

## REVIEW ARTICLE OPEN



# Predictive and prognostic transcriptomic biomarkers in soft tissue sarcomas

Eve Merry<sup>1</sup>, Khin Thway<sup>1,2</sup>, Robin L. Jones<sup>1,3,4</sup> and Paul H. Huang<sup>1,2,4</sup>✉

Soft tissue sarcomas (STS) are rare and heterogeneous tumours comprising over 80 different histological subtypes. Treatment options remain limited in advanced STS with high rates of recurrence following resection of localised disease. Prognostication in clinical practice relies predominantly on histological grading systems as well as sarcoma nomograms. Rapid developments in gene expression profiling technologies presented opportunities for applications in sarcoma. Molecular profiling of sarcomas has improved our understanding of the cancer biology of these rare cancers and identified potential novel therapeutic targets. In particular, transcriptomic signatures could play a role in risk classification in sarcoma to aid prognostication. Unlike other solid and haematological malignancies, transcriptomic signatures have not yet reached routine clinical use in sarcomas. Herein, we evaluate early developments in gene expression profiling in sarcomas that laid the foundations for transcriptomic signature development. We discuss the development and clinical evaluation of key transcriptomic biomarker signatures in sarcomas, including Complexity Index in SARComas (CINSARC), Genomic Grade Index, and hypoxia-associated signatures. Prospective validation of these transcriptomic signatures is required, and prospective trials are in progress to evaluate reliability for clinical application. We anticipate that integration of these gene expression signatures alongside existing prognosticators and other Omics methodologies, including proteomics and DNA methylation analysis, could improve the identification of 'high-risk' patients who would benefit from more aggressive or selective treatment strategies. Moving forward, the incorporation of these transcriptomic prognostication signatures in clinical practice will undoubtedly advance precision medicine in the routine clinical management of sarcoma patients.

*npj Precision Oncology* (2021)5:17; <https://doi.org/10.1038/s41698-021-00157-4>

## INTRODUCTION

Soft-tissue sarcomas (STS) are a group of rare, heterogeneous tumours showing variable mesenchymal differentiation, with >80 histological subtypes<sup>1</sup>. At diagnosis, prognostic value is predominantly given to histological grade, tumour size and depth<sup>2–5</sup>. Despite complete surgical resection of localised disease, 40–50% of patients develop locally recurrent or metastatic disease within 5 years<sup>2,6,7</sup>. Treatment options are limited for advanced or metastatic STS<sup>6</sup>. Palliative anthracycline-based chemotherapy has remained first-line therapy for decades, despite low response rates (15–25%)<sup>2,8,9</sup>. Given the propensity of STS to metastasise with significant associated mortality rate, defining tumour prognosis, and new therapeutic targets to enable selective, personalised therapeutic strategies is of considerable interest<sup>2,6</sup>. Molecular biomarkers including transcriptomic signatures and selected genetic markers may provide an avenue for personalised medicine and improved management in these rare diseases. Herein, we provide a comprehensive review of transcriptomic biomarkers currently in development for STS. We conclude by offering a perspective on future avenues for the use of transcriptomic signatures to enhance sarcoma treatment.

## Early developments in gene expression analysis of STS

Histological grading has been used since the late 1970s to predict tumour aggressiveness and prognosis in STS<sup>10</sup>. The most prominent grading systems used are the National Cancer Institute (NCI)<sup>11</sup> and the French Federation of Cancer Centres Sarcoma Group (FNCLCC)<sup>3</sup> systems defined in 1984. The latter was shown to be superior in a comparative study within the same

population<sup>12</sup>, and this three-grade system is most commonly used in practice<sup>13</sup>. FNCLCC uses three independent histological factors to allocate tumour grade; tumour differentiation, mitotic index, and necrosis<sup>3</sup>. In the majority of STS subtypes, histologic grade is considered the most important prognostic factor, as demonstrated by a multivariate analysis of 1240 patients with localised disease<sup>5</sup>. However, grade is not always a reliable parameter. Limitations of grading systems include the indeterminate prognosis of 'grade 2' consisting of approximately 50% of STS assessed with FNCLCC<sup>12</sup>, as well as concerns surrounding reproducibility between independent pathologists, with 75% concordance in FNCLCC grading of STS samples in a study<sup>14</sup>. Finally, grade interpretation from core biopsy samples should be approached with caution since STS are not uniform throughout<sup>15</sup>, and grading systems were developed on whole tumour samples<sup>3,11</sup>.

Introduction of gene expression profiling technologies like DNA microarrays presented an opportunity to develop molecular profiling as a complementary tool to histological grading for improved sarcoma prognostication. Pioneered by Patrick Brown in the late 1990s, DNA microarrays were introduced to probe gene expression alterations at, what was at that time, an unprecedented genomics scale<sup>16</sup>. The ability to accurately measure multiple genes simultaneously ushered an era of molecular tumour analysis with potential for deriving new molecular classification approaches and prognostic tools.

The power of this technology as applied in sarcomas was shown by the Van de Rijn group who reported gene expression profiles of 41 STS with cDNA microarrays<sup>17</sup>. They identified clusters of genes showing specific expression for synovial sarcoma (SS),

<sup>1</sup>Sarcoma Unit, The Royal Marsden Hospital, London, UK. <sup>2</sup>Division of Molecular Pathology, The Institute of Cancer Research, London, UK. <sup>3</sup>Division of Clinical Studies, The Institute of Cancer Research, London, UK. <sup>4</sup>These authors jointly supervised this work: Robin L. Jones, Paul H. Huang ✉email: paul.huang@icr.ac.uk

gastrointestinal stromal tumours (GIST) and two subgroups of leiomyosarcoma (LMS), highlighting subtype-specific gene signatures with potential to identify genes involved in sarcomagenesis. Two studies in the early 2000s used cDNA microarray technology to develop gene expression profiles associated with poor outcome in LMS<sup>18,19</sup>. Lee et al. compared gene expression profiles of 20 primary and 7 metastatic LMS to reveal differential expression of 335 genes, with a subset of 80 'discriminating genes' highly expressed in the metastatic group<sup>19</sup>. Similarly, Ren et al. identified 92 genes that were differentially expressed between low grade, well-differentiated LMS and less well-differentiated, high-grade and metastatic LMS, suggesting that gene expression data could be used to identify clinically aggressive tumours within a specific subtype<sup>18</sup>. Work by Francis et al. identified a 244-gene signature in 89 primary, high-grade STS, which split the cohort into two prognostic subsets<sup>20</sup>. Several hypoxia-related genes, notably *HIF1A* and its targets, alongside genes promoting chromosomal instability, were upregulated in the signature. Although undertaken in small sample sizes and lacking independent validation, these studies provided early promise that gene expression profiling may have prognostic value in sarcomas.

By far the most widespread use of transcriptomics in sarcoma to date has been in molecular classification of distinct sarcoma subtypes. Multiple studies have demonstrated that transcriptomics can readily distinguish between different histological subtypes<sup>17,21–23</sup>. More recently, transcriptomic signatures have shown utility in delineating heterogeneous molecular subgroups within histological subtypes. For instance, Guo et al.<sup>24</sup>, found that LMS is composed of three molecular subtypes with distinct clinical outcomes and biological pathways. A recent study also showed that small round cell sarcomas with distinct fusions (e.g., *CIC*-fused and *BCOR*-rearranged tumours) had unique transcriptomic profiles<sup>25</sup>. Further reinforcing the idea that transcriptomics has a role in the molecular classification of fusion driven sarcomas, the same group showed that consensus clustering of gene expression data in endometrial stromal sarcomas (ESS) was capable of identifying a high grade group comprised of tumours harbouring *BCOR* rearrangements, and a low grade group composed of tumours with a fusion of a *PRC2* zinc finger protein (such as *JAZF1* and *PHF1*)<sup>26</sup>. Interestingly, *YWHAE-NUTM2* positive ESS which is typically considered a clinically aggressive subtype was found to be split in the high grade and low grade clusters, indicative of molecular heterogeneity within this fusion positive ESS<sup>26</sup>.

Taken together, these studies laid the foundation for comprehensive molecular analyses of sarcomas and showed that gene expression profiles may have potential to inform on tumour grade, subtype classification, molecular biology, and relapse risk/prognosis. Our review will focus on the development of prognostic transcriptomic signatures in STS.

### Predictive and prognostic transcriptomic signatures in STS

Advances in RNA sequencing and related profiling technologies and the resulting decrease in the cost of comprehensive transcriptomic analysis has led to the development of several different transcriptomic signatures for STS prognostication. These include Complexity INDEX in SARCOMAs (CINSARC), Genomic Grade Index (GGI) and hypoxia-associated signatures (Table 1). We describe the development of these biomarkers, their underlying biological basis and the current status in clinical development.

**Complexity index in sarcomas.** The most advanced transcription signature candidate under clinical evaluation is CINSARC. Developed by the group of Frédéric Chibon, CINSARC is a 67-gene signature built from a genomic and transcriptomic analysis of 183 primary non-translocation-related sarcomas<sup>27</sup>. The underlying rationale for characterising these tumours is that non-

translocation-related sarcomas, including LMS, undifferentiated pleomorphic sarcomas (UPS) and dedifferentiated liposarcomas, are prone to metastasis<sup>6,7</sup>. A three-step approach was used to define the signature. In the first step, significantly expressed genes were selected according to i) comparative genomic hybridisation (CGH) imbalances ii) FNCLCC grade (specifically grade 3 versus 2) and iii) a previously reported 70-gene chromosome instability signature developed by Carter et al.<sup>28</sup>. In the second step, gene ontology analysis identified pathways associated with histologic grade or genomic imbalance. Finally, in the third step, genes involved in the most overrepresented pathways were selected. These 67 genes comprised i) 37 genes related to CGH imbalance ii) 18 associated with histologic grade and iii) 22 additional genes, not yet included, from the Carter signature. When gene ontology analysis was performed, the 67 genes were found to be related to control of chromosome integrity and mitosis. Specifically, chromosome biogenesis, condensation, alignment and segregation, cell cycle/mitosis and cytokinesis control, and the microtubule-kinesin complex.

The STS training set ( $n = 183$ ), alongside breast cancer and lymphoma samples, was used to generate gene expression centroids for CINSARC grading. There are two CINSARC grades, C1, which is a low CINSARC score comprising 'good prognosis' patients, and C2, a high CINSARC score comprising 'poor prognosis' patients. In the training cohort, patients in the C1 and C2 groups had 5-year metastasis-free survival (MFS) rates of 75% and 35%, respectively ( $p = 1 \times 10^{-7}$ ), which was validated in an independent patient cohort ( $n = 127$ )<sup>27</sup>. While metastasis risk could be predicted by CINSARC (training set: HR 3.7; 95% CI 2.2–6.3 and validation set: HR 2.7; 95% CI 1.02–7.2); FNCLCC grade could only predict metastatic outcome in the validation set and not the training set. Multivariate analysis showed that CINSARC was an independent prognostic factor for metastasis when adjusted for FNCLCC grade, subtype, and vascular/bone involvement. A notable advantage of the CINSARC signature is its ability to split FNCLCC grade 2 into good (C1) or poor (C2) prognosis groups based on metastatic potential. This study was the first to describe a gene expression-based risk classifier, that could potentially identify patients with high metastatic potential. Since its initial development in gene expression microarrays, CINSARC has been extended to other platforms, namely RNA-sequencing and Nanostring probe-based technology in both frozen and formalin-fixed paraffin embedded (FFPE) tissue<sup>29,30</sup>.

CINSARC has also been evaluated in other sarcoma subtypes. GISTs are the most frequent mesenchymal tumours of the gastrointestinal tract, making up 25% of STS<sup>1</sup>. Unlike genomically-complex STS, GISTs are characterised by activating point mutations, most commonly in *KIT* and *PDGFRA*<sup>31</sup>. However, despite optimal surgical resection, 20–40% patients relapse<sup>32</sup>. In the original study, CINSARC categorised a GIST cohort ( $n = 32$ ) into two groups<sup>27</sup>. In C1 no cases developed metastasis, whereas in C2, MFS rates at 5 and 10 years were 61% and 30%, respectively. In a follow-up study, CINSARC stratified 60 primary untreated GISTs into two distinct groups with significantly different MFS<sup>32</sup>. The C1 group ( $n = 32$ ) were metastasis/relapse-free at 5 years, with a 5-year MFS rate of 38% in C2 ( $n = 28$ ).

Synovial sarcoma is characterised by a specific translocation t(X;18), and account for 5–10% of STS<sup>1</sup>. CINSARC has been shown to stratify 58 primary untreated SS into two prognostic groups, C1 and C2, with 5-year MFS rates of 78% and 33%, respectively<sup>33</sup>. This was validated in an independent series of 40 primary untreated SS with similar outcomes. On multivariate analysis, CINSARC was shown to be an independent prognosticator to FNCLCC, and able to split FNCLCC grades 2 and 3 into good and poor prognostic groups in SS. It should be noted that evaluation of CINSARC in translocation-driven sarcomas has been limited to synovial sarcomas and further studies in additional translocation-driven sarcomas is required.

**Table 1.** Summary of key studies in CINSARC, GGI and hypoxia-based transcriptomic signature development.

Signature	Year	Authors	Title	Journal	Contribution to signature development in sarcomas	References
Complexity Index in SARComas (CINSARC)	2010	Chibon, F. et al.	Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity	<i>Nat. Med.</i> <b>16</b> , 781–787 (2010).	Development of CINSARC, 67-gene prognostic signature in STS, with validation in STS, GIST, breast carcinoma and lymphoma	<a href="#">27</a>
	2012	Lagarde, P. et al.	Mitotic checkpoints and chromosome instability are strong predictors of clinical outcome in gastrointestinal stromal tumours	<i>Clin. Cancer Res.</i> <b>18</b> , 826–838 (2012).	CINSARC stratification of GIST cohort into high-risk and low-risk (C2 and C1) prognostic groups	<a href="#">32</a>
	2013	Lagarde, P. et al.	Chromosome instability accounts for reverse metastatic outcomes of paediatric and adult synovial sarcomas	<i>J. Clin. Oncol.</i> <b>31</b> , 608–615 (2013).	CINSARC stratification of Synovial sarcoma cohort into high-risk and low-risk (C2 and C1) prognostic groups	<a href="#">33</a>
	2013	Italiano, A. et al.	Genetic profiling identifies two classes of soft-tissue leiomyosarcomas with distinct clinical characteristics	<i>Clin. Cancer Res.</i> <b>19</b> , 1190–1196 (2013).	CINSARC stratification of leiomyosarcoma cohort into high-risk and low-risk (C2 and C1) prognostic groups	<a href="#">42</a>
	2020	Croce, S. et al.	The nanocind signature is an independent prognosticator of recurrence and death in uterine leiomyosarcomas	<i>Clin. Cancer Res.</i> <b>26</b> , 855–861 (2020)	CINSARC stratification of uterine leiomyosarcoma cohort into high-risk and low-risk (C2 and C1) prognostic groups	<a href="#">41</a>
Genomic Grade Index (GGI)	2012	Bertucci, F. et al.	Genomic Grade Index predicts postoperative clinical outcome of GIST	<i>Br. J. Cancer</i> <b>107</b> , 1433–1441 (2012)	GGI stratification of GIST cohort into GGI-high and GGI-low prognostic groups	<a href="#">49</a>
	2018	Bertucci, F. et al.	The Genomic Grade Index predicts postoperative clinical outcome in patients with soft-tissue sarcoma	<i>Ann. Oncol.</i> <b>29</b> , 459–465 (2018)	GGI stratification of localised STS cohort into GGI-high and GGI-low prognostic groups	<a href="#">51</a>
Hypoxia-based signatures	2016	Aggerholm-Pedersen, N. et al.	A prognostic profile of hypoxia-induced genes for localised high-grade soft tissue sarcoma	<i>Br. J. Cancer</i> <b>115</b> , 1096–1104 (2016).	Application of a 15-gene hypoxia signature in a cohort of STS to produce high- and low-hypoxia prognostic groups	<a href="#">62</a>
	2018	Yang, L. et al.	Validation of a hypoxia-related gene signature in multiple soft tissue sarcoma cohorts	<i>Oncotarget</i> <b>9</b> , 3946–3955 (2018).	Development of a STS-specific 24-gene hypoxia signature	<a href="#">66</a>

LMS are smooth muscle tumours that represent 10–15% STS<sup>1</sup>. They can occur in almost any location of the body, commonly abdomen, retroperitoneum, large blood vessels, and uterus<sup>34</sup>. Uterine LMS (uLMS) account for 7% of STS and 1–3% of uterine malignancies<sup>34,35</sup>. These are aggressive tumours, with 5-year overall survival (OS) rate of 41% for all International Federation of Gynaecology and Obstetrics (FIGO) stages<sup>36</sup>. FNCLCC grading has failed to predict outcome in uLMS<sup>37</sup>, and WHO could not identify an appropriate grading system for uLMS in 2014<sup>38</sup>. FIGO staging remains the primary prognostic factor, alongside an uLMS-specific nomogram for predicting post-resection 5-year OS<sup>39,40</sup>. There is a need for improved prognosticators to identify patients who might benefit from adjuvant chemotherapy<sup>41</sup>. Italiano et al. showed that CINSARC classified 73 primary LMS, of multiple anatomical locations, into two groups with significantly different MFS<sup>42</sup>. In a retrospective series of 60 uLMS CINSARC divided this cohort, of all FIGO stages, into high-risk (C2) and low-risk (C1) groups<sup>41</sup>. C2 had 5-year relapse-free survival (RFS) rate of 9% compared to 51% in C1. OS rate at 5 years was 29% in C2 and 86% in C1. These findings were validated in an independent series of 32 uLMS from The Cancer Genome Atlas Consortium. FIGO Stage I uLMS are localised, but have a propensity to relapse<sup>43</sup>. CINSARC was able to divide stage I tumours into good and poor prognostic groups, the latter with high risk of relapse and death<sup>41</sup>.

Overall, CINSARC has been shown to be an effective independent risk classifier across a spectrum of subtypes including those harbouring complex karyotypes, point mutations, and translocations<sup>32,33,41,42</sup>.

**Genomic grade index.** Gene expression signatures have reached routine use for risk stratification to aid clinical decision-making in breast cancer<sup>44,45</sup>. Given that there are some similarities in the morphological criteria used to grade both STS and breast cancer [mitotic index in both FNCLCC sarcoma grading and Nottingham (breast cancer) grading systems and tumour differentiation (according to subtype: well, moderately, and poorly in sarcomas; tubule differentiation in breast cancer)], application of breast cancer gene expression tools in STS was of interest<sup>27,46,47</sup>. GGI is a 108-gene mRNA signature developed in a cohort of 64 early-stage oestrogen receptor-positive breast cancers by comparing gene expression profiles of histologic grade 3 and 1 tumours<sup>48</sup>. GGI was shown to reclassify grade 2 breast cancers into two prognostic groups, with high GGI expression (GGI-high) associated with higher risk of recurrence and worse prognosis compared to low GGI expression (GGI-low)<sup>48</sup>.

Bertucci et al. evaluated GGI in 86 non-metastatic resected GISTs<sup>49</sup>. They showed that the GGI-high subgroup had a 5-year RFS rate of 46% (95% CI 28–77,  $n = 20$ ) compared to 91% (95% CI 82–100;  $n = 66$ ) in the GGI-low group ( $p = 1.4 \times 10^{-6}$ )<sup>49</sup>. GGI also independently predicted RFS in an independent series of GIST samples ( $n = 60$ ). At present GIST prognostication relies on the Armed Forces Institute of Pathology (AFIP) classification based on tumour size, site, and mitotic rate<sup>50</sup>. GGI-high samples were more frequently associated with poor prognosis variables, including AFIP high-risk tumours<sup>49</sup>. The prognostic performance of GGI was compared to AFIP and both had independent prognostic value on multivariate analysis for predicting relapse risk. GGI was able to further define the intermediate/high-risk AFIP samples into two groups; high-risk and low-risk, with a 5-year RFS rate of 35% (95% CI 17–70) and 73% (95% CI 52–100), respectively ( $p = 8.5 \times 10^{-3}$ ). To explore the relationship between GGI classification and imatinib response, the authors analysed a small cohort of pre-treatment GISTs ( $n = 28$ ) from patients with advanced primary or recurrent operable GIST treated with 8–12 week neoadjuvant imatinib in a phase II trial (RTOG0132). There was greater tumour shrinkage (evaluated on CT with Response Evaluation Criteria in Solid Tumours [RECIST]) in GGI-high tumours than GGI-low, suggesting that GGI-high tumours were more sensitive to

imatinib, with need for larger prospective trials to confirm this.

The same group set out to identify correlations between GGI-based classification and clinicopathological variables in a cohort of localised STS ( $n = 678$ )<sup>51</sup>. The most frequent histological subtypes were liposarcoma (38%), UPS (30%), and LMS (38%), thus 433 (65%) were defined as 'genomically complex'. 56% were FNCLCC grade 3, and 41% were classified as GGI-low with 59% GGI-high. The GGI-high subgroup had a poor prognosis with 5-year MFS rate of 53% (95% CI 47–59) compared to 78% (95% CI 72–85) in the GGI-low group ( $p = 3.02 \times 10^{-11}$ ). Additionally, the GGI-high group was associated with STS in the extremities versus the trunk, complex genetic profiles and FNCLCC grade 3 tumours. GGI was able to stratify STS patients with histologic grades 1 and 2 into two prognostic groups, GGI-high and GGI-low, with different 5-year MFS rates; 59% (95% CI 46–76) and 74% (95% CI 62–87), respectively.

For comparison, CINSARC was applied to the Bertucci et al. GIST cohort ( $n = 86$ ) and C2 and C1 groups defined with a 5-year RFS rate of 67% (95% CI 53–86;  $n = 38$ ) and 92% (95% CI 84–100;  $n = 48$ ), respectively ( $p = 0.01$ )<sup>49</sup>. Ontology analysis showed that, similar to the C2 subgroup, overexpressed genes in the GGI-high subgroup included those associated with cell cycle control and genome stability. When comparing GGI and CINSARC, there were 39 genes in common<sup>51</sup>. Additionally, strong correlation was observed between GGI and CINSARC classes in a further study by Bertucci et al. of 678 STS, with 71% concordance in allocation to low-risk and high-risk groups<sup>51</sup>. Similar to CINSARC, the GGI signature is composed of genes involved in cell cycle regulation with further work required to determine regulators of GGI gene expression and scope for applications in targeted therapy.

**Hypoxia-based signatures.** Intratumoural hypoxia is considered an adverse prognostic factor for metastatic spread in multiple malignancies<sup>52–55</sup>. An association between tumour hypoxia levels and poor outcome has been observed in STS<sup>54,56–58</sup>. Nordmark et al. evaluated 28 STS tumours and stratified cases according to tumour oxygenation; tumours with median  $pO_2 > 19$  mmHg were classified as well-oxygenated and  $\leq 19$  mmHg as hypoxic. The hypoxic group had significantly lower disease-specific survival of 40% compared to 77% for well-oxygenated tumours ( $p = 0.05$ )<sup>54</sup>. Hypoxic tumours also had significantly poorer 5-year OS probability of 28% versus 77% ( $p = 0.01$ ). Another small study of primary, high-grade STS ( $n = 30$ ), treated with neoadjuvant radiotherapy and hyperthermia found that more hypoxic tumours, defined as  $pO_2$  median value  $< 10$  mmHg pre-treatment, had disease-free survival (DFS) rate of 35% compared to 70% in tumours with median  $pO_2 > 10$  mmHg pre-treatment<sup>56</sup>. In addition, metastatic recurrence in eight patients was associated with significantly lower median pre-treatment  $pO_2$  compared to those that did not metastasise. Work to identify the hypoxia biomarkers followed; Forker et al. utilised STS specimens from the phase III adjuvant radiotherapy Vortex trial, which assessed whether reduced adjuvant radiotherapy volume could improve limb function in adults with extremity STS without compromising local control<sup>59</sup>. Immunohistochemistry for hypoxia protein marker expression, specifically HIF-1 $\alpha$ , CAIX, and GLUT1, was performed on tissue microarrays from histologically heterogeneous STS specimens ( $n = 203$ ). CAIX was a significant prognostic biomarker, with worse DFS at 5 years in tumours with  $> 10\%$  CAIX staining compared to  $\leq 10\%$  staining (HR 1.75; 95% CI 1.04–2.94;  $p = 0.037$ ). However, there was concern that there was limited overlap expression of these three hypoxia markers across tumour samples, with co-expression absent in hypoxic samples. This may be a result of significant molecular heterogeneity of sarcoma limiting use of single biomarkers. Robust interpretation of whole tumour hypoxia may be best achieved by assessing multiple-gene response with hypoxia gene expression signatures<sup>59,60</sup>.

Aggerholm-Pedersen et al. applied a 15-gene hypoxia-induced signature, previously developed for head and neck cancer<sup>61</sup>, to

stratify STS patients into 'more hypoxic' (high hypoxia gene expression) and 'less hypoxic' groups in FFPE tissue samples<sup>62</sup>. Fifteen genes were selected from a panel of 30 validated hypoxia-responsive, pH-independent genes, including *LOX*, *P4HA1*, and *P4HA2* involved in extracellular matrix modulation, and genes involved in glycolysis such as *SLC2A1*, *PFKB3*, and *PDK1*; both biological processes influenced by hypoxia<sup>61,63–65</sup>. The signature was used in a test ( $n = 55$ ) and validation cohort ( $n = 77$ ), each made up predominantly of liposarcomas and UPS<sup>62</sup>. In the test cohort, HR for disease-specific mortality was 4.09 (95% CI 1.34–12.46;  $p = 0.013$ ) for 'more hypoxic' tumours compared with 'low hypoxic' tumours. Recurrent disease was also higher in the 'more hypoxic' group compared with the 'less hypoxic' group (odds ratio 3.96; 95% CI 0.98–14.7;  $p = 0.03$ ). Similar outcomes were observed in the validation cohort. This study suggests that hypoxia gene signatures could be used as prognostic biomarkers for STS, but was limited by wide confidence intervals in the survival analysis and lower hazard ratio in the validation cohort, requiring further validation in larger cohorts<sup>62</sup>. Furthermore, 16 samples, with corresponding oxygen tension measurements available, were analysed with an unexpected finding: cases with low hypoxia levels (high pO<sub>2</sub>) were associated with high hypoxic gene expression. Therefore, the study could not conclude that expression of these 15 genes in STS is hypoxia driven.

In another study, Yang et al. developed a STS-specific 24-gene hypoxia signature<sup>66</sup>. To define this signature they first identified 33 genes induced by hypoxia in seven cell lines representative of common STS subtypes in adults. These 33 seed genes were used to derive a hypoxia signature in clinical specimens; the training cohort ( $n = 182$ ) was separated into two groups; high-hypoxia and low-hypoxia, based on unsupervised clustering of the 33 genes, with 24/33 seed genes (73%) significantly upregulated in the high-hypoxia group. Gene set enrichment analysis was undertaken of the entire transcriptomic data for cases in the high-hypoxia group revealing upregulation of 16 hypoxia pathways. Survival analysis showed that 'high-hypoxia' tumours had worse 5-year distant metastatic-free survival (DMFS) rate than 'low-hypoxia' tumours (HR 2.43; 95% CI 1.49–3.96;  $p = 0.00036$ ), which was validated in two independent cohorts of heterogeneous STS subtypes. The hypoxic signature retained its prognostic significance in multivariable analysis adjusted for histological diagnosis, tumour site, site, gender, and age. However, tumour grade was omitted from the multivariable analysis. Given histologic grade is considered the most reliable prognostic factor for the majority of STS, this is a major study limitation<sup>5</sup>. The 24-gene signature was deemed a superior prognosticator to the earlier 15-gene signature<sup>62</sup>, since the latter only achieved prognostic significance for DMFS on multivariable analysis in 2 of the 3 STS cohorts used in Yang et al's study<sup>66</sup>. Notably there was overlap of 8 common genes between the signatures.

It has been hypothesised that tumour hypoxia drives genomic instability, and thus promotes tumour aggressiveness and distant spread<sup>67</sup>. Interestingly in Yang et al. study, CINSARC was used as a measure of genome instability, and in a combined analysis of training and validation cohorts, more C2 tumours were found in the high-hypoxia than the low-hypoxia group (76 and 48%, respectively)<sup>66</sup>. Prognostic value of the hypoxic gene signature was enhanced by CINSARC with significantly worse DMFS in combined high-hypoxia/C2 patients than low-hypoxia/C1 tumours (HR 6.74, 95% CI 3.84–11.84,  $p = 3.13 \times 10^{-11}$ ). This highlights the potential of integrating different transcriptomic biomarker signatures to derive better risk classifiers for sarcoma prognostication.

### Clinical evaluation of transcriptomic biomarkers in STS

Having described transcriptomic biomarker development in STS, here we consider their clinical relevance. Thus far, all published studies have been carried out retrospectively in heterogeneous

STS populations. There is a need for prospective clinical trials, ideally randomised, to evaluate these transcriptomic signatures in STS to achieve the necessary level of evidence (LoE) for incorporation into international guidelines and use in the clinical setting<sup>68</sup>.

Transcriptomic signatures could contribute to the much-debated topic of peri-operative chemotherapy in STS. Adjuvant chemotherapy following 'gold-standard' surgical resection of localised STS remains a very controversial topic<sup>69–71</sup>. Since the 1980s, a number of adjuvant chemotherapy trials have been published, but most are limited by the inclusion of heterogeneous subtypes, small numbers and suboptimal schedules. In the neoadjuvant setting, a multicentre phase III randomised trial assigned patients ( $n = 328$ ) to 3 or 5 cycles of combined epirubicin/ifosfamide neoadjuvant chemotherapy, concluding that prolonged chemotherapy did not benefit OS<sup>72</sup>. Another phase III randomised trial of 287 patients with localised, high risk (grade 3, size  $\geq 5$  cm) STS of five histologic subtypes (LMS, high-grade myxoid LPS, SS, MPNST, UPS) found that histology-tailored neoadjuvant chemotherapy was not associated with better DFS or OS compared to standard anthracycline-ifosfamide chemotherapy<sup>73</sup>. Therefore, current international guidelines advise neoadjuvant chemotherapy can be considered in 'high-risk' STS<sup>2</sup>. Transcriptomic signatures may have utility in further defining 'high-risk' STS patients that might benefit from peri-operative chemotherapy.

At present, of the signatures described, only CINSARC is being evaluated in prospective clinical trials (Table 2). The first is a phase III randomised trial (NCT03805022) assessing whether more intensive peri-operative chemotherapy improves outcome of patients with resectable STS and high-risk CINSARC signature (C2). Control arm A will assess C2 patients treated with three cycles neoadjuvant chemotherapy (doxorubicin/ifosfamide) followed by surgery +/- radiotherapy. In experimental arm B, C2 patients will receive an additional three cycles of chemotherapy, followed by surgery +/- radiotherapy, with a 3rd prospective arm for low-risk CINSARC patients (C1) receiving treatment at the discretion of the investigator. The primary endpoint is metastatic progression-free survival after 3 years of follow-up. The second is a prospective single-arm observational study (NCT02789384) in non-metastatic STS, that aims to validate the prognostic value of CINSARC and correlation with chemotherapy efficacy. Patients will be classified with CINSARC prior to neo-adjuvant anthracycline-based chemotherapy. Following chemotherapy, patients will proceed to surgery +/- radiotherapy, with follow-up for the duration of treatment. The primary outcome of chemotherapy efficacy is RECIST v1.1 response, with further analysis to determine association of CINSARC grading with response and survival.

Another phase III multicentre randomised trial (NCT04307277) will evaluate peri-operative chemotherapy in C2 patients with resectable FNCLCC grade 1/2 STS. Control arm C2 patients will be randomised to surgical excision +/- radiotherapy, whilst experimental arm C2 patients will receive four cycles of peri-operative anthracycline-based chemotherapy in addition to standard management. A third prospective cohort will include C1 patients treated at the discretion of the investigators. The primary endpoint is MFS. This trial aims to determine whether CINSARC can identify chemo-sensitive tumours within the FNCLCC 'low/indeterminate' risk category, which could assist identifying patients likely to benefit from peri-operative chemotherapy.

### OUTSTANDING QUESTIONS

The developments in prognostic transcriptomic signatures are very promising but there remain outstanding questions to be addressed. The vast majority of transcriptomic signatures in sarcoma have been developed from bulk measurements of tumour specimens. However, all the studies described to date

**Table 2.** CINSARC prospective clinical trials.

Trial number	Phase	Title	Allocation	Intervention model	Primary outcomes	Secondary outcomes
NCT03805022	III	Benefit of intensified peri-operative chemotherapy within High-risk CINSARC patients with resectable soft-tissue sarcomas (CINSARC)	Randomised	Parallel assignment	- Metastasis progression-free survival [3 years]	- Loco-regional relapse-free survival [3 years] - Progression-free survival [3 years] - Overall survival [3 years] - Best overall response as per RECIST v1.1 [6 months] - Histological response [6 months] - Adverse events/Toxicity as per NCI-CTCAE v5.0 [6 months]
NCT02789384	N/A	Prognostic value of the CINSARC signature and correlation with chemotherapy efficacy in soft-tissue sarcomas: A Biomarker Study. (NEOSarcomics)	N/A	Observational- Single Group Assignment	- Best overall response as per RECIST v1.1 [6 months]	- Histological response [6 months] - Metastasis-free survival [3 years] - Overall survival [3 years] - Adverse events/Toxicity as per NCI-CTCAE v4.0 [5 years]
NCT04307277	III	Interest of peri-operative chemotherapy in patients with CINSARC high-risk localised Grade 1 or 2 soft tissue sarcoma	Randomised	Parallel assignment	- Metastasis-free survival [5 years]	- Disease-free survival [5 years] - Toxicity as per NCI-CTCAE v5.0 - Overall survival [5 years]

have not systematically established if the performance of these prognostic signatures are influenced by the intratumoural heterogeneity inherent within STS<sup>15,74</sup>, for instance by evaluation of distinct heterogeneous tumour regions with spatial transcriptomics. Furthermore, there is a diversity of distinct cell types in the tumour microenvironment (e.g., immune cells, fibroblasts and endothelial cells) which can be readily assessed by histopathology review but is lost in bulk transcriptomic data. Emerging data from single cell RNA sequencing has started to shed light on the intratumoural heterogeneity inherent in some subtypes<sup>75</sup>, but the cost of deploying this technology in routine clinical setting is prohibitive. Alternatively, deconvolution methodologies may be used to establish distinct immune cell types from bulk transcriptomic data<sup>76</sup>. Given that these transcriptomic signatures have been established and validated in bulk measurements, whether the use of single cell or deconvolution analysis will lead to a loss in prognostic value remains an open question. From a biological perspective, it is also currently unknown if transcriptomic subgroups are retained during disease progression or treatment. For instance, it is unclear if CINSARC subgroups remain stable in patient-matched primary, locally relapsed and metastatic tumours and if the subgroup classification alters in response to therapy. This stability of subgroup has profound implications for implementation in clinical practice as it determines which point in the patient journey the prognostic assay should be applied. Finally, interlaboratory benchmarking for reliability and reproducibility of these complex transcriptomic assays needs to be undertaken across multiple laboratories prior to use as routine clinical assays.

## FUTURE PERSPECTIVES AND CONCLUSIONS

Use of transcriptomic biomarkers in sarcomas provides an exciting opportunity to improve the limited treatment options and dismal prognosis for patients with advanced metastatic STS<sup>6</sup>. Personalised medicine may be achievable, with allocation of therapy based on prognostic transcriptomic biomarker expression. Further prospective validation of CINSARC, GGI, and hypoxia gene signatures is required to advance these signatures for routine use. In order to improve STS prognostication, histology-specific and site-specific nomograms combining clinical and pathological tumour and patient characteristics, have been used since 2002<sup>77,78</sup>. Combining existing sarcoma nomograms with transcriptomic signatures may further improve accuracy of risk assessment<sup>79</sup>. Additionally, integration of gene expression analysis with other Omic profiling methodologies, such as proteomics and DNA methylation analysis, is also likely to improve prognostication with new biomarker-directed treatment opportunities<sup>80–83</sup>.

Beyond their use in the peri-operative setting, transcriptomic signatures may have utility in the context of next generation therapies. There are several biomarkers for targeted therapies such as PD-L1 expression for immune checkpoint inhibitors and PARP-1 levels for PARP inhibitors<sup>84,85</sup>. Integration of these biomarkers with transcriptomic signatures may be of prognostic benefit. Bertucci et al. assessed expression of PD-L1 and PARP-1 in separate retrospective series of genomically-complex STS, and found that tumours with high PD-L1 or PARP-1 expression had worse MFS<sup>86,87</sup>. These studies found that expression of these biomarkers could enhance prognostication with transcriptomic signatures. For example, PARP-1 expression complemented the prognostic value of CINSARC, with a 5-year MFS rate of 77% in C1/PARP1-low tumours compared to 43% in C2/PARP1-high cohort ( $p = 2.55 \times 10^{-11}$ ). Through identifying high-risk patients, these integrated biomarkers could select patients who would benefit from targeted therapy and more aggressive treatment strategies.

A better understanding of orthogonal molecular features (such as copy number alterations and DNA methylation status) as well as the regulatory mechanisms driving the observed transcriptional signatures will be key to delineating sarcomagenesis and

identifying potential therapeutic targets. Lesluyes and Chibon recently applied the CINSARC signature to the multi-Omics dataset from The Cancer Genome Atlas sarcoma cohort<sup>23</sup>. Focusing on the sarcomas with complex genetics, they showed that copy number alterations were significantly increased in C2 compared with the C1 group ( $p = 2.33 \times 10^{-2}$ ), as was whole genome-doubling ( $p = 1.82 \times 10^{-2}$ ) and higher ploidies ( $p = 3.94 \times 10^{-3}$ )<sup>88</sup>. Low whole-genome DNA methylation was associated with more aggressive C2 tumours. However, DNA methylation was not observed in CINSARC promoter regions which is indicative that this is not a direct regulator of CINSARC expression. Further, most miRNAs (67%) overexpressed in the C2 group were known oncomirs. Understanding the regulation, epigenetic or otherwise, of CINSARC gene expression is an exciting avenue of future translational research.

Much can be learned from progress in other tumour types where precision medicine is widely used in clinical practice<sup>89</sup>. In order to be approved for clinical use, gene expression signatures need to undergo a complex series of steps to demonstrate the required LoE in prospective clinical trials. Promising early steps have been undertaken with CINSARC, and we anticipate that wide adoption of transcriptomic signatures in sarcoma clinical practice will require international collaboration for robust prospective validation of their prognostic and clinical utility.

Received: 2 November 2020; Accepted: 4 February 2021;

Published online: 05 March 2021

## REFERENCES

- WHO Classification of Tumours Editorial Board. *Soft Tissue and Bone Tumours*. (IARC Press, 2020).
- Casali, P. G. et al. Soft tissue and visceral sarcomas: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **29**, iv51–iv67 (2018).
- Trojani, M. et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int. J. Cancer* **33**, 37–42 (1984).
- Callegaro, D. et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. *Lancet Oncol.* **17**, 671–680 (2016).
- Coindre, J. M. et al. Predictive value of grade for metastasis development in the main histologic types of adult soft tissue sarcomas: a study of 1240 patients from the French Federation of Cancer Centers Sarcoma Group. *Cancer* **91**, 1914–1926 (2001).
- Weitz, J., Antonescu, C. R. & Brennan, M. F. Localized extremity soft tissue sarcoma: improved knowledge with unchanged survival over time. *J. Clin. Oncol.* **21**, 2719–2725 (2003).
- Zagars, G. K. et al. Prognostic factors for patients with localized soft-tissue sarcoma treated with conservation surgery and radiation therapy: an analysis of 1225 patients. *Cancer* **97**, 2530–2543 (2003).
- Judson, I. et al. Doxorubicin alone versus intensified doxorubicin plus ifosfamide for first-line treatment of advanced or metastatic soft-tissue sarcoma: a randomised controlled phase 3 trial. *Lancet Oncol.* **15**, 415–423 (2014).
- Antman, K. et al. An intergroup phase III randomized study of doxorubicin and dacarbazine with or without ifosfamide and mesna in advanced soft tissue and bone sarcomas. *J. Clin. Oncol.* **11**, 1276–1285 (1993).
- Russell, W. O. et al. A clinical and pathological staging system for soft tissue sarcomas. *Cancer* **40**, 1562–1570 (1977).
- Costa, J., Wesley, R. A., Glatstein, E. & Rosenberg, S. A. The grading of soft tissue sarcomas. Results of a clinicohistopathologic correlation in a series of 163 cases. *Cancer* **53**, 530–541 (1984).
- Guillou, L. et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. *J. Clin. Oncol.* **15**, 350–362 (1997).
- Golouh, R. & Bracko, M. What is current practice in soft tissue sarcoma grading? *Radiol. Oncol.* **35**, 47–52 (2001).
- Coindre, J. M. et al. Reproducibility of a histopathologic grading system for adult soft tissue sarcoma. *Cancer* **58**, 306–309 (1986).
- Schneider, N. et al. The adequacy of core biopsy in the assessment of smooth muscle neoplasms of soft tissues: implications for treatment and prognosis. *Am. J. Surg. Pathol.* **41**, 923–931 (2017).
- Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci.* **95**, 14863 LP–14814868 (1998).
- Nielsen, T. O. et al. Molecular characterisation of soft tissue tumours: a gene expression study. *Lancet* **359**, 1301–1307 (2002).
- Ren, B. et al. Gene expression analysis of human soft tissue leiomyosarcomas. *Hum. Pathol.* **34**, 549–558 (2003).
- Lee, Y.-F. et al. A gene expression signature associated with metastatic outcome in human leiomyosarcomas. *Cancer Res.* **64**, 7201–7204 (2004).
- Francis, P. et al. Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. *BMC Genomics* **8**, 73 (2007).
- Linn, S. C. et al. Gene expression patterns and gene copy number changes in dermatofibrosarcoma protuberans. *Am. J. Pathol.* **163**, 2383–2395 (2003).
- West, R. B. et al. Determination of stromal signatures in breast carcinoma. *PLoS Biol.* **3**, e187 (2005).
- Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell* **171**, 950–965.e28 (2017).
- Guo, X. et al. Clinically relevant molecular subtypes in leiomyosarcoma. *Clin. Cancer Res.* **21**, 3501–3511 (2015).
- Watson, S. et al. Transcriptomic definition of molecular subgroups of small round cell sarcomas. *J. Pathol.* **245**, 29–40 (2018).
- Brahmi, M. et al. Molecular classification of endometrial stromal sarcomas using RNA sequencing defines nosological and prognostic subgroups with different natural history. *Cancers* **12**, 2604 (2020).
- Chibon, F. et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781–787 (2010).
- Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N. & Szallasi, Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat. Genet.* **38**, 1043–1048 (2006).
- Lesluyes, T. et al. RNA sequencing validation of the complexity INDEX in SARComas prognostic signature. *Eur. J. Cancer* **57**, 104–111 (2016).
- Le Guellec, S. et al. Validation of the complexity INDEX in SARComas prognostic signature on formalin-fixed, paraffin-embedded, soft-tissue sarcomas. *Ann. Oncol.* **29**, 1828–1835 (2018).
- Rubin, B. P., Heinrich, M. C. & Corless, C. L. Gastrointestinal stromal tumour. *Lancet* **369**, 1731–1741 (2007).
- Lagarde, P. et al. Mitotic checkpoints and chromosome instability are strong predictors of clinical outcome in gastrointestinal stromal tumors. *Clin. Cancer Res.* **18**, 826–838 (2012).
- Lagarde, P. et al. Chromosome instability accounts for reverse metastatic outcomes of pediatric and adult synovial sarcomas. *J. Clin. Oncol.* **31**, 608–615 (2013).
- George, S., Serrano, C., Hensley, M. L. & Ray-Coquard, I. Soft tissue and uterine leiomyosarcoma. *J. Clin. Oncol.* **36**, 144–150 (2018).
- Toro, J. R. et al. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978–2001: An analysis of 26,758 cases. *Int. J. Cancer* **119**, 2922–2930 (2006).
- Burghaus, S. et al. Outcome and prognosis in uterine sarcoma and malignant mixed Mullerian tumor. *Arch. Gynecol. Obstet.* **294**, 343–351 (2016).
- Pautier, P. et al. Analysis of clinicopathologic prognostic factors for 157 uterine sarcomas and evaluation of a grading score validated for soft tissue sarcoma. *Cancer* **88**, 1425–1431 (2000).
- Robert J. Kurman, Maria Luisa Carcangiu, C. Simon Herrington, R. H. Y. *WHO Classification of Tumours of Female Reproductive Organs* 4th edn (IARC WHO Classification Tumours, 2014).
- Amant, F., Coosemans, A., Debiec-Rychter, M., Timmerman, D. & Vergote, I. Clinical management of uterine sarcomas. *Lancet Oncol.* **10**, 1188–1198 (2009).
- Zivanovic, O. et al. A nomogram to predict postresection 5-year overall survival for patients with uterine leiomyosarcoma. *Cancer* **118**, 660–669 (2012).
- Croce, S. et al. The nanocind signature is an independent prognosticator of recurrence and death in uterine leiomyosarcomas. *Clin. Cancer Res.* **26**, 855–861 (2020).
- Italiano, A. et al. Genetic profiling identifies two classes of soft-tissue leiomyosarcomas with distinct clinical characteristics. *Clin. Cancer Res.* **19**, 1190–1196 (2013).
- Seagle, B.-L. L. et al. Prognosis and treatment of uterine leiomyosarcoma: A National Cancer Database study. *Gynecol. Oncol.* **145**, 61–70 (2017).
- Duffy, M. J. et al. Clinical use of biomarkers in breast cancer: updated guidelines from the European Group on Tumor Markers (EGTM). *Eur. J. Cancer* **75**, 284–298 (2017).
- Cardoso, F. et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **30**, 1194–1220 (2019).
- Bertucci, F., Finetti, P., Sabatier, R. & Birnbaum, D. The CINSARC signature: prognostic and predictive of response to chemotherapy? *Cell Cycle* **9**, 4025–4027 (2010).

47. Lesluyes, T., Delespaul, L., Coindre, J.-M. & Chibon, F. The CINSARC signature as a prognostic marker for clinical outcome in multiple neoplasms. *Sci. Rep.* **7**, 5480 (2017).
48. Sotiriou, C. et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl Cancer Inst.* **98**, 262–272 (2006).
49. Bertucci, F. et al. Genomic Grade Index predicts postoperative clinical outcome of GIST. *Br. J. Cancer* **107**, 1433–1441 (2012).
50. Miettinen, M. & Lasota, J. Gastrointestinal stromal tumors: review on morphology, molecular pathology, prognosis, and differential diagnosis. *Arch. Pathol. Lab. Med.* **130**, 1466–1478 (2006).
51. Bertucci, F. et al. The Genomic Grade Index predicts postoperative clinical outcome in patients with soft-tissue sarcoma. *Ann. Oncol.* **29**, 459–465 (2018).
52. Sullivan, R. & Graham, C. H. Hypoxia-driven selection of the metastatic phenotype. *Cancer Metastasis Rev.* **26**, 319–331 (2007).
53. Hockel, M. et al. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. *Cancer Res.* **56**, 4509–4515 (1996).
54. Nordmark, M. et al. Hypoxia in human soft tissue sarcomas: adverse impact on survival and no association with p53 mutations. *Br. J. Cancer* **84**, 1070–1075 (2001).
55. Brizel, D. M., Sibley, G. S., Prosnitz, L. R., Scher, R. L. & Dewhirst, M. W. Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. *Int. J. Radiat. Oncol. Biol. Phys.* **38**, 285–289 (1997).
56. Brizel, D. M. et al. Tumor oxygenation predicts for the likelihood of distant metastases in human soft tissue sarcoma. *Cancer Res.* **56**, 941–943 (1996).
57. Evans, S. M. et al. Hypoxia in human intraperitoneal and extremity sarcomas. *Int. J. Radiat. Oncol. Biol. Phys.* **49**, 587–596 (2001).
58. Evans, S. M. et al. EF5 binding and clinical outcome in human soft tissue sarcomas. *Int. J. Radiat. Oncol. Biol. Phys.* **64**, 922–927 (2006).
59. Forker, L. et al. The hypoxia marker CAIX is prognostic in the UK phase III Vortex-Biobank cohort: an important resource for translational research in soft tissue sarcoma. *Br. J. Cancer* **118**, 698–704 (2018).
60. Harris, B. H. L., Barberis, A., West, C. M. L. & Buffa, F. M. Gene expression signatures as biomarkers of tumour hypoxia. *Clin. Oncol.* **27**, 547–560 (2015).
61. Toustrup, K. et al. Gene expression classifier predicts for hypoxic modification of radiotherapy with nimorazole in squamous cell carcinomas of the head and neck. *Radiother. Oncol.* **102**, 122–129 (2012).
62. Aggerholm-Pedersen, N. et al. A prognostic profile of hypoxia-induced genes for localised high-grade soft tissue sarcoma. *Br. J. Cancer* **115**, 1096–1104 (2016).
63. Sørensen, B. S. et al. The usability of a 15-gene hypoxia classifier as a universal hypoxia profile in various cancer cell types. *Radiother. Oncol.* **116**, 346–351 (2015).
64. Eisinger-Mathason, T. S. K. et al. Hypoxia-dependent modification of collagen networks promotes sarcoma metastasis. *Cancer Discov.* **3**, 1190–1205 (2013).
65. Payen, V. L., Brisson, L., Dewhirst, M. W. & Sonveaux, P. Common responses of tumors and wounds to hypoxia. *Cancer J.* **21**, 75–87 (2015).
66. Yang, L. et al. Validation of a hypoxia related gene signature in multiple soft tissue sarcoma cohorts. *Oncotarget* **9**, 3946–3955 (2018).
67. Luoto, K. R., Kumareswaran, R. & Bristow, R. G. Tumor hypoxia as a driving force in genetic instability. *Genome Integr.* **4**, 5 (2013).
68. Simon, R. M., Paik, S. & Hayes, D. F. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J. Natl Cancer Inst.* **101**, 1446–1452 (2009).
69. Woll, P. J. et al. Adjuvant chemotherapy with doxorubicin, ifosfamide, and lenograstim for resected soft-tissue sarcoma (EORTC 62931): a multicentre randomised controlled trial. *Lancet Oncol.* **13**, 1045–1054 (2012).
70. Adjuvant chemotherapy for localised resectable soft tissue sarcoma in adults. *Cochrane database Syst. Rev.* <https://doi.org/10.1002/14651858.CD001419> (2000).
71. Pervaiz, N. et al. A systematic meta-analysis of randomized controlled trials of adjuvant chemotherapy for localized resectable soft-tissue sarcoma. *Cancer* **113**, 573–581 (2008).
72. Gronchi, A. et al. Short, full-dose adjuvant chemotherapy in high-risk adult soft tissue sarcomas: a randomized clinical trial from the Italian Sarcoma Group and the Spanish Sarcoma Group. *J. Clin. Oncol.* **30**, 850–856 (2012).
73. Gronchi, A. et al. Neoadjuvant chemotherapy in high-risk soft tissue sarcomas: final results of a randomized trial from Italian (ISG), Spanish (GEIS), French (FSG), and Polish (PSG) Sarcoma Groups. *J. Clin. Oncol.* **38**, 2178–2186 (2020).
74. Lee, A. T. J. et al. The adequacy of tissue microarrays in the assessment of inter- and intra-tumoural heterogeneity of infiltrating lymphocyte burden in leiomyosarcoma. *Sci. Rep.* **9**, 14602 (2019).
75. Aynaoud, M.-M. et al. Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. *Cell Rep.* **30**, 1767–1779.e6 (2020).
76. Sturm, G. et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445 (2019).
77. Kattan, M. W., Leung, D. H. Y. & Brennan, M. F. Postoperative nomogram for 12-year sarcoma-specific death. *J. Clin. Oncol.* **20**, 791–796 (2002).
78. Callegaro, D., Miceli, R., Mariani, L., Raut, C. P. & Gronchi, A. Soft tissue sarcoma nomograms and their incorporation into practice. *Cancer* **123**, 2802–2820 (2017).
79. Gobble, R. M. et al. Expression profiling of liposarcoma yields a multigene predictor of patient outcome and identifies genes that contribute to liposarcomagenesis. *Cancer Res.* **71**, 2697–2705 (2011).
80. Kondo, T. Current status of proteomics of soft tissue sarcomas. *Expert Rev. Proteom.* **14**, 1131–1140 (2017).
81. Liu, P. et al. Aberrant DNA methylations in chondrosarcoma. *Epigenomics* **8**, 1519–1525 (2016).
82. Burns, J., Wilding, C. P., L Jones, R., Huang, H. & Proteomic, P. research in sarcomas—current status and future opportunities. *Semin. Cancer Biol.* **61**, 56–70 (2020).
83. Nacev, B. A. et al. The epigenomics of sarcoma. *Nat. Rev. Cancer* <https://doi.org/10.1038/s41568-020-0288-4> (2020).
84. Darwin, P., Toor, S. M., Sasidharan Nair, V. & Elkord, E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp. Mol. Med.* **50**, 1–11 (2018).
85. Thomas, A., Murai, J. & Pommier, Y. The evolving landscape of predictive biomarkers of response to PARP inhibitors. *J. Clin. Invest.* **128**, 1727–1730 (2018).
86. Bertucci, F. et al. PDL1 expression is a poor-prognosis factor in soft-tissue sarcomas. *Oncoimmunology* **6**, e1278100 (2017).
87. Bertucci, F. et al. PARP1 expression in soft tissue sarcomas is a poor-prognosis factor and a new potential therapeutic target. *Mol. Oncol.* **13**, 1577–1588 (2019).
88. Lesluyes, T. & Chibon, F. A global and integrated analysis of CINSARC-associated genetic defects. *Cancer Res.* **80**, 5282–5290 (2020).
89. Krzyszczyk, P. et al. The growing role of precision and personalized medicine for cancer treatment. *Technology* **6**, 79–100 (2018).

## ACKNOWLEDGEMENTS

Grants from Cancer Research UK (C56167/A29363) and the Royal Marsden/Institute of Cancer Research National Institute for Health Research Biomedical Research Centre.

## AUTHOR CONTRIBUTIONS

E.M., R.L.J., and P.H.H. contributed to the conception, drafting and design of the paper. K.T., contributed to the drafting and design of the paper. All authors contributed to the writing, preparation and approved the final submitted manuscript.

## COMPETING INTERESTS

R.L.J.: Consulting or Advisory Role: Lilly, Immune Design, Merck Serono, Adaptimmune, Daiichi Sankyo, Eisai, Morphotek, TRACON Pharmaceuticals, Immodulon Therapeutics, Deciphera Pharmaceuticals, PharmaMar, Blueprint Medicines, Clinigen Group, Epizyme, Boehringer Ingelheim, Upto Date. The remaining authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to P.H.H.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021