Journal of Pathology Informatics 15 (2024) 100368

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Utility of artificial intelligence in a binary classification of soft tissue tumors

Jing Di [a], Caylin Hickey [a], Cody Bumgardner [a], Mustafa Yousif [b], Mauricio Zapata [c], Therese Bocklage [a], Bonnie Balzer [d], Marilyn M. Bui [e], Jerad M. Gardner [f], Liron Pantanowitz [g], Shadi A. Qasem [a,h,*]

[a] University of Kentucky College of Medicine, Lexington, KY, United States
[b] University of Michigan, Ann Arbor, MI, United States
[c] Northside Hospital, Georgia, United States
[d] Cedars-Sinai Medical Center, Los Angeles, CA, United States
[e] Moffitt Cancer Center & Research Institute, Tampa, FL, United States
[f] Geisinger Medical Center, Danville, PA, United States
[g] University of Pittsburgh Medical Center, Pittsburgh, PA, United States
[h] Baptist Health Jacksonville, Jacksonville, FL, United States

ARTICLE INFO

Keywords:
Artificial intelligence
Digital pathology
Sarcoma
Soft tissue tumors
Whole-slide images
Deep learning
Diagnosis

ABSTRACT

Soft tissue tumors (STTs) pose diagnostic and therapeutic challenges due to their rarity, complexity, and morphological overlap. Accurate differentiation between benign and malignant STTs is important to set treatment directions, however, this task can be difficult. The integration of machine learning and artificial intelligence (AI) models can potentially be helpful in classifying these tumors. The aim of this study was to investigate AI and machine learning tools in the classification of STT into benign and malignant categories. This study consisted of three components: (1) Evaluation of whole-slide images (WSIs) to classify STT into benign and malignant entities. Five specialized soft tissue pathologists from different medical centers independently reviewed 100 WSIs, representing 100 different cases, with limited clinical information and no additional workup. The results showed an overall concordance rate of 70.4% compared to the reference diagnosis. (2) Identification of cell-specific parameters that can distinguish benign and malignant STT. Using an image analysis software (QuPath) and a cohort of 95 cases, several cell-specific parameters were found to be statistically significant, most notably cell count, nucleus/cell area ratio, nucleus hematoxylin density mean, and cell max caliper. (3) Evaluation of machine learning library (Scikit-learn) in differentiating benign and malignant STTs. A total of 195 STT cases (156 cases in the training group and 39 cases in the validation group) achieved approximately 70% sensitivity and specificity, and an AUC of 0.68. Our limited study suggests that the use of WSI and AI in soft tissue pathology has the potential to enhance diagnostic accuracy and identify parameters that can differentiate between benign and malignant STTs. We envision the integration of AI as a supportive tool to augment the pathologists' diagnostic capabilities.

Introduction

Soft tissue tumors (STT) represent a complex diagnostic area within oncology due to their rarity and heterogeneity. The diagnostic process is fraught with challenges, often resulting from the limited availability of specialized expertise and the broad spectrum of tumor subtypes, which can result in delayed or incorrect diagnoses.[1] A study reviewing second opinion diagnosis in STT found a discordance rate of 38%, with 25% of these cases being classified as major diagnostic error impacting patient management. This has also been shown to increase the number and cost of malpractice cases in sarcoma care.[2,3] Emerging technologies such as digital pathology and radiomics show promise in improving the accuracy of cancer diagnosis, characterization, and monitoring.[4–8]

The utilization of whole-slide images (WSIs) has made it easier to obtain consultations, even across different institutions.[9–11] Soft tissue pathology, particularly, presents significant challenges for most pathologists, often necessitating expert consultations. Molecular techniques, including next-generation sequencing, have contributed significantly to better classification of certain soft tissue sarcomas.[12] However, getting a timely and accurate initial working diagnosis mainly based on histology will expedite the initiation of appropriate management; hence the need for additional tools to aid in that initial assessment.

Recent advancements in digital pathology and artificial intelligence (AI) have begun to show potential in enhancing diagnostic precision. Studies such as that by Foersch et al[13] have illustrated significant advancements in the performance of AI-assisted diagnoses in soft tissue sarcoma,

* Corresponding author at: Baptist Health Jacksonville, 800 Prudential Drive, Jacksonville, FL 32256, USA.
E-mail address: s.qasem@uky.edu (S.A. Qasem).

http://dx.doi.org/10.1016/j.jpi.2024.100368
Received 2 December 2023; Received in revised form 25 January 2024; Accepted 9 February 2024
Available online 15 February 2024

emphasizing the progressive nature of this field. This underscores the need for ongoing research to refine AI applications in pathology, ensuring they are robust across various studies and datasets. Distinguishing between benign/reactive processes and malignant ones is a complex task in soft tissue pathology. The large number of entities involved, and the rarity of these tumors make it challenging to apply existing machine learning and AI models to this field.[4,6,13]

The current investigation aims at exploring the application of AI in the classification of STT into benign and malignant entities, as a step towards the incorporation of AI into the clinical workflow. The goals of this study are: (1) identify cell-specific parameters that can aid in the classification of STTs as benign vs malignant, and (2) explore the capabilities of AI, in comparison to expert pathologists, in evaluating benign and malignant STT. By addressing these goals, the study aims to develop AI techniques that aid in accurately diagnosing STT.

Methods

For all the experiments below, the ground truth was the original diagnosis obtained utilizing all glass slide, immunostains and molecular techniques, if necessary (Table 4).

Expert pathologist review (study arm 1)

Five soft tissue pathologists from five different medical centers independently reviewed 100 WSI of hematoxylin and eosin (H&E)-stained slides representing 100 different soft tissue cases, following the institutional research protocol approved by the Institutional Review Board. Only one slide per case was provided with limited clinical information (patient age, gender, and anatomic location). Immunohistochemical and molecular information was not provided to the reviewing pathologists. Deidentified slides were scanned at 20 × magnification using an Aperio scanner (Aperio AT2, Leica Biosystems, Illinois), and acquired digital files were converted from .svs to a DICOM format. They were uploaded to a locally hosted, externally accessible compute node for review using a web-based WSI viewer system (Orthanc v1.3.2 WSI Plugin v0.5). Pathologists were asked to choose one of four diagnostic categories (benign, intermediate/borderline, malignant, and uncertain) and to provide up to three differential diagnoses. A REDCap survey (https://www.project-redcap.org/) was used to record the answers.

Answers were compared to the original "ground-truth" diagnoses. Major and minor discordances, as well as uncertainty were defined as follows. Major discordance referred to discrepancies that could change patient management or prognosis, such as mistaking a benign tumor for a malignant one, or vice versa. Minor discordance involved differences that would not affect overall treatment, such as subclassification within the same category of benign or malignant tumors. Uncertainty was used for cases where a definitive diagnosis could not be reached due to insufficient information or ambiguous histological features.

Identifying cell specific parameters (study arm 2)

A cohort of 95 STT cases was utilized from the "expert pathologist review" experiment described above, including 60 benign and 35 malignant cases, encompassing 68 distinct STT entities. The cases were scanned at 40 × magnification using a high-throughput scanner (Aperio AT2) and uploaded to an OMERO server[14] for annotation. The regions of interest (ROIs) were marked on all slides by the pathologist (Fig. 1). At the highest layer of resolution, the slide images were divided into 768-pixel tiles (Fig. 2). Each tile was evaluated for tissue percentage and color factors using pre-processing software developed as part of the Deep HistoPath project (https://github.com/CODAIT/deep-histopath). The metadata associated with tile and cell metrics was used to select the top 500 tiles per case within ROI. Cell detection was performed on each tile with at least 25% tissue using QuPath,[15] providing 38 cell-specific parameters, which were averaged per tile. The cell-based metrics were averaged both across the top tiles per case and at the case-level. Welch's t test was used, with $p < 0.05$ considered significant.

Employing an AI model to differentiate benign vs malignant soft tissue tumors (study arm 3)

A total of 195 soft tissue cases were collected from the files of one of the authors (SQ), encompassing the 95 cases from the "cell-specific parameters" experiment described above. These cases were divided into a training group (156) and validation group (39). A free software machine learning

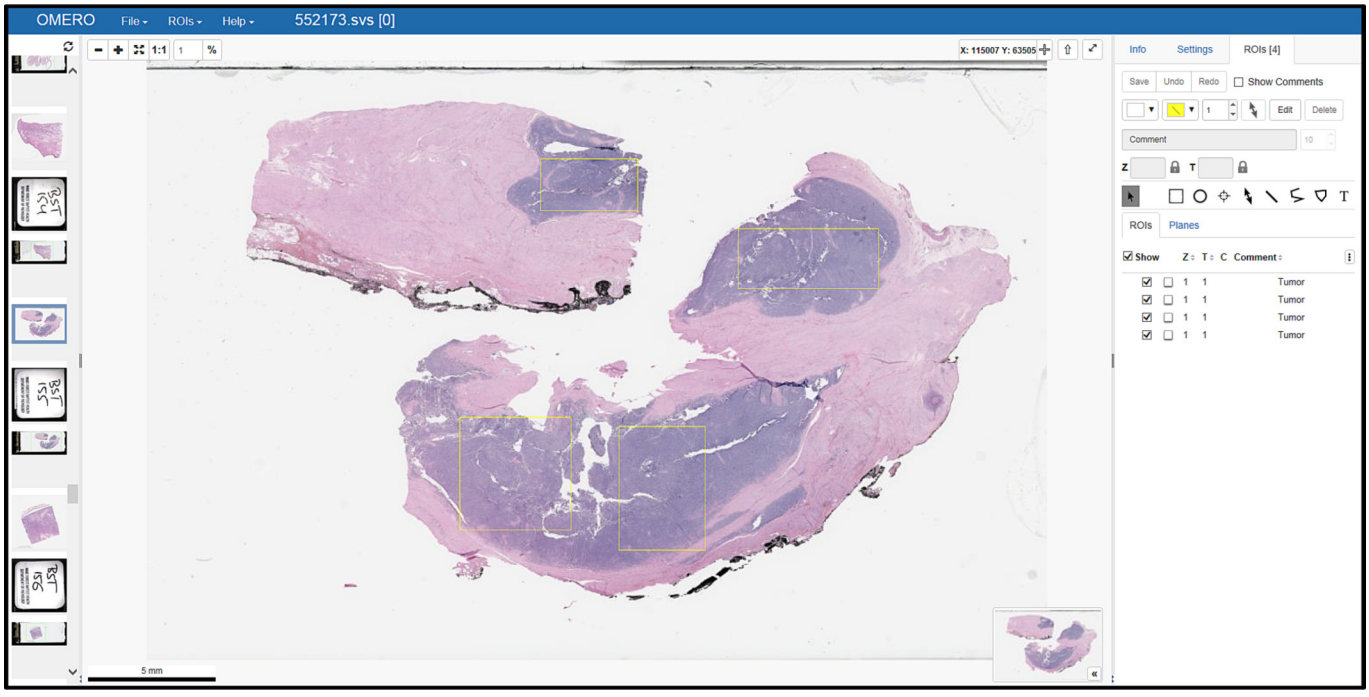

Fig. 1. Representative whole-slide image uploaded to OMERO for annotation of ROI.

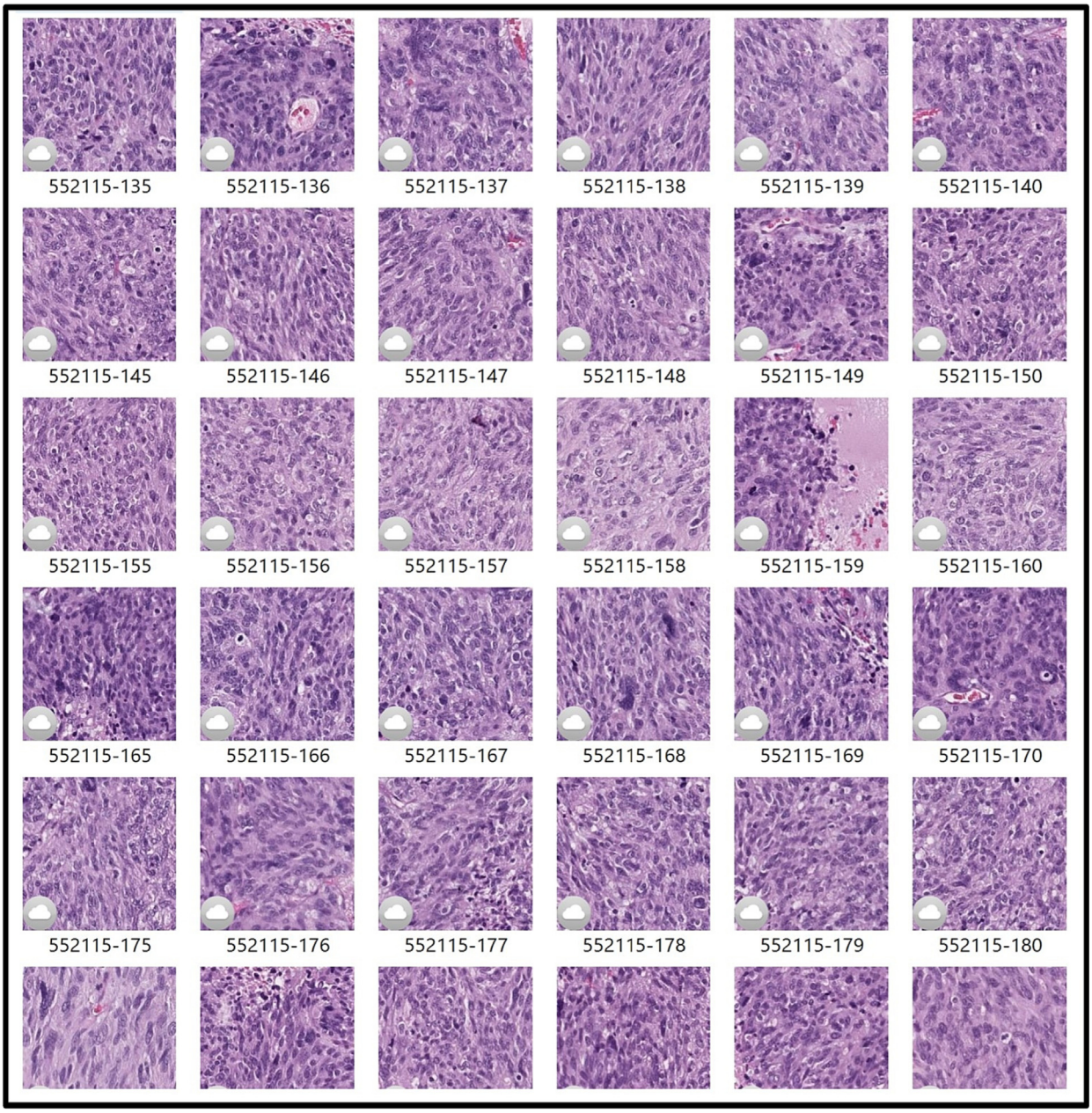

Fig. 2. Representative case divided into 768-pixel tiles at the highest layer of resolution. 38 cell-specific parameters were detected using QuPath, which were averaged per tile.

library for Python programming language (Scikit-learn) was employed to analyze these cases into benign and malignant STT (https://scikit-learn.org/stable/). Scikit-learn, an open-source machine learning library for Python programming language, offers a wide range of classification, regression, and clustering algorithms, including support-vector machines, random forests, and gradient boosting.

Results

Pathologist diagnostic concordance rate using WSI

The concordance rate of the pathologists when assessed against the reference diagnosis, made using traditional microscope, was 70.4% across the four diagnostic categories. In detail, minor discordances were observed in 11.6% of the cases, where the pathologists' diagnoses were close but not exactly the same as the reference. Major discordances, wherein the diagnoses substantially differed, occurred in 5% of the cases. Additionally, there was a 13% rate of uncertainty where the pathologists could not reach a definitive diagnosis (Table 1). Analyzing the cases further, malignant tumors were most accurately diagnosed with an 81.7% concordance rate, whereas benign cases exhibited the highest rate of major discordance at 7.7%. Intermediate cases showed the highest rates of minor discordance (28%) and uncertainty (22.4%) (Fig. 3). Overall, a correct differential diagnosis was established in 63% of the cases. Additionally, the pathologists reported either excellent or satisfactory quality of the scanned images for 96.4% of the cases.

Cell-specific parameters and their significance in benign vs malignant STT

Within the scope of 95 cases, encompassing 52 females and 43 males aged from 9 months to 90 years (with an average age of 41 and a median of 42), image analysis highlighted several cell-specific parameters that showed statistically significant differences ($p < 0.05$). These parameters included cell count, nucleus/cell area ratio, nucleus hematoxylin OD* mean,

Table 1
Summary of study findings (study arm 1). "Concordance" indicates agreement with the reference diagnosis with regards to the diagnostic categories (benign, intermediate/borderline, and malignant). Correct diagnosis is recorded when one of the differentials provided by the pathologist matches the reference diagnosis (e.g., myxoid liposarcoma). Please note that this arm of the study was performed using one slide only, no ancillary testing, and limited clinical data.

Pathologists (P)	Overall concordance rate (%)	Major discordance (%)	Minor discordance (%)	Uncertain diagnosis (%)	Correct DX (%)
P1	70	5	10	15	66
P2	71	3	8	18	70
P3	70	4	9	17	63
P4	76	9	15	0	37
P5	65	4	16	15	79
Overall	70.4	5	11.6	13	63

cell max caliper, cell area, cell perimeter, cell circularity, and cell min caliper (Table 2). The Welch's t-test confirmed significant distinctions in the mean values of these parameters between benign and malignant groups, suggesting their potential utility in creating machine learning models for aiding with a soft tissue diagnosis.

AI model performance relative to expert pathologists

The machine learning models that achieved the best performance in distinguishing benign from malignant STTs included gradient boosting, neural network, xgboost, random forest, bagging, histgradientboosting, sgdclassifier, and logistic regression. For instance, logistic regression exhibited a sensitivity of 0.737 and a specificity of 0.8, whereas random forest showed a sensitivity of 0.864 and a specificity of 0.75. The average sensitivity among these models was 0.60, specificity was 0.75, and overall accuracy was 0.68. The area under the receiver operating characteristic curve (AUC) for these models was also 0.68 (Table 3). These results indicate that the performance of the AI models is on par with that of the expert pathologists, suggesting a promising role for AI in supporting diagnostic processes in soft tissue pathology.

Discussion

Accurate diagnosis of STT and their subtypes is crucial in determining effective personalized oncology treatment plans for the best patient outcomes. There is scarcity of studies focusing on diagnostic discrepancies, especially in cases that have received second opinion in soft tissue pathology.[1,16] One notable study examining the diagnosis of sarcoma through histopathology review revealed a substantial 24% discordance in diagnoses between community pathologists and an expert sarcoma reference pathology group. Sixty-six percent of these discordant cases had clinically significant implications for treatment recommendations.[16] Interestingly, for all major discordant cases, excluding non-mesenchymal lesions, the diagnosis could have been made through conventional H&E-stained slides. The primary reason for diagnostic errors was the limited experience of non-specialized surgical pathologists with uncommon and atypical neoplasms.

Whole-slide imaging has emerged as a powerful tool for enhancing the care of cancer patients, accelerated by the COVID-19 pandemic and the CMS approval of remote sign out, thereby increasing the adoption of the technology.[17] This technology enables timely assessment of tumor tissue by experts in STT, it fosters collaboration among specialists in sarcoma management, and it provides pathologists in underserved regions, without access to sarcoma centers, the opportunity to consult with a sarcoma specialist.[18] Previous studies have also highlighted the utility of WSI technology in diagnosing STT. For instance, Sargen et al reported a notable diagnostic accuracy of 89% using WSI for STT by two experienced soft tissue pathologists.[19] In another study, nine pathologists, with different levels of expertise, assessed 291 STT using WSI, and demonstrated a substantial increase in accuracy from 46.3% (±15.5%) to 87.1% (±11.1%) with the assistance of deep machine learning.[13] These findings emphasize the pivotal role of specialized pathology assessment in sarcoma diagnosis, and the need to leverage WSI for this purpose.

The expert review experiment in this project aimed to establish a baseline of what is achievable with limited clinical information and lack of ancillary studies, for comparison with an AI-assisted scenario using H&E-stained slides only. It is important to recognize that a comprehensive diagnostic evaluation necessitates the review of all case slides,

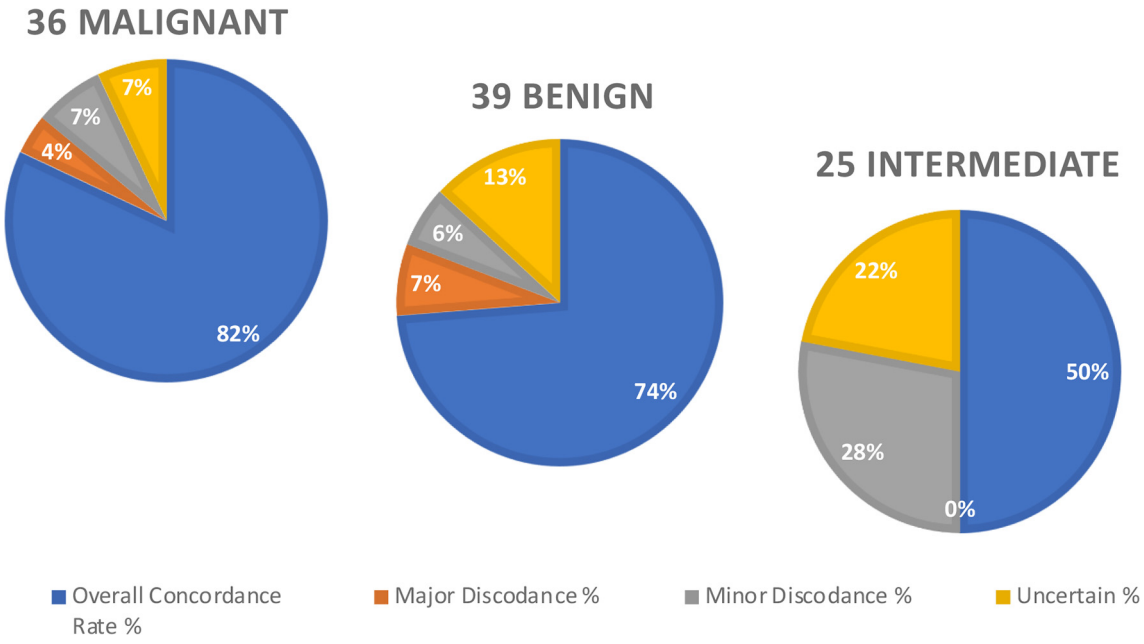

Fig. 3. Expert pathologists' review (study arm 1). This was performed using one slide only, no ancillary testing, and limited clinical data. Malignant cases had the highest concordance rate (82%), benign cases had the highest major discordance rate (7%), and intermediate cases had the highest minor discordance (28%) and uncertainty rates (22%).

Table 2
Cell-specific parameters of soft tissue tumors and their *p* values (Study arm 2).

	Mean		Difference between means ± SEM	95% confidence interval	*p* value
	0	1			
Cases (n)	60	35			
Nucleus/Cell area ratio	0.218	0.2731	0.05504 ± 0.008515	0.03804 to 0.07204	**<0.0001**
Cell count	171.9	278.9	107.0 ± 18.51	70.08 to 143.9	**<0.0001**
Nucleus area	28.41	29.7	1.287 ± 0.7150	−0.1327 to 2.708	0.075
Nucleus perimeter	23.16	23.21	0.04822 ± 0.3044	−0.5563 to 0.6527	0.8745
Nucleus: Circularity	0.665	0.6748	0.009753 ± 0.008050	−0.006236 to 0.02574	0.2288
Nucleus: Max caliper	8.984	8.748	−0.2359 ± 0.1266	−0.4874 to 0.01556	0.0656
Nucleus: Hematoxylin OD mean	0.3127	0.3741	0.06141 ± 0.02067	0.02032 to 0.1025	**0.0039**
Cell: Max caliper	16.33	14.95	−1.381 ± 0.2853	−1.949 to −0.8130	**<0.0001**
Nucleus: Eccentricity	0.809	0.7885	−0.02050 ± 0.005593	−0.03161 to −0.009397	**0.0004**
Nucleus: Min caliper	4.579	4.854	0.2756 ± 0.06869	0.1391 to 0.4120	**0.0001**
Nucleus: Hematoxylin OD sum	36.71	47.15	10.44 ± 2.362	5.738 to 15.15	**<0.0001**
Nucleus: Hematoxylin OD std dev	0.1034	0.1144	0.01096 ± 0.004120	0.002775 to 0.01914	**0.0092**
Nucleus: Hematoxylin OD max	0.5784	0.6721	0.09370 ± 0.02779	0.03846 to 0.1489	**0.0011**
Nucleus: Hematoxylin OD min	0.1047	0.1376	0.03292 ± 0.01320	0.006608 to 0.05923	**0.0149**
Nucleus: Hematoxylin OD range	0.4736	0.5344	0.06078 ± 0.01833	0.02436 to 0.09719	**0.0013**
Nucleus: Eosin OD mean	0.2425	0.2352	−0.007304 ± 0.01187	−0.03094 to 0.01633	0.5401
Nucleus: Eosin OD sum	27.54	28.55	1.012 ± 1.403	−1.803 to 3.827	0.4738
Nucleus: Eosin OD std dev	0.06607	0.0649	−0.001174 ± 0.003765	−0.008651 to 0.006303	0.7559
Nucleus: Eosin OD max	0.3925	0.3862	−0.006242 ± 0.01772	−0.04147 to 0.02899	0.7255
Nucleus: Eosin OD min	0.08946	0.08085	−0.008608 ± 0.006045	−0.02072 to 0.003505	0.1601
Nucleus: Eosin OD range	0.303	0.3054	0.002360 ± 0.01477	−0.02697 to 0.03169	0.8734
Cell: Area	135.1	110.7	−24.42 ± 5.009	−34.38 to −14.46	**<0.0001**
Cell: Perimeter	44.33	40.5	−3.831 ± 0.7991	−5.422 to −2.240	**<0.0001**
Cell: Circularity	0.8166	0.7968	−0.01988 ± 0.004069	−0.02796 to −0.01180	**<0.0001**
Cell: Min caliper	10.88	9.898	−0.9780 ± 0.2126	−1.401 to −0.5548	**<0.0001**
Cell: Eccentricity	0.6993	0.7058	0.006489 ± 0.003267	1.171e-006 to 0.01298	**0.05**
Cell: Eosin OD mean	0.179	0.1718	−0.007139 ± 0.009776	−0.02658 to 0.01230	0.4672
Cell: Eosin std dev	0.08495	0.08461	−0.0003421 ± 0.005141	−0.01056 to 0.009875	0.9471
Cell: Eosin OD max	0.4168	0.4017	−0.01506 ± 0.01941	−0.05364 to 0.02351	0.4397
Cell: Eosin OD min	0.01473	0.007824	−0.006906 ± 0.002861	−0.01259 to −0.001224	**0.0178**
Cytoplasm: Hematoxylin OD mean	0.08647	0.122	0.03549 ± 0.009522	0.01644 to 0.05454	**0.0004**
Cytoplasm: Hematoxylin OD std dev	0.06128	0.07984	0.01855 ± 0.004686	0.009220 to 0.02789	**0.0002**
Cytoplasm: Hematoxylin OD max	0.3381	0.4205	0.08239 ± 0.02081	0.04094 to 0.1238	**0.0002**
Cytoplasm: Hematoxylin OD min	−0.01809	−0.00513	0.01296 ± 0.004507	0.003989 to 0.02194	**0.0052**
Cytoplasm: Eosin OD mean	0.1613	0.1477	−0.01362 ± 0.009514	−0.03254 to 0.005288	0.1558
Cytoplasm: Eosin OD std dev	0.07641	0.074	−0.002417 ± 0.004904	−0.01216 to 0.007325	0.6233
Cytoplasm: Eosin OD max	0.3729	0.35	−0.02283 ± 0.01892	−0.06043 to 0.01477	0.2309
Cytoplasm: Eosin OD min	0.01692	0.009997	−0.006924 ± 0.002960	−0.01280 to −0.001046	**0.0215**

The bolded entities represent statistically significant values.

along with access to clinical information, imaging data, and often additional tests like immunohistochemistry and molecular studies.[20] In particular, intermediate (borderline) lesions continue to pose challenges in classification, underscoring the need for supplementary tools in the diagnostic process.

This study identified several cell-specific parameters as being statistically significant ($p<0.05$) in distinguishing between benign and malignant STT. Recent research has also shown the potential of nuclear morphology as a deep learning biomarker for cellular senescence, which can be applied to cancer.[21] In another study, deep learning algorithms significantly enhanced pathologists' accuracy in diagnosing leiomyosarcomas and predicting outcomes, increasing accuracy from 46.3% to 87.1%.[8] These findings suggest that cell-specific parameters hold promise in aiding pathologists in distinguishing benign from malignant STT and hence improving diagnostic accuracy.

Table 3
Representative AI models metrics, classifying benign and malignant soft tissue tumors (Study arm 3).

Metric	AUC	Accuracy	Sensitivity	Specificity
Gradientboosting	0.664	0.667	0.579	0.75
Neuralnetwork	0.638	0.641	0.526	0.75
Xgboost	0.639	0.641	0.579	0.7
Randomforest	0.717	0.718	0.684	0.75
Bagging	0.666	0.667	0.632	0.7
Tabpfn	0.743	0.744	0.737	0.75
Histgradientboosting	0.664	0.667	0.579	0.75
Sgdclassifier	0.584	0.59	0.368	0.8
Logisticregression	0.768	0.769	0.737	0.8

In this study, Scikit-learn[22] AI software, was utilized, employing gradient boosting, neural networks, xgboost, and logistic regression algorithms. The successful classification of benign and malignant STT, in our study, underscores the potential of an AI-based approach to enhance diagnostic accuracy in soft tissue pathology. A major limitation of this study is the relatively small number of cases analyzed and the inclusion of a large number of STT entities. Unfortunately, this is an expected challenge in studies involving STT, as the majority of these tumors are rare, and the differential diagnosis can be quite broad. In addition, the boundaries between benign, reactive, intermediate-, and low-grade malignancy can be blurry. Another limitation is the lack of validation using an independent dataset.

Despite the aforementioned success yielded in our pilot study, it is crucial to acknowledge that the role of AI in STT diagnosis and treatment is still premature, and that further research is imperative to validate its practical applicability in the clinical setting. Nevertheless, we believe that expanding the dataset with more cases has the potential to significantly improve the performance of the AI model. Finally, the implementation of AI in pathology mandates careful consideration of various technical and ethical aspects, such as patient data privacy, data interpretability, model transparency, and potential biases.

Table 4
List of cases used for all study arms (Study arms 1, 2, and 3). Cases with * are not considered soft tissue tumors but have either presented in a location and/or showed morphological overlap with soft tissue tumors warranting them being submitted to the soft tissue service for consultation.

Case	Age/Sex	Location	Diagnosis
1	45 F	Esophagus	Ewing/PNET
2	88 F	Chest wall	Elastofibroma
3	3m M	Flank	Kaposiform hemangioendothelioma
5	8m M	Thigh	Juvenile xanthogranuloma
6	63 F	Vulva	Angiomyofibroblastoma
7	53 M	Leg	Cutaneous leiomyosarcoma
8	54 M	Intestinal	Malignant GIST
9	26 M	Rectum	*Balloon cell melanoma
11	31 M	Thumb	Perineurioma
12	49 M	Mediastinal	*Type-A thymoma
14	34 F	Retroperitoneal	Ewing/PNET
15	76 M	Elbow	*Late stage erythema elevatum diutinum
17	42 M	Arm	Nodular fasciitis
20	78 M	Shoulder	Ischemic fasciitis
21	95 F	Intra-abdominal	*Granulosa cell tumor
22	15 F	Subcutaneous	Malignant giant cell tumor of soft parts
26	42 M	Hip	Schwannoma
27	45 F	Arm	MPNST arising in a neurofibroma
28	36 M	Ankle	Clear cell sarcoma
29	2 F	Thigh	Sclerosing rhabdomyosarcoma
31	57 F	Vagina	Benign genital stromal polyp
34	50 F	Thigh	Biphasic synovial sarcoma, grade 2
35	55 M	Shoulder	Pleomorphic rhabdomyosarcoma, high grade
36	66 F	Scapula	HPC/SFT with malignant potential
38	79 F	Groin	*Metastatic melanoma
39	13 M	Leg	DFSP
40	47 M	Retroperitoneal	Schwannoma
42	7 F	Thigh	Pleomorphic sarcoma with giant cells (malignant giant cell tumor of soft parts)
45	8 M	Shoulder	Granular cell tumor
46	62 M	Nose	Well differentiated fibrosarcoma (grade 1)
48	29 M	Preauricular	solitary fibrous tumor/hemangiopericytoma
49	16 M	Tongue	Granular cell tumor
51	19 M	Intra-ventricular	Hemangioma, cavernous/capillary type
52	59 F	Pelvis	*High grade carcinosarcoma
53	49 F	Uterus	Epithelioid leiomyoma
56	24 M	Trunk	Dermatofibrosarcoma protuberans
58	38 F	Foot	Marked Stasis changes
60	42 F	Uterine serosa	Mesenchymal tumor, favor unusual smooth muscle tumor of uncertain malignant potential
63	78 M	Arm	Kaposi sarcoma
68	66 M	Abdomen	*Diffuse follicle center lymphoma
69	8 F	Small bowel	Reactive changes
70	39 F	Forearm	Sarcoma with myofibroblastic features
74	55 F	Axillary	*Malignant S100 positive tumor, favor melanoma
77	23 F	Retroperitoneal	Angiolymphoid hyperplasia (with eosinophilia)
78	59 M	Back	*Adnexal neoplasm of at least low grade malignancy
79	52 M	Arm	Myxoid variant of hemangiopericytoma/SFT
80	66 M	Parotid	Pleomorphic sarcoma with myoid differentiation
81	42 F	Arm	Epithelioid sarcoma
82	25 M	Periumbilical	Myoepithelial tumor of soft tissue, histologically benign
84	69 F	Omentum	*Low grade endometrial stromal sarcoma
86	42 M	Mesenteric	Follicular dendritic cell tumor
87	52 F	Breast	Histiocytic tumor of uncertain malignant potential
90	1 F	Abdominal wall	Juvenile xanthogranuloma
92	6 F	Finger	Cellular juvenile aponeurotic fibroma
93	84 M	Temple	*Ulcerating carcinoma
94	20 F	Chest wall	Monophasic synovial sarcoma, high grade
95	5 F	R foot	Fibrous histiocytoma
97	41 F	R tongue	Myofibroma
99	67 M	Toe	Low grade pleomorphic sarcoma
103	8 M	Hand	*Cellular blue nevus
104	30 M	Shoulder	*Soft tissue chordoma with bone erosion
105	42 F	Lower abdomen	*Malignant tumor c/w myoepithelial carcinoma, high grade
107	1 M	R chest	Infantile fibromatosis
108	35 M	Spine (T1)	Metastatic malignant peripheral nerve sheath tumor (MPNST), epithelioid type
109	43 M	Flank	Angiosarcoma, high grade
120	41 F	Breast	Leiomyoma
122	39 M	Stomach	Malignant gastrointestinal stromal tumor (GIST)
124	41 M	Rectum	Gastrointestinal stromal tumor
127	47 M	Subcutaneous	Low grade fibromyxoid sarcoma
129	12 M	Toe	Benign mesenchymal tumor with features of angiomatosis and myofibromatosis
130	67 F	Vulva	Benign genital stromal tumor
135	43 F	Wrist	Pleomorphic undifferentiated sarcoma, high grade
139	42 F	Arm	Extranodal Rosai Dorfman disease
141	75 M	Sinonasal	Hemangiopericytoma-like tumor of the nasal passages

Table 4 (*continued*)

Case	Age/Sex	Location	Diagnosis
142	45 M	Kidney	Malignant glomus tumor
143	16 M	Knee	Necrobiotic granuloma
145	51 M	Neck	Benign mesenchymal tumor, favor spindle cell lipoma
148	85 M	Scalp	*Poorly differentiated carcinoma, probably metastatic
151	45 M	Retroperitoneal	Monophasic synovial sarcoma, high grade
153	47 M	Upper leg	Dermatofibrosarcoma protuberans
157	51 F	Mediastinal	Myxoid/round cell liposarcoma, high grade
158	64 F	Toe	Ewing sarcoma
160	30 F	Paraspinal	Sclerosing epithelioid fibrosarcoma
164	14 M	L4 vertebra	Langerhans cell histiocytosis
165	1 M	Hip	Calcifying aponeurotic fibroma
167	31 M	Scalp	Alveolar soft part sarcoma. Rule out metastasis
168	5 M	Tongue	Reactive myofibroblastic proliferation
169	58 F	Sternocleidomastoid muscle	Soft tissue myoepithelioma, histologically benign
171	26 M	Omentum	Benign fibroblastic proliferation, favor reactive
174	71 F	Neck	Malignant hemangiopericytoma/solitary fibrous tumor, high grade
176	92 M	Neck	*Desmoplastic melanoma
177	33 F	Calf	Neuroblastoma-like schwannoma (schwannoma with collagen rosettes)
178	66 M	Patella	Glomus tumor
179	32 F	Knee	Angiomatoid fibrous histiocytoma
180	39 M	Shoulder	Kaposi sarcoma
184	59 F	Orbit	Hemangiopericytoma
186	18 M	Scrotum	Embryonal rhabdomyosarcoma
188	62 F	Leg	Fibrous histiocytoma with atypical (monster) cells
191	34 F	Calf	*Paraganglioma-like dermal melanocytic tumor
193	47 F	Uterine serosa	Leiomyoma
195	3 F	labia majora	Lipoblastoma
196	19 F	Brachial plexus	Epithelioid nerve sheath tumor, probably of low grade malignancy
198	60 F	Mesenteric/small bowel	Bacillary angiomatosis
199	44 F	Calf	Extraskeletal myxoid chondrosarcoma
200	50 F	Anus	*Malignant melanoma

Conclusion

This study demonstrates the potential of AI techniques to enhance the diagnosis and classification of STTs, with promising results in distinguishing benign from malignant cases and highlighting the relevance of cell-specific parameters. A comprehensive diagnostic evaluation, encompassing all slides and access to immunohistochemical and molecular studies, remains indispensable for accurate soft tissue diagnosis. The use of large heterogeneous, well-curated and annotated/labelled datasets will be essential to bolster AI model training and accuracy in the field of soft tissue pathology. Nevertheless, the integration of AI into the diagnostic process offers substantial promise, poised to elevate our capacity to differentiate between these tumors, ultimately leading to heightened diagnostic precision and improved patient outcomes.

We would like to emphasize that this is a proof-of-concept small study, and larger studies are needed to validate and expand on the findings. The current study provides a foundation upon which future studies can be built. Future research should focus on expanding the dataset and refining AI algorithms to potentially improve diagnostic sensitivity and specificity. Additionally, ongoing efforts must be made to address technical and ethical considerations, such as ensuring data privacy and the accuracy of AI-generated results.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledg the generous support of the Sanders-Brown Center at the University of Kentucky through their staff and equipment for whole slide imaging.

References

1. Arbiser ZK, Folpe AL, Weiss SW. Consultative (expert) second opinions in soft tissue pathology. Analysis of problem-prone diagnostic situations. Am J Clin Pathol 2001;116(4):473–476. https://doi.org/10.1309/425H-NW4W-XC9A-005H.
2. Rupani A, Hallin M, Jones RL, Fisher C, Thway K. Diagnostic differences in expert second-opinion consultation cases at a tertiary sarcoma center. Sarcoma 2020;2020:9810170. https://doi.org/10.1155/2020/9810170.
3. Mesko NW, Mesko JL, Gaffney LM, Halpern JL, Schwartz HS, Holt GE. Medical malpractice and sarcoma care–a thirty-three year review of case resolutions, inciting factors, and at risk physician specialties surrounding a rare diagnosis. J Surg Oncol 2014;110(8):919–929. https://doi.org/10.1002/jso.23770.
4. Ailia MJ, Thakur N, Abdul-Ghafar J, Jung CK, Yim K, Chong Y. Current trend of artificial intelligence patents in digital pathology: a systematic evaluation of the patent landscape. Cancers (Basel) 2022;14(10). https://doi.org/10.3390/cancers14102400.
5. Borowsky AD, Glassy EF, Wallace WD, et al. Digital whole slide imaging compared with light microscopy for primary diagnosis in surgical pathology. Arch Pathol Lab Med 2020;144(10):1245–1253. https://doi.org/10.5858/arpa.2019-0569-OA.
6. Crombé A, Roulleau-Dugage M, Italiano A. The diagnosis, classification, and treatment of sarcoma in this era of artificial intelligence and immunotherapy. Cancer Commun (Lond) 2022;42(12):1288–1313. https://doi.org/10.1002/cac2.12373.
7. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646–674. https://doi.org/10.1016/j.cell.2011.02.013.
8. Kantidakis G, Litière S, Neven A, et al. New benchmarks to design clinical trials with advanced or metastatic liposarcoma or synovial sarcoma patients: an EORTC - Soft Tissue and Bone Sarcoma Group (STBSG) meta-analysis based on a literature review for soft-tissue sarcomas. Eur J Cancer 2022;174:261–276. https://doi.org/10.1016/j.ejca.2022.07.010.
9. Kusta O, Rift CV, Risør T, Santoni-Rugiu E, Brodersen JB. Lost in digitization - a systematic review about the diagnostic test accuracy of digital pathology solutions. J Pathol Inform 2022;13, 100136. https://doi.org/10.1016/j.jpi.2022.100136.
10. Mukhopadhyay S, Feldman MD, Abels E, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). Am J Surg Pathol 2018;42(1):39–52. https://doi.org/10.1097/PAS.0000000000000948.
11. Saini T, Bansal B, Dey P. Digital cytology: current status and future prospects. Diagn Cytopathol 2023;51(3):211–218. https://doi.org/10.1002/dc.25099.
12. Rottmann D, Abdulfatah E, Pantanowitz L. Molecular testing of soft tissue tumors. Diagn Cytopathol 2023;51(1):12–25. https://doi.org/10.1002/dc.25013.
13. Foersch S, Eckstein M, Wagner DC, et al. Deep learning for diagnosis and survival prediction in soft tissue sarcoma. Ann Oncol 2021;32(9):1178–1187. https://doi.org/10.1016/j.annonc.2021.06.007.
14. Allan C, Burel JM, Moore J, et al. OMERO: flexible, model-driven data management for experimental biology. Nat Methods 2012;9(3):245–253. https://doi.org/10.1038/nmeth.1896.

15. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. Sci Rep 2017;7(1):16878. https://doi.org/10.1038/s41598-017-17204-5.
16. Al-Ibraheemi A, Folpe AL. Voluntary second opinions in pediatric bone and soft tissue pathology: a retrospective review of 1601 cases from a single mesenchymal tumor consultation service. Int J Surg Pathol 2016;24(8):685–691. https://doi.org/10.1177/1066896916657591.
17. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Post-Public Health Emergency (PHE) Guidance CMS, Ed. 2021.
18. Rajasekaran RB, Whitwell D, Cosker TDA, Gibbons CLMH, Carr A. Will virtual multidisciplinary team meetings become the norm for musculoskeletal oncology care following the COVID-19 pandemic? - experience from a tertiary sarcoma centre. BMC Musculoskelet Disord 2021;22(1):18. https://doi.org/10.1186/s12891-020-03925-8.
19. Sargen MR, Luk KM, Stoff BK, et al. Diagnostic accuracy of whole slide imaging for cutaneous, soft tissue, and melanoma sentinel lymph node biopsies with and without immunohistochemistry. J Cutan Pathol 2018;45(8):597–602. https://doi.org/10.1111/cup.13268.
20. Tsagkaris C, Trygonis N, Spyrou V, Koulouris A. Telemedicine in care of sarcoma patients beyond the COVID-19 pandemic: challenges and opportunities. Cancers (Basel) 2023;15(14). https://doi.org/10.3390/cancers15143700.
21. Heckenbach I, Mkrtchyan GV, Ezra MB, et al. Nuclear morphology is a deep learning biomarker of cellular senescence. Nat Aging 2022;2(8):742–755. https://doi.org/10.1038/s43587-022-00263-3.
22. Abraham A, Pedregosa F, Eickenberg M, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14. https://doi.org/10.3389/fninf.2014.00014.