

Article

# OSmfs: An Online Interactive Tool to Evaluate Prognostic Markers for Myxofibrosarcoma

Huimin Li <sup>1,†</sup>, Longxiang Xie <sup>1,†</sup>, Qiang Wang <sup>2,†</sup>, Yifang Dang <sup>1</sup>, Xiaoxiao Sun <sup>1</sup>, Lu Zhang <sup>1</sup>, Yali Han <sup>1</sup>, Zhongyi Yan <sup>1</sup>, Huan Dong <sup>1</sup>, Hong Zheng <sup>1</sup>, Yongqiang Li <sup>1</sup>, Wan Zhu <sup>3</sup> and Xiangqian Guo <sup>1,\*</sup>

<sup>1</sup> Cell Signal Transduction Laboratory, Bioinformatics Center, Henan Provincial Engineering Center for Tumor Molecular Medicine, Institute of Biomedical Informatics, School of Medical Sciences, Henan University, Kaifeng 475001, China; lihuimin0202@163.com (H.L.); xielongxiang123@126.com (L.X.); dangyifanga@163.com (Y.D.); sunxiaoxiao1654@163.com (X.S.); zhanglu9128@126.com (L.Z.); hangreen5508@163.com (Y.H.); yanzhongyi2008@163.com (Z.Y.); donghuan8626@163.com (H.D.); medhenu@126.com (H.Z.); liyongqiang@vip.henu.edu.cn (Y.L.)

<sup>2</sup> School of Software, Henan University, Kaifeng 475001, China; qiangwang@henu.edu.cn

<sup>3</sup> Department of Anesthesia, Stanford University, Stanford, CA 94305, USA; ms.wanzhu@gmail.com

\* Correspondence: xqguo@henu.edu.cn

† These authors have contributed equally to this work.

Received: 28 October 2020; Accepted: 17 December 2020; Published: 19 December 2020



**Abstract:** Myxofibrosarcoma is a complex genetic disease with poor prognosis. However, more effective biomarkers that forebode poor prognosis in Myxofibrosarcoma remain to be determined. Herein, utilizing gene expression profiling data and clinical follow-up data of Myxofibrosarcoma cases in three independent cohorts with a total of 128 Myxofibrosarcoma samples from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases, we constructed an easy-to-use web tool, named Online consensus Survival analysis for Myxofibrosarcoma (OSmfs) to analyze the prognostic value of certain genes. Through retrieving the database, users generate a Kaplan–Meier plot with log-rank test and hazard ratio (HR) to assess prognostic-related genes or discover novel Myxofibrosarcoma prognostic biomarkers. The effectiveness and availability of OSmfs were validated using genes in ever reports predicting the prognosis of Myxofibrosarcoma patients. Furthermore, utilizing the cox analysis data and transcriptome data establishing OSmfs, seven genes were selected and considered as more potentially prognostic biomarkers through overlapping and ROC analysis. In conclusion, OSmfs is a promising web tool to evaluate the prognostic potency and reliability of genes in Myxofibrosarcoma, which may significantly contribute to the enrichment of novel potential prognostic biomarkers and therapeutic targets for Myxofibrosarcoma.

**Keywords:** Myxofibrosarcoma; OSmfs; tool; prognostic markers

## 1. Introduction

Based on the morphologic and clinicopathologic classification criteria established in 2002 and redefined in 2013, the myxoid variant of malignant fibrous histiocytoma (MFH) with a predominant myxoid component (>50%) was renamed as Myxofibrosarcoma (MFS) by World Health Organization (WHO) [1,2]. MFS is a common adult sarcoma with the traits of curvilinear vessels in the variable myxoid stroma and multinodular growth of spindle to polygonal cells [3,4]. MFS tumors harbor highly complex karyotypes, often sharing the aberrations observed in leiomyosarcoma (LMS) and undifferentiated pleomorphic sarcoma (UPS) [5]. The overall recurrence rate of MFS is up to 50–60% [6]. To guide the clinical management of MFS, prognostic biomarkers are essential to be explored and developed.

Previous MFS studies by array comparative genomic hybridization (CGH) have demonstrated that MFS is the most highly complicated sarcoma and exhibits a large degree of molecular heterogeneity. The most common copy number gain/amplification was on 5p, occurring in 60% of MFS [7]. The genes in 5p include S-Phase Kinase Associated Protein 2 (*SKP2*) and  $\alpha$ -Methylacyl-CoA Racemase (*AMACR*), both of which have been reported to have a cumulative effect and are potential oncogenes involving MFS tumorigenesis [8,9]. Additionally, overexpression of MET Proto-Oncogene, Receptor Tyrosine Kinase (*MET*), CD109 molecule (*CD109*), and *Ezrin* (*EZR*) have been reported as potential biomarkers for the aggressive behavior of MFS [10]. However, the prognosis biomarker and risk stratification for MFS has been less investigated up to now.

Recently, the employment of public omics data has gradually become an important means to discover new prognostic biomarkers. Xie et al. found that *KRT8* is an importantly potential prognostic biomarker of lung adenocarcinoma according to public omics data [11]. Furthermore, multiple public resources, such as OncoLnc [12], UALCAN [13] and Kaplan–Meier Plotter [14], provide an available interface to explore the association of survival with gene expression using data from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) database. Our previous studies also depicted web tools to determine the prognostic value of genes in leiomyosarcoma [15] and bladder cancer [16]. However, until now, there is no effective tool to enforce the evaluation of the prognostic value of genes in MFS. Collectively, the aim of this study is to develop a specialized tool, named Online consensus Survival analysis (OSmfs) for MyxoFibroSarcoma, to evaluate the prognostic value of genes in MFS.

## 2. Materials and Methods

### 2.1. Data Collection

There were three independent gene expression datasets with clinical follow-up information of MFS collected from GEO and TCGA by searching with the keywords of “Myxofibrosarcoma” and “survival” or “prognosis”. In total, 128 unique MFS cases were chosen for OSmfs construction; the clinical characteristics of each dataset used in OSmfs were listed and summarized in Table 1. The three independent datasets of TCGA, GSE71118 [17], and GSE72545 [18] from TCGA and GEO databases include 25, 39, and 64 myxofibrosarcoma (MFS) samples, with 7, 10 and 21 death events, respectively. Furthermore, only GSE71118 has 10 metastatic MFS samples. Data in TCGA of MFS come from RNA sequencing; however, GSE71118 and GSE72545 belong to microarrays data. The data from different datasets may possess batch effects. The combined datasets mean that each cohort was divided separately into subgroups (based on high vs. low expression of an inputted gene), which are then pooled for survival analysis.

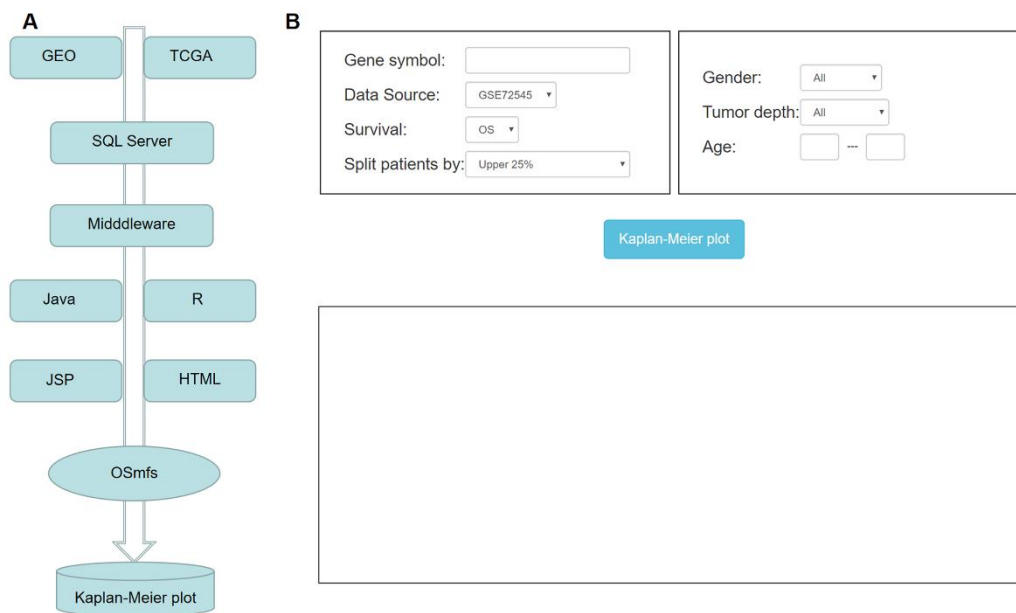
### 2.2. Design of OSmfs

The method to develop OSmfs has been previously described [15]. In short, the dynamic web interfaces are developed in HTML 5.0 and hosted by Tomcat in a Windows server system. The server-side scripts developed in Java control the output of the analysis results, the R package “R serve,” acts as a middleware to connect R and Java. The SQL Server is used to store and integrate gene expression profiles and clinical data. As the webserver is “out-of-the-box”, when users input the official gene symbol, the prognosis analyses will be performed by the R package “survival” to generate the Kaplan–Meier (KM) curves with a hazard ratio (HR, 95% confidence interval (CI)) and calculate log-rank *p*-value (Figure 1A). OSmfs is freely accessible at <http://bioinfo.henu.edu.cn/MFS/MFSList.jsp>.

**Table 1.** Datasets used in Online consensus Survival analysis for Myxofibrosarcoma (OSmfs).

Dataset	Platform	Clinical Outcomes	No. of Samples	Death Event	Data Sources	Gender (M/F)	Age (Median $\pm$ SD)	Metastasis	Tumor Depth (Deep/Superficial)
TCGA	RNAseq	OS, DFI, PFI, DSS, PFS	25	7	TCGA	11/14	60.0 $\pm$ 14.78	NA	21/4
GSE71118	GPL570-55999	<i>MFS</i>	39	10	GEO	NA	NA	10	NA
GSE72545	GPL96-57554	OS	64	21	GEO	23/41	63.5 $\pm$ 16.67	NA	52/12

OS: Overall survival; DFI: Disease free interval; PFI: Progression free interval; DSS: disease specific survival; PFS: progression free survival; *MFS*: Metastasis free survival (The Italic is aimed to differentiate Metastasis free survival and Myxofibrosarcoma); NA: Not available. SD: Standard Deviation.



**Figure 1.** The System flow diagram and operation interface of OSmfs. (A) System flow diagram of OSmfs construction. (B) The input interface of prognosis analysis in OSmfs.

### 2.3. Venny Analysis

Utilizing the data of Cox analysis according to transcription profile data and clinical information in TCGA, GSE71118 and GSE72545, we picked out genes with prognostic significance ( $p < 0.05$ ) in MFS in each of the three datasets, and then compared the three gene lists to look for the common prognostic genes in the three cohorts by Venny 2.0.2.

### 2.4. Receiver Operating Characteristic (ROC) Analysis

Utilizing overall survival (OS) event state (0: Alive; 1: Dead) information, we divided the transcriptional data of one gene into two groups. The area under the curve (AUC) scores represents the capacity for certain genes to predict alive and dead state in overall survival. An AUC of 0.5 represents a test with no discriminating ability, whereas an AUC of 1.0 represents a test with perfect discrimination.

## 3. Results

### 3.1. Application of OSmfs

OSmfs is a web tool that assesses the prognostic value of genes in MFS. To apply OSmfs, users need to input an official gene symbol or a gene signature (one gene per line), specify the “Data Source” including “TCGA”, “GSE71118”, “GSE72545” and “Combined” (the combination of the above three datasets), specify the “Survival” including OS, disease-free interval (DFI), progression-free interval (PFI), progression-free survival (PFS), disease-specific survival (DSS) and metastasis-free survival (MFS), and select one cutoff (Upper 25%, Upper 30%, Upper 50%, Upper 25% vs. Lower 25%, Upper 30% vs. Lower 30%, Upper 50% vs. Lower 50%, Lower 25%, Lower 30%, Lower 50%, Trichotomy and Quartile) in “Split patients by” item. With regard to follow-up clinical information, when one selects TCGA or GSE72545 in the “Data Source” column, three clinical factors named Gender (All, Male and Female), Tumor depth (All, Deep, Superficial) and Age (Any scope) can be chosen (Figure 1B) when one selects GSE71118 dataset, clinical factor Metastasis (Metastasis, No and All) can be selected to perform multivariate analysis whose results are illustrated through the Kaplan–Meier curves with HR (95% CI) and log-rank  $p$ -value, which allow users to evaluate the validity and reliability of prognostic biomarker candidates.

### 3.2. Validation of Prior MFS Biomarkers in OSmfs

To assess the property and reliability of OSmfs, we quarried 12 genes acting as poor prognostic biomarkers in either mRNA or protein level in MFS patients through retrieving Pubmed, then, cleared that the overexpression of 7 genes in the 12 genes, including Integrin Subunit  $\alpha$  10 (*ITGA10*) [18], *CD109* [6], Cyclin Dependent Kinase 6 (*CDK6*) [19], Cyclin Dependent Kinase Inhibitor 2A (*CDKN2A*) [20], *MET* [21], *Cyclin D1* (*CCND1*) [20] and *EZR* [22] predict adverse survival for MFS patients according to OSmfs online analysis (Table 2), however, the other 5 genes including *AMACR* [8], *SKP2* [9], KRAS Proto-Oncogene G TPase (*KRAS*) [23], Epidermal Growth Factor Receptor (*EGFR*) [24], and Argininosuccinate Synthase 1 (*ASS1*) [25] were not identified as prognosticators in MFS possibly due to difference of the size or clinical information of datasets used in previous studies and that of the datasets adopted in OSmfs (Table 2). Moreover, the mRNA-level expression data were utilized in OSmfs while the prognostic role in MFS for genes in ever reports were concluded based on mRNA or protein-level data. Hence, under the absence of prognostic tools for MFS, OSmfs may be an efficient web tool to evaluate the prognostic value of genes and explore new biomarkers in MFS patients.

### 3.3. Identification of Potentially Novel Prognostic Biomarkers in MFS

To identify novel potential prognostic biomarkers in MFS patients, synthesized cox analysis of patients' survival was performed with data derived from TCGA, GSE71118 and GSE72545 listed in Table 1. The overlapping results of genes potentially and significantly presenting adverse prognosis illustrated that 7 genes were all associated with poor prognosis of MFS in the three cohorts (Figure 2).

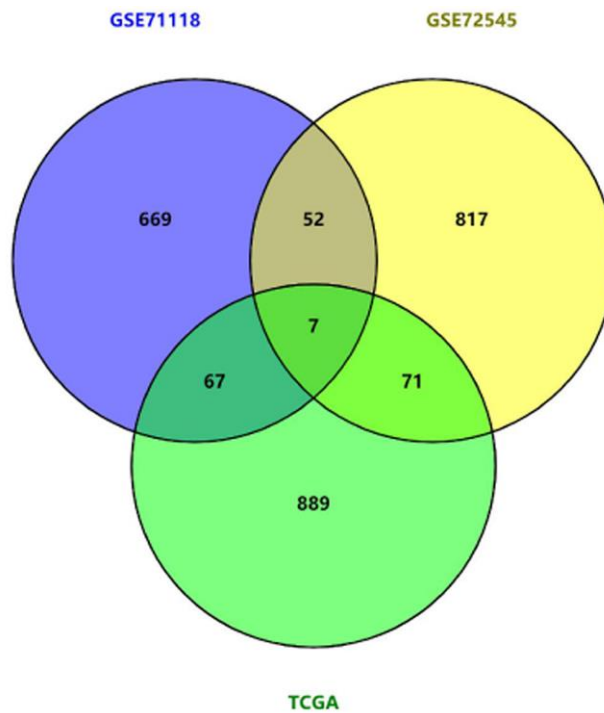
The gene overexpression of the seven genes named Lysophospholipase 1 (*LYPLA1*), DBF4 Zinc Finger B (*DBF4B*), Matrix Metalloproteinase 13 (*MMP13*), Polo Like Kinase 1 (*PLK1*), Trans-membrane Protein 158 (*TMEM158*), Wnt Family Member 5B (*WNT5B*), and RUNX Family Transcription Factor2 (*RUNX2*) potentially predicted poor overall survival ( $p < 0.05$ ,  $HR > 1$ ) in MFS based on the three independent cohorts (Table 3). To be specific, *LYPLA1* overexpression independently correlated with worse overall survival ( $p = 0.0223$ ,  $HR = 5.8292$ ) (Figure 3A) according to TCGA, with adverse metastasis-free survival ( $p = 0.0108$ ,  $HR = 5.1703$ ) according to GSE71118 (Figure 3B), with worse overall survival ( $p = 0.0067$ ,  $HR = 3.3368$ ) according to GSE72545 (Figure 3C), and with worse overall survival ( $p = 0.0004$ ,  $HR = 3.8115$ ) according to combined analysis (Figure 3D), respectively. The role as poor prognostic markers for the other 6 genes including *DBF4B*, *MMP13*, *PLK1*, *TMEM158*, *WNT5B* and *RUNX2* was also reflected using Kaplan–Meier curves in Figure 4A–F. Novelty and in conclusion, in Pubmed searching, none of the 7 genes have been reported to have any association with the prognosis of MFS, suggesting the seven genes may be novelty and potentially prognostic markers in MFS.

To further validate the sensitivity and specificity of the 7 genes distinguishing alive and dead state of overall survival in MFS, we performed ROC analyses using transcription data and clinical information in TCGA and GSE72545, the AUC Scores of *LYPLA1*, *MMP13*, *DBF4B*, *PLK1*, *TMEM158*, *WNT5B* and *RUNX2* according to TCGA were 0.6270, 0.8413, 0.7619, 0.7460, 0.6746, 0.8016, 0.5952, respectively (Figure 5), the AUC Scores of *LYPLA1* (203007\_x\_at, 212449\_s\_at), *MMP13* (205959\_at), *DBF4B* (206661\_at), *PLK1* (202240\_at), *TMEM158* (213338\_at), *WNT5B* (221029\_s\_at), *RUNX2* (221282\_x\_at, 221283\_at, 216994\_s\_at) according to GSE72545 were respectively 0.6667, 0.6600, 0.6955, 0.6069, 0.5648, 0.7486, 0.6567, 0.6401, 0.6190, 0.6755 (Figure 6), further indicating that the 7 genes may be equipped with the capacity to be prognostic markers.

**Table 2.** Analysis of known prognostic biomarkers in OSmfs.

Literature Results							Validation Results				
Genes	Sample (n)	Detection Level	Clinical Outcomes	Valida-Tion	References	Issuing Time	Clinical Outcomes	HR(95%CI)	p Value	Probe ID	Datasets
<i>ITGA10</i>	64	RNA	DSS	Yes	18	2016	OS	6.66 (2.76–16.08)	<0.0001 <sup>a</sup>	206766_at	GSE72545
<i>CD109</i>	37	Protein	OS	Yes	6	2015	OS PFS	5.03 (1.11–22.92) 4.26 (1.40–12.96)	0.0366 <sup>a</sup> 0.0105 <sup>c</sup>		TCGA TCGA
<i>CDK6</i>	77	Protein	<i>MFS</i> , DSS	Yes	19	2012	<i>MFS</i>	4.14 (1.19–14.44) 5.04 (1.41–17.98)	0.0258 <sup>b</sup> 0.0127 <sup>b</sup>	224847_at 221198_at	GSE71118
<i>CDKN2A</i>	116	mRNA protein	OS	Yes	20	2017	OS <i>MFS</i>	2.86 (1.20–6.83) 3.52 (1.01–12.29) 3.52 (1.01–12.29)	0.0177 <sup>a</sup> 0.0483 <sup>b</sup> 0.0483 <sup>b</sup>	211156_at 207039_at 209644_x_at	GSE72545 GSE71118
<i>MET</i>	86	Protein	<i>MFS</i> , OS	Yes	21	2010	OS	5.45 (2.29–12.98) 4.12 (1.73–9.77) 3.36 (1.42–7.95) 5.07 (2.10–12.21)	0.0001 <sup>a</sup> 0.0013 <sup>a</sup> 0.0057 <sup>a</sup> 0.0003 <sup>a</sup>	203510_at 213816_s_at 211599_x_at 213807_x_at	GSE72545
<i>CCND1</i>	116	mRNA	OS	Yes	20	2017	OS	4.00 (1.68–9.54) 4.59 (1.93–10.90)	0.0018 <sup>a</sup> 0.0006 <sup>a</sup>	208712_at 208711_s_at	GSE72545
<i>EZR</i>	78	Protein	<i>MFS</i> , DSS	Yes	22	2010	OS	10.29 (1.22–86.65)	0.032 <sup>a</sup>		TCGA
<i>AMACR</i>	105	Protein	DSS, <i>MFS</i>	Yes	8	2014	<i>MFS</i> OS	1.45 (0.37–5.60) 0.54 (0.21–1.43)	0.5935 0.2186 <sup>d</sup>	Average	GSE71118 Combined
<i>SKP2</i>	82	mRNA	<i>MFS</i> , DSS, OS	Yes	9	2006	<i>MFS</i> OS	2.65 (0.74–9.46) 1.53 (0.71–3.32)	0.1328 0.2797 <sup>d</sup>	Average	GSE71118 Combined
<i>KRAS</i>	35	mRNA protein	OS	Yes	23	2009	OS	0.54 (0.40–2.21)	0.8897 <sup>d</sup>		Combined
<i>EGFR</i>	47	Protein	OS	Yes	24	2004	OS	1.72 (0.79–3.72)	0.1704 <sup>d</sup>		Combined
<i>ASS1</i>	90	Protein mRNA	DSS, <i>MFS</i>	Yes	25	2013	OS	1.50 (0.69–3.25)	0.3086 <sup>d</sup>		Combined

*ITGA10*: Integrin Subunit  $\alpha$  10; *CD109*: CD109 Molecule; *CDK6*: Cyclin Dependent Kinase 6; *CDKN2A*: Cyclin Dependent Kinase Inhibitor 2A; *MET*: MET Proto-Oncogene, Receptor Tyrosine Kinase; *CCND1*: Cyclin D1; *EZR*: Ezrin; *AMACR*:  $\alpha$ -Methylacyl-CoA Racemase; *SKP2*: S-Phase Kinase Associated Protein 2; *KRAS*: KRAS Proto-Oncogene GTPase; *EGFR*: Epidermal Growth Factor Receptor; *ASS1*: Argininosuccinate Synthase 1. Statistic Significance:  $p < 0.05$ . <sup>a</sup>: Significant  $p$ -value of OS; <sup>b</sup>: Significant  $p$ -value of *MFS*; <sup>c</sup>: Significant  $p$ -value of PFS; <sup>d</sup>: No significant  $p$ -value validated in any cohorts for OS. OS: overall survival; PFS: progression-free survival; *MFS*: metastasis-free survival; DSS: disease-specific survival.



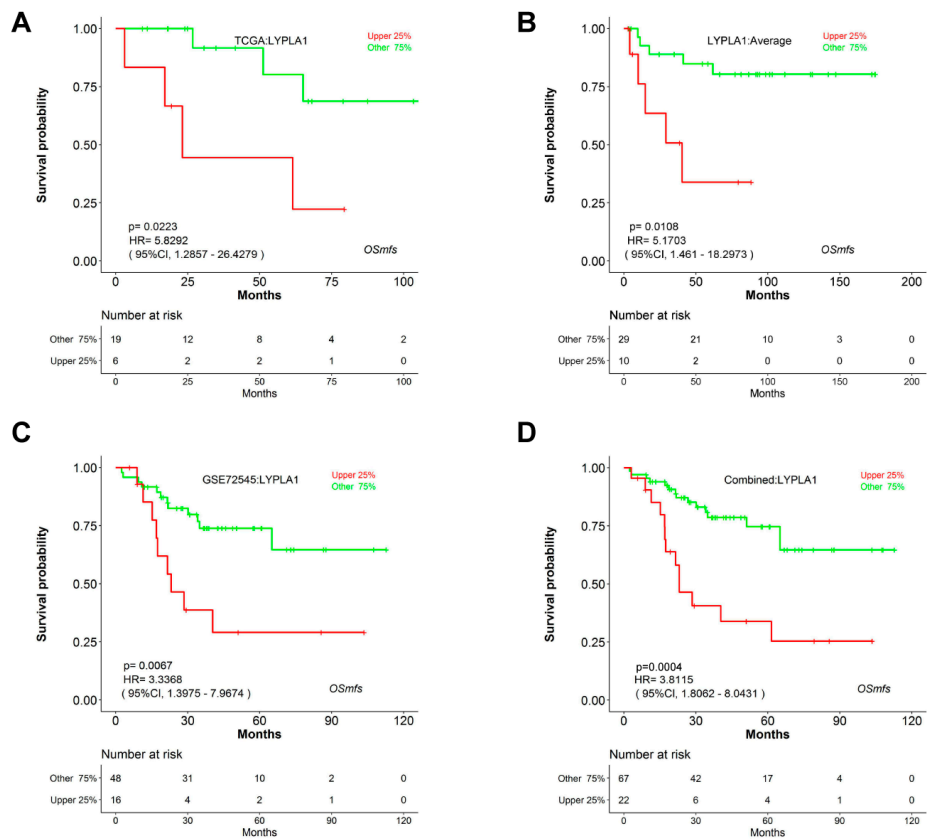
**Figure 2.** The Venn diagram analysis of the prognostic markers according to Cox analysis ( $p < 0.05$ ) using transcriptome data of MFS from TCGA, GSE71118, GSE72545 cohorts. Seven genes were selected out using the Venny online tool.

**Table 3.** The novel biomarker candidates of MFS.

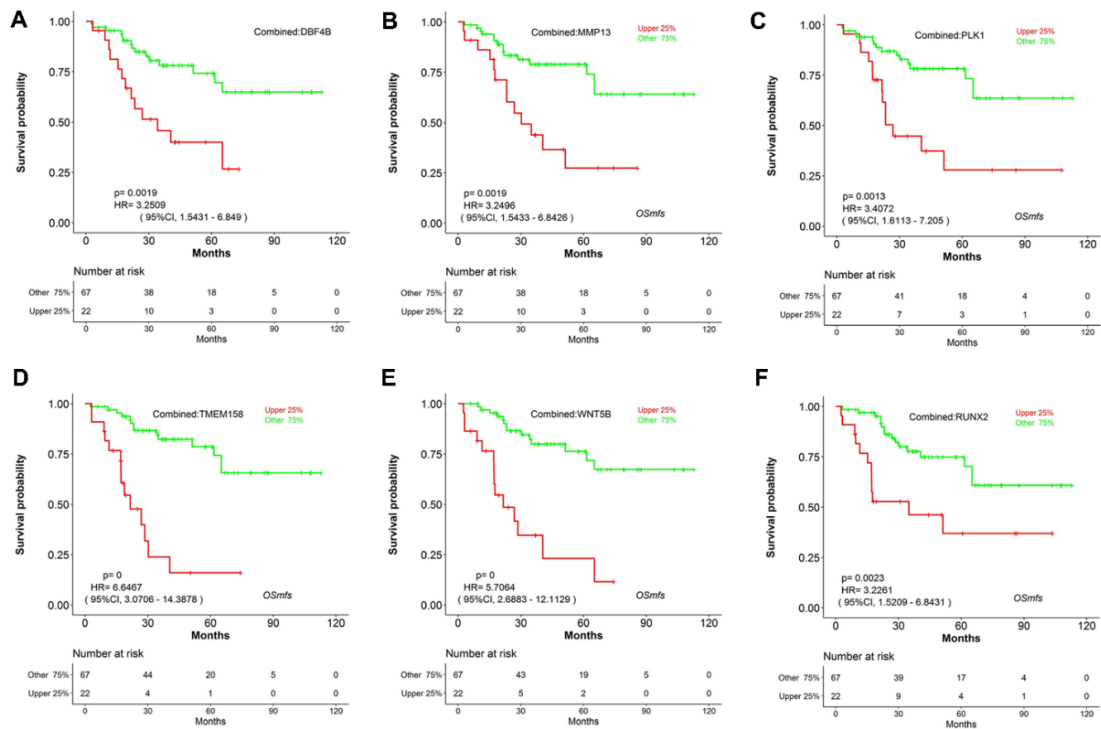
Genes	Data Source	Outcome	$p$ Value	HR (95%CI)	Cut-Off
<i>LYPLA1</i>	TCGA	OS	0.0223	5.83 (1.29–26.43)	Upper 25%
	GSE71118	MFS	0.0108	5.17 (1.46–18.30)	Upper 25%
	GSE72545	OS	0.0067	3.34 (1.40–7.97)	Upper 25%
<i>DBF4B</i>	TCGA	OS	0.0099	7.42 (1.62–34.00)	Upper 25%
	GSE71118	MFS	0.0111	5.18 (1.46–18.47)	Upper 25%
	GSE72545	OS	0.0438	2.44 (1.03–5.79)	Upper 25%
<i>MMP13</i>	TCGA	OS	0.0264	5.62 (1.22–25.84)	Upper 25%
	GSE71118	MFS	0.0003	13.13 (3.31–52.13)	Upper 25%
	GSE72545	OS	0.0216	2.76 (1.16–6.55)	Upper 25%
<i>PLK1</i>	TCGA	OS	0.0088	19.61 (2.12–181.55)	Upper 25%
	GSE71118	MFS	0.0368	3.77 (1.08–13.14)	Upper 25%
	GSE72545	OS	0.0296	2.62 (1.10–6.25)	Upper 25%
<i>TMEM158</i>	TCGA	OS	0.0143	18.57 (1.79–192.47)	Upper 25%
	GSE71118	MFS	0.0293	4.03 (1.15–14.12)	Upper 25%
	GSE72545	OS	<0.0001	5.98 (2.50–14.30)	Upper 25%
<i>WNT5B</i>	TCGA	OS	0.0099	7.42 (1.62–34.00)	Upper 25%
	GSE71118	MFS	0.012	5.02 (1.43–17.68)	Upper 25%
	GSE72545	OS	0.0002	5.14 (2.16–12.24)	Upper 25%
<i>RUNX2</i>	TCGA	OS	0.024201	8.17 (1.32–50.71)	Upper 25%
	GSE71118	MFS	0.006546	5.92 (1.64–21.36)	Upper 25%
	GSE72545	OS	0.015861	2.92 (1.22–6.97)	Upper 25%

*LYPLA1*: Lysophospholipase 1; *DBF4B*: DBF4 Zinc Finger B; *MMP13*: Matrix Metalloproteinase 13; *PLK1*: Polo Like Kinase 1; *TMEM158*: Transmembrane Protein 158; *WNT5B*: Wnt Family Member 5B; *RUNX2*: RUNX Family Transcription Factor 2.



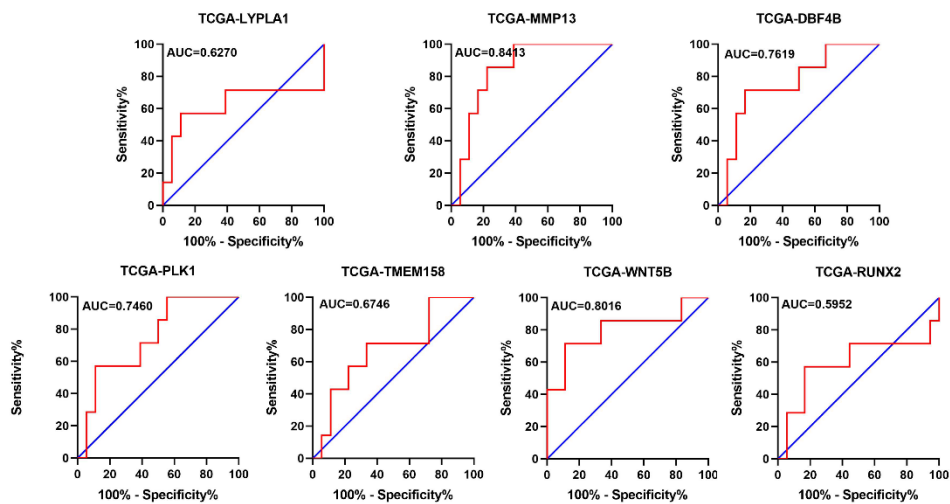


**Figure 3.** Prognostic analysis of *LYPLA1* in OSmfs. The Kaplan–Meier survival analyses of *LYPLA1* using data in TCGA (A), GSE71118 (B), GSE72545 (C) and Combined (D) were shown. *LYPLA1* displays prognostic significance for MFS patients based on survival analysis of the four cohorts.

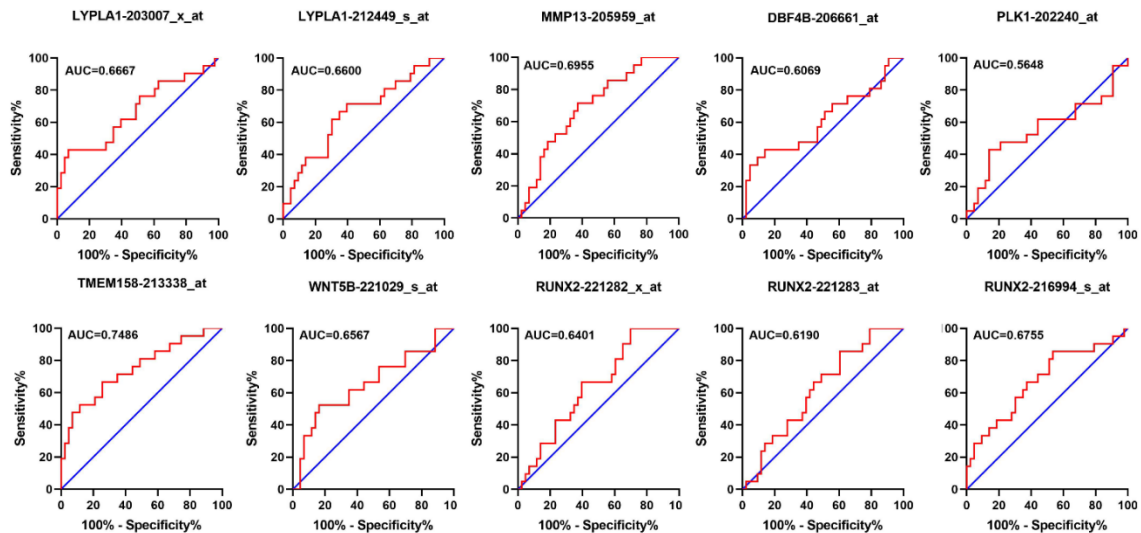


**Figure 4.** The combined overall survival analysis of *DBF4B* (A), *MMP13* (B), *PLK1* (C), *TMEM158* (D), *WNT5B* (E), *RUNX2* (F) according to TCGA, GSE71118, GSE72545 cohorts in OSmfs.





**Figure 5.** The ROC analyses of overall survival for *LYPLA1*, *MMP13*, *DBF4B*, *PLK1*, *TMEM158*, *WNT5B*, *RUNX2* according to transcription data in TCGA. ROC: receiver operator characteristic, AUC: area under the curve.



**Figure 6.** The ROC analyses of overall survival for *LYPLA1* (203007\_x\_at, 212449\_s\_at), *MMP13* (205959\_at), *DBF4B* (206661\_at), *PLK1* (202240\_at), *TMEM158* (213338\_at), *WNT5B* (221029\_s\_at), *RUNX2* (221282\_x\_at, 221283\_at, 216994\_s\_at) according to transcription data in GSE72545. ROC: receiver operator characteristic, AUC: area under the curve.

#### 4. Discussion

OSmfs, the first web tool to assess the prognostic potency and reliability of genes in MFS, may be a significant tool for the working scientific community and further contribute to the enrichment of MFS prognostic biomarkers, supplying assistance to doctors or medical oncologists; more exact diagnosis reference needs other validations. To investigate the credibility and specificity of OSmfs, we totally collected 12 genes clarified to possess worse survival in MFS patients. With regard to these markers, the prognostic significance of 7 genes analyzed in OSmfs was consistent with the results ever reported, however, the other 5 genes were not identified as prognosticators in MFS, possibly due to difference in the size or clinical information of datasets used in previous studies and that of the datasets employed in OSmfs. Moreover, the mRNA-level expression data were utilized in OSmfs while the prognostic role in MFS for genes in ever reports were concluded based on protein-level or mRNA data. In addition, the renaming of MFS in 2013 WHO classification which changes the scope of cases designated as MFS,

may be also one reason why not all biomarkers candidates previously identified to evaluate patient survival can be verified in our web tool. Interestingly, 3 of the 5 biomarkers unverified in our tool were identified as prognostic biomarkers before 2013. To be highlighted, new potential prognostic biomarkers of MFS patients can also be explored using Cox regression analysis data constructing OSmfs, such as *LYPLA1*, *DBF4B*, *MMP13*, *PLK1*, *TMEM158*, *WNT5B* and *RUNX2*, they were here detected to be potential biomarkers to evaluate poor prognosis for MFS patients. *LYPLA1* acting as a homodimer, exhibits both depalmitoylating as well as lysophospholipase activity, plays a tumor-promotor role in non-small cell lung cancer cells [26]. *DBF4B* is a serine-threonine kinase linking cell cycle regulation to genome duplication. *DBF4B-FL* is required for colon cancer cell proliferation and maintenance of genomic stability [27]. *MMP13* is involved in the breakdown of extracellular matrix in normal physiological processes and disease processes [28], they play a vital role in the prognosis of various cancers such as gastric cancer [29], colorectal cancer [30], and oral squamous cell carcinoma [31]. *PLK1* plays an important role in the initiation, maintenance, and completion of mitosis. Dysfunction of *PLK1* may promote cancerous transformation and drive tumor progression. *PLK1* overexpression was reported to be associated with poor prognoses in a variety of cancers [32]. *TMEM158* facilitates the progression of several carcinomas such as pancreatic cancer [33]. *WNT5B* has been implicated in oncogenesis and developmental processes, including regulation of cell fate and patterning during embryogenesis. Ever report indicated *WNT5B* could serve as a prognostic biomarker in hepatocellular carcinoma [34]. *RUNX2*, a transcription factor, acts as an essential factor in osteoblast differentiation and bone development and regulates a much wider tissue range [35], it could promote breast cancer bone metastasis by increasing integrin $\alpha$ 5-mediated colonization [36].

Receiver operating characteristic (ROC) curve analysis acting as an efficient tool is used extensively in medicine to present diagnostic accuracy. Recently, ROC analysis has been commonly used for characterizing the accuracy of medical imaging techniques, non-imaging diagnostic tests, and prediction/risk scores in various settings involving screening, prognosis, diagnosis, staging, and treatment [37]. The area under the ROC curve (AUC) is a global measure of the ability of a test to discriminate whether a specific condition is present or not [38]. ROC curve can also reflect the prognostic ability of the markers. AUC scores in ROC analyses represent the discriminative capacity of the markers [39]. The ROC analyses provide complementary information compared with Cox regression analysis of the potential and novel markers. In this study, the AUC scores which are all greater than 0.5 for the 7 genes indicated the discriminative capacity for the alive and dead state of overall survival in MFS.

## 5. Conclusions

Hence, OSmfs is a potential and significant prognostic analysis tool to evaluate the prognostic value of one gene or a signature in MFS. To be highlighted, the limitations for the tool including that the datasets used to build the tool are small, the datasets come from different platforms and the diagnosis cannot be 100% guaranteed for the included samples are really present. Based on the cancer statistics in 2019, the incidence of sarcomas was not high [40]. Myxofibrosarcoma (MFS) is a unique subtype of soft tissue sarcoma, approximately only 5% of soft tissue sarcoma is diagnosed to be Myxofibrosarcoma [1]. Hence, there are not much public data with regard to Myxofibrosarcoma. To collect as much data as possible, we selected data from different platforms including TCGA and GEO, which may be one reason why the diagnosis cannot be 100% guaranteed for the included samples. OSmfs will be gradually improved and updated to pull in new MFS patients' data available from TCGA or GEO datasets according to the new criterion defined by WHO in 2013.

**Author Contributions:** Conceptualization, X.G.; software, Q.W.; validation, H.L., Z.Y. and H.D.; formal analysis, L.Z.; investigation, L.X., Y.D., X.S., Y.H. and H.Z.; resources, Y.D., X.S. and Z.Y.; data curation, H.L., L.X.; writing—original draft preparation, H.L.; writing—review and editing, L.X., Q.W., L.Z., Y.H., Y.L., W.Z.; visualization, Q.W.; supervision, X.G.; funding acquisition, X.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by Program for Innovative Talents of Science and Technology in Henan Province (No. 18HASTIT048), Yellow River Scholar Program (No.H2016012), Program for Young Key Teacher of Henan Province (2016GGJS-214).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Roland, C.; Wang, W.; Lazar, A.; Torres, K. Myxofibrosarcoma. *Surg. Oncol. Clin. N. Am.* **2016**, *25*, 775–788. [[CrossRef](#)]
2. Jo, V.; Fletcher, C. WHO classification of soft tissue tumours: An update based on the 2013 (4th) edition. *Pathology* **2014**, *46*, 95–104. [[CrossRef](#)]
3. Haglund, K.; Raut, C.; Nascimento, A.; Wang, Q.; George, S.; Baldini, E. Recurrence Patterns and Survival for Patients With Intermediate- and High-Grade Myxofibrosarcoma. *Int. J. Radiat. Oncol. Biol. Phys.* **2012**, *82*, 361–367. [[CrossRef](#)]
4. Sanfilippo, R.; Miceli, R.; Grosso, F.; Fiore, M.; Puma, E.; Pennacchioli, E.; Barisella, M.; Sangalli, C.; Mariani, L.; Casali, P.; et al. Myxofibrosarcoma: Prognostic Factors and Survival in a Series of Patients Treated at a Single Institution. *Ann. Surg. Oncol.* **2010**, *18*, 720–725. [[CrossRef](#)]
5. Ogura, K.; Hosoda, F.; Arai, Y.; Nakamura, H.; Hama, N.; Totoki, Y.; Yoshida, A.; Nagai, M.; Kato, M.; Arakawa, E.; et al. Integrated genetic and epigenetic analysis of myxofibrosarcoma. *Nat. Commun.* **2018**, *9*, 2765. [[CrossRef](#)]
6. Emori, M.; Tsukahara, T.; Murata, K.; Sugita, S.; Sonoda, T.; Kaya, M.; Soma, T.; Sasaki, M.; Nagoya, S.; Hasegawa, T.; et al. Prognostic impact of CD109 expression in myxofibrosarcoma. *J. Surg. Oncol.* **2015**, *111*, 975–979.
7. Barretina, T.; Taylor, B.; Banerji, S.; Ramos, A.; Lagos-Quintana, M.; DeCarolis, P.; Shah, K.; Socci, N.; Weir, B.; Ho, A.; et al. Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. *Nat. Genet.* **2010**, *42*, 715–721. [[CrossRef](#)]
8. Li, C.; Fang, F.; Lan, J.; Wang, J.; Kung, H.; Chen, L.; Chen, T.; Li, S.; Wang, Y.; Tai, H.; et al. AMACR Amplification in Myxofibrosarcomas: A Mechanism of Overexpression That Promotes Cell Proliferation with Therapeutic Relevance. *Clin. Cancer Res.* **2014**, *20*, 6141–6152. [[CrossRef](#)]
9. Huang, H.; Kang, H.; Li, C.; Eng, H.; Chou, S.; Lin, C.; Hsiung, C. Skp2 overexpression is highly representative of intrinsic biological aggressiveness and independently associated with poor prognosis in primary localized myxofibrosarcomas. *Clin. Cancer Res.* **2006**, *12*, 487–498. [[CrossRef](#)]
10. Scoccianti, G.; Ranucci, V.; Frenos, F.; Greto, D.; Beltrami, G.; Capanna, R.; Franchi, A. Soft tissue myxofibrosarcoma: A clinico-pathological analysis of a series of 75 patients with emphasis on the epithelioid variant. *J. Surg. Oncol.* **2016**, *114*, 50–55. [[CrossRef](#)]
11. Xie, L.; Guo, J.; Sun, X.; Xie, T.; Zhang, L.; Yan, Z.; Amin, H.; Guo, X. High KRT8 Expression Independently Predicts Poor Prognosis for Lung Adenocarcinoma Patients. *Genes* **2019**, *10*, 36. [[CrossRef](#)]
12. Anaya, J. OncoLnc: Linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput. Sci.* **2016**, *2*, e67. [[CrossRef](#)]
13. Chandrashekar, D.; Bashel, B.; Balasubramanya, S.; Creighton, C.; Ponce-Rodriguez, I.; Chakravarthi, B.; Varambally, S. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* **2017**, *19*, 649–658. [[CrossRef](#)]
14. Menyhárt, O.; Nagy, Á. Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. *R. Soc. Open Sci.* **2018**, *5*, 181006. [[CrossRef](#)]
15. Wang, Q.; Xie, L.; Dang, Y.; Sun, X.; Xie, T.; Guo, J.; Han, Y.; Yan, Z.; Zhu, W.; Wang, Y.; et al. OSlms: A Web Server to Evaluate the Prognostic Value of Genes in Leiomyosarcoma. *Front. Oncol.* **2019**, *9*, 190. [[CrossRef](#)]
16. Zhang, G.; Dong, H.; Yang, M.; Xie, L.; Yuan, Q.; Zhu, W.; Dang, Y.; Sun, X.; Wang, Y.; Guo, X. OSblca: A Web Server for Investigating Prognostic Biomarkers of Bladder Cancer Patients. *Front. Oncol.* **2019**, *9*, 466.
17. Tom, L.; Gaëlle, P.; Marine, R.; Céline, B.; Pauline, L.; Valérie, D.; Carlo, L.; Agnès, N.; Philippe, T.; Dominique, V.; et al. RNA sequencing validation of the Complexity INDEX in SARComas prognostic signature. *Eur. J. Cancer* **2016**, *57*, 104–111.

18. Okada, T.; Lee, A.; Qin, L.; Agaram, N.; Mima, T.; Shen, Y.; O'Connor, R.; López-Lago, M.; Craig, A.; Miller, M.; et al. Integrin- $\alpha$ 10 Dependency Identifies RAC and RICTOR as Therapeutic Targets in High-Grade Myxofibrosarcoma. *Cancer Discov.* **2016**, *6*, 1148–1165. [[CrossRef](#)]
19. Tsai, J.; Li, C.; Kao, Y.; Wang, J.; Fang, F.; Wang, Y.; Wu, W.; Wu, L.; Hsing, C.; Li, S.; et al. Recurrent Amplification at 7q21.2 Targets CDK6 Gene in Primary Myxofibrosarcomas and Identifies CDK6 Overexpression as an Independent Adverse Prognosticator. *Ann. Surg. Oncol.* **2012**, *19*, 2716–2725. [[CrossRef](#)]
20. Heitzer, E.; Sunitsch, S.; Gilg, M.; Lohberger, B.; Rinner, B.; Kashofer, K.; Stündl, N.; Ulz, P.; Szkandera, J.; Leithner, A.; et al. Expanded molecular profiling of myxofibrosarcoma reveals potentially actionable targets. *Mod. Pathol.* **2017**, *30*, 1698–1709. [[CrossRef](#)]
21. Lee, J.; Li, C.; Fang, F.; Wang, J.; Jeng, Y.; Yu, S.; Lin, Y.; Wu, J.; Tsai, J.; Li, S.; et al. prognostic implication of MET overexpression in myxofibrosarcomas: An integrative array comparative genomic hybridization, real-time quantitative PCR, immunoblotting, and immunohistochemical analysis. *Mod. Pathol.* **2010**, *23*, 1379–1392. [[PubMed](#)]
22. Huang, H.; Li, C.; Fang, F.; Tsai, J.; Li, S.; Lee, Y.; Wei, H. Prognostic Implication of Ezrin Overexpression in Myxofibrosarcomas. *Ann. Surg. Oncol.* **2010**, *17*, 3212–3219. [[CrossRef](#)] [[PubMed](#)]
23. Willems, S.; Mohseny, A.; Balog, C.; Sewrajsing, R.; Bruijn, I.; Knijnenburg, J.; Cleton-Jansen, A.; Sciote, R.; Fletcher, C.; Deelder, A.; et al. Cellular/intramuscular myxoma and grade I myxofibrosarcoma are characterized by distinct genetic alterations and specific composition of their extracellular matrix. *J. Cell. Mol. Med.* **2009**, *13*, 1291–1301. [[CrossRef](#)]
24. Sato, O.; Wada, T.; Kawai, A.; Yamaguchi, U.; Makimoto, A.; Kokai, Y.; Yamashita, T.; Chuman, H.; Beppu, Y.; Tani, Y.; et al. Expression of epidermal growth factor receptor, ERBB2 and KIT in adult soft tissue sarcomas. *Cancer* **2005**, *103*, 1881–1890.
25. Huang, H.; Wu, W.; Wang, Y.; Wang, J.; Fang, F.; Tsai, J.; Li, S.; Hung, H.; Yu, S.; Lan, J.; et al. ASS1 as a Novel Tumor Suppressor Gene in Myxofibrosarcomas: Aberrant Loss via Epigenetic DNA Methylation Confers Aggressive Phenotypes, Negative Prognostic Impact, and Therapeutic Relevance. *Clin. Cancer Res.* **2013**, *19*, 2861–2872. [[CrossRef](#)]
26. Mohammed, A.; Zhang, C.; Zhang, S.; Shen, Q.; Li, J.; Tang, Z.; Liu, H. Inhibition of cell proliferation and migration in non-small cell lung cancer cells through the suppression of LYPLA1. *Oncol. Rep.* **2019**, *41*, 973–980. [[CrossRef](#)]
27. Chen, L.; Luo, C.; Shen, L.; Liu, Y.; Wang, Q.; Zhang, C.; Guo, R.; Zhang, Y.; Xie, Z.; Wei, N.; et al. SRSF1 Prevents DNA Damage and Promotes Tumorigenesis through Regulation of DBF4B Pre-mRNA Splicing. *Cell Rep.* **2017**, *21*, 3406–3413.
28. Pittayapruerk, P.; Meehansan, J.; Prapapan, O.; Komine, M.; Ohtsuki, M. Role of Matrix Metalloproteinases in Photoaging and Photocarcinogenesis. *Int. J. Mol. Sci.* **2016**, *17*, 868. [[CrossRef](#)]
29. Ren, J.; Liu, J.; Sui, X. Correlation of COX-2 and MMP-13 expressions with gastric cancer and their effects on prognosis. *J. BUON Off. J. Balk. Union Oncol.* **2018**, *23*, 665–671.
30. Huang, M.; Chang, H.; Chung, F.; Yang, M.; Yang, Y.; Wang, J.; Lin, S. MMP13 is a potential prognostic marker for colorectal cancer. *Oncol. Rep.* **2010**, *24*, 1241–1247.
31. Vincent-Chong, V.; Salahshourifar, I.; Karen-Ng, L.; Siow, M.; Kallarakkal, T.; Ramanathan, A. Overexpression of MMP13 is associated with clinical outcomes and poor prognosis in oral squamous cell carcinoma. *Sci. World J.* **2014**, *2014*, 897523. [[CrossRef](#)] [[PubMed](#)]
32. Liu, Z.; Sun, Q.; Wang, X. PLK1, A Potential Target for Cancer Therapy. *Transl. Oncol.* **2017**, *10*, 22–32. [[CrossRef](#)] [[PubMed](#)]
33. Fu, Y.; Yao, N.; Ding, D.; Zhang, X.; Liu, H.; Ma, L.; Shi, W.; Zhu, C.; Tang, L. TMEM158 promotes pancreatic cancer aggressiveness by activation of TGF $\beta$ 1 and PI3K/AKT signaling pathway. *J. Cell Physiol.* **2020**, *235*, 2761–2775. [[CrossRef](#)] [[PubMed](#)]
34. Dong, J.; Ying, L.; Shi, K. Expression of the Wnt ligands gene family and its relationship to prognosis in hepatocellular carcinoma. *Cancer Cell Int.* **2019**, *19*, 34. [[CrossRef](#)]
35. Ferrari, N.; McDonald, L.; Morris, J.S.; Cameron, E.R.; Blyth, K. RUNX2 in mammary gland development and breast cancer. *J. Cell Physiol.* **2013**, *228*, 1137–1142. [[CrossRef](#)]
36. Li, X.; Lu, J.; Tan, C.; Wang, Q.; Feng, Y. RUNX2 promotes breast cancer bone metastasis by increasing integrin  $\alpha$ 5-mediated colonization. *Cancer Lett.* **2016**, *380*, 78–86. [[CrossRef](#)]

37. Obuchowski, N.A.; Bullen, J.A. Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. *Phys. Med. Biol.* **2018**, *63*, 07tr01.
38. Zhe, H.; Jane, C.; Dawn, T. What is an ROC curve? *Emerg. Med. J.* **2017**, *34*, 357–359.
39. Combesure, C.; Thomas, V.P.; Damien, C.W.; Jean-Pierre, D.; Yohann, F. Prognostic ROC curves: A method for representing the overall discriminative capacity of binary markers with right-censored time-to-event endpoints. *Epidemiology* **2014**, *25*, 103–109. [[CrossRef](#)]
40. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics. *CA A Cancer J. Clin.* **2019**, *69*, 7–34. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).