In recent years, as diagnosing thyroid nodules increase, thyroid cancer incidence has been increasing across the world (14-16). The most common thyroid cancers are papillary and, then, follicular TCs that, as the studies indicate, include 85% to 90% and 10%to 15% of total TCs, respectively (17); they cover 94.1% and 3% of total cancers, respectively. As the studies in the US suggest, the raise of TCs outbreak in the US has been 6.6% in both genders from 1997 to 2009. In the present research, until 2006, the age bracket has been 30 to 60 years and in 2003, those over 60 years have had the highest TCs. According to a review paper by Pellegriti et al. in 2004 to 2008, in men, the highest TCs outbreak has been seen in the one over 65 years and in

Table 2. The Results From the Logistic Regression Model in Patients with Thyroid Nodules

Variable	Odds Ratio	Parent Statistics	S.D	Sig. Level
1nd half of 2005	0.53	5.98	0.225	0.014
2nd half of 2005	0.45	9.55	0.269	0.002
2nd half of 2007	0.64	5.07	0.194	0.024
2nd half of 2011	1.4	5.59	0.143	0.018
2nd half of 2012	1.36	4.51	0.144	0.034
2nd half of 2013	1.45	6.49	0.147	0.011
Hyperthyroidism	0.23	56.55	0.196	< 0.001
Hypothyroidism	0.13	163.02	0.161	< 0.001
Diabetes	0.26	26.62	0.267	< 0.001
Right lobe nodule	0.065	179.2	0.205	< 0.001
Left lobe nodule	0.24	96.93	0.146	< 0.001
Bilateral nodule	2.38	8.95	0.348	0.003
Multi Nodular goiter	0.095	52.05	0.326	< 0.001
Cholesterol	0.45	40.86	0.124	< 0.001
T4	1.29	4.49	0.12	0.034
T3	0.43	12.94	0.237	< 0.001
TSH	2.98	181.95	0.081	< 0.001
HbA1C	0.33	6.11	0.45	0.013
Calcium	4.96	48.87	0.229	< 0.001
Phosphorus	4.77	17.06	0.378	< 0.001
Vitamin D	0.41	14.2	0.237	< 0.001
Thyroglobulin	12.95	82.2	0.283	< 0.001
Sleep disorders	0.044	9.49	1.02	< 0.001
Sore throat	0.86	5.8	1.02	0.002
Severely obese people	1.34	11.79	0.086	0.016

women, in 55 to 64 years [10]. According to the results from the logistic regression finesse, it can be concluded that between 2011 and 2013, the patients have had the maximum TCs odds, which is true regarding the increasing trend of cancer thorough the world in recent years. TSH serves as an independent risk factor in predicting thyroid nodules malignancy [18]. Various experiments on animals suggest the raise of iodine deficiency induced TCs as TSH goes up.

In the current study, the patients with high TSH will come up with 2.98% times further TC odds. The results extracted from several prospective studies display the increased risk of TC in obese individuals, consistent with this research derived results [10, 19]. In the present study, the TCs odds in severely obese individuals can be 1.34 times than others. The ones with multi-nodular and single nodules goiters face the same risk in thyroid malignancy; on the other hand, the prevalence of malignancy may be lower in the individuals with multi-nodular goiter, compatible with his study results [20, 21]. Having bilateral nodule has been identified of other risk factors behind thyroid cancers affliction, in line with the current research [5]. According to the results of this study, the patients with bilateral nodules are exposed to 2.38 times more odds to get TCs than other ones.

Analyzing other variables has not been proved in TC malignancy so far. Of such variables gained as meaningful in the final fitted model and with the odds ratio higher than 1, we can mention calcium, phosphorus, thyroglobulin, and T4 hormone. It is hoped that through conducting more studies in this field in future, we access more fulfilling results on thyroid cancers risk factors in patients with thyroid nodules.

5.1. Conclusions

The incidence of TCs follows a rising trend; thus, in order to identify the risk factors in getting this disease, it is essential to identify the most critical factors. Of such factors identified in thyroid malignancy in thyroid nodules patients, severe obesity, bilateral nodules, abnormal TSH hormone, T4 hormone, calcium, phosphorus, and thyroglobulin can be pointed out.

Acknowledgments

We have acknowledged the committee of research at Mazandaran University of Medical Sciences for their supporting of this study.

Footnotes

Author's Contribution: None declared.

Conflict of Interests: There was no conflict of interests to be declared.

Financial disclosure: None declared.

References

1. Ghani H, Akbari Jami M. Diagnostic value of FNAc in Thyroidal nodules in Kashan from 1998 to 2000. *KAUMS J.* 2001;**5**(2):71–5.
2. Esteghamati AR, Yousefizadeh A, Rakhshan M, Mehrdad R. Diagnostic value of FNA in thyroid nodules. *IranJ Endocrinol Metab.* 2001;**3**(3):193–203.
3. Gharib H, Papini E, Valcavi R, Baskin HJ, Crescenzi A, Dottorini ME, et al. American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocr Pract.* 2006;**12**(1):63–102. doi: 10.4158/EP.12.1.63. [PubMed: 16596732].
4. Larijani B, Pajouhi M, Bastanhagh MH, Sadjadi A, Sedighi N, Eshraghian MR. Evaluation of suppressive therapy for cold thyroid nodules with levothyroxine: double-blind placebo-controlled clinical trial. *Endocr Pract.* 1999;**5**(5):251–6. doi: 10.4158/EP.5.5.251. [PubMed: 15251662].
5. Kasper D, Fauci A, Hauser S, Longo D, Jameson J, Loscalzo J. *Harrison's principles of internal medicine.* 2015.
6. Curtis RE, Rowlings PA, Deeg HJ, Shriner DA, Socie G, Travis LB, et al. Solid cancers after bone marrow transplantation. *N Engl J Med.* 1997;**336**(13):897–904. doi: 10.1056/NEJM199703273361301. [PubMed: 9070469].
7. Lee JI, Kim SH, Tan AH, Kim HK, Jang HW, Hur KY, et al. Decreased health-related quality of life in disease-free survivors of differentiated thyroid cancer in Korea. *Health Qual Life Outcomes.* 2010;**8**:101. doi: 10.1186/1477-7525-8-101. [PubMed: 20840792].
8. Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin.* 2014;**64**(2):104–17. doi: 10.3322/caac.21220. [PubMed: 24639052].
9. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. *Cancer incidence in five continents, Volume IX.* IARC Press, International Agency for Research on Cancer; 2007.
10. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. *J Cancer Epidemiol.* 2013;**2013**:965212. doi: 10.1155/2013/965212. [PubMed: 23737785].
11. Lim HJ, Sun J. Nonparametric Tests for Interval-Censored Failure Time Data. *Biometr J.* 2003;**45**(3):263–76. doi: 10.1002/bimj.200390011.
12. Bashiri MKRR. Parameter estimation for improving communication using binary logistic regression. *J Manufactur Operat.* 2011;**2**(1):135–54.
13. Mosure DJ, Berman S, Kleinbaum D, Halloran ME. Predictors of Chlamydia trachomatis infection among female adolescents: a longitudinal analysis. *Am J Epidemiol.* 1996;**144**(10):997–1003. [PubMed: 8916511].
14. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Waldron W, et al. *SEER Cancer Statistics Review, 1975–2008. National Cancer Institute; Bethesda, MD: 2011.* Bethesda: National Cancer Institute; 2011.
15. Hundahl SA, Fleming ID, Fremgen AM, Menck HR. A National Cancer Data Base report on 53,856 cases of thyroid carcinoma treated in the U.S., 1985-1995 [see commetns]. *Cancer.* 1998;**83**(12):2638–48. [PubMed: 9874472].
16. Leenhardt L, Grosclaude P, Cherie-Challine L, Thyroid Cancer C. Increased incidence of thyroid carcinoma in france: a true epidemic or thyroid nodule management effects? Report from the French Thyroid Cancer Committee. *Thyroid.* 2004;**14**(12):1056–60. doi: 10.1089/thy.2004.14.1056. [PubMed: 15650358].
17. Wang F, Wang Y, Wang L, Wang X, Sun C, Xing M, et al. Strong association of high urinary iodine with thyroid nodule and papillary thyroid cancer. *Tumour Biol.* 2014;**35**(11):11375–9. doi: 10.1007/s13277-014-2397-8. [PubMed: 25119588].
18. Zhang Y, Hua LEI, Huang J. A Spline-Based Semiparametric Maximum Likelihood Estimation Method for the Cox Model with Interval-Censored Data. *Scandinavia J Statistics.* 2010;**37**(2):338–54. doi: 10.1111/j.1467-9469.2009.00680.x.
19. Duick DS, Klopper JP, Diggans JC, Friedman L, Kennedy GC, Lanman RB, et al. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Thyroid.* 2012;**22**(10):996–1001. doi: 10.1089/thy.2012.0180. [PubMed: 22873825].
20. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2006;**16**(2):109–42. doi: 10.1089/thy.2006.16.109. [PubMed: 16420177].
21. Ravangard R, Arab M, Rashidian A, Akbarisari A, Zare A, Zeraati H. Comparison of the results of Cox proportional hazards model and parametric models in the study of length of stay in a tertiary teaching hospital in Tehran, Iran. *Acta Med Iran.* 2011;**49**(10):650–8. [PubMed: 22071639].