Journal of High School Science

Comparison of two representatives of bone tumor classifications

Jeon MW[1], Sung ZH[2]

Submitted: August 2, 2024, Revised: version 1, September 13, 2024
Accepted: October 4, 2024

Abstract

This study compares the histological method versus the Memorial Sloan Kettering Cancer Center (MSKCC) method for categorizing cancer types using Machine Language algorithms. Its objective is to assess the potential of a newly introduced approach (MSKCC) for oncology classification. Development in computing algorithms has advanced systematic disease classification ability. These algorithms help clinical professionals make appropriate diagnoses and design rapid-responding treatment procedures. Bone tumors are uncommon oncological diseases, showing an encouraging five-year survival rate of ~70%. The histological classification method for Bone tumors follows the conventional sorting out criteria, depending on tumor location, patient age and grade, among other factors. Recently, MSKCC has launched a new simplified oncology classification based on a homogenized diagnosis after incorporating feedback from public tumor associations and consortia. The MSKCC system simplifies classification in hard-to-classify heterogenous tumors and addresses a pressing and evolving need of accurate diagnosis. We aimed to evaluate whether the MSKCC classification resulted in similar categorization of cancers that were conventionally classified using the histological method. In this comparative analysis, we concluded that the MSKCC method possessed the classifying power equivalent to the histological method with the added advantage of much-simplified categorization. This new method to classify cancers can lead to improved cancer treatment strategies for favorable outcomes.

Keywords

Bone tumor, Comparative analysis, Prediction analysis, Machine Learning, Memorial Sloan Kettering Cancer Center, Oncology classification, Histological classification, MSKCC

[1]Matthew Woomin Jeon, Asia Pacific International School (APIS), 57 Wolgye-ro 45ga-gil, Nowon-gu, Seoul, 01874, South Korea. wmjeon@apis.org
[2]Corresponding author: Zu Hawn Sung, Superconducting Quantum Materials and Systems Center, Fermilab, Batavia, IL 60510, USA. zsung@fnal.gov

1. Introduction

Systematic cancer classification is essential for patient care and research, reflecting the evolving need for advanced oncology treatments (1, 2). A conventional hierarchical classification method has been widely accepted based on histological criteria, such as tumor location, type, age, the extent of spread, and other features (2, 3). However, recent advanced computing technology significantly improves the sorting ability and capability related to specific characteristics of a cancer type (4). It helps medical practitioners establish dynamic treatment strategies for favorable outcomes, which may lead to a significant increase in long-term survival rates. It is, therefore, desirable to compare the performance of these methods to clarify the potential of a new approach.

Bone tumor is generally defined as cancer that develops in the bone itself rather than from metastasis from elsewhere in the body (5). A primary bone tumor is a rarely diagnosed disease, constituting less than 1% of overall malignancies. Its 5-year relative survival rate is 68.9%, and the estimates of new cases in 2023 are ~ 2% (6). Children and young adults at early ages are commonly diagnosed (5, 7). Like other malignant tumors (or cancer), the World Health Organization (WHO) periodically updates a classification system. It classifies bone tumors depending on presumed origin tissues, stage related to histology, size, and presence or absence of metastases (8, 9). In most cases, specific bone tumors emerge in the long bones, such as those found in the legs and arms. It sometimes occurs in the pelvis, ribs, legs, and spine (5, 7). However, cancers that start in the bone marrow cannot be designated as bone tumors because cancer derived from the bone marrow is defined as leukemia, multiple myeloma, and lymphoma (10).

Primary bone tumors are classified as benign, malignant, and metastatic, depending on their histologic features (5). Benign tumors are non-life threatening. Malignant tumors are defined as cancer that grows and expands at a rapid rate and can be converted to metastatic or spread to other parts of the body (11). Metastatic tumors are a type of cancer that spreads from other parts of the body in which cancer emerges and spreads to the bone. Malignant tumors are divided into Osteosarcoma, Chondrosarcoma, Ewing Sarcoma, and Osteochondromas (5,12). Osteosarcoma is the most common type, making up about one-third of the cases. Chondrosarcoma is the second most frequently diagnosed bone cancer, making up about a quarter of all primary cases. It mainly starts in cartilages at the pelvis, knees, shoulders, or upper part of thighs. Ewing sarcoma is the third most common malignant cancer, comprising ~15% of the primary cases. It usually originates in the middle part of the bones at the hips, ribs, upper arms, and thighs. Osteochondromas is another commonly diagnosed type of bone cancer. It also accounts for 30-35% of all diagnosed cases. Osteochondromas mainly occur at the end of the bones near the growth plate when the cartilage and bone overgrow (13). This is the most common non-cancerous bone growth and occurs between the ages of 10 and 30 (14).

Symptoms of bone cancers vary depending on the type of bone tumor (7). The most common

symptom is pain. For diagnosis, computer tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET) are usually employed due to their convenience and patient comfort. If these imaging techniques reveal signs of bone cancer, physicians then perform immunohistochemistry to identify antigens in the bone tissues. Blood tests and biopsies are also performed. The history of bone infection is another critical parameter that is considered in making a diagnosis (5, 7). In the past, treatment options for bone cancers were minimal. Amputation of the affected part was the only available method. However, advanced technology allows the use of multiple targeted treatment methods. Surgery, chemotherapy, and radiation therapy are the primary treatment methods that are applied either alone or in combination with other methods. By combining chemotherapy with surgical methods, the survival rate has increased to more than 70% for both osteosarcoma and Ewing sarcoma (9). Cryosurgery (the freezing and killing of cancer cells) has recently been employed in addition to surgery for some bone cancer patients (15).

The conventional classification of bone tumors follows the histological criteria (1, 2). The Memorial Sloan Kettering Cancer Center (MSKCC) (7) recently launched a new tumor classification standard associated with a homogenized diagnosis of cancers by adapting feedback from public tumor associations and consortia. This system relies on dynamic classification powered by computing technology. It aims to provide a well-established treatment procedure derived from precision diagnosis in order to achieve a favorable outcome in cancer treatment. In this study, we accessed the different features of the MSKCC method and the conventional histological method by applying statistical comparative and accuracy prediction analysis based on Machine Learning technology. We hypothesized that investigating an extraordinary case would be beneficial for improving the understanding of the MSKCC classifying versatility and robustness. Thus, we focused on an uncommon cancer, bone tumor, that has shown a favorable 5-year survival rate as described above. The goal was to clarify further the accuracy and potential of this newly introduced oncology classification approach supported by advanced computing resources.

2. Data

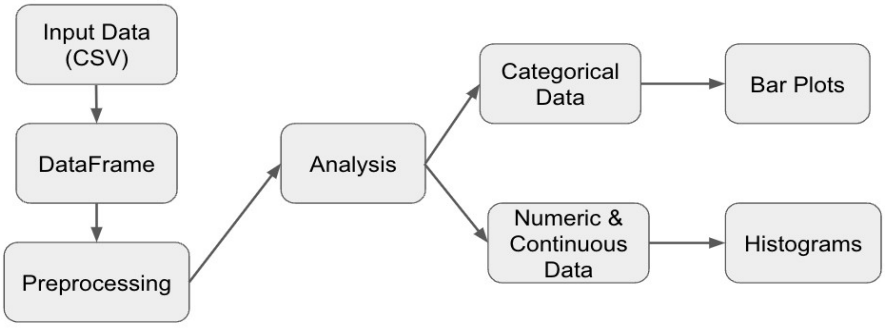
We used the dataset that monitored the progress status of 500 patients who suffered from bone tumors. It was collected at MSKCC between 2010 and 2020 (7) and classified the bone tumors' features based on the histological features as well as the MSKCC types. This data is publicly shareable *via* the MSK Data Catalog (16) or available at the Kaggle open source platform (17). The histological method classified the bone cancer into 13 types: Osteosarcoma, Ewing sarcoma, Pleiomorphic leiomyosarcoma and MFH (Malignant fibrous histiocytoma), among others (see Figure 4 for a complete list). In contrast, the MSKCC system simplifies the classification into only three categories, MFH, Leiomyosarcoma and Synovial sarcoma. The patients' final status is described as NED (no evidence of disease), AWD (alive with disease), or D (death). The treatments are either combined surgery, radiation therapy, or chemotherapy.

3. Methods

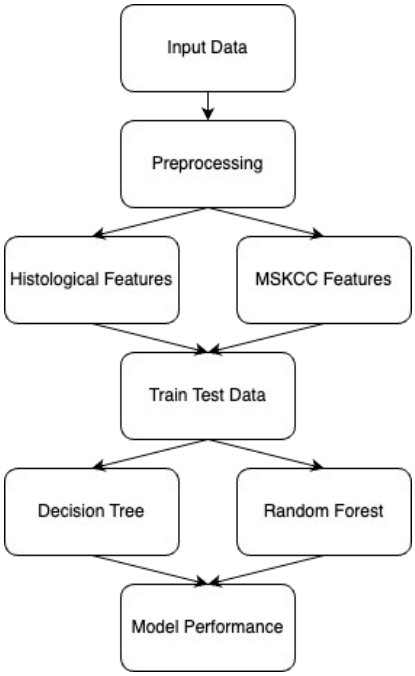
The raw data is in the form of comma-separated values (CSV), composed of numbers and texts. We applied the Google Colab, a cloud-based notebook-style data analysis framework, as a working platform, which provides a convenient way to code, share, and present data analysis (18). We pre-processed the data after loading it into a tabular data frame using the panda's Python package (19). In the initial preprocessing, we examined the data to check for any missing values or obvious outliers. The data comprised 9 columns and 500 rows. However, we actually used 6 features; 'Sex', 'Age', 'Grade', 'Histological type', 'site of primary STS', 'Treatment', and the 'Histological type'. These six features were replaced by 'MSKCC type' for the MSKCC prediction analysis, as explained later.

Figure 1a shows the flow chart for the comparative analysis process. The analysis began by plotting all the features to ascertain if there were any unique or anomalous behaviors. A bar type plot or a histogram was used if the data type was categorical, or numeric and continuous respectively. A multiple bar graph, which presents multiple sets of related data on the same graph was used to depict additional information layers for categorical data. The advantage of using multiple bar graphs is that they compare the same values for differently sliced data. For numeric and continuous data, additional layers of information can be presented by slicing data along the information of interest. For example, if the age distribution was of interest along the different statuses, the age data was filtered (sliced) by different status values, and then the histograms of the age distribution for specific status values were plotted separately. The advantage of slicing the information was that it isolated its effect on different statuses.

After the comparative analysis, we performed the accuracy prediction on the status of patients using two different supervised machine learning models (decision tree and random forest). The two ML models compared the prediction estimates of the two classification systems and evaluated their efficiencies. The process is presented in Fig. 1b. In the preprocessing stage, we converted the features (attributes) to a vector form which served as input for the models. The sklearn library (22) was used to provide a convenient function to split data into train and test samples randomly. We used the default 75:25 split ratio, as 75% for training and 25% for testing. Some parameters, which are called hyper-parameters, cannot be learned from model training steps. We tuned the acquired hyper-parameters by a 5-fold cross-validation using the grid search method. Among many hyper-parameters to be tuned, we only focused on max_depth and min_samples_leaf. These hyper-parameters exist in both ML models. We applied the same cross-validation procedure and measured performance using test samples by checking accuracies and F1-scores (23). In addition, we calculated the area under the ROC curve (AUC) and the area under the precision recall curve (AUCPR) to compare the performances of the models.



(a) The first part of the analysis: statistical comparative analysis



(b) The second part of the analysis: predictive machine learning analysis

Figure 1. Schematic diagram of two analyses applied in this study. The arrows describe the sequence of the process. (a) shows the first part of the analysis process, a statistical comparative analysis, and (b) shows the second part of the analysis, which is the prediction process of two types using two different machine learning models (decision tree and random forest)

4. Results

We first analyzed the entire data structure based on a statistical comparative analysis. Figure 2 shows the bone tumor patients' overview and each attribute (features) term: sex, site of primary STS (location of bone

tumor), status, and treatment. This countplot provides an overall insight into the structure of the bone tumor dataset applied. There were more females than males in the patient cohort. The status bar graph with different sexes is presented on the same plot so that status data can be easily compared for different cases based on the sex of the patient. Major bone tumor occurrence sites were divided into seven different body parts. The thigh is the more noticeable site, followed by the right buttock, as shown in Fig 2b. In Fig. 2c, NED (no evidence of disease) constituted almost half of all cases, and to be consistent with others, the number of D (death) was slightly higher than the number of AWD (alive with disease). Surgery, chemotherapy, and radiotherapy were combined for the significant treatment purpose rather than applied as individual attributes. Radiotherapy + surgery produced the most favorable result.

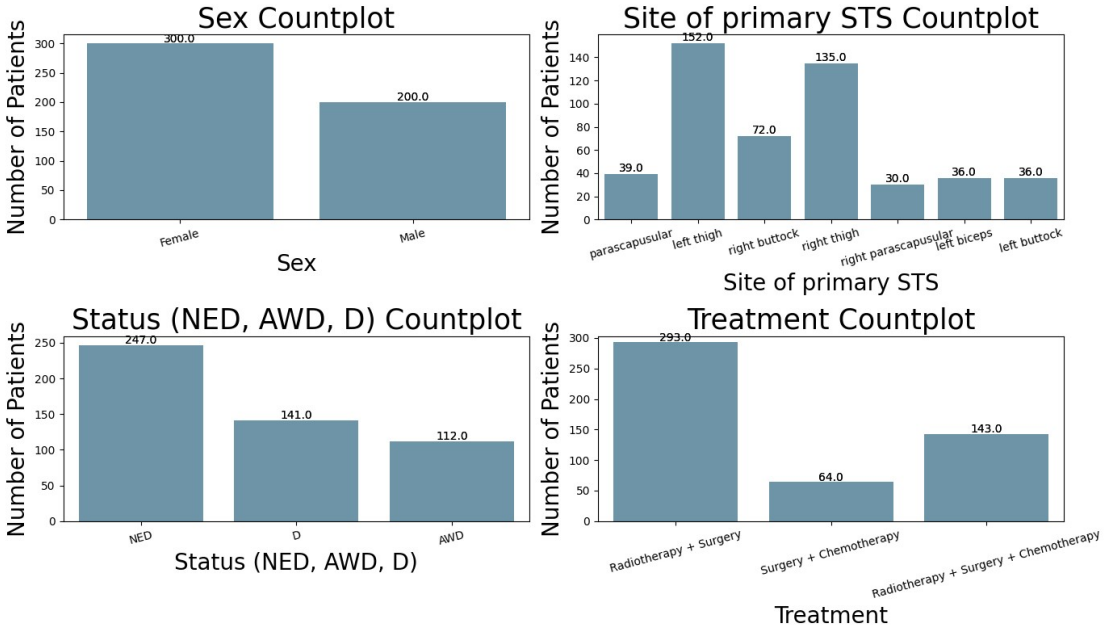


Figure 2. The count plots for each category in the bone tumor dataset applied in this study. (a) Sex, (b) Site of Primary STS (location of bone tumor), (c) Status, and (d) Treatments. The x-axis represents the dataset's attributes, and the y-axis shows the total counts. The number on the top of the bars exhibits the total counts.

Figure 3 compares the sliced histograms for the age distributions of the three statuses. Among the total deaths (counts: 141), patients aged 40 to 60 presented with the most deaths. However, for the total number (counts: 247) of patients with no evidence of disease (NED), patients between the ages of 60 and 80 showed the highest number, suggesting that some of them likely fully recovered from the bone tumor. Those who live with the disease (counts: 112) showed a similar age distribution to the Death histogram. Based on these trends, it may be inferred that persons aged 40 to 60 are more vulnerable to bone tumors, but if they survive, they do so into their 80s.

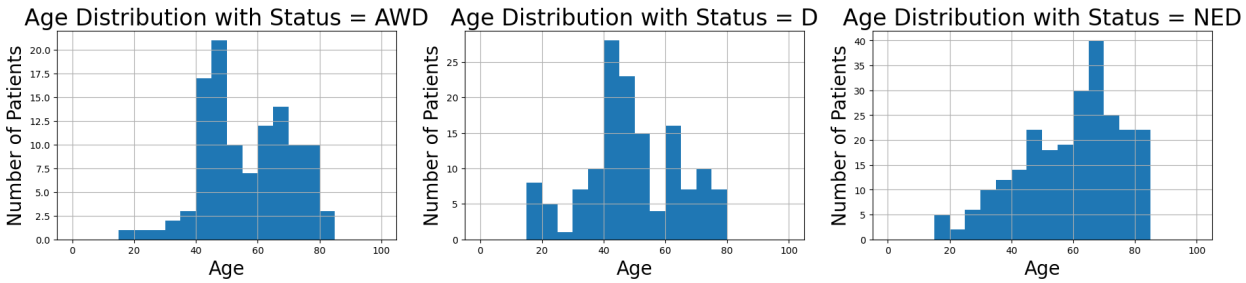


Figure 3. Age distribution histogram based on the three statuses: AWD (alive with disease), D (death), and NED (no evidence of disease). The X and Y axes present the patients' ages and total number of patients.

After evaluating the structural relationship between individual features (attributes), a direct comparison of the two bone tumor classifications was made. Figure 4 shows the number of patients based on the histological and MSKCC classifications. The histologic method (20) exhibits 13 different types of bone tumors. The most common type was pleomorphic leiomyosarcoma, which constituted the most number of patients compared to synovial sarcoma. Pleomorphic leiomyosarcoma constituted ~ 40%, and the others constituted ~ 60%. In contrast, MSKCC (21) shows the only three major types. MFH presented with the highest distribution, with over 50 more patients than leiomyosarcoma. Leiomyosarcoma and Synovial sarcoma afflicted approximately the same number of patients.

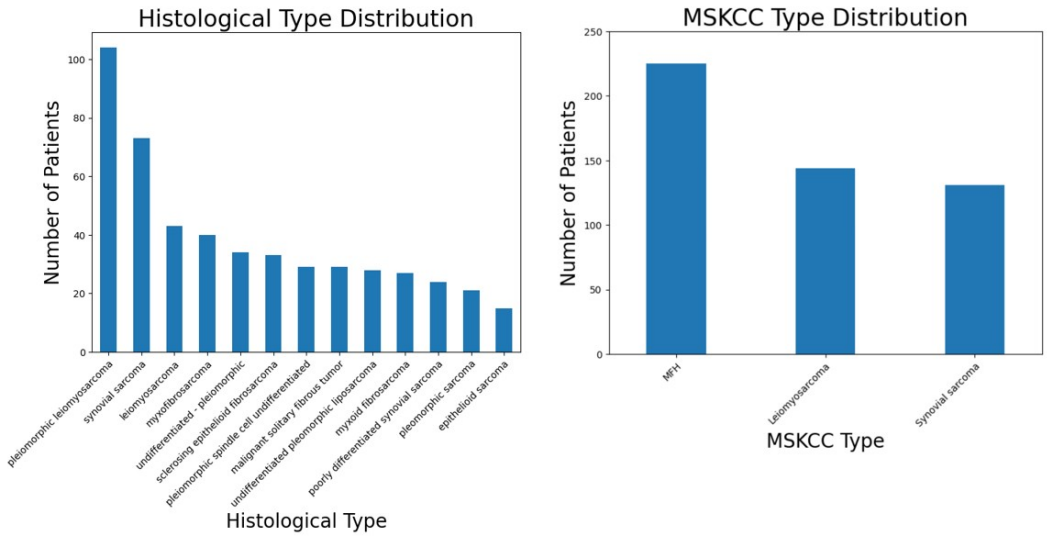


Figure 4. The number of patients bar plots for comparison of the histological and MSKCC bone tumor type classification. Histological bone tumor types are classified based on the cytological features with cell types (20). MSKCC-type classifications are structured along with patient care and oncology studies (21).

To understand the features that the MSKCC classification method uses (21), the number of patients related to the status was compared to those diagnosed using the histological method, as shown in Figure 5. As expected, the MSKCC yields a simple distribution. Leiomyosarcoma is predominant in the case of NED, and MFH is a prominent contributor to AWD. In contrast, 13 different bone tumor diseases are assorted for each status using the histological method, but the extent of individual importance appears vague and ambiguous. Noticeably, synovial sarcoma and pleiomorphic leiomyosarcoma are the most common cases in the D and NED, respectively. This comparison presents the simplicity of the classifying ability shown in the MSKCC method.

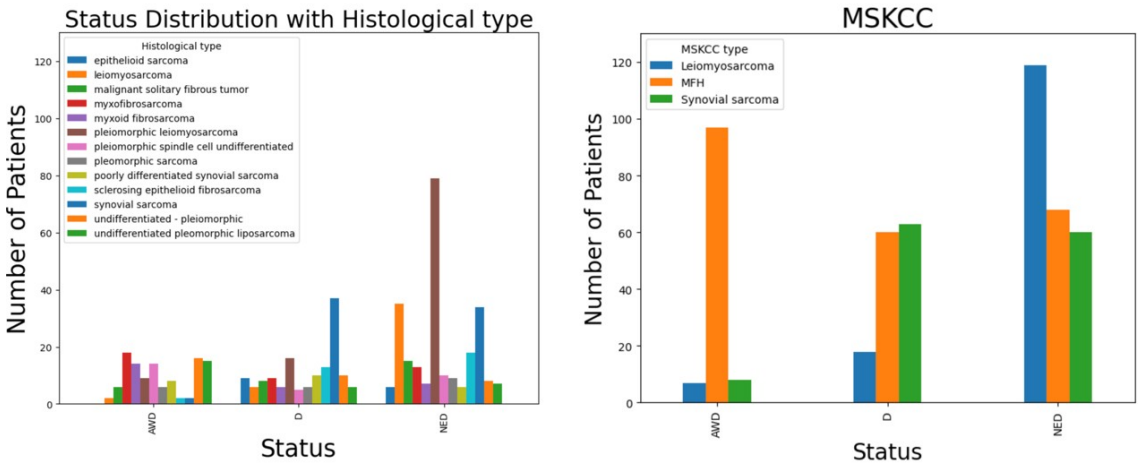
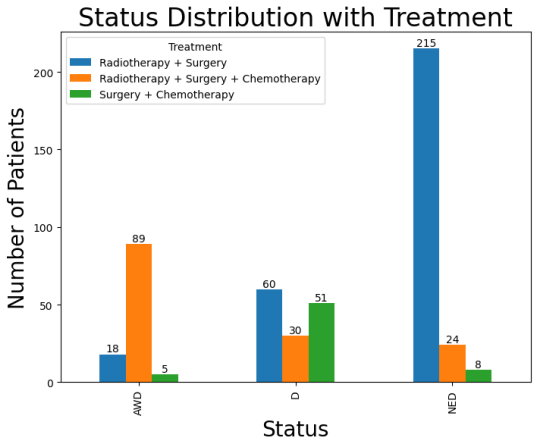


Figure 5. Comparison of the bone tumor patients' status for the histological and MSKCC classifications. The patients' numbers are plotted for AWD (alive with disease), D (death), and NED (no evidence of disease), respectively.



Treatment	Survival rate
Radiotherapy + Surgery	0.795
Radiotherapy + Surgery + Chemotherapy	0.790
Surgery + Chemotherapy	0.203

Figure 6. left) Survival rates related to the treatment. Right) The status distribution with treatments for the patients

Figure 6 shows the status associated with cancer treatments. Combining surgery with radiotherapy led to a more favorable survival rate, almost three times higher than when surgery was combined with chemotherapy. Unfortunately, this data shows the survival rate for the entire population of patients, regardless of the way the tumor is classified. Thus, it is hard to determine which treatment method is more effective based on a comparison of the survival rates of the two treatment methods.

Table 1. F1 score, Precision, and Recall values on the prediction for the two classification methods using supervised machine learning algorithms. (the range of the scale, the best: 1 for 100%, the worst: 0 for 0%)

		Histological type	MSKCC type
Decision Tree (DT)	F1 score	0.71	0.69
	Precision	0.72	0.75
	Recall	0.72	0.74
Random Forest (RF)	F1 score	0.71	0.72
	Precision	0.69	0.78
	Recall	0.70	0.74

Subsequent to the comparative analysis, we evaluated the prediction accuracy of the status (NED, AWD, D) for the two classification types using the decision tree (DT) and random forest (RF) supervised learning algorithms. Table 1 presents the resultant F1 scores, Precision, and Recall values. F1 score is the harmonic mean of precision and recall; a higher F1 score means the model has both a high precision and a high recall (19). Precision and recall are analogous to specificity and sensitivity. Specificity is a measure of how well a test can identify true negatives while sensitivity is a measure of how well a test can identify true positives. The DT model ranked the histological method of classification slightly higher in terms of the F1, precision and recall while the RF ranked the MSKCC classfication method slightly higher based on the same attributes, although the difference was not statistically significant. Conclusively, this predictive comparison analysis suggested that the expected result arrived *via* the MSKCC method would be equivalent to the result based on the histological method. The classifying MSKCC method was thus found likely comparable to the classification accuracy, sensitivity and specificity of the conventional histological method. Figure 7 presents the precision and recall for the two classifications using confusion matrices.

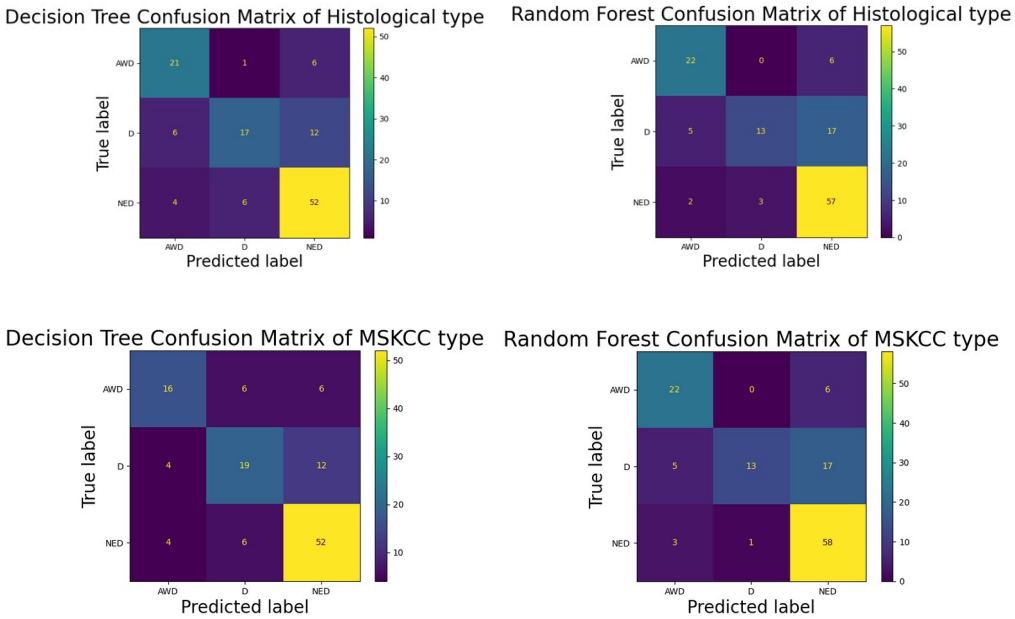


Figure 7. Comparison of the confusion matrix for the two classification methods (Top: the histological and bottom: MSKCC types). AWD (alive with disease), D (death), and NED (no evidence of disease), respectively.

5. Discussion

Current oncology research demands a new systematic classification platform that can assist in making decisions based on computation-based protocols. This claim has been further magnified with the advance of computation-based data analysis. MSKCC recently launched a dynamic and flexible cancer classification platform, OncoTree (21), which incorporated feedback from the cancer physician and research community. In this study, we attempted to make a point-by-point comparison of bone tumor classifications by visualizing the statistical variations of the MSKCC method against the conventional histological method. Subsequently, we compared the prediction accuracy of the expected status from the two different classifications with supervised machine learning algorithms.

In Fig. 5, the two classifications showed specific features for certain types of bone tumors regarding the classification status of the patient. For the NED status, leiomyosarcoma was distinguishable in the MSKCC classification method, while pleiomorphic leiomyosarcoma was primarily predominant in the histological method. Leiomyosarcoma is a morphologic variant of pleomorphic leiomyosarcoma, suggesting that the MSKCC incorporates a broad and straightforward range of classifying features. The two classification methods showed a distinct difference in the death status. MFH (malignant fibrous histiocytoma) and synovial sarcoma from the MSKCC method showed a similar contribution rate, but only synovial sarcoma was significant in the histological method. It may be that the MSKCC takes into consideration unknown parameters not disclosed in this dataset.

Unfortunately, this lack of information presented challenges in understanding how the MSKCC classification provides further simplified sorting capability covering a broad range of bone tumor types. Nevertheless, our statistical comparative analysis elucidated a rapid and easy distinguishable classifying feature of the MSKCC method, compared to the histological method.

Either radiotherapy with surgery or radiotherapy and surgery with chemotherapy significantly improved the survival rate, as presented in Figure 6. Since the detailed clinical information about the treatment was unknown, we could not evaluate the extent to which the treatments determined the status depending on the two classification methods. However, we predicted the accuracy that each classification could provide for the final status based on the provided features such as sex, age, grade, site of the primary STS, and treatment. We hypothesized that the expected accuracy should be commensurate with the capability and ability of each classifying method. As described in Table 1, the prediction accuracy of the MSKCC classification method was comparable to that of the histological method. Thus, the MSKCC classification was found to be equally accurate, sensitive and precise to classify bone tumors as the conventional hierarchical classification system. To some extent, the MSKCC method may be preferred to the histological method because it can provide a similar prediction accuracy, and more features can be incorporated as they become available (such as genetic data), potentially increasing the predicting power of this method. Bone tumors are rare in oncology. This allowed us to directly access the potential of a new classifying approach developed upon advanced computing technology. Our study shows that the new approach can effectively reflect the current needs of oncology care treatments for a high survival rate. The code for this study is linked in reference (24).

6. Conclusion

Cancer classification using the current histological method is cumbersome and onerous. Recently, Memorial Sloan Kettering Cancer Center (MSKCC) has launched a new simplified community-based oncology classification – OncoTree - based on a homogenized diagnosis after incorporating feedback from physicians, and the cancer research community and consortia. The MSKCC system simplifies classification in hard-to-classify heterogenous tumors and addresses a pressing and evolving need of accurate diagnosis. We aimed to evaluate whether the MSKCC classification resulted in similar categorization of cancers that were conventionally classified using the histological method; and just as importantly; if this simpler classification led to similar patient outcomes as those from the histological classification method. We used two Machine Language algorithms; Decision Tree and Random Forest; to predict patient status using both the classification methods. In this comparative analysis, we concluded that the MSKCC method possessed the classifying power equivalent to the histological method with the added advantage of much-simplified categorization. This new method to classify cancers can lead to improved cancer treatment strategies for favorable outcomes.

Acknowledgment

I would like to acknowledge Dr. Z Sung for advising on the design and analysis of this project.

7. References

1. "Cancer Classification", National Cancer Institute, SEER training Modules, https://training.seer.cancer.gov/disease/categories/classification.html, and "Cancer Staging", https://www.cancer.gov/about-cancer/diagnosis-staging/staging

2. International Classification of Disease for Oncology, 3rd edition, ICD-O-3.2, World Health Organization, updated in 2019. https://www.who.int/standards/classifications/other-classifications/international-classification-of-diseases-for-oncology

3. Carbone A. Cancer Classification at the Crossroads. Cancers.12(4), 980, 2020. https://doi.org/10.3390/cancers12040980

4. Gupta S, Gupta MK, Shabaz M, Sharma A. Deep learning techniques for cancer classification using microarray gene expression data. Front Physiol. 13, 952709, 2022. https://doi.org/10.3389/fphys.2022.952709

5. Franchi A. Epidemiology and classification of bone tumors. Clin Cases Miner Bone Metab. 9(2), 92-5, 2012. https://pmc.ncbi.nlm.nih.gov/articles/PMC3476517/

6. Cancer Stat Facts: Bone and Joint Cancer, NIH, National Cancer Institute, Surveillance, Epidemiology, and End Results Program, https://seer.cancer.gov/statfacts/html/bones.html

7. Bone Cancer, Memorial Sloan Kettering Cancer Center, https://www.mskcc.org/cancer-care/types/bone

8. Choi JK, Ro JY. The 2020 WHO Classification of Tumors of Bone: An Updated Review. Advances In Anatomic Pathology. 28(3), 119-138, 2021. https://doi.org/10.1097/pap.0000000000000293

9. Bartelstein MK, Boland PJ. Fifty years of bone tumors. J Surg Oncol. 126(5), 906-912, 2022. https://doi.org/10.1002/jso.27027

10. Bone marrow cancer, NIH, National Cancer Institute, NCI Dictionary of Cancer Terms, https://www.cancer.gov/publications/dictionaries/cancer-terms/def/bone-marrow-cancer

11. Gibbs CP, Weber K, ; Scarborough MT. Malignant Bone Tumors. The Journal of Bone and Joint Surgery 83(11), 1728-1745, 2001. https://pubmed.ncbi.nlm.nih.gov/12064130/

12. Kindblom LG. Bone tumors: epidemiology, classification, pathology. Eds. Davies AM, Sundaram M, James SLJ. Imaging of bone tumors and tumor-like lesions: techniques and applications, Springer Berlin, Heidelberg, Germany, 2009, 1-15. https://link.springer.com/chapter/10.1007/978-3-540-77984-1_1

13. He Y, Pan I, Bao B, Halsey K, Chang M, Liu H, et al. Deep learning-based classification of primary bone tumors on radiographs: A preliminary study. EBioMedicine. 62,103121, 2020. https://doi.org/10.1016/j.ebiom.2020.103121

14. Mohi AMB, Alzahrani AA, Reda BR, A case report of Femur Osteochondroma in 22 years old female patient. International journal of advanced research. 8(10), 1263-1267, 2020. http://dx.doi.org/10.21474/IJAR01/xxx

15. Bickels J, Robert CK, Eller I, Malawer MM, Cryosurgery in the treatment of bone tumors. Operative Techniques in Orthopaedics. 9(2), 79-83, 1999. https://doi.org/10.1016/S1048-6666(99)80025-1

16. MSK Data Catalog, Memorial Sloan Kettering Cancer Center, https://datacatalog.mskcc.org/

17. Bone Tumor, https://www.kaggle.com/datasets

18. Google Python web-based platform, google colab, https://colab.research.google.com/

19. Pandas Python data analysis library, https://pandas.pydata.org/

20. Bertoni F, Bacchini P, Classification of bone tumors. European Journal of Radiology. 27, Supplement 1, S74-S76, 1998. https://doi.org/10.1016/S0720-048X(98)00046-1.

21. Kundra R, Zhang H, Sheridan R, Sirintrapun SJ, Wang A, Ochoa A, et al., OncoTree: A Cancer Classification System for Precision Oncology. JCO Clin Cancer Inform. 5:221-230, 2021. https://doi.org/10.1200/cci.20.00108

22. Scikit-lean machine learning program, https://scikit-learn.org/stable/index.html

23. Podder P, Bharati S, Mondal MRH, Kose U. Application of machine learning for the diagnosis of COVID-19. Eds. Kose U, Gupta D, de Albuquerque VHC, Khanna A. Data Science

for COVID-19, Academic Press, Cambridge, MA, USA, 2021, 175-194, https://doi.org/10.1016/B978-0-12-824536-1.00008-3

24. https://colab.research.google.com/drive/1iZ9kn5MDu-qHSXBJQvizoidj9qzmaSUM?usp=sharing