



**NAZARBAYEV UNIVERSITY SCHOOL OF MEDICINE
LONGITUDINAL SCHOLARLY PROJECT**

**Prevalence, Incidence and Mortality of Thyroid Cancer in
Kazakhstan: Data from Unified National Electronic Health System
2014-2021**

Student: Ramina Abdymananova

PI: Abduzhappar Gaipov

Co-PI: Kuralay Atageldiyeva

Date of submission: 01/03/2025

Abstract

Although a rare type of cancer overall, thyroid cancer is one of the most common endocrine malignancies. According to global studies, its prevalence has been increasing worldwide and continues to grow. Although there are global epidemiological studies on thyroid cancer that involved many regions, there is limited data on the epidemiology of thyroid cancer in Central Asian countries, including Kazakhstan.

This project was done to provide information on demographic characteristics of diagnosed patients, determine prevalence, incidence and mortality rates of thyroid cancer in Kazakhstan and conduct survival analysis for such patients. According to the data provided from the Unified National Electronic Health System, in total, 4,877 cases of thyroid cancer have been identified during the period from 2014 to 2021 in Kazakhstan. Based on the results, most of the diagnosed patients throughout the given period were females. The highest incidence and prevalence rates were found in the age group of 51-70 years old while the highest mortality rate was among patients older than 70. Increasing age and male sex were reported as the major predictors of mortality among thyroid cancer patients.

Introduction

Thyroid gland, located in the neck, is an organ of the endocrine system that produces such hormones as thyroxine (T4) and triiodothyronine (T3) and by their means regulates body metabolism (Hiller-Sturmhöfel and Bartke, 1998). Thyroid cancer is a malignancy of the thyroid gland that is considered rare among other types of cancer, however, is one of the most common endocrine malignancies (Khodamoradi et al., 2018).

According to many studies, the prevalence and incidence of thyroid cancer has been increasing worldwide. The study by Bao et al. (2021) involved population data from 204 countries during the period from 1990 to 2019 and concluded that there were more than 233,000 new cases of thyroid cancer in 2019. Pizzato et al. (2022) also reported that there was a global increase in thyroid cancer incidence rates but no major changes in mortality rates.

There have been several studies investigating thyroid cancer epidemiology in Kazakhstan with a focus on incidence rates. The study by Igissinov et al. (2019) has shown that there were 5,559 new cases of thyroid cancer in Kazakhstan during the period from 2009 to 2018. The analysis

has also revealed that the incidence was much higher in women than men, and peak incidence of thyroid cancer among Kazakhstani was at the age of 50-59 years in both sexes (Igissinov et al., 2019). Overall, the study by Igissinov leaves room for further research. For example, other patients characteristics such as ethnicity, country region, area of living, social status should be considered. It would be also reasonable to assess the geographical distribution of thyroid cancer cases across Kazakhstan as well as to compare the burden of thyroid cancer between the urban and rural population.

Although certain data on thyroid cancer incidence in Kazakhstan is available, further research that would include an extended descriptive analysis is needed. To achieve this, the Unified National Electronic Health System (UNEHS) that was established in Kazakhstan to collect and store electronic healthcare data can be used as a major source for the research. (Arynova and Baiguzhinova, 2019). As the database contains inpatient and outpatient registries of all patients within the period from 2014 to 2021, it will be used as a main data source in this project. The data will be used for descriptive analysis, and such characteristics as sex, age, ethnicity, area of residence, urban/rural region, social status will be included. Furthermore, crude and adjusted prevalence, incidence and mortality rates will be included into the analysis in the settings of this project. Finally, survival analysis for patients diagnosed with thyroid cancer will be conducted to assess the overall prognosis and identify mortality predictors.

Methods

Study design and population

This is a retrospective study that utilized the UNEHS data on thyroid cancer patients. Data management included careful cohort selection based on medical records of the patients diagnosed with thyroid cancer, more specifically based on International Classification of Diseases 10 (ICD-10). Patients with ICD-10 code C73 were defined to have thyroid cancer. Originally, 58,314 cases have been identified in the dataset as cases of thyroid cancer. Data cleaning was then performed to remove duplicate admissions based on population registry numbers (RPN ID). After data management, 4,877 unique cases have been used for further analysis. Data regarding the total population in the country including age and gender distribution was provided in the documents of the Statistics Committee under the Ministry of National Economy of the Republic of Kazakhstan.

Statistical analysis

Descriptive analysis of patients with thyroid cancer was performed based on such variables as age at diagnosis, gender, ethnicity, residence, mortality status, hospital admission and social status. Age at diagnosis was calculated by using the birth date and diagnosis date of the patient based on their RPN ID. After calculations, 5 age groups have been formed. Other variables were extracted directly from the information provided per each RPN ID. Ethnicity was divided into 3 groups of Kazakhs, Russians and Others, taking into account the frequency of each. Residence was defined as a patient living either in an urban or rural area. Hospital admission type was either planned, urgent or not specified. Patients were also classified as employed, unemployed, retiree, disabled and others. Additionally, for calculating the number of cases in each region of Kazakhstan, old regions distribution was used as it was used in the original UNEHS dataset. Incidence, prevalence and mortality rates were calculated per 100,000 population as it is a standard choice for statistical analysis of cancer cases.

Survival analysis for this project was conducted using Kaplan-Meier method of plotting the survivor function and cox regression modelling. Statistical significance of the difference was assessed via log-rank test. Cox regression analysis included 2 major models, one providing crude values for hazard ratios (HR) and another one being adjusted for such factors as age group, sex, living area and ethnicity group. P-value cutoff was set at 0.05 for survival analysis.

Statistical analysis was conducted on STATA Notebook as a part of the server running on JupyterLab for confidentiality reasons.

Results

Demographic characteristics

Overall, after data management and cleaning, there were 4,877 cases of thyroid cancer identified in Kazakhstan in the period of 2014-2021 based on the retrieved data. Table 1 demonstrates distribution of the given cases based on such characteristics as age at diagnosis, gender, ethnicity, residence, hospital admission and social status. It was identified that the mean age at diagnosis was around 60.5 years. Additionally, almost half (51.59%) of the thyroid cancer patients were 51-70 years old, while only 1 patient younger than 18 was found to have thyroid cancer.

Among all identified patients, the majority were female (86.75%). Most of the patients with thyroid cancer were Kazakh (57.47%), while slightly more than quarter of them were Russian (27.89%). Almost 15% of the patients were of a different ethnicity. 75% of the diagnosed patients were found to be living in an urban area, while the rest were residents of rural regions.

Although 28% of hospital admissions were unspecified, the vast majority of them were planned (69%). 2% of admissions were urgent and half of the urgently admitted patients were attributed as mortality cases.

Most of the patients described their social status as other (41%), while there were 21% of employed and 21% of retired patients in the cohort. 14% patients were unemployed at the moment of diagnosis and 1.25% were disabled.

Table 1. Socio-demographic characteristics of patients with thyroid cancer in the period of 2014-2021.

Demographic characteristics	Total n = 4,877 (100%)	Alive n = 4,473 (91.72%)	Dead n = 404 (8.28%)
Age, mean +-SD	60.5+-13.4	59.5+-13.1	72.4+-10.7
Age groups, n (%)			
<18	1 (0.02%)	1 (0.02%)	0 (0%)
18-34	494 (10.13%)	492 (11%)	2 (0.5%)
35-50	1,409 (28.89%)	1368 (30.58%)	41 (10.15%)
51-70	2,516 (51.59%)	2290 (51.20%)	226 (55.94%)
>70	457 (9.37%)	322 (7.20%)	135 (33.42%)
Gender, n (%)			
Female	4231 (86.75%)	3,951 (88.33%)	280 (69.31%)
Male	646 (13.25%)	522 (11.67%)	124 (30.69%)
Ethnicity, n (%)			
Kazakh	2,803 (57.47%)	2,573 (57.52%)	230 (56.93%)
Russian	1,360 (27.89%)	1,234 (27.59%)	126 (31.19%)
Other	714 (14.64%)	666 (14.89%)	48 (11.88%)
Residence, n (%)			
Rural	1,209 (24.79%)	1,084 (24.23%)	125 (30.94%)
Urban	3,668 (75.21%)	3,389 (75.77%)	279 (69.06%)
Hospital admission, n (%)			
Planned	3,372 (69.14%)	3,019 (67.49%)	353 (87.37%)
Urgent	101 (2.07%)	52 (1.16%)	49 (12.13%)
Unspecified	1,404 (28.79%)	1,402 (31.34%)	2 (0.5%)
Social status, n (%)			
Employed	1,035 (21.22%)	985 (22.02%)	50 (12.38%)
Unemployed	688 (14.11%)	646 (14.44%)	42 (10.40%)
Retiree	1,053 (21.59%)	812 (18.15%)	241 (59.65%)
Disabled	61 (1.25%)	52 (1.16%)	9 (2.23%)
Other	2040 (41.83%)	1978 (44.22%)	62 (15.35%)

Figure 1 represents all thyroid cancer cases identified between 2014 and 2021 distributed by Kazakhstani regions and counting 3 biggest cities as separate regions. Therefore, the vast

majority of the cases were identified in Almaty city (1,499 cases), followed by Karaganda (1,031 cases) and Eastern Kazakhstan region (843 cases). The least number of cases were found to be in the Atyrau (58 cases) and Mangystau (148 cases) regions.

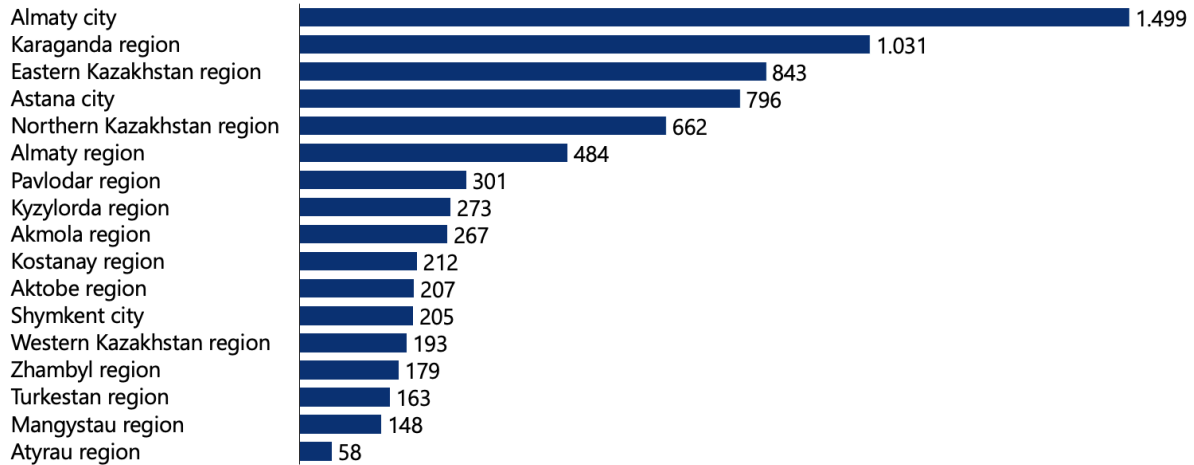


Figure 1. Total number of thyroid cancer cases in Kazakhstan in the period of 2014-2021 distributed by region.

Figures 2-4 represent incidence, prevalence and mortality rates of thyroid cancer based on gender distribution, respectively. It was demonstrated that the highest incidence rates were in 2016 and 2019, and both were favoring females. Prevalence rates were also favoring females during the whole period of 2014-2021 and were exhibiting an upward trend throughout the time. Mortality rate was the highest in 2019 but then started to go downward. Additionally, the mortality rate per 100,000 population was never found to be more than 1.

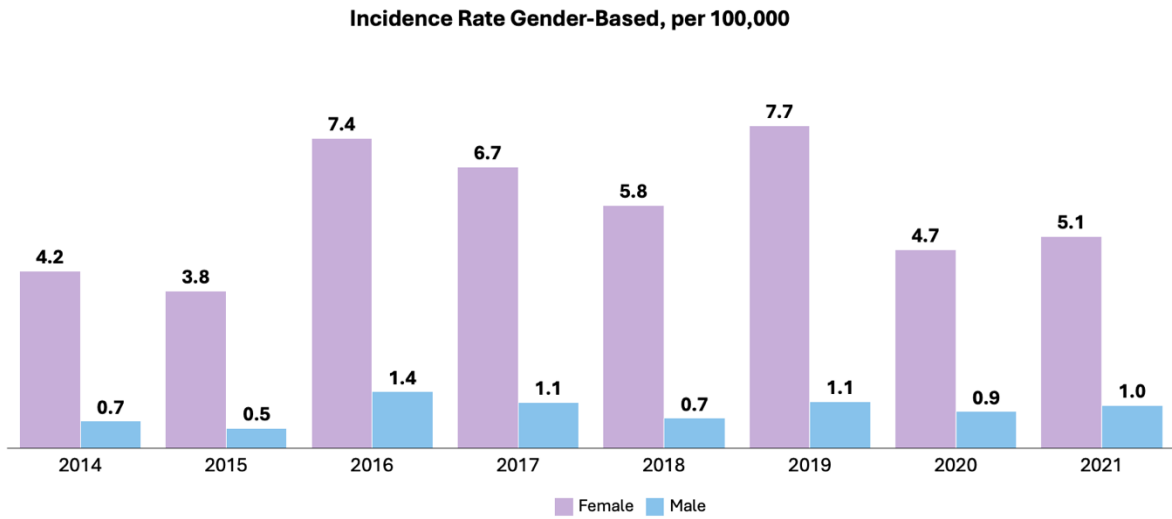


Figure 2. Incidence rate of thyroid cancer in Kazakhstan during the period of 2014-2021 based on gender distribution.

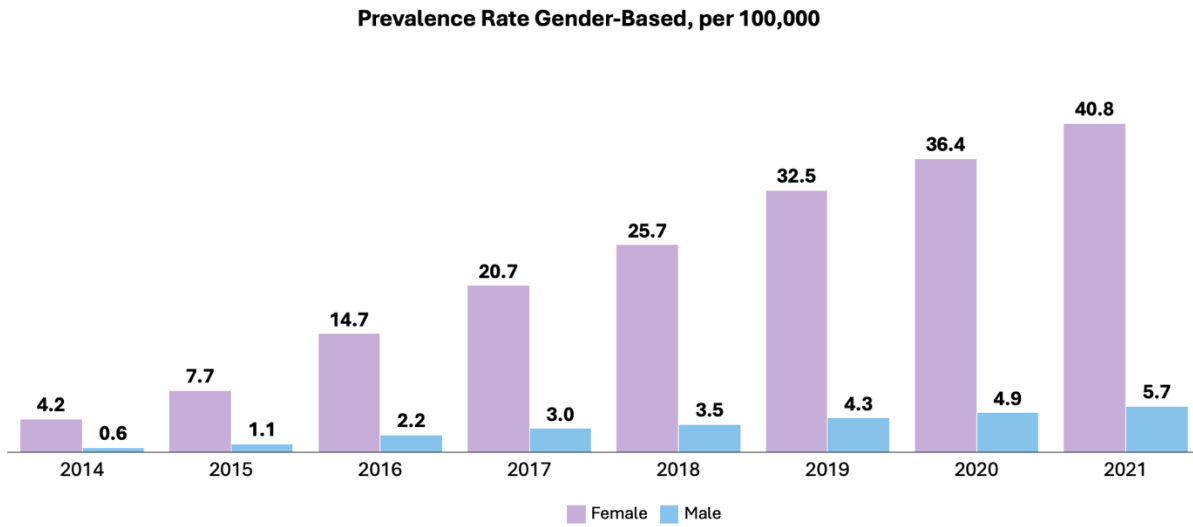


Figure 3. Prevalence rate of thyroid cancer in Kazakhstan during each year in the period of 2014-2021 based on gender distribution.

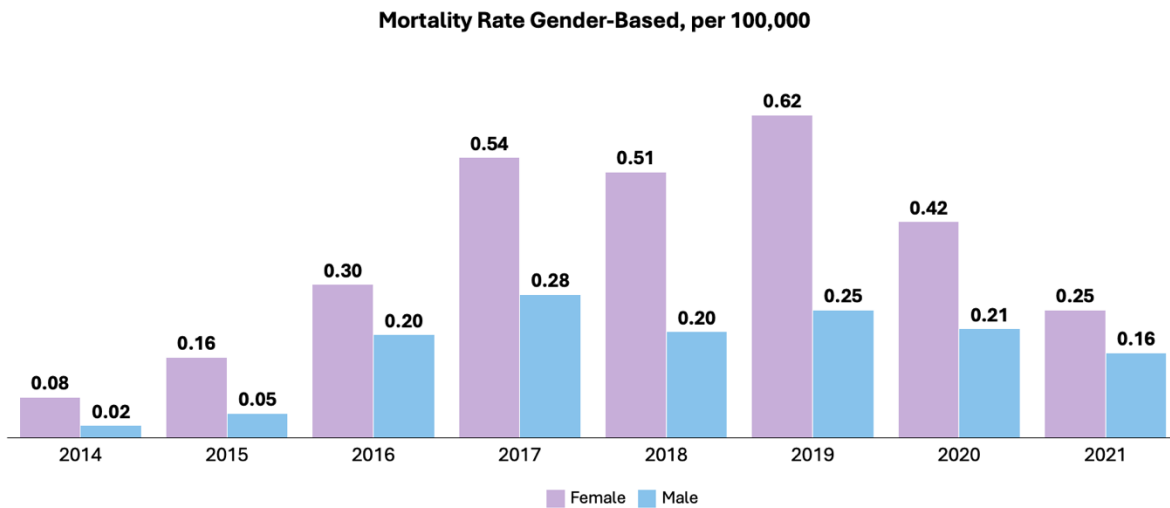


Figure 4. Mortality rate of thyroid cancer in Kazakhstan during each year in the period of 2014-2021 based on gender distribution.

Incidence, prevalence and mortality rates calculated based on age groups in the period of 2014-2021 are shown on Figures 5-7. Throughout the whole period, incidence rates were the highest among the 51-70 years old age group. As the prevalence rate was going upwards each year, it was also the highest in the same age group as incidence rate. As seen on Figure 7, mortality rate was overall 0 for the patients younger than 18, and it was the highest among patients older than 70 years old.

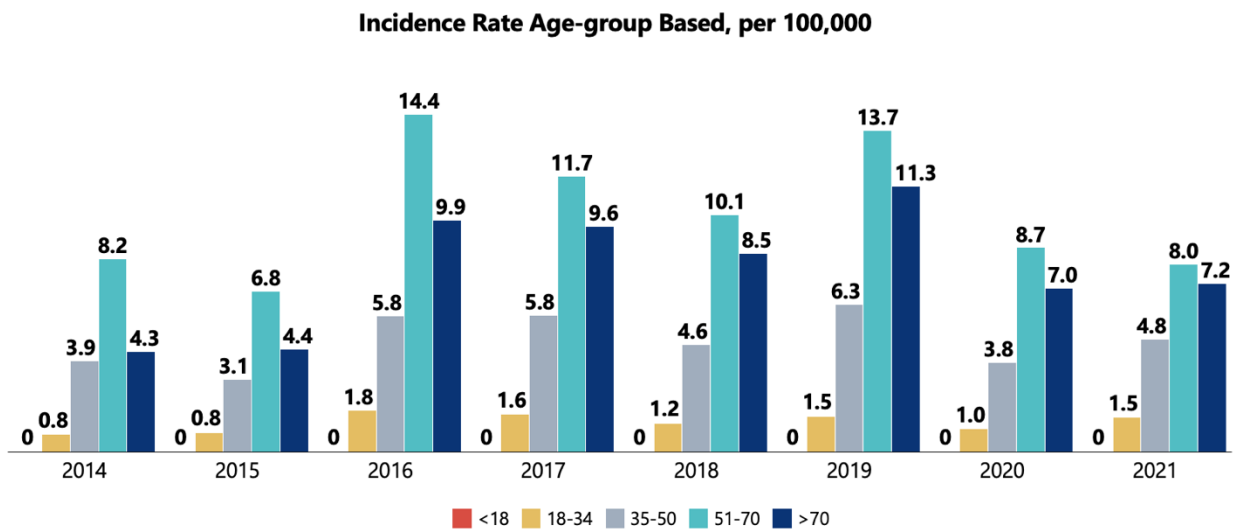


Figure 5. Incidence rate of thyroid cancer in Kazakhstan during each year in the period of 2014-2021 based on age groups.

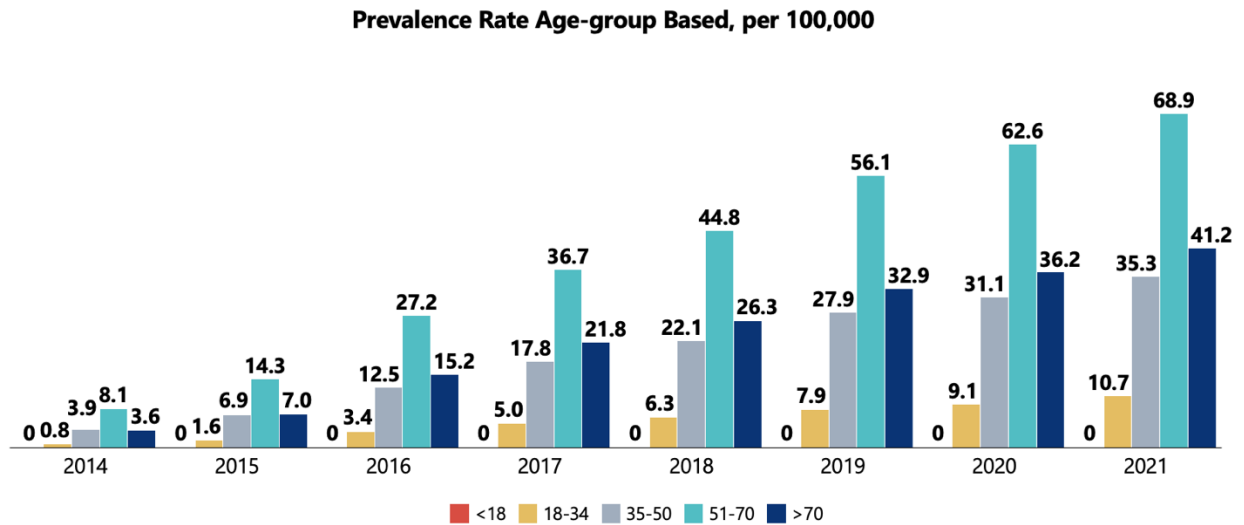


Figure 6. Prevalence rate of thyroid cancer in Kazakhstan during each year in the period of 2014-2021 based on age groups.

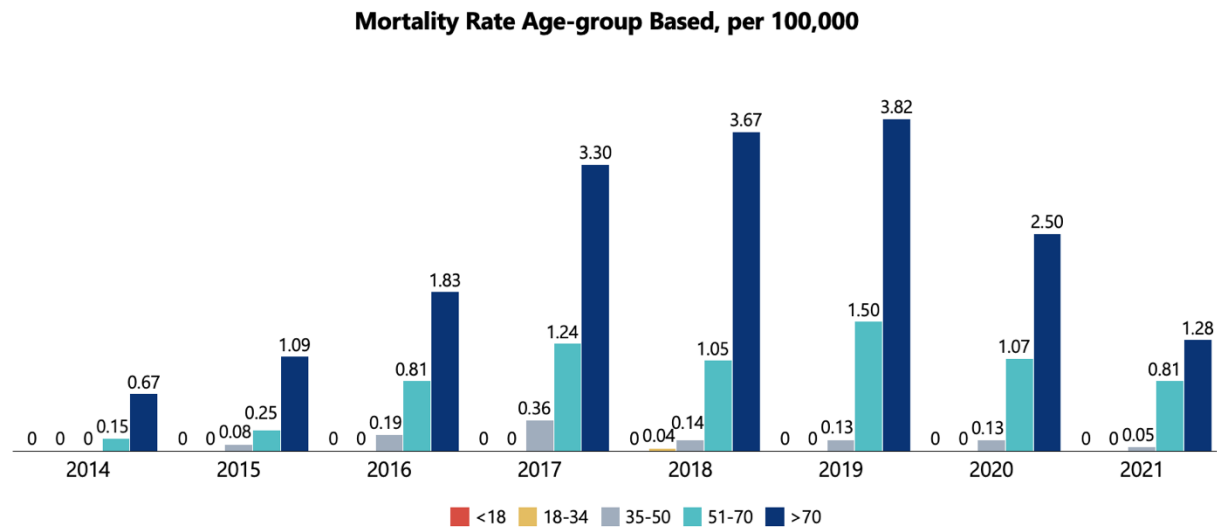


Figure 7. Mortality rate of thyroid cancer in Kazakhstan during each year in the period of 2014-2021 based on age groups.

Survival analysis

Figure 8 demonstrates the overall survival plot for the entire cohort of thyroid cancer patients with 95% confidence interval. As can be seen from the survivor function, it does not go below 0.75 throughout the whole follow-up period of 8 years.

Figures 9-12 show Kaplan-Meier survival graphs that were stratified by socio-demographic factors such as sex, age group, living area and ethnicity. Log-rank test was conducted to assess the significance of the difference in survivor function between the given groups. Based on the obtained results, the difference in survival was significant between males and females, people residing in urban and rural areas, and age groups. Survivor function stratified by ethnicity groups did not show statistically significant difference.

Table 1 provides hazard ratios (HR) with 95% confidence intervals and corresponding p-values for Model 1 of cox regression analysis which reports crude values, and Model 2 which corresponds to the analysis adjusted to demographic parameters. Overall, it shows that among the age groups, patients between 51-70 years old and patients older than 70 demonstrate a 3-fold and 12-fold higher risk of death, respectively, relative to the reference group. Younger individuals had a much lower risk of death, according to the analysis. At the same time, males demonstrate a higher risk of death compared to females, their hazard ratio being 3.18. Patients with thyroid cancer living in a rural area show a slightly higher risk of death compared to urban citizens. As for ethnicity groups, there were no statistically significant differences between the groups in terms of hazard ratio.

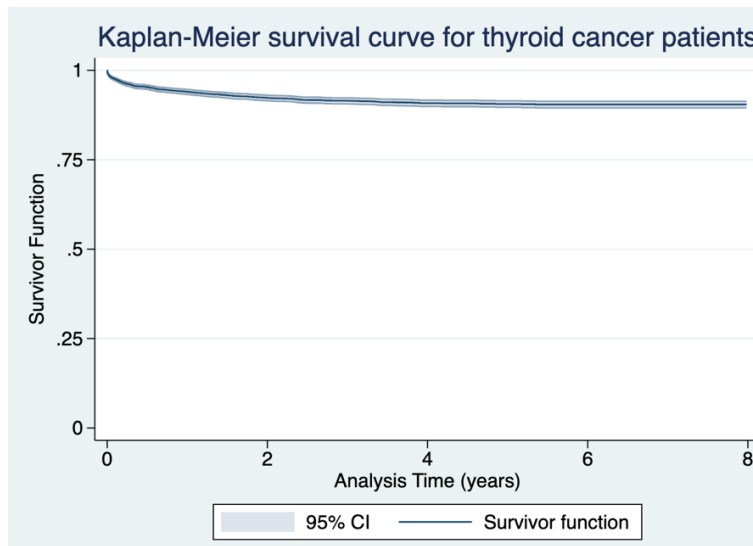


Figure 8. Kaplan-Meier survival curve for the entire cohort with 95% confidence intervals.

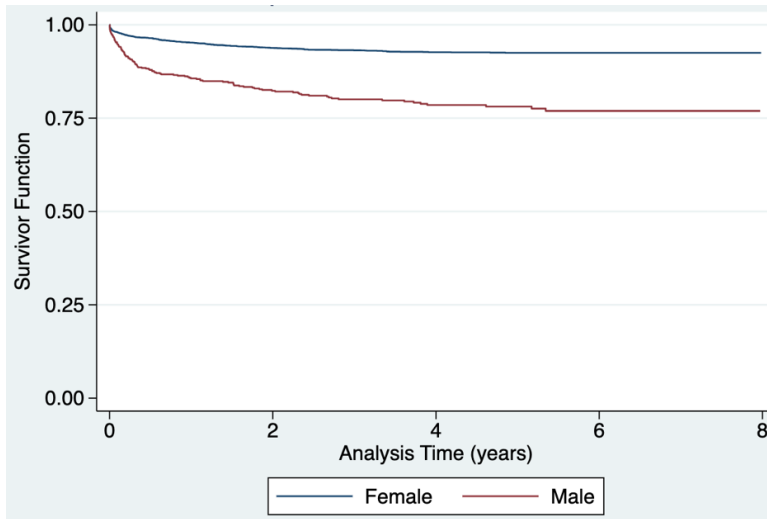


Figure 9. Kaplan-Meier survival curve stratified by sex, log-rank test $p < 0.05$.

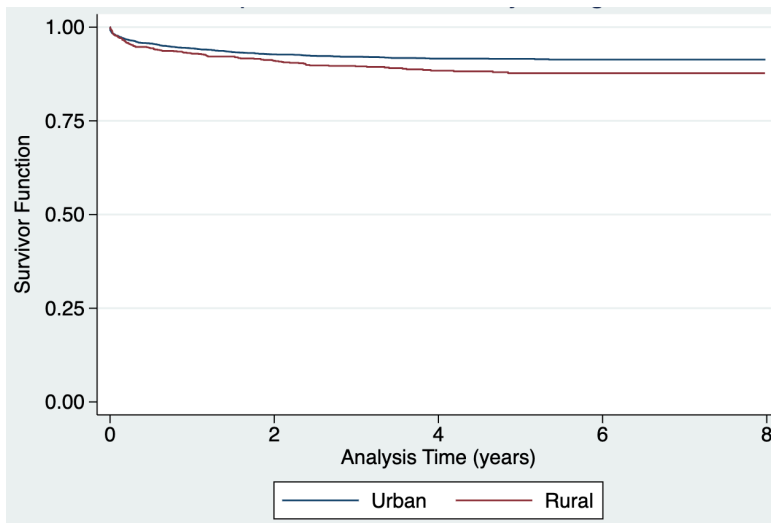


Figure 10. Kaplan-Meier survival curve stratified by living area, log-rank test $p < 0.05$.

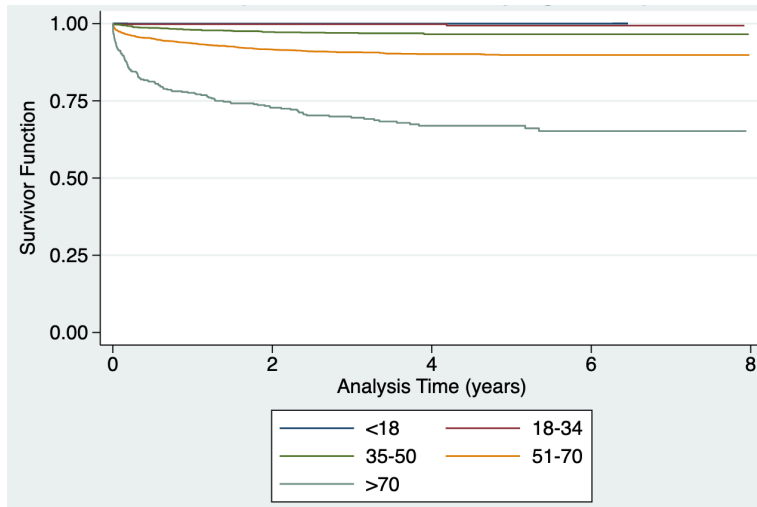


Figure 11. Kaplan-Meier survival curve stratified by age group, log-rank test $p < 0.05$.

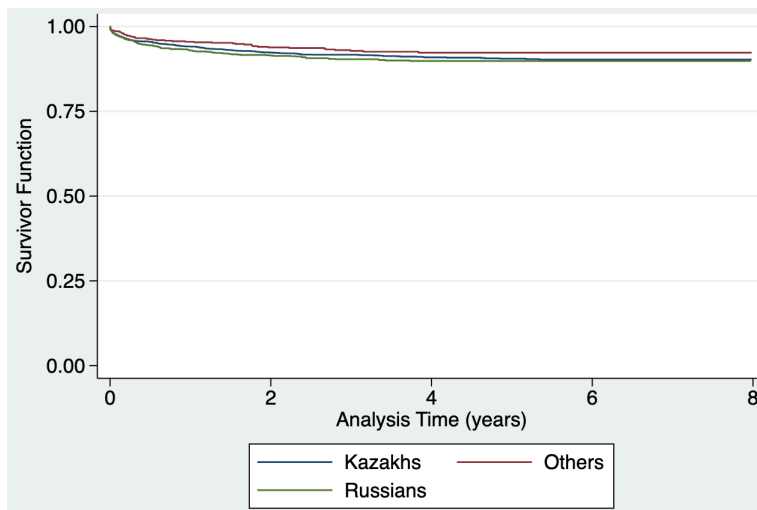


Figure 12. Kaplan-Meier survival curve stratified by ethnicity, log-rank test $p = 0.16$.

Table 2. Association between age, sex, residence, ethnicity and all-cause mortality rates for thyroid cancer between 2014 and 2021.

Variable	Unadjusted HR (95% CI)	p-value	Adjusted to demographics HR (95% CI)	p-value
Demographics				
Age groups				
18-34	0.14 [0.03 - 0.57]	<0.05	0.15 [0.04 - 0.60]	<0.05
35-50	ref.		ref.	
51-70	3.15 [2.26 - 4.40]	<0.05	3.10 [2.22 - 4.32]	<0.05
>70	12.11 [8.54 - 17.18]	<0.05	11.77 [8.27 - 16.73]	<0.05
Sex				
Female	ref.		ref.	
Male	3.18 [2.57 - 3.92]	<0.05	2.86 [2.31 - 3.54]	<0.05
Residence				
Urban	ref.		ref.	
Rural	1.37 [1.11 - 1.70]	<0.05	1.39 [1.13 - 1.73]	<0.05
Ethnicity				
Kazakh	ref.		ref.	
Russian	1.10 [0.89 - 1.37]	0.39	0.98 [0.78 - 1.23]	0.863
Other	0.80 [0.58 - 1.09]	0.154	0.76 [0.55 - 1.04]	0.081

Discussion

Study observations

This study took into account various demographic characteristics which allowed to perform an extended descriptive analysis of the given cohort. Overall, the statistics retrieved from the dataset aligned with global trends regarding thyroid cancer.

For example, it has been shown that the majority of thyroid cancer patients were females, and most of the patients were 51-70 years old at the age of diagnosis. Similar statistics is prevalent globally, and also was described earlier by Igissinov et al. (2019) in their study. Although the time period for the given study and their research is different, the descriptive characteristics based on age and gender coalign.

It was demonstrated that the majority of patients were residing in urban areas, while only a quarter were rural residents. Such results can be explained by a higher level of educated citizens in the big cities, better quality of diagnostics compared to rural regions and more diagnostic institutions to be referred to.

As for hospital admissions and social status of the patients, it was hard to receive a complete picture due to large numbers of patients not specifying these characteristics. Therefore, complete comparison of these variables was not entirely possible for this study.

Half of the urgently admitted patients contributed to the mortality rates of thyroid cancer, likely due to complications that caused the urgency of the admission. Although, the data was also quite limited due to the high number of unspecified admissions.

Regional distribution showed the biggest number of cases in Almaty city, Karaganda region and Eastern-Kazakhstan region. Although such numbers of cases being diagnosed in Almaty can be explained by better diagnostics and, likely, a bigger population in the city overall, it remains relatively unclear why Karaganda and Eastern-Kazakhstan regions exhibited such trends despite their total population.

Incidence, prevalence and mortality rates were constantly the highest among females throughout the whole period which also corresponds to the global observation of thyroid cancer favoring females. The same variables were calculated for each of the age groups present in the study showing the highest incidence and prevalence rates among 51-70 years old and the highest mortality rate among people older than 70. Since these were all-cause mortality rates, it is unclear whether such mortality rates are related to the diagnosis of thyroid cancer or other comorbidities and age.

As per survival analysis, the overall Kaplan-Meier survival curve for the entire cohort indicated a generally good prognosis for thyroid cancer patients as the survivor function over 8 years of follow-up showed a slight decline throughout the period.

For stratified Kaplan-Meier survival analysis, there was a statistically significant difference in survival trends among patients from different age groups, male and female patients, patients living in urban and rural areas.

The analysis has shown that older patients, in particular, patients older than 50, had lower survival probabilities. Patients older than 70 had the lowest survival rates throughout the follow-up period. These results were further confirmed in the cox regression model where the hazard ratios were the highest for the subgroups of 51-70 years old patients and patients older than 70. Therefore, it can be assumed that increasing age can be an indicator of poor prognosis in patients with thyroid cancer.

Moreover, the results of survival analysis using cox regression model are generally aligned with the trends observed in Kaplan-Meier analysis. Kaplan-Meier plot stratified by sex showed that males had a lower survival probability than females throughout the follow-up period.

Additionally, stratification by living area also showed a slightly less survival rate for patients residing in rural regions. In cox regression analysis, it was confirmed that male sex was a strong predictor for poor prognosis in thyroid cancer patients based on the given hazard ratio. As for the patients living in rural areas, they also had a slightly higher risk of death than urban citizens, based on the hazard ratios. This could be explained by limited access to healthcare facilities, lower education levels among the citizens and therefore, later diagnosis and more predisposition to treatment adherence issues.

Both Kaplan-Meier analysis and cox regression analysis did not report any statistically significant differences in survival among patients with different ethnicities. Such results could be explained by limited additional data. The patients were grouped into 3 major categories of Kazakhs, Russians and Others. It is possible that the results would be slightly different if there was more data available in terms of ethnicities of the patients and the group of Others would be split into several others.

Limitations

Although this was a rather extended study in terms of descriptive analysis, there were certain pitfalls during the research related to the amount of available data. It was hard to assess the exact numbers of planned and urgent hospital admissions due to the numbers of unspecified ones. Social status distribution was also not totally reliable due to not specified data. Overall, there is a

possibility of data unavailability which could also affect the socio-demographic and survival analysis for this cohort. Missing data could be a major contributor to the selection bias in this analysis as some patients could have been excluded from the cohort during the data cleaning step of the project. This way, it is possible the cohort does not represent the entire Kazakhstani population diagnosed with thyroid cancer for the given period.

Additionally, data extracted from the UNEHS only included hospital admissions in the period of 2014-2021 making it impossible to find out existing cases before 2014. Therefore, prevalence in 2014 was calculated without consideration of the possible thyroid cancer cases before that. This could have contributed to significant underestimation of prevalence and prevalence rates in 2014 and 2015. Therefore, further analysis is needed while taking into account cases of thyroid cancer existing before the given period.

Survival analysis is significantly limited due to the period of follow-up. As the available data covered the period of time between 2014 and 2021, it is insufficient to make major conclusions on survival trends due to the period length. The analysis also could be further improved if more data was available on the treatment that patients received and their comorbidities. Examples of comorbidities could include hypertension, presence of other cancers, cardiovascular diseases and diabetes mellitus as those were proposed to be the most common comorbidities in a study by Kuijpers et al. (2006). Such factors need to be assessed in cox regression analysis since it is common that they contribute significantly to the higher risk of death among patients. The last but not least, distribution of different types of thyroid cancer among the diagnosed patients could also be helpful in further survival analysis as some types are more aggressive than others and could affect the overall mortality trends.

Further healthcare implications

Based on the obtained results, it can be assumed that increasing age, male sex and living in rural areas can be predictors of poor prognosis among the patients with thyroid cancer. Therefore, it would be beneficial for the healthcare system to integrate certain interventions targeting the older population and male patients as those are more likely to have a higher death risk. Similarly, there should be an increase in healthcare resources in rural regions, considering the difference in survival between urban and rural citizens. For example, establishing more regional cancer centers could improve diagnostic rates in the regions as well as provide more treatment

options. Since such changes require financial support and time, telemedicine could also be handled as a useful tool in monitoring thyroid cancer patients.

In conclusion, this study was mainly focusing on key demographic factors affecting survival of thyroid cancer patients in Kazakhstan. In general, the obtained data was coalignining with previous research done by Igissinov et al., 2019 and global data on thyroid cancer. The strongest predictors of mortality were age and sex, although living area also had an impact on prognosis. The study was limited mostly by scarcity of original data such as no information on comorbidities and treatment types that the patients in this cohort have received. Therefore, further research is needed to assess the impact of such significant factors on patients' survival. These findings will provide additional data needed to analyze possible future healthcare implications needed to improve survival in patients with thyroid cancer.

Reference List

Arynova, Z., & Baiguzhinova, L. (2019). Development of electronic healthcare in Kazakhstan as a factor of improving the quality of medical services. *Fundamental and Applied Researches in Practice of Leading Scientific Schools*, 31(1), 3-6. <https://doi.org/https://doi.org/10.33531/farplss.2019.1.01>

Bao, W. Q., Zi, H., Yuan, Q. Q., Li, L. Y., & Deng, T. (2021). Global burden of thyroid cancer and its attributable risk factors in 204 countries and territories from 1990 to 2019. *Thoracic cancer*, 12(18), 2494–2503. <https://doi.org/10.1111/1759-7714.14099>

Hiller-Sturmhöfel, S., & Bartke, A. (1998). The endocrine system: an overview. *Alcohol health and research world*, 22(3), 153–164.

Igissinov, N., Kozhakhmetov, S., Zhantubetova, M., Igissinova, G., Bilyalova, Z., Akpolatova, G., Lyust, V., Koblandin, S., Turebayev, D., Adaibayev, K., Omarbekov, A., Tarzhanova, D., Zhantureyeva, A., & Esseyev, A. (2019). Thyroid Cancer in Kazakhstan: Component Analysis of Incidence Dynamics. *Asian Pacific journal of cancer prevention : APJCP*, 20(9), 2875–2880. <https://doi.org/10.31557/APJCP.2019.20.9.2875>

Khodamoradi F., Ghoncheh M., Mehri A., Hassanipour S., Salehiniya H. (2018). Incidence, mortality, and risk factors of thyroid cancer in the world: a review. *World Cancer Research Journal*, 5 (2): e1093 DOI: 10.32113/wcrj_20186_1093

Kuijpers, J. L., Janssen-Heijnen, M. L., Lemmens, V. E., Haak, H. R., Heijckmann, A. C., & Coebergh, J. W. (2006). Comorbidity in newly diagnosed thyroid cancer patients: a population-based study on prevalence and the impact on treatment and survival. *Clinical endocrinology*, 64(4), 450–455. <https://doi.org/10.1111/j.1365-2265.2006.02492.x>

Pizzato, M., Li, M., Vignat, J., Laversanne, M., Singh, D., La Vecchia, C., & Vaccarella, S. (2022). The epidemiological landscape of thyroid cancer worldwide: GLOBOCAN estimates for incidence and mortality rates in 2020. *The lancet. Diabetes & endocrinology*, 10(4), 264–272. [https://doi.org/10.1016/S2213-8587\(22\)00035-3](https://doi.org/10.1016/S2213-8587(22)00035-3)