ASSOCIAÇÃO BRASILEIRA DE ENERGIA NUCLEAR – ABEN
2019 International Nuclear Atlantic Conference - INAC 2019
Santos, SP, Brazil, October 21-25, 2019

Agência Brasileira do ISBN
ISBN 978-85-99141-08-3

INVESTIGATION OF THE PARAMETERS AFFECTING PATIENTS WAITING TIME IN THE RADIOTHERAPY TREATMENT BY USING ALGORITHMS TO EVALUATE ELECTRONIC HEALTH RECORDS

Caroline Z. S. Emiliozzi[1,2], Mario O.de Menezes[1] and Edson G. Moreira[1]

[1] Instituto de Pesquisas Energéticas e Nucleares (IPEN / CNEN - SP)
Av. Professor Lineu Prestes 2242
05508-000 São Paulo, SP, Brazil
caroline.zep@gmail.com

[2]Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
Departamento Radioterapia
Av. Dr. Enéas de Carvalho Aguiar, 255
05403-000 São Paulo, SP, Brazil

ABSTRACT

Cancer is the second leading cause of death worldwide and Radiotherapy (RT) is an important modality in the treatment of these patients, which consists in deploying ionising radiation to destroy or damage cancer cells. With this growing global burden, demand for RT has been increasing continuously and supply-demand imbalances have become a major concern. The reason is that delays in radiotherapy can affect the outcome by permitting local proliferation of clonogenic cells and spread of the cancer beyond the treatment volume. Studies show a common cause of anxiety for radiotherapy patients is the fact that they do not know how long they will have to wait for treatment to start. In this study, we analyze the data of electronic health records to attempt to provide a better understanding of the problem and provide an initial estimate of radiotherapy patient's waiting time. The data for this project comes from a subset of MOSAIQ, a relational database system developed by Elekta and used as an electronic health record system by the Radiation Oncology Department at the Hospital das Clínicas de São Paulo (HCFMUSP). The dataset consists of real historical data collected between January 2016 and December 2018. Visual Basic for application (VBA) and RSTUDIO Software were used to extract and analyze the data. Our work goal is to investigate a set of factors and verify their influence on patient waiting time. Factors as diagnosis, patient's age, priority of the diagnosis, and the season in which treatment planning has initiated may reveal crucial information about overall efficiency and guide us to improve clinical procedures and practices.

1. INTRODUCTION

Cancer is the second leading cause of death worldwide and Radiotherapy (RT) is an important modality in its treatment that deploys ionising radiation to destroy or damage cancer cells [1]. The goal of radiation therapy is to deliver the maximum amount of radiation to the cancerous cells, while minimizing the amount of radiation to the surrounding healthy areas. About 60% of cancer patients will receive this treatment modality at some point in their journey. In other words, radiation therapy plays a vital role in their overall treatment.
With this growing global burden, demand for RT has been increasing continuously and supply-demand imbalances have become a major concern. According to the Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), 600,000 new cases of cancer are expected to be diagnosed during 2018-2019 [2] and data provided by Sociedade Brasileira de

Radioterapia (SBRT) show that the current deficit of radiotherapy machines for adequate care for patients from public health system called Sistema Único de Saúde (SUS) is around 162 units. As a result, about 90,000 cancer patients are waiting for treatment [3].

Delays in healthcare services are one of the most challenging problems facing the healthcare system today. Evidence has been published of the negative impact of treatment delays on measures such as tumor progression, persistence of cancer symptoms, psychological distress and decreased cancer control and survival rates [4,5]. In his work, Chen et al (2008) concluded that while the evidence did not support a relationship between waiting time and risk of distant metastasis, an increased risk of local recurrence of 1.14 per month of waiting time across all investigated cancer types was found [4].

Due to the negative impact of treatment delays (time between the diagnosis and the initiation of the cancer treatment), in 2012, the Federal Government decreed the "Law of 60 days" (Federal Law number 12.732/12). This law came into effect in 2013 and regulates the maximum period a patient with cancer has to wait, in order to initiate his/her treatment, in an attempt to decrease the time between diagnosis and treatment and the consequences thereafter [6]. However, in many cases, that is not obeyed.

The reason for delays in Radiotherapy is not only due to imbalance between capacity and demand, but also due to inefficiency of workflow, for instance, scheduling problems. Consequently, as there is pressure to contain costs, it is often not possible to solve the problem of the lack of equipment. However, the problem of inefficient processes can be attacked. Seeking to reduce the potential impact of delays on radiation therapy cancer patients and motivated by inefficiency in the use of expensive resources, in this work, we analyze the data of electronic health records to attempt to provide a better understanding of the problem and provide an initial estimate of radiotherapy patients waiting time.

2. METHODS

2.1. Data collection

This was a retrospective study of patients data of electronic health records performed in the Radiation Oncology Department at the Hospital das Clínicas de São Paulo (HCFMUSP) between January 2016 and December 2018. Our center performs treatments on 3 linear accelerators (linac), each linac working for 12 hour per day. The data for this project comes from a subset of MOSAIQ, a relational database system developed by Elekta. Members of the radiation oncology team use this system for every aspect of day-to-day clinical workflows, including tasking, scheduling, treatment planning, imaging and treating. The MOSAIQ does not come with any tools to analyze the collected data. Therefore, a data-mining technique was used. Visual Basic for application (VBA) and RSTUDIO Software were used to extract information and/or patterns. The dataset, which was properly anonymized, consists of real historical data from 5543 patients. We extracted parameters as: diagnosis, patient's age, data and time of each task at the department, number of treatments fields, treatment dose, number of sessions of treatment and the season in which treatment planning has initiated. Patients were categorized by age into kids (ages 0-18), young adults (ages 19-35 years), middle-aged adults (ages 36-60 years) and older adults (aged older than 60 years) and the international classification of diseases (ICD) was used to classify the diagnosis [7]. Furthermore, a subdivision of the ICDs by region was performed. 30 groups of ICDs were formed:

- Abdomen (C22-C25, K70-K77 and K80-87),
- Appendix (K35-K38),
- Articulation (M00-M22)
- Bone (C40-C41, D16 and M926)
- Breast (C50, D05, D24, N60-N64)
- Digestive (C16-C17, D01, D13, K90 K93, K20-K31, K50-K52, K55-K59 and K53)
- Disorder (F00-F99)
- Ear (H60-H75, H80-H83 and H90-H95)
- Endocrine (E00-E07, E10-E16 ,E20-E35, E40-E46, E50-E79)
- Epithelials (M800-M801, M805, M809, M12, M14, M839, M843, M844, M850, M855-M856, M859, M868, M872, M949)
- Esophagus (C15)
- Gland(C73-C76, D34-D36 and D44)
- Gyneco (C51-C58, D06-D09, D25-D28, D39, N70-N99, M910 and M911)
- Head and Thorax (D02, D14 and D38)
- Head/Face/Neck (C00-C14, C30-C33 and D10-D11)
- Hemato (C42, C88, C90-C96 and D45-D51)
- Lymphoma (C81-C85, M959, M965, M967 and M969-M972)
- Malformation (Q00-Q45, Q50-56, Q60-Q99)
- Metastasis (C77-C80)
- Moles (C48, C49, D20, D21, M880-M881, M884, M885, M889, M893, M900, M904-M906 and M958)
- Neck and Abdomen (D00 and D37)
- NS (C47, C69-C72, D31-D33, D42-D43, G10-G13, G20-G26, G30-G47, G50-G64, G70-G73, G80-G83, G90-G99, M938, M953, M954 and R24)
- Ocular (H00-H06 and H10-H59)
- Oral (K00-K14 and M927)
- Pelvis (C18-C21, D12 and K60-K62)
- Peritoneal (K65-K67)
- Skin (C43-C44, C46, D03-D04, D17-D19, D22, D23, L00-L14, L20-L30, L40-L45, L50-L75 and L80-L99)
- Thorax (C34, C37-C39, C45, D15)
- Uro (C60-C68, D29-D30, D40-D41 and N00-N51)
- Vascular (I00-I02, I105-I15, I20-I52, I60-I74, I77-I89 and I95-I99).

2.2 Delay Time Calculation

In the context of radiation oncology, there are two main types of waiting time. The first type is the interval between the time the patient receives their CT-scan and their first treatment appointment and the second one occurs on-site in the clinical waiting room, after the patient has been called to start daily radiotherapy treatment. These studies focus on the first type called pre-treatment events time. The radiotherapy pathway is shown in Figure 1. The waiting time data was analyzed to understand and explain the factors which affect delays in order to decrease waiting times for radiation therapy patients.

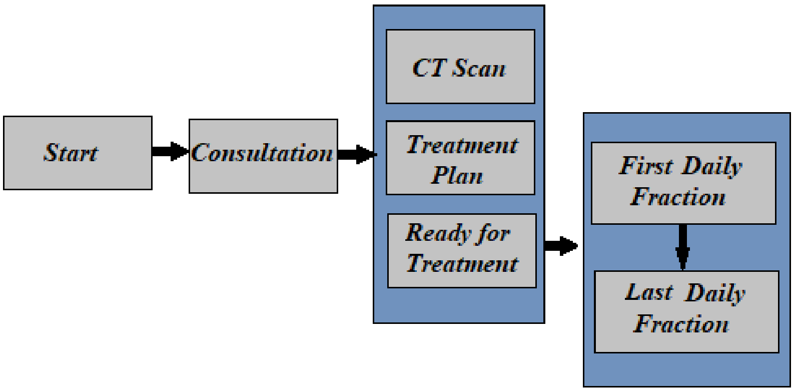

Figure 1: An overview of the patient's radiotherapy pathway

3. RESULTS

A preliminary analysis of the data was performed as shown in figures 2-4. Fig. 2 and Fig. 3 shows that the distribution of age and patients by group are not homogeneous. 50% of patients are over 60 years of age and the most frequent diagnoses are related to cancer of the breast, prostate, central nervous system and head and neck.

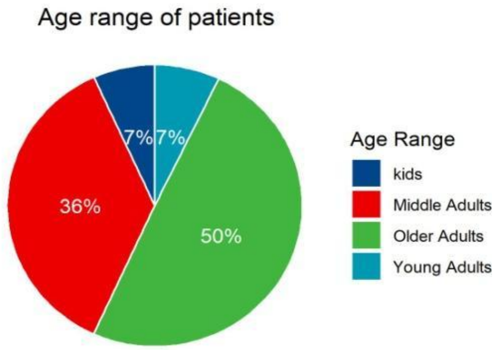

Figure 2: Percentage of patients in each age range.

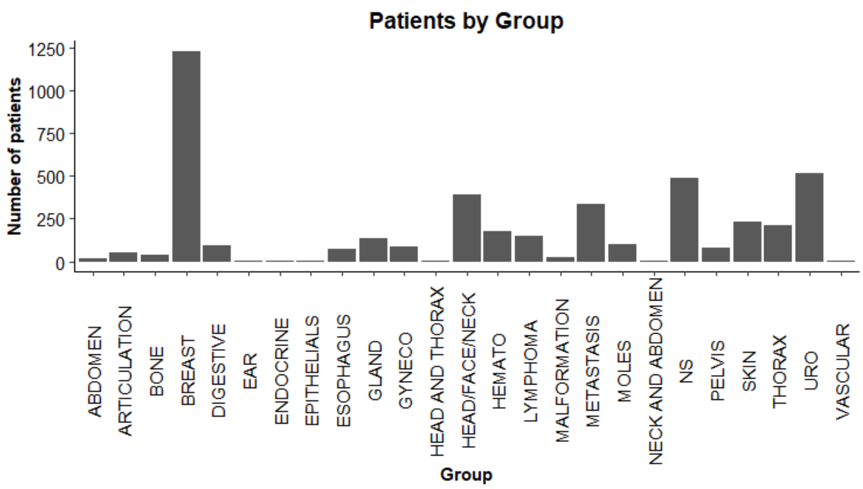

Figure 3: The number of patients in each group. Five groups have no patients.

The more frequent ICDs in Hospital das Clínicas de São Paulo are Breast malignant neoplasm (C508,C509, C500 and C504), Prostate malignant neoplasm (C61), Secondary malignant neoplasm (C793 and C795), Brain malignant neoplasm (C719), Bronchi and lungs malignant neoplasm (C349) and Benign neoplasm of the meninges (D320),e. g. Fig. 4. These data are similar to the global estimate showing that the most common cancers in 2012 were lung (1.8 million), breast (1.7 million), intestine (1.4 million), and prostate (1.1 million) [2].

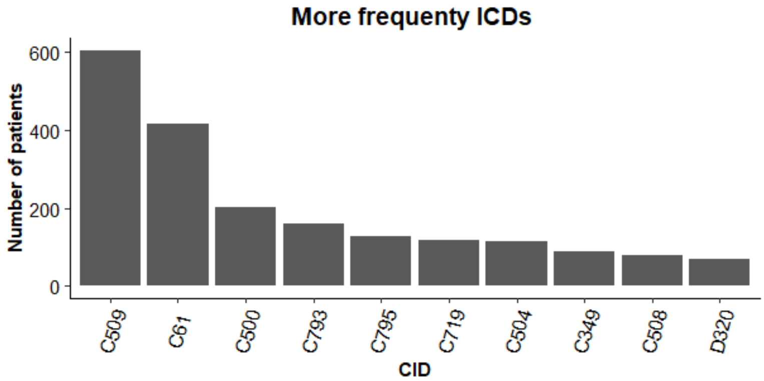

Figure 4: The 10 most frequent ICDs.

By the evaluation of the boxplot graph of treatment data, e.g. Fig. 5,, it can be seen that only the data of the number of sessions didn't have a large variation of values (1-60 sessions). The UM median per treatment was 423, the dose median was 270 cGy and the number of fields median was 4.

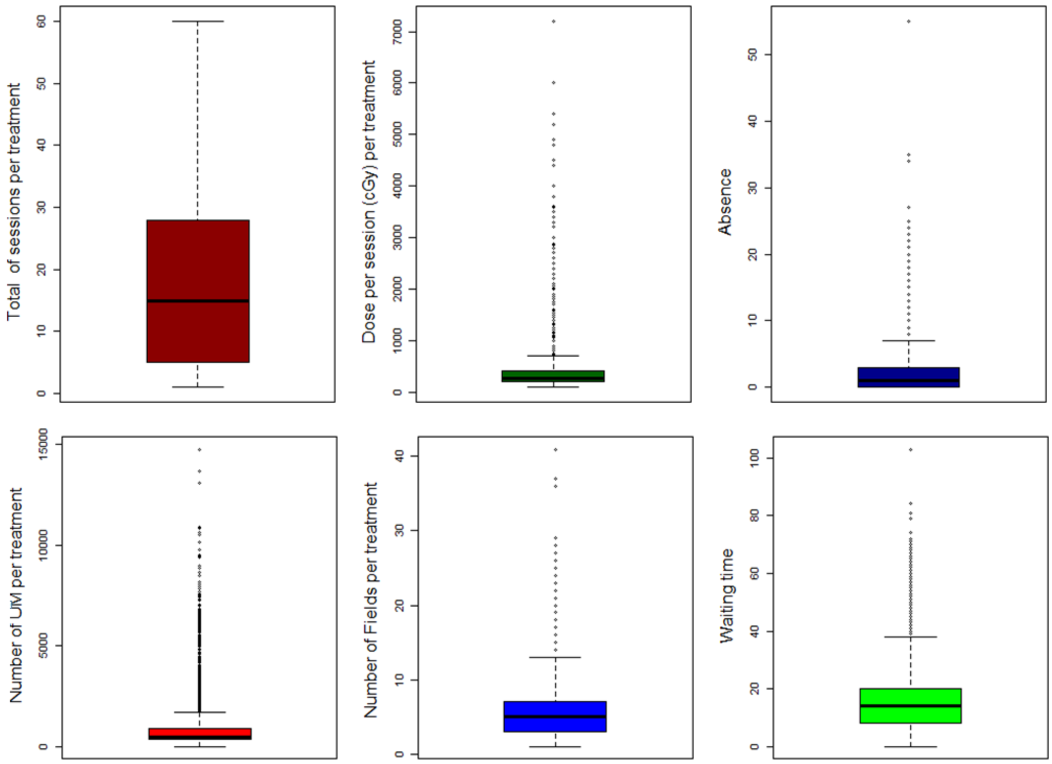

Figure 5: **Evaluation of the empirical distribution of treatment data.**

The average waiting time found was 15 days (the interval between the time the patient receives their CT-scan and their first treatment appointment). The maximum value found was 103 days, the reason for which was the lack of an exam result. Furthermore, the average absence per patient during the treatment was 2 days. The maximum value found was 55 days that happens when the patient has a clinical worsening and needs to be hospitalized, not being able to attend for the treatment. For the evaluation of the seasonality of patients, the number of new patients per month was plotted in Fig. 6. The number of new patients doesn't have a significant variation over the year. A small decrease in the number of patients in December seems to be observed.

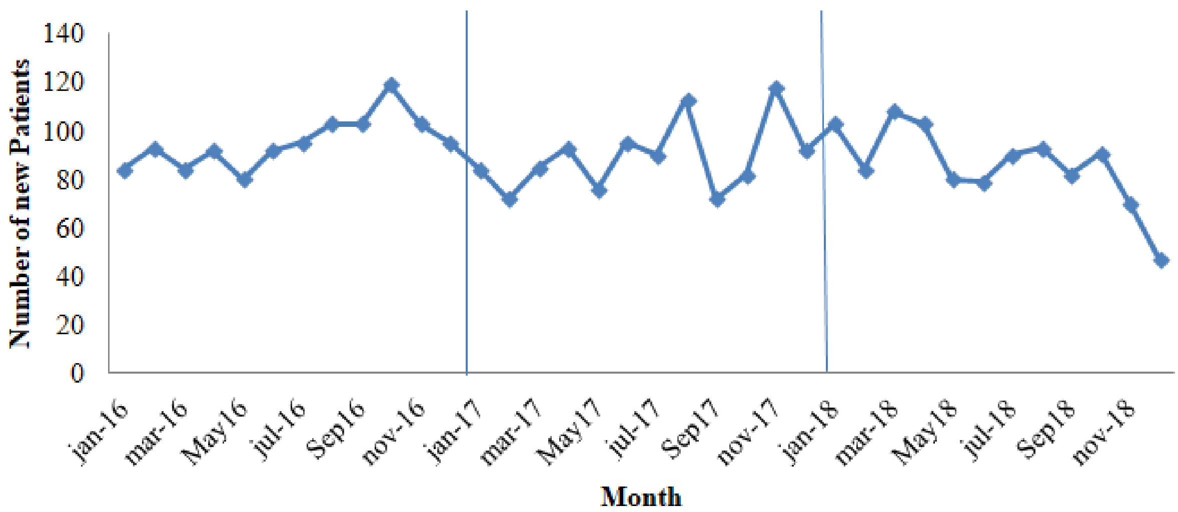

Figure 6: Number of new patients per month. Blue lines separate the data by year.

During the analysis of the patients' data, the number of relapses caught our attention. To try to understand this, some tests were performed and shown in Fig.7-9. The Relapse percentage was 29%. In the department, 20% of patients were treated twice and less than 0.5% were treated seven times. From Fig. 8, the thorax, bone and metastasis group had the greatest number of recurrences. The patients of endocrine, ear, head and thorax, malformation and vascular group hadn't had recurrences.

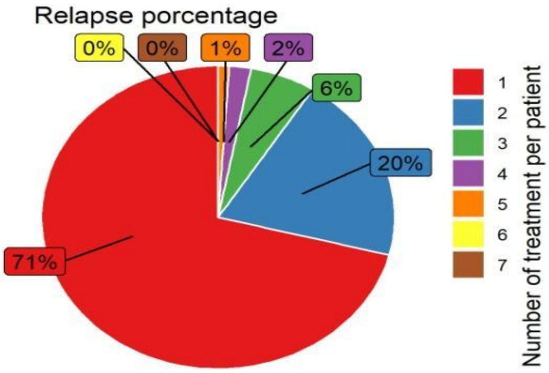

Figure 7: **The chart shows patients' percentage who returned to a new treatment in the department. If the treatment number is equal to one, the patient treated once.**

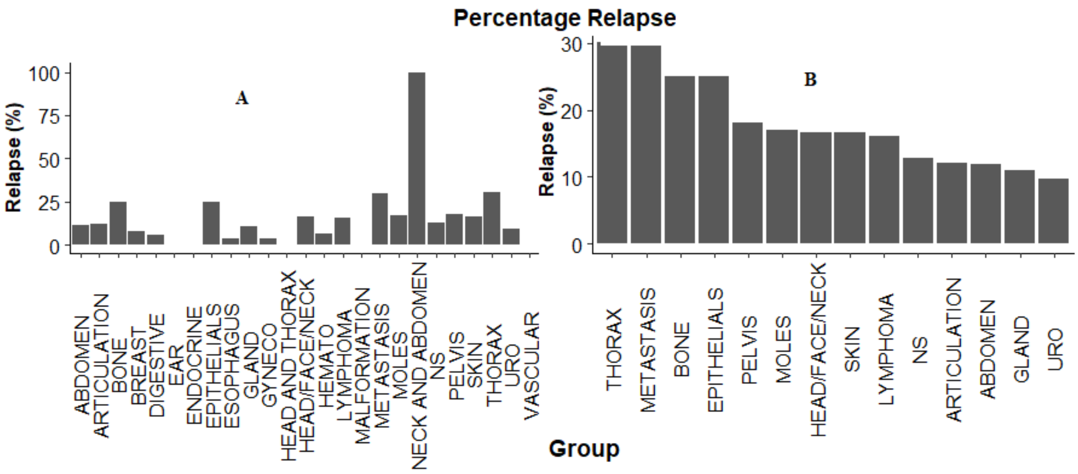

Figure 8: **Percentage of patients who returned to a new treatment per group. The letter A shows the percentage who have relapse by group and the letter B shows the group with higher percentage of relapse.**

Fig. 9 suggests there is a weak link between recurrence and age, since the relapse percentage by age group almost did not change. However, Kids appear to be more likely to relapse.

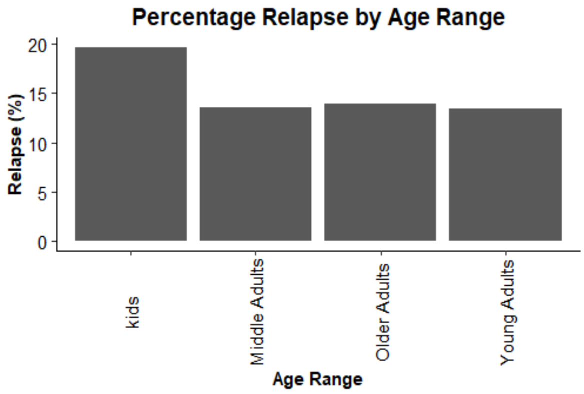

Figure 9: **Percentage of patients who returned to a new treatment per age range.**

Another analysis to understand the factors that affect the recurrence rate was to calculate the correlation coefficient between the variables, e. g. Fig 10. However, no data presented a strong correlation coefficient. We observe by the correlation coefficient that the group with the largest number of patients also presented the highest number of recurrences. However, as the group is a categorical variable, it was not shown in the figure below.

To try to understand the waiting time, we analyzed the relationship of waiting time with the group, age and month of admission as shown in Tables 1-3, where **N** is sampling, **WT** is average waiting time (days), **Max** and **Min** are the maximum and minimum waiting time respectively and **SD** is the standard deviation of the data.

Evaluating the waiting time per month, it is noticed that the value didn't have significant difference, ranging from 13-17 days. However, each month had at least one case, with a waiting time longer than the Federal Law number 12.732/12. Comparing to the age range, the waiting time also didn't have significant difference, ranging from 15-16 days. Only Kids didn't exceed 60 days. The factor which most affected the waiting time was the diagnosis, as shown in Table-3. The Neck and Abdomen Group had an average waiting time of 45 days and Head/Face/Neck had an average waiting time of 20 days. Some groups had average waiting time of 0 days, which are urgencies cases. The other groups had a waiting time value of approximately 15 days.

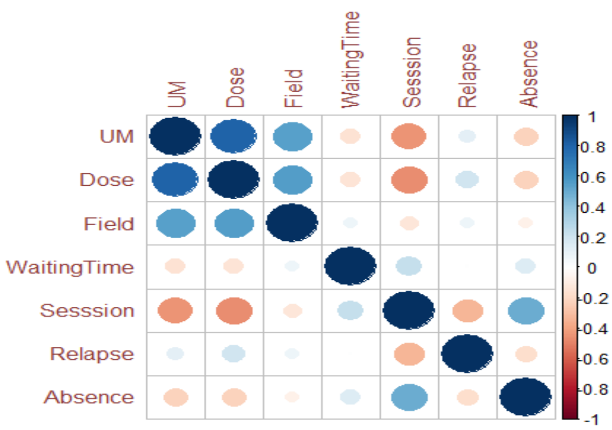

Figure 10: Calculation of the correlation coefficient between the numerical variables.

Table 1: Analysis of the relationship between waiting time and month of admission

Month	N	WT	Max	Min	SD
April	335	16.89	84.00	0.00	11.79
August	371	13.50	72.00	0.00	9.40
December	312	14.50	69.00	0.00	10.85
February	288	15.40	65.00	0.00	10.79
January	309	14.82	56.00	0.00	8.43
July	319	13.12	81.00	0.00	10.40
June	326	15.44	79.00	0.00	10.73
March	357	16.79	103.00	0.00	12.45
May	381	17.01	74.00	0.00	11.86
November	288	15.62	67.00	0.00	10.54
October	378	13.99	72.00	0.00	10.68
September	277	14.31	60.00	0.00	9.36

Table 2: Analysis of the relationship between waiting time and the age range

Age Range	N	WT	Max	Min	SD
kids	243	14.78	49.00	0.00	11.26
Middle Adults	1311	15.40	103.00	0.00	10.70
Older Adults	1785	15.19	84.00	0.00	10.49
Young Adults	257	16.40	79.00	0.00	12.98

Table 3: Analysis of the relationship between waiting time and the group

Group	N	WT	Max	Min	SD
ABDOMEN	14	12.71	25.00	0.00	9.22
ARTICULATION	50	9.74	32.00	0.00	9.95
BONE	47	15.09	79.00	0.00	15.92
BREAST	961	15.58	81.00	0.00	8.27
DIGESTIVE	67	10.13	33.00	0.00	9.55
EAR	2	26.00	41.00	11.00	21.21
ENDOCRINE	2	0.00	0.00	0.00	0.00
EPITHELIALS	8	13.13	27.00	0.00	9.00
ESOPHAGUS	50	14.96	33.00	0.00	5.97
GLAND	144	15.62	53.00	0.00	10.53
GYNECO	22	18.55	67.00	0.00	17.19
HEAD AND THORAX	1	0.00	0.00	0.00	
HEAD/FACE/NECK	426	20.34	71.00	0.00	12.81
HEMATO	163	7.64	52.00	0.00	8.22
LYMPHOMA	157	17.29	65.00	0.00	9.40
MALFORMATION	20	19.40	103.00	0.00	25.06
METASTASIS	372	9.42	50.00	0.00	7.55
MOLES	85	15.95	64.00	0.00	11.15
NECK AND ABDOMEN	5	45.00	72.00	14.00	28.28
NS	471	17.60	61.00	0.00	9.93
PELVIS	68	12.88	42.00	0.00	8.98
SKIN	196	15.88	84.00	0.00	13.57
THORAX	218	11.22	57.00	0.00	9.90
URO	387	15.98	74.00	0.00	10.50
VASCULAR	1	0.00	0.00	0.00	

4. CONCLUSIONS

The analyze of large amount of data provided by electronic health records to reveal statistical patterns and trends are difficult without computational help and at the same time, these data need to be transformed into knowledge to improve the quality and efficiency of treatment. This work wanted to understand better the waiting time to bring solutions to decrease it. By the data evaluated, the factor which most affected the waiting time was the diagnosis, suggesting we need to improve the workflow for some pathologies, since, for some regions, the time waiting difference was significant. With this information, a more detailed study of the workflow of these problematic groups should be carried out, for instance, checking the doctor's and medical physicist's schemes and evaluating the number of specialists part of this treatment group. Although the waiting time is on average 15 days, there are patients who are waiting more than 60 days and we need to make sure the problem is not an internal workflow one.

By analyzing the data, we noticed that the percentage of patients who return to a new treatment is significant (30%), so through the analysis of the correlation coefficient, we try to find factors affecting this number. With this analysis method, only the diagnostic appeared to be promising. For a future study, an analysis with more complex machine learning algorithm, like logistic regression, will be performed.

REFERENCES

1. International Atomic Energy Agency, *Planning national radiotherapy services: a practical tool*, International Atomic Energy Agency, Vienna & Austria (2010).
2. Instituto Nacional de Cancer José Alencar Gomes da Silva, *Estimate/2018 – Cancer Incidence in Brazil*, Instituto Nacional de Cancer José Alencar Gomes , Rio de Janeiro & Brazil (2018)
3. P. A. Luciane, M. S. Natan and T. M. A. Adriana, "Necessidades Atuais de Radioterapia no SUS e Estimativas para o Ano de 2030", *Revista. Brasileira De Cancerologia*, **62.1,** pp. 35, https://rbc.inca.gov.br/revista/index.php/revista/article/view/177 (2019).
4. Z. Chen, W. King, R. Pearcey, M. Kerba and W. J. Mackillop, "The relationship between waiting time for radiotherapy and clinical outcomes: a systematic review of the literature", *Radiotherapy Oncology,* **87** pp. 3–16 (2018).
5. M. A. Ebert,"Utilitarian prioritization of radiation oncology patients based on maximization of population tumour control" , *Physics in Medicine & Biology,* **58**.12, pp. 4013 (2013)
6. Brasil. Lei nº 12.732, de 22 de novembro de 2012. Diário Oficial da República Federativa do Brasil, Brasília, DF. Seção 1, p.1. (2012)
7. "Tabelas CID-10", http://www.datasus.gov.br/cid10/V2008/download.htm (2008).