Neuro-Oncology

23(4), 557–571, 2021 | doi:10.1093/neuonc/noaa280 | Advance Access date 16 December 2020

A Bayesian approach for diagnostic accuracy of malignant peripheral nerve sheath tumors: a systematic review and meta-analysis

Enrico Martin,[†] Ritchie T. J. Geitenbeek,[†] J. Henk Coert, David F. Hanff, Laura H. Graven, Dirk J. Grünhagen, Cornelis Verhoef and Walter Taal

Department of Plastic and Reconstructive Surgery, University Medical Center Utrecht, Utrecht, the Netherlands (E.M., R.T.J.G., J.H.C.); Department of Surgical Oncology, Erasmus Medical Center, Rotterdam, the Netherlands (E.M., R.T.J.G., D.J.G., C.V.); Department of Radiology and Nuclear Medicine, Erasmus Medical Center, Rotterdam, the Netherlands (D.F.H., L.H.G.); Department of Neuro-Oncology/Neurology, Erasmus Medical Center Cancer Institute, Rotterdam, the Netherlands (W.T.)

[†]These authors contributed equally to this work.

Corresponding Author: Enrico Martin, MD, Department of Plastic and Reconstructive Surgery G04.126, University Medical Center Utrecht, PO Box 85060, 3508 AB Utrecht, the Netherlands (e.martin-3@umcutrecht.nl).

Abstract

Background. Malignant peripheral nerve sheath tumors (MPNST) carry a dismal prognosis and require early detection and complete resection. However, MPNSTs are prone to sampling errors and biopsies or resections are cumbersome and possibly damaging in benign peripheral nerve sheath tumor (BPNST). This study aimed to systematically review and quantify the diagnostic accuracy of noninvasive tests for distinguishing MPNST from BPNST.

Methods. Studies on accuracy of MRI, FDG-PET (fluorodeoxyglucose positron emission tomography), and liquid biopsies were identified in PubMed and Embase from 2000 to 2019. Pooled accuracies were calculated using Bayesian bivariate meta-analyses. Individual level-patient data were analyzed for ideal maximum standardized uptake value (SUV_{max}) threshold on FDG-PET.

Results. Forty-three studies were selected for qualitative synthesis including data on 1875 patients and 2939 lesions. Thirty-five studies were included for meta-analyses. For MRI, the absence of target sign showed highest sensitivity (0.99, 95% CI: 0.94-1.00); ill-defined margins (0.94, 95% CI: 0.88-0.98); and perilesional edema (0.95, 95% CI: 0.83-1.00) showed highest specificity. For FDG-PET, SUV_{max} and tumor-to-liver ratio show similar accuracy; sensitivity 0.94, 95% CI: 0.91-0.97 and 0.93, 95% CI: 0.87-0.97, respectively, specificity 0.81, 95% CI: 0.76-0.87 and 0.79, 95% CI: 0.70-0.86, respectively. SUV_{max} ≥3.5 yielded the best accuracy with a sensitivity of 0.99 (95% CI: 0.93-1.00) and specificity of 0.75 (95% CI: 0.56-0.90).

Conclusions. Biopsies may be omitted in the presence of a target sign and the absence of ill-defined margins or perilesional edema. Because of diverse radiological characteristics of MPNST, biopsies may still commonly be required. In neurofibromatosis type 1, FDG-PET scans may further reduce biopsies. Ideal SUV_{max} threshold is ≥3.5.

Key Points

1. Biopsies may be omitted in nerve sheath tumors in the presence of a target sign and the absence of ill-defined margins or perilesional edema.
2. In neurofibromatosis type 1, FDG-PET scans may further reduce biopsies.
3. Ideal maximum standardized uptake value threshold is >3.5.

Importance of the Study

Distinguishing benign from malignant peripheral nerve sheath tumors (MPNST) can be troublesome and accuracy of imaging is debated which frequently results in the need for biopsies. MPNSTs are notorious for sampling errors and biopsies may cause nerve damage. By means of Bayesian meta-analyses, the accuracy of common MRI and PET characteristics has been quantified which guides current diagnostics for MPNST. Using the target sign on MRI, biopsies may be obviated in 40% of symptomatic patients. The presence of perilesional edema or ill-defined margins is highly suspicious for MPNST. In NF1 patients, FDG-PET scans may further decrease the need for biopsies. SUV_{max} and tumor-to-liver ratio showed equal efficacy and the SUV_{max} threshold of ≥3.5 seems to be the ideal threshold, in general, and can reduce the biopsies in another 65%. Other imaging characteristics and liquid biopsies are of interest as well, but require further research to obtain an ideal diagnostic algorithm.

Peripheral nerve sheath tumors are relatively common and include both benign and malignant tumors. Schwannomas are the most common benign peripheral nerve sheath tumors (BPNSTs) and neurofibromas make up the largest proportion of remaining BPNSTs.[1,2] Nerve sheath tumors may arise sporadically or in association with neurofibromatoses. Malignant peripheral nerve sheath tumors (MPNSTs) may, in contrast to schwannomas, arise from neurofibromas and are rare and aggressive soft-tissue sarcomas (STS), accounting for 2%-3% of all STS.[3,4] Although MPNSTs are very rare in the common population, neurofibromatosis type 1 (NF1) patients have an 8%-13% lifetime risk of developing an MPNST being the leading cause of mortality in these patients.[5,6] Prognosis of MPNSTs is poor with median survival ranging between 5 and 6 years, demanding aggressive treatment.[7,8] Adequate and timely recognition is paramount as surgical resection is key in improving survival.[7–9] This is in contrast to BPNST treatment where resection should only be considered in selected cases but which can be removed by intracapsular resections, minimizing neurologic damage.[10,11]

Unfortunately, BPNSTs and MPNSTs are difficult to distinguish based on presenting symptoms.[12,13] Computed tomography and ultrasound play a limited role in the diagnostic work-up and are mainly used to guide biopsies. Magnetic resonance imaging (MRI) should be used to further characterize lesions, but several studies argue that MRIs alone are insufficiently reliable to detect MPNSTs.[14,15] Biopsies are therefore commonly used, but may be needlessly cumbersome and because of their origin in nerve tissue biopsies are often painful and, in some cases, may lead to persisting nerve damage.[16] Additionally, MPNSTs commonly arise within neurofibromas and harbor significant intratumoral heterogeneity making them prone to sampling errors possibly more so than other sarcomas.[17,18] Lastly, not all lesion sites are approachable for biopsy.[19] In NF1 patients, the use of ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET-CT) has gained popularity as several researches have suggested a high sensitivity of detecting MPNSTs using maximum standardized uptake values (SUV_{max}) as a quantitative metabolic imaging marker. However, ideal threshold values remain unknown and suggested thresholds may yield high false positive rates leading to unnecessary biopsies or even surgeries.[20,21] It is thus far difficult to find a balance in NF1 patients between prevention and overdiagnosis.

Over the past decades, biomarkers have established their key role in diagnosis and treatment of numerous cancers, including prostate cancer,[22] breast cancer,[23] and lung cancer.[24] Noninvasive liquid biopsies are therefore of interest as well in the diagnosis of malignant transformation in nerve sheath tumors. Percutaneous biopsies are ideally avoided, but given current uncertainties of accurately distinguishing MPNSTs and BPNSTs with noninvasive diagnostic tools, this study aimed to find diagnostic accuracies of MRI, FDG-PET (fluorodeoxyglucose positron emission tomography), and liquid biopsies by means of a systematic review and meta-analyses. These findings may result in the characterization of lesions that obviate the need for biopsies.

Methods

Literature Search

A systematic search was performed in both PubMed and Embase databases according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines, in order to identify all potentially relevant articles between January 2000 and November 2019. The search string was built with the help of a professional librarian using search terms related to "MRI," "PET," "liquid biopsy," and "MPNST." The exact search syntaxes for PubMed and Embase are shown in Supplementary Table 1. Studies were included when both extracranial MPNSTs and BPNSTs were evaluated and described their differences using MRI, FDG-PET, and/or liquid biopsy. Exclusion criteria were lack of full text, case reports, conference abstracts, and reviews. The initial review was conducted by two independent authors (E.M. and R.T.J.G.). Disagreements were solved through discussion, in which two additional authors were involved (D.F.H. and L.H.G.). By cross-referencing included articles, additional studies not initially included in our search were added.

Data Extraction

Study, patient, and diagnostic test characteristics were extracted from included studies by two independent authors (R.T.J.G. and E.M.). Values of true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) were extracted per study for all mentioned diagnostic tests. Whenever this was not directly available, the rate of MPNST and provided sensitivity and specificity were used to recalculate TP, FN, FP, and TN. A true positive corresponded to an MPNST, a true negative to a BPNST. A true negative was concluded whenever the lesion was resected, there had been a biopsy with adequate follow-up or in NF1 patients, the lesion was suspected to be benign and there had been adequate follow-up to exclude potential malignant transformation. Additionally, individual SUV_{max} values were collected when available. When the available data were insufficient for recalculation or individual SUV_{max} were missing, the corresponding authors of the study were requested for additional data. A reminder email was sent up to three times.

Statistical Analysis

Using TP, FN, FP, and TN, sensitivity, specificity, and likelihood ratios were calculated for all available diagnostic tests. Sensitivity and specificity were plotted in forest plots with 95% credible intervals (95% CI). Accuracy was based on determining the presence of an MPNST. Bayesian bivariate meta-analyses were performed on imaging characteristics included in at least three independent studies using the package "meta4diag" in R.[25] A further description for choosing this approach can be found in Supplementary File 2. In case of overlapping data between studies, data from the largest and most appropriate study were chosen for inclusion in quantitative synthesis. Penalized complexity priors were used for prior distributions.[26] Summary data were presented using summary receiver-operating characteristic (SROC) plots. Heterogeneity was assessed visually. Sources of heterogeneity were searched through subgroup analyses categorizing both FDG-PET and MRI studies in: large number of lesions (≥50 lesions), large proportion of MPNST (>33%), symptomatic lesions included only, and histologically proven lesions included only (either by biopsy or resection). MRI studies were additionally categorized for inclusion of NF1 patients only or mixed cases. Using the individual patient data of SUV_{max} values, Bayesian bivariate meta-analyses of diagnostic accuracy were performed for thresholds at 3.0, 3.5, 4.0, and 4.5. The best threshold was obtained by evaluating significant differences in sensitivity first, after which lowest sensitivity thresholds were excluded and highest specificity was evaluated. For all comparisons made, significant differences were concluded whenever the lower bound of the 95% CI of the highest accuracy did not include the mean of the lower accuracy. We anticipated only a few studies on liquid biopsies and functional MRI sequences which would exclude them from meta-analyses, thus characteristics found in these studies would be assessed qualitatively. All statistical analyses were performed using R version 3.6.0 (R Core Team, 2019).

Quality Assessment

The quality of included studies was appraised by two independent authors (R.T.J.G. and E.M.) using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Supplementary File 3). Disagreements were solved through discussion. For patient selection, case-control studies, exclusion of patients with difficult diagnosis, or inclusion of histologically proven lesions only were deemed at high risk of bias. For index testing, studies were assessed at high risk of bias when radiologists and nuclear medicine physicians were not blinded for pathology results or when new thresholds were used in results that were previously not determined in their method section. The reference standard was at high risk of bias when the pathologist was not blinded for results of the index test or if the lesion was found a BPNST without histological confirmation and a follow-up of less than 6 months. Risk of bias regarding flow and timing was only present if studies changed their reference standard during the study period. Applicability concerns were raised whenever a study was at high risk of bias.

Results

After removal of duplicates, a total of 8463 citations were identified in PubMed and Embase databases. One hundred potentially relevant articles were selected through title/abstract screening. After full-text screening 43 studies were selected for qualitative synthesis (Fig. 1). These studies included data on 1875 patients and 2939 lesions. Amongst the included studies were 12 studies on MRI characteristics, 21 studies on FDG-PET characteristics, 7 studies on both MRI and FDG-PET, and 3 studies on liquid biopsies (Table 1). Twenty-eight studies included NF1 patients only. In the remaining studies, the percentage of NF1 patients ranged from 12% to 65%. The proportion of MPNST compared to BPNST varied from 2:1 to 1:29. Thirty-five studies were included for quantitative synthesis. Diagnostic accuracies of characteristics not included in quantitative synthesis are shown in Supplementary Table 2.

Conventional MRI Characteristics

Sixteen studies describing a total of 12 conventional MRI characteristics were included for quantitative synthesis.[12,14,15,20,21,27–37] These studies included a total of 1041 tumors in 925 patients (48% NF1). Eight studies included in meta-analyses were at high risk of bias, mainly due to the inclusion of histologically confirmed lesions only or the exclusion of patients who had received treatment prior to imaging (Supplementary File 3).[20,21,27,28,30,31,33,37]

Nine studies reported on ill-defined margins.[12,14,21,28,31–33,35,36] Pooled sensitivity and specificity were 0.52 (95% CI: 0.40-0.65) and 0.94 (95% CI: 0.88-0.98), respectively (Table 2). Pooled pLR was 11.03 (3.83-31.62)

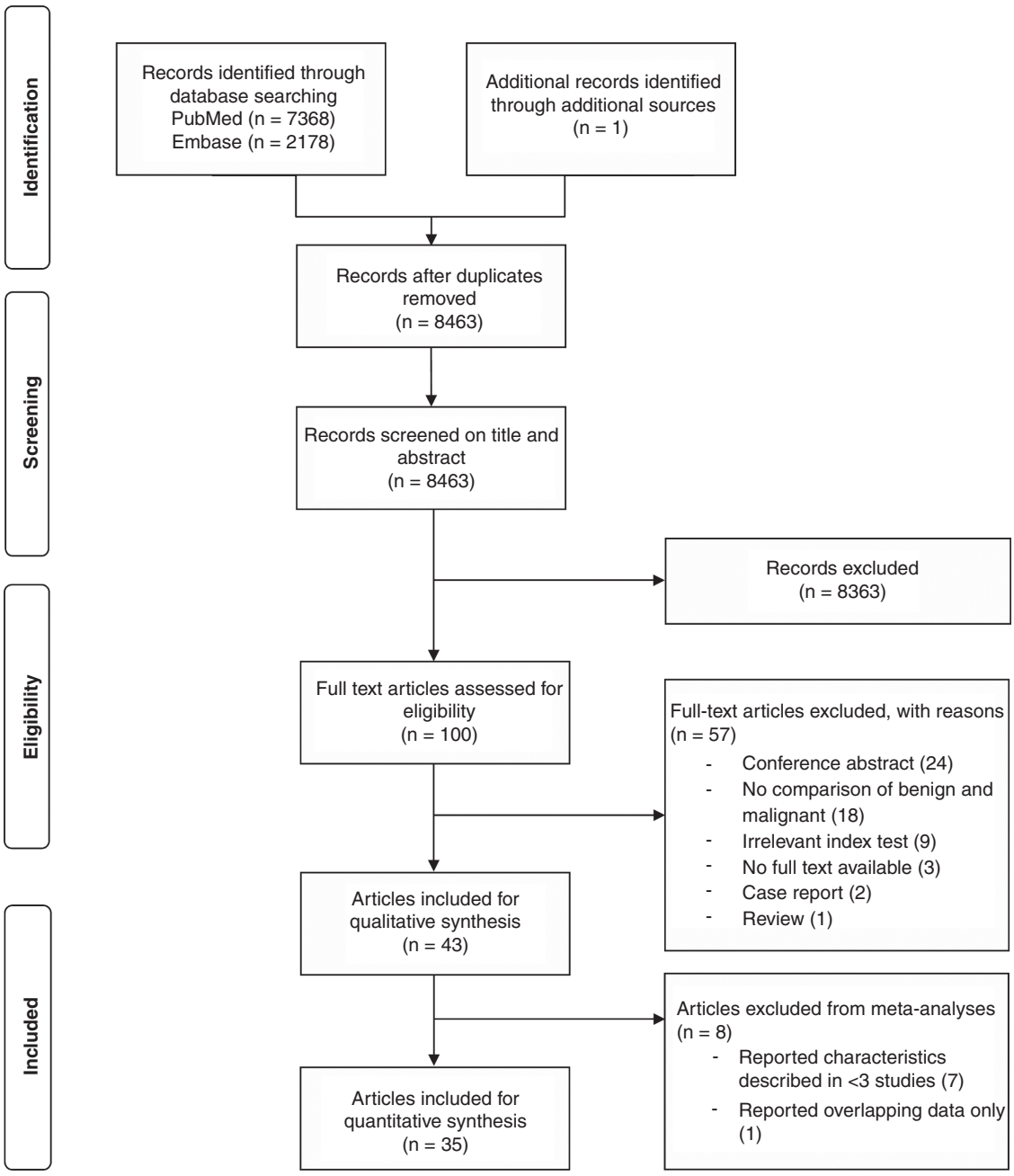

Fig. 1 Flowchart depicting study selection.

and nLR was 0.51 (0.36-0.66). The forest plot and the 95% prediction region in the SROC plot (Fig. 2, Supplementary Figure 1) demonstrated moderate heterogeneity between studies. Sensitivity was higher in studies with a smaller total sample of lesions (Supplementary Table 3). Specificity was lower in studies with a higher proportion of MPNSTs, those that included symptomatic lesions only or histologically proven lesions only.

Five studies reported on perilesional edema.[12,21,27,35,36] Pooled sensitivity and specificity were 0.65 (95% CI: 0.38-0.87) and 0.95 (95% CI: 0.83-1.00), respectively. Pooled pLR was 3415.18 (3.15-5948.77) and nLR was 0.38 (0.12-0.69). There was moderate heterogeneity between studies. Sensitivity was higher in studies with a smaller proportion of MPNST and when only NF1 patients were included and was lower when only histologically proven lesions were included.

Seven studies reported on cystic degeneration or necrosis.[12,14,21,28,34–36] Pooled sensitivity and specificity were 0.48 (95% CI: 0.23-0.71) and 0.86 (95% CI: 0.61-0.98), respectively. Pooled pLR was 5.75 (1.27-23.69) and nLR was 0.61 (0.34-0.91). There was moderate heterogeneity between

Table 1 Study Characteristics of Included Studies

Author, Year	Study Period	Mean Age (Range)	Np	NI	NF1, %	Noninvasive Diagnostics								
						Tumor Type			MRI			PET-CT		
						MPNST	NF	S	NA	Conventional	Functional	Features	TTS	Biomarkers
Ahlawat, 2018	2010-2016	40 (8-68)	42	48	50	15	11	11	11	T1/T2	DWI/ADC			
Ahlawat, 2019	2010-2017	30 (8-53)	21	55	100	19	*NA*	*NA*	36	T1/T2	DWI/ADC	SUV_{max}, size	60	
Azizi, 2018	2003-2014	14 (3-23)	41	114	100	16	*NA*	*NA*	98			SUV_{max}, T/L	60	
Bensaid, 2007	2000-2006	31 (7-77)	38	49	100	6	20	1	22			T/L	60	
Benz, 2010	2005-2008	46 (21-82)	34	40	41	17	9	14	0			SUV_{max}, size	60	
Bredella, 2007	2000-2006	37 (17-73)	45	50	100	16	8	*NA*	26			SUV_{max}, SUV_{mean}, qual	60	
Broski, 2016	2002-2014	38 (16-79)	38	43	61	20	17	6	0	T1/T2		SUV_{max}, T/L, MTV, TLG, heterogeneity	60	
Cardona, 2003	*NA*	45 (18-81)	13	25	38	13	4	2	6			SUV_{mean}, qual	60	
Chhabra, 2011	1996-2010	50 (15-92)	56	56	21	21	24	11	0	T1/T2				
Chirindel, 2015	2003-2013	36 (8-77)	41	93	100	24	*NA*	*NA*	69				60, 240	
Combemale, 2014	2000-2012	31 (2-77)	113	145	100	40	55	*NA*	50			SUV_{max}, T/L	60	
Cook, 2017	*NA*	35 (9-86)	54	54	100	24	*NA*	*NA*	30			SUV_{max}, SUV_{mean}, SUV_{peak}, GLCM, NGTDM, delayed	90, 240	
Demehri, 2014	2008-2013	38 (18-54)	29	31	45	9	6	14	2	T1/T2	DWI/ADC, DCE			
Derlin, 2013	2006-2011	30 (2-63)	31	99	100	9	*NA*	*NA*	90	T1/T2		SUV_{max}, SUV_{mean}, HI_{suv}	60	
Fayad, 2014	2009-2012	42 (11-78)	20	24	50	4	3	6	11	T1/T2	MRS		*NA*	
Ferner, 2000	1996-1998	24 (12-62)	18	23	100	7	5	*NA*	11			SUV_{mean}	60	
Ferner, 2008	1996-2005	31 (5-71)	105	116	100	29	28	*NA*	59			SUV_{max}	60	
Furniss, 2007	1995-2004	43 (3-87)	30	205	54	32	100	73	0	*NA*	*NA*			
Hummel, 2010	*NA*	21 (5-50)	32	22	100	5	*NA*	*NA*	17					ADM, HGF
Johansson, 2014	*NA*	36 (12-69)	124	*NA*	63	*NA*	*NA*	*NA*	*NA*					sAXL
Karabatsou, 2009	*NA*	38 (19-63)	9	9	100	5	4	0	0			SUV_{max}	60	
Karsy, 2016	2007-2015	41 (*NA*)	127	127	46	24	82	17	4	T1/T2				
Lerman, 2019	2005-2015	35 (15-73)	107	408	100	39	67	*NA*	302			SUV_{max}, T/L, HI_{suv}, TLG, TMTV	60	
Li, 2008	*NA*	47 (20-82)	26	26	*NA*	9	1	16	0	T1/T2				
Matsumine, 2008	*NA*	43 (14-80)	37	37	100	19	18	0	0	T1/T2				
Matsumoto, 2015	*NA*	43 (2-71)	23	23	*NA*	8	0	15	0	T1/T2				
Mautner, 2007	*NA*	25.5 (8-47)	4	4	100	1	3	0	0			SUV_{mean}	*NA*	
Meany, 2013	*NA*	18 (10-45)	14	60	100	2	8	*NA*	50			SUV_{max}	60	
Moharir, 2010	2006-2008	9 (2-14)	11	16	100	2	*NA*	*NA*	14			SUV_{max}	45	
Nose, 2013	2006-2011	52 (15-88)	*NA*	22	*NA*	10	2	10	0			SUV_{max}	60	

Table 1 Continued

Author, Year	Study Period	Mean Age (Range)	Np	NI	NF1, %	Noninvasive Diagnostics								
						Tumor Type			MRI			PET-CT		
						MPNST	NF	S	NA	Conventional	Functional	Features	TTS	Biomarkers
Park, 2013	*NA*	33 (14-63)	104	*NA*	100	30	*NA*	*NA*	39					IGFBP1, RANTES
Razek, 2018	*NA*	34 (9-64)	34	34	100	11	6	17	0	T1/T2	DWI/ADC			
Reinert, 2018	2012-2018	20 (2-44)	28	83	100	8	*NA*	*NA*	75	T1/T2	DWI/ADC	SUV_{max}, T/L, SUV_{mean}	60	
Salamon, 2013	2006-2011	33 (2-69)	50	159	100	19	*NA*	*NA*	140			SUV_{max}, SUV_{mean}, HI_{suv}, size, heterogeneity	60	
Salamon, 2014	2006-2011	33 (2-69)	49	147	100	18	*NA*	*NA*	129			SUV_{max}, T/L, SUV_{mean}, TTF, TTM	60	
Salamon, 2015	2006-2014	37 (17-69)	36	*NA*	100	19	*NA*	*NA*	*NA*			SUV_{max}, T/L, SUV_{mean}, WB-TLG, WB-MTV, size	60	
Schwabe, 2019	2007-2016	30 (9-62)	41	70	100	36	*NA*	*NA*	34	T1/T2		SUV_{max}, T/L, SUV_{peak}	60	
Tsai, 2012	2000-2011	15 (1-20)	18	26	100	10	16	0	0			SUV_{max}	60	
Van der Gucht, 2016	2006-2012	33 (*NA*)	49	149	100	16	24	*NA*	109			SUV_{max}, T/L, HI_{suv}, TLG, TMTV	60	
Warbey, 2009	2004-2008	31 (9-86)	62	85	100	21	18	*NA*	46			SUV_{max}, delayed	90, 240	
Wasa, 2010	1990-2007	42 (16-83)	61	61	56	41	20	0	0	T1/T2				
Well, 2018	2014-2017	34 (17-54)	26	67	100	12	30	*NA*	25	T1/T2	DWI/ADC			
Yu, 2016	2011-2015	53 (23-78)	34	34	12	6	26	2	0	T1/T2				

Abbreviations: ADC, apparent diffusion coefficient; ADM, adrenomedullin; BPNST, benign peripheral nerve sheath tumor; DCE, dynamic contrast enhancement; DWI, diffusion-weighted imaging; HGF, hepatocyte growth factor; HI_{suv}, SUV-based heterogeneity index obtained by dividing intratumoral SUV_{max} by SUV_{mean} of that lesion; IGFBP1, insulin-like growth factor binding protein 1; MPNST, malignant peripheral nerve sheath tumor; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; MTV, metabolic tumor volume; *NA*, not available; NF, neurofibroma; NI, number of lesions; Np, number of patients; PET-CT, positron emission tomography-computed tomography; qual, qualitative PET-CT assessment; RANTES, Regulated upon Activation Normal T-cell Expressed and Secreted; S, schwannoma; sAXL, soluble fraction from the extracellular domain of AXL; SUV_{max}, maximum standardized uptake value; SUV_{mean}, mean standardized uptake value; T/L, tumor-to-liver ratio; TLG, total lesion glycolysis; TMTV, total metabolic tumor volume; TTF, tumor-to-fat ratio; TTM, tumor-to-muscle ratio; TTS, time to scan in minutes; T1, T1-weighted imaging; T2, T2-weighted imaging.

Table 2 Pooled Diagnostic Performance of Bayesian Meta-Analyses

Characteristic	N	Median Cutoff (Range)	Pooled Accuracies (95% CI)			
			Sensitivity	Specificity	Positive LR	Negative LR
Conventional MRI						
Ill-defined margins	9	*NA*	0.52 (0.40-0.65)	0.94 (0.88-0.98)	11.03 (3.83-31.62)	0.51 (0.36-0.66)
Perilesional edema	5	*NA*	0.65 (0.38-0.87)	0.95 (0.83-1.00)	3415.18 (3.15-5948.77)	0.38 (0.12-0.69)
Cystic changes	7	*NA*	0.48 (0.23-0.71)	0.86 (0.61-0.98)	5.75 (1.27-23.69)	0.61 (0.34-0.91)
Heterogeneity onT1	3	*NA*	0.85 (0.56-1.00)	0.48 (0.03-0.96)	9.23 (0.81-31.82)	1.60 (0.01-5.42)
Heterogeneity onT2	5	*NA*	0.78 (0.64-0.90)	0.52 (0.23-0.80)	1.94 (0.90-4.82)	0.49 (0.15-1.37)
Tumor enhancement	6	*NA*	0.63 (0.50-0.76)	0.81 (0.60-0.95)	4.81 (1.44-16.60)	0.46 (0.28-0.72)
Irregular shape	4	*NA*	0.33 (0.04-0.73)	0.82 (0.71-0.90)	2.03 (0.18-5.42)	0.81 (0.26-1.22)
Intratumoral lobulation	5	*NA*	0.57 (0.41-0.72)	0.89 (0.83-0.93)	5.38 (2.87-9.31)	0.49 (0.30-0.68)
Absence of target sign	7	*NA*	0.99 (0.94-1.00)	0.33 (0.15-0.54)	1.51 (1.13-2.25)	0.04 (0.00-0.30)
Absence of split-fat sign	3	*NA*	0.76 (0.57-0.91)	0.44 (0.16-0.78)	1.67 (0.82-4.56)	0.68 (0.15-1.94)
Tumor size	3	5.0 (4.7-6.25)	0.72 (0.47-0.92)	0.85 (0.69-0.94)	5.63 (2.05-13.65)	0.34 (0.07-0.67)
Qualitative assessment	3	*NA*	0.71 (0.53-0.85)	0.92 (0.81-0.98)	12.44 (2.13-39.05)	0.32 (0.15-0.56)
PET-CT						
SUV_{max}	13	3.96 (2.35-6.1)	0.94 (0.91-0.97)	0.81 (0.76-0.87)	5.22 (3.74-7.51)	0.07 (0.03-0.12)
Tumor-to-liver ratio	7	1.77 (1.38-3.0)	0.93 (0.87-0.97)	0.79 (0.70-0.86)	4.69 (2.89-7.41)	0.09 (0.03-0.18)
Qualitative assessment	5	*NA*	0.94 (0.88-0.98)	0.82 (0.71-0.91)	5.86 (3.00-11.24)	0.07 (0.02-0.16)
PET-CT individual patient-level data						
SUV_{max}	11	3.0	0.99 (0.94-1.00)	0.61 (0.40-0.80)	2.82 (1.59-5.64)	0.02 (0.00-0.13)
SUV_{max}	11	3.5	0.99 (0.93-1.00)	0.75 (0.56-0.90)	4.66 (2.12-11.39)	0.02 (0.00-0.13)
SUV_{max}	11	4.0	0.97 (0.86-1.00)	0.78 (0.62-0.91)	5.17 (2.44-11.76)	0.04 (0.00-0.21)
SUV_{max}	11	4.5	0.86 (0.63-0.98)	0.88 (0.77-0.97)	9.59 (3.24-30.96)	0.16 (0.01-0.46)

Abbreviations: CI, credibility interval; LR, likelihood ratio; MRI, magnetic resonance imaging; N, number of studies; NA, not applicable; PET-CT, positron emission tomography-computed tomography; SUV_{max}, maximum standardized uptake value.

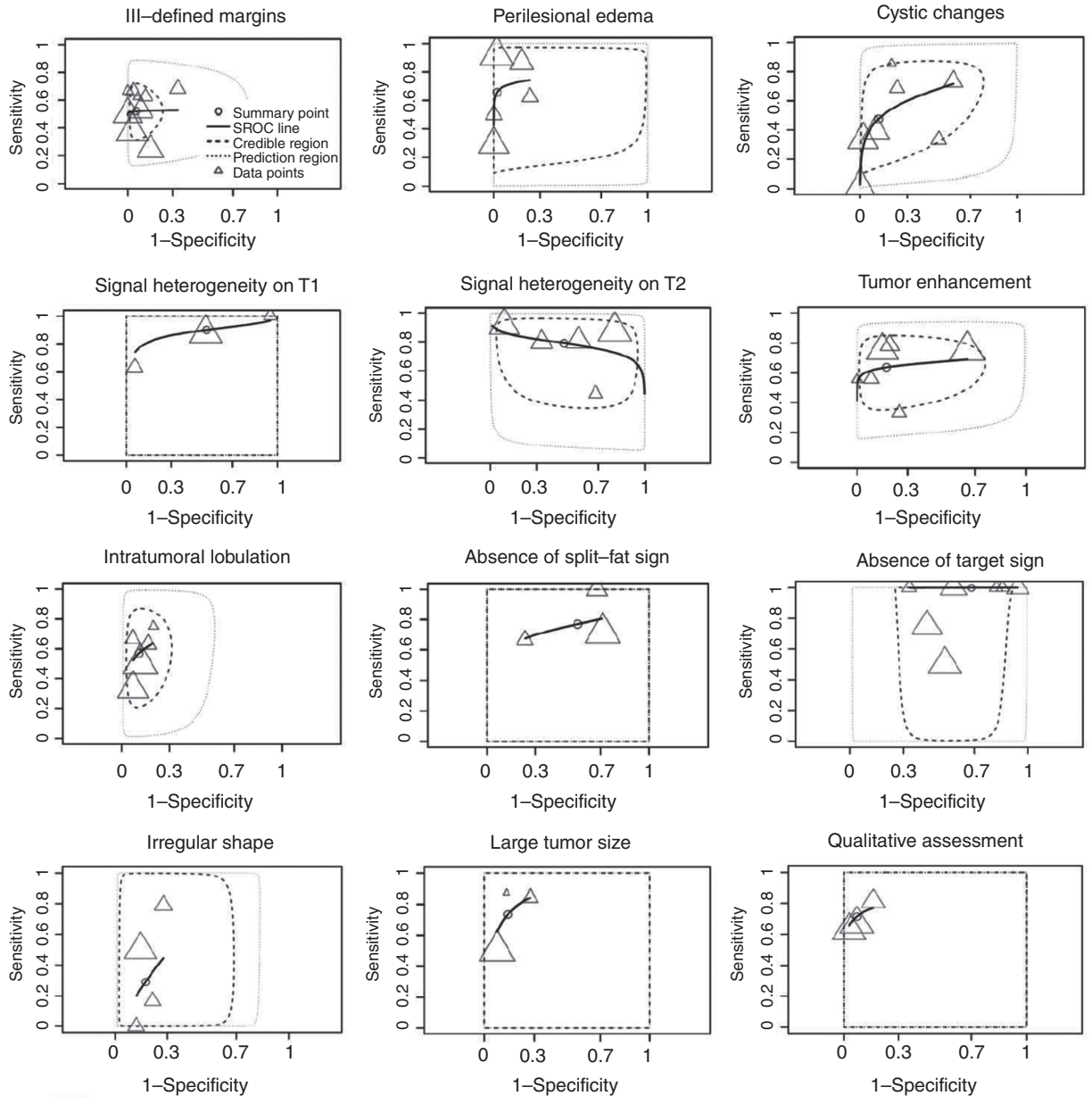

Fig. 2 SROC plots of MRI characteristics.

studies. Sensitivity was higher in studies with smaller sample of lesions, smaller proportion of MPNST, and when only histologically proven lesions were included. Specificity was higher among studies with larger sample of lesions and lower in studies including NF1 patients only or histologically proven lesions only.

Three studies reported on signal heterogeneity on T1 sequences.[14,29,32] Pooled sensitivity and specificity were 0.85 (95% CI: 0.56-1.00) and 0.48 (95% CI: 0.03-0.96), respectively. Pooled pLR was 9.23 (0.81-31.82) and nLR was 1.60 (0.01-5.42). There was substantial heterogeneity between studies. Sensitivity was lower in studies including NF1 patients only. Specificity was higher in studies including NF1 patients only and those with a higher proportion of MPNST.

Five studies reported on signal heterogeneity on T2 sequences.[14,27–29,35] Pooled sensitivity and specificity were 0.78 (95% CI: 0.64-0.90) and 0.52 (95% CI: 0.23-0.80), respectively. Pooled pLR was 1.94 (0.90-4.82) and nLR was 0.49 (0.15-1.37). There was substantial heterogeneity between studies. Sensitivity was lower in studies with a smaller sample of lesions and in those that included histologically proven lesions only.

Six studies reported on irregular or peripheral tumor enhancement after contrast administration.[12,14,15,27,31,36] Pooled sensitivity and specificity were 0.63 (95% CI:

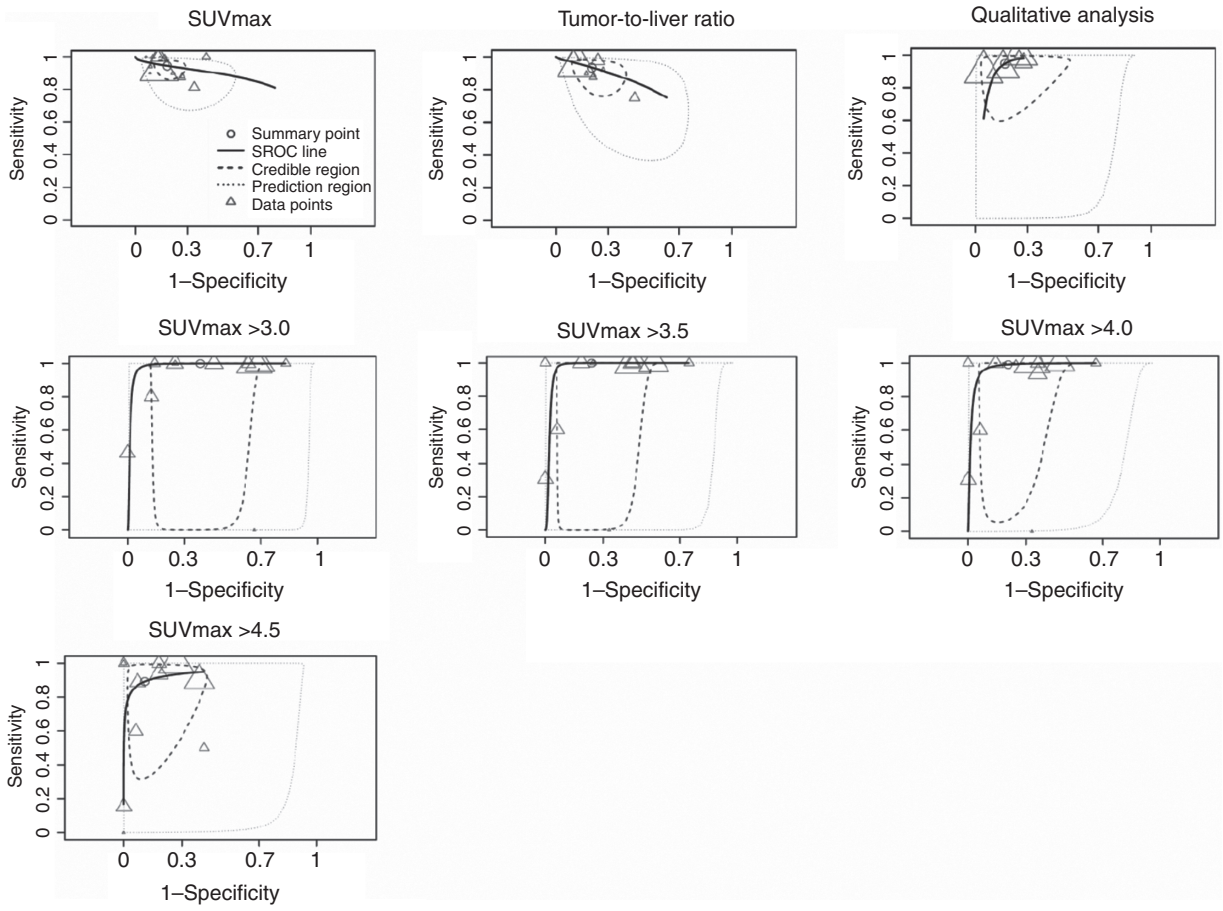

Fig. 3 SROC plots of FDG-PET characteristics.

0.50-0.76) and 0.81 (95% CI: 0.60-0.95), respectively. Pooled pLR was 4.81 (1.44-16.60) and nLR was 0.46 (0.28-0.72). There was moderate heterogeneity between studies. Sensitivity was lower in studies including histologically proven lesions only. Specificity was higher in studies with a smaller sample of lesions and higher prevalence of MPNST.

Five studies reported on intratumoral lobulation.[32,33,35,36] Pooled sensitivity and specificity were 0.57 (95% CI: 0.41-0.72) and 0.89 (95% CI: 0.83-0.93), respectively. Pooled pLR was 5.38 (2.87-9.31) and nLR was 0.49 (0.30-0.68). There was limited heterogeneity between studies. Heterogeneity in sensitivity may be caused by studies with higher total number of lesions and including NF1 patients only. No sources were found explaining heterogeneity in specificity.

Three studies reported on the absence of split-fat sign.[28,31,36] The split-fat sign represents fat deposition around the lesion and is usually seen as a tapered rim of fat signal near the proximal and distal ends of the lesion. Pooled sensitivity and specificity were 0.76 (95% CI: 0.57-0.91) and 0.44 (95% CI: 0.16-0.78), respectively. Pooled pLR was 1.67 (0.82-4.56) and nLR was 0.68 (0.15-1.94). There was limited heterogeneity between studies. Sensitivity was higher in studies with smaller proportion of MPNSTs. Specificity was higher in studies including NF1 patients only.

Seven studies studied the use of absence of target sign, a classic sign in neurogenic tumors on T2-weighted imaging referring to a high signal intensity ring peripherally surrounding an area of low signal intensity centrally.[12,15,28,32,35,36] Pooled sensitivity and specificity were 0.99 (95% CI: 0.94-1.00) and 0.33 (95% CI: 0.15-0.54), respectively. Pooled pLR was 1.51 (1.13-2.25) and nLR was 0.04 (0.00-0.30). There was substantial heterogeneity between studies. Sensitivity was higher in studies with smaller amount of lesions, higher proportion of MPNSTs, including symptomatic lesions only, and histologically proven lesions only. Sensitivity was lower in those including NF1 patients only. Specificity was higher in studies including NF1 patients only.

Four studies reported on irregular shape.[29,32,35,36] Pooled sensitivity and specificity were 0.33 (95% CI: 0.04-0.73) and 0.82 (95% CI: 0.71-0.90), respectively. Pooled pLR was 2.03 (0.18-5.42) and nLR was 0.81 (0.26-1.22). There was substantial heterogeneity between studies. Sensitivity was higher in studies including a larger proportion of MPNST and those including NF1 patients only.

Three studies reported on tumor size.[20,30,33] Thresholds varied from 4.7 to 6.3 cm. Pooled sensitivity and specificity were 0.72 (95% CI: 0.47-0.92) and 0.85 (95% CI: 0.69-0.94), respectively. Pooled pLR was 5.63 (2.05-13.65) and nLR was 0.34 (0.07-0.67). There was moderate heterogeneity between studies. Sensitivity was higher in studies with a higher proportion of MPNST. Specificity was higher in studies with smaller sample of lesions, lower proportion of MPNST, and those including histologically proven lesions only. Specificity was lower in studies including NF1 patients only.

Three studies reported on qualitative MRI assessment.[14,30,37] Pooled sensitivity and specificity were 0.71 (95% CI: 0.53-0.85) and 0.92 (95% CI: 0.81-0.98), respectively. Pooled pLR was 12.44 (2.13-39.05) and nLR was 0.32 (0.15-0.56). There was limited heterogeneity between studies, but no source of heterogeneity was found.

Functional MRI Characteristics

Six studies reported on 16 functional MRI characteristics (Supplementary Table 2).[20,27,29,35,38,39] No characteristic was evaluated in more than 2 different populations. Mean apparent diffusion coefficient (ADC_{mean}) was evaluated in two studies.[35,38] Sensitivity ranged from 0.91 to 0.92 and specificity from 0.91 to 0.98. pLR ranged from 10.46 to 50.42 and nLR from 0.09 to 0.10. Minimal ADC (ADC_{min}) was evaluated in two studies as well.[20,35] Sensitivity ranged from 0.89 to 0.98 and specificity from 0.93 to 0.94. pLR ranged from 14.15 to 14.43 and nLR from 0.03 to 0.12. One study used diffusion coefficient D and perfusion fraction f to investigate a number of characteristics.[35] Sensitivities ranged from 0.81 to 0.96 and specificity from 0.55 to 0.98. pLR ranged from 2.11 to 99.08 and nLR from 0.04 to 0.22. D_{min} and f_{center} yielded highest sensitivities (0.96), and dark and D_{margin} highest specificity (0.99). One study reported on using the target sign on ADC and diffusion-weighted imaging (DWI).[27] Sensitivity ranged from 0.80 to 0.97 and specificity from 0.39 to 0.63. pLR ranged from 1.32 to 2.64 and nLR from 0.05 to 0.51. One study evaluated early arterial enhancement on dynamic contrast enhancement MRI.[29] Sensitivity was 0.50 and specificity 0.89, with a pLR of 4.50 and nLR of 0.56. Accuracy was highest when evaluating target sign on ADC mapping with higher specificity compared to static T1-weighted imaging. One study reported on trimethylamine (TMA) peak and TMA fraction.[39] Sensitivity was 0.90 for both and specificity was 0.50 for TMA peak and 0.62 for TMA fraction. pLR ranged from 1.8 to 2.35 and nLR from 0.16 to 0.20.

FDG-PET Characteristics

Twenty studies describing a total of 3 FDG-PET characteristics were included for quantitative synthesis.[14,15,19–21,34,40–53] These studies included a total of 1850 tumors in 924 patients. Most studies scanned 60 minutes after FDG injection, except for two studies that scanned at 45- and 90-minute postinjection, respectively. Seven studies included in meta-analyses were at high risk of bias for patient selection, mainly because they included histologically confirmed lesions only or patients who had received treatment prior to imaging were excluded (Supplementary File 3).[19–21,43,46,51,53] Two studies were at high risk of bias for the use of their reference standard which was a follow-up period of ≤6 months.[42,47] One study scored a high risk of bias for index test because the nuclear medicine physician was not blinded to the pathology report.[19]

Thirteen studies reported on SUV_{max} (Table 2).[15,19–21,34,40–47] Thresholds varied from 2.35 to 6.1. Pooled sensitivity and specificity were 0.94 (95% CI: 0.91-0.97) and 0.81 (95% CI: 0.76-0.87), respectively. Pooled pLR was 5.22 (3.74-7.51) and nLR was 0.07 (0.03-0.12). The forest plot and the 95% prediction region in the SROC plot demonstrated moderate heterogeneity between studies (Fig. 3, Supplementary Figure 2). Higher specificity was found in studies that included a higher proportion of MPNST (Supplementary Table 3).

Seven studies reported on the tumor SUV_{max} to liver SUV_{mean} ratio (T/L ratio).[15,21,34,40,42,48,49] Thresholds varied from 1.4 to 3.0. Pooled sensitivity and specificity were 0.93 (95% CI: 0.87-0.97) and 0.79 (95% CI: 0.70-0.86), respectively. Pooled pLR was 4.69 (2.89-7.41) and nLR was 0.09 (0.03-0.18). There was moderate heterogeneity between studies, but no source for heterogeneity was found.

Five studies reported on qualitative FDG-PET analysis.[14,50–53] Pooled sensitivity and specificity were 0.94 (95% CI: 0.88-0.98) and 0.82 (95% CI: 0.71-0.91), respectively. Pooled pLR was 5.86 (3.00-11.24) and nLR was 0.07 (0.02-0.16). There was moderate heterogeneity between studies. Higher sensitivity was found in studies including a smaller sample of lesions. Higher specificity was found in studies that included symptomatic lesions only.

Eleven studies reported individual patient-level data of SUV_{max} on 246 patients.[20,21,43,44,46,50,51,54–57] Highest sensitivities were found for thresholds at 3.0 and 3.5 (0.99) and highest specificity was found for a threshold at 4.5 (0.88, Table 2). Accuracy was not significantly different between thresholds of 3.0, 3.5, and 4.0. However, sensitivity at a threshold of 3.5 was nonsignificantly higher than 4.0 (0.99 vs 0.97) and specificity was higher at 3.5 compared to 3.0 (0.75 vs 0.61). There was substantial heterogeneity between studies (Fig. 3, Supplementary Figure 2). Sensitivity was higher in studies including a larger amount of lesions and a higher proportion of MPNST (Supplementary Table 3). Sensitivity was lower in studies that included symptomatic lesions only. Specificity was higher in studies including a smaller amount of lesions and symptomatic lesions only. Specificity was lower for studies including histologically proven lesions only.

Liquid Biopsies

Three studies reported on liquid biopsies, identifying 4 potential circulating biomarkers. One study used microarray analysis to identify genes that encode putative secreted proteins in 22 patients with BPNSTs and/or MPNSTs.[58] They found elevated serum levels of adrenomedullin (ADM) as a potential biomarker for malignant transformation of PNST with significantly higher mean ADM concentrations in NF1 patients with MPNST compared to NF1 patients with plexiform neurofibromas only (0.24 vs

0.18 ng/mL; $P = .03$). The diagnostic accuracy was not provided. A second study found that soluble fraction from the extracellular domain of AXL (sAXL) serum levels was higher in NF1 patients with plexiform neurofibromas and MPNSTs compared to those with dermal neurofibromas only, sAXL could not differentiate MPNST from others.[59] A third study performed screening for 56 potential serum biomarkers in 104 NF1 patients (with and without MPNST) compared with 41 controls.[60] Insulin growth factor binding protein 1 (IGFBP1) was elevated in MPNST patients and was able to discriminate them with a sensitivity of 0.90 and specificity of 0.50. Regulated upon Activation Normal T-cell Expressed and Secreted (RANTES) was also elevated and had a sensitivity of 0.90 and a specificity of 0.26 to discriminate MPNST patients.

Discussion

MRI characteristics could varyingly detect MPNST, but the absence of a target sign was highly sensitive. Ill-defined margins and perilesional edema could adequately distinguish MPNSTs from BPNST. FDG-PET has the highest diagnostic accuracy for detecting MPNST in NF1 patients, with equal efficacy when using SUV_{max} or T/L ratio. Functional MRI and liquid biopsies may be useful tools as well, but do require more research.

MRI in Nerve Sheath Tumors

Both MPNSTs and BPNSTs can exhibit rather different characteristics on MRI, highlighted by findings in this study. The presence of a target sign was the only MRI characteristic that might rule out MPNSTwith a nLR of <0.1.[61] Based on this finding, biopsies could be obviated for tumors with target signs. However, two studies reported 6/94 MPNSTs in this meta-analysis with a target sign.[14,15] One may argue that in order to omit a biopsy, in addition to the presence of a target sign, perilesional edema and ill-defined margins should be absent as well. Moreover, many BPNSTs do not show a target sign; 59.9% (range: 43.3%-94.3%) in this meta-analysis. Nevertheless, in the remaining 40.1%, a biopsy may possibly be omitted. The presence of perilesional edema and ill-defined margins can adequately detect MPNST as the pLRs are more than 10, but biopsies may still be needed because these features can be present in a minority of BPNST as well. Unfortunately, perilesional edema and ill-defined margins are only present in 29%-92% and 25%-68% of MPNSTs, respectively. Other characteristics that only have a moderate ability to differentiate MPNST and BPNST should therefore also be considered, including cystic changes, heterogeneity on T1, intratumoral lobulation, and large tumor size. An ideal combination of moderately specific characteristics adjacent to ill-defined margins and perilesional edema is still lacking, but may further reduce the need for biopsies. This is partially reflected by the diagnostic accuracy of qualitative assessment of MRIs which could not outperform either sensitivity or specificity of single characteristics.[14,30,37] Likewise, studies that reported diagnostic algorithms combining features decreased in sensitivity, albeit a rise in specificity.[12,21] Hence, conventional MRIs are imperfect and further diagnostics including FDG-PET in NF1 and biopsies may still be necessary in many cases. Few studies report on interobserver agreement of MRI characteristics in PNST, but two studies showed very good to excellent agreement between radiologists.[27,29] Functional MRI sequences may provide additional value in MPNST as DWI and ADC mapping yielded higher accuracy of detecting malignancy than conventional MRI characteristics.[20,35] MPNSTs show increased cellularity which makes ADC_{min} values relevant. Its use has, however, only been tested in two distinct populations and warrants further investigation.

FDG-PET in NF1 Patients

FDG-PET scans are increasingly being applied to detect malignancy in NF1 patients with varying frequency of use across centers. Many efforts have been made to find ideal semi-quantitative parameters that adequately detect MPNSTs as well as exclude benign neurofibromas. SUV_{max} is the most commonly used characteristic, but ideal thresholds vary across studies. The threshold of ≥3.5 has been proposed most commonly as the ideal threshold.[14,47,53] This has been debated as several authors claim the threshold should be higher for it to be useful. Nonetheless, the threshold of ≥3.5 yielded highest accuracies across 11 different populations, which strengthens the belief that this threshold should be used. Indeed the characteristic remains imperfect as it is only has a moderately good positive likelihood ratio (4.7), meaning biopsies still play an important role as neurofibromas may also exhibit SUV_{max} values of ≥3.5 in 34.6% of patients in this meta-analyses. Nevertheless, the remaining 65.4% with SUV_{max} values of <3.5 do not require biopsies if they do not present ill-defined margins or perilesional edema on MRI. Delayed scans have been proposed to increase the accuracy of detecting MPNSTs, but it has not yet repeatedly been proven.[41,47,52] Besides, this method requires more resources and exposes patients to additional radiation. SUV measurement may additionally vary across scanners due to differing reasons. The use of proportional SUV values of tumor to tissue may be more reproducible as it reduces measurement variations. Most commonly, the T/L ratio is used, but ideal thresholds are still missing. The T/L ratio did provide equal diagnostic accuracy compared to SUV_{max}. To diminish variations across scanners and increase reproducibility of thresholds, the European Association of Nuclear Medicine Research Ltd. (EARL) set up criteria to which scanners should adhere.[62] To our knowledge, none of the studies in this review reported on a population scanned with a PET scanner that adheres to these criteria. Qualitative assessment of FDG-PET scans is also not subjected to variation in measurements and although interobserver agreement is good within studies, standardized criteria are currently lacking. Besides the use of FDG-PET scans to identify malignant transformation, it may also facilitate CT-guided biopsies and increase accuracies.[63] MPNSTs arising from plexiform neurofibromas can show heterogeneous degrees of malignancy within one tumor and are notorious for sampling errors,[17,18] thus,

PET-CT-guided biopsies may be beneficial. Several studies in this review have shown that PET-MRI may adequately be used in the NF1 population and is particularly interesting in these patients as it combines the accuracy of both diagnostic modalities.[15,21] Moreover, replacing the CT with an MRI scan diminishes radiation exposure, which may accumulate due to numerous follow-up scans necessary in NF1 populations.[64,65]

Strengths, Limitations, and Future Perspectives

Limitations to this study include the relatively high proportion of studies included to be at high risk of bias, most commonly due to concerns regarding patient selection. There was heterogeneity among study populations which led to heterogeneity of diagnostic accuracy as evaluated by subgroup analyses. Studies could be too strict in patient selection when only histologically proven lesions are included, possibly representing a group of lesions that are considered high risk of malignancy based on imaging. Contrarily, when non-symptomatic lesions are included the proportion of low-risk lesions rises. Subgroup analyses in this study should however be interpreted with caution as many were performed on a small number of studies. Most studies were also retrospective of nature further diminishing quality of evidence. Lastly, MPNSTs can be difficult to distinguish histologically from BPNSTs and a central review of pathology would be ideal. Despite these limitations using a Bayesian approach, the quantification of diagnostic accuracy and uncertainty of common MRI and FDG-PET characteristics were reliable even when a total number of studies or patients was small and there was heterogeneity in thresholds.[25] Unfortunately many features of interest, such as delayed scanning in FDG-PET and functional MRI are thus far infrequently studied which excluded them from meta-analyses. Yet these features seem promising, possibly providing higher accuracies compared to features analyzed in meta-analyses. Based on the findings of this study, future research should investigate several knowledge gaps. First, the MRI characteristics found in this study should be validated in a large series of patients to distinguish a patient group at high risk for malignant transformation which minimizes the need for further diagnostics in low-risk patients. Second, only symptomatic or growing lesions should generally undergo imaging. More regular radiological follow-up could be considered of large plexiform neurofibromas at tumor sites that may cause symptoms only in a later stage, such as retroperitoneal or mediastinal sites. In younger children, more frequent radiological follow-up may be indicated as well in case of plexiform neurofibromas at tumor sites that may cause morbidity even without malignant transformation. The value of cystic changes, heterogeneity on T1- and T2-weighted images, large tumor size, and intratumoral lobulation should be studied for additional value too. DWI and ADC imaging seem of interest as well and might be of particular interest in the sporadic patient population. Schwannomas are the most common form of BPNST in sporadic patients and cannot be reliably distinguished on FDG-PET as schwannomas commonly have high levels of FDG uptake.[66] Also, schwannomas with cystic changes are common (ancient schwannomas) and may exhibit heterogeneous features.[29] MRI characteristics need to be assessed between sporadic and NF1 patients to explore possible variations in diagnostic accuracy which may necessitate different diagnostic guidelines. In NF1, the use of a SUV_{max} threshold of ≥3.5 should be replicated in a large database of patients who underwent scans that adhere to EARL criteria. Additionally, late scanning and other semi-quantitative parameters should be evaluated in the same population to find one with higher specificity. Altogether, these findings may enable proper diagnostic algorithms to arise for evaluating MRI scans and using distinct threshold values of FDG-PET characteristics in NF1 populations. This way unnecessary imaging, biopsies, and harmful resections will diminish. In sporadic patients, suspect lesions should then undergo biopsy based on MRI findings. In NF1 patients, suspect lesions should be evaluated with additional FDG-PET imaging. Lesions with SUV_{max} >3.5 or high T/L ratio should have a PET-guided biopsy. Whenever biopsies of suspect lesions are negative one may consider nerve-sparing resection or a wait-and-scan approach. Furthermore, the use of radiomics and deep learning has not yet been studied in nerve sheath tumors, but may be useful when studies are performed correctly including sufficient MPNST images. It may even help stratifying low- and high-grade MPNSTs.[67–69] The search for an ideal liquid biopsy should be stimulated as well since its use may diminish the need for FDG-PET scans and decrease radiation exposure in the NF population who is already prone to tumorigenesis.

Conclusion

MRI characteristics are varyingly present in MPNSTs. The presence of ill-defined margins or perilesional edema is highly suspect of malignant transformation and requires biopsies or FDG-PET scans in NF1 for further characterization. Conventional MRI may rule out MPNSTs and may obviate the need for biopsies or additional FDG-PET scans in the presence of a target sign and absence of ill-defined margins or perilesional edema as this is suggestive of a benign lesion. Cystic changes, heterogeneity on T1-weighted images, intratumoral lobulation, and large tumor size should nevertheless be taken into account as well. FDG-PET scans should be offered to NF1 patients with suspect lesions on MRI that are symptomatic to further reduce the need for biopsies. SUV_{max} and T/L ratio have similar accuracies. Ideal threshold for SUV_{max} seems to be ≥3.5. Functional MRI sequences may be useful as well, but require more research for their exact implementation. Liquid biopsies have not yet proven higher diagnostic accuracy than available imaging techniques.

Supplementary Material

Supplementary material is available at *Neuro-Oncology* online.

Keywords

liquid biopsy | MPNST | neurofibromatosis type 1 | PET-CT | radiology

Funding

No funding was received from any extramural sources.

Acknowledgments

We would kindly like to thank Dr S.M. Broski, Dr M. Schwabe, and Dr J. Wasa for providing additional patient information to be included in this meta-analysis.

Conflict of interest statement. No author has any conflicts of interest to declare.

Authorship statement. Conceptualization: E.M., R.T.J.G., J.H.C., D.J.G., C.V., and W.T. Data curation: E.M., R.T.J.G., L.H.G., and D.F.H. Formal analysis: E.M. and R.T.J.G. Funding acquisition: N/A. Investigation: E.M., R.T.J.G., J.H.C., D.F.H., L.H.G., D.J.G., C.V., and W.T. Methodology: E.M., R.T.J.G., J.H.C., D.J.G., C.V., and W.T. Project administration: E.M. and R.T.J.G. Resources: J.H.C., D.J.G., and C.V. Software: E.M. and R.T.J.G. Supervision: J.H.C., D.J.G., C.V., and W.T. Validation: J.H.C., D.F.H., L.H.G., D.J.G., C.V., and W.T. Visualization: E.M. and R.T.J.G. Writing of manuscript: E.M. and R.T.J.G. Review and editing of manuscript: J.H.C., D.F.H., L.H.G., D.J.G., C.V., and W.T.

References

1. Kim DH, Murovic JA, Tiel RL, Moes G, Kline DG. A series of 146 peripheral non-neural sheath nerve tumors: 30-year experience at Louisiana State University Health Sciences Center. *J Neurosurg*. 2005;102(2):256–266.
2. Montano N, D'Alessandris QG, D'Ercole M, et al. Tumors of the peripheral nervous system: analysis of prognostic factors in a series with long-term follow-up and review of the literature. *J Neurosurg*. 2016;125(2):363–371.
3. Le Guellec S, Decouvelaere AV, Filleron T, et al. Malignant peripheral nerve sheath tumor is a challenging diagnosis: a systematic pathology review, immunohistochemistry, and molecular analysis in 160 patients from the French Sarcoma Group Database. *Am J Surg Pathol*. 2016;40(7):896–908.
4. Brennan MF, Antonescu CR, Moraco N, Singer S. Lessons learned from the study of 10,000 patients with soft tissue sarcoma. *Ann Surg*. 2014;260(3):416–421; discussion 421.
5. Evans DG, Baser ME, McGaughran J, Sharif S, Howard E, Moran A. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J Med Genet*. 2002;39(5):311–314.
6. Ferner RE, Huson SM, Thomas N, et al. Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. *J Med Genet*. 2007;44(2):81–88.
7. Valentin T, Le Cesne A, Ray-Coquard I, et al. Management and prognosis of malignant peripheral nerve sheath tumors: the experience of the French Sarcoma Group (GSF-GETO). *Eur J Cancer*. 2016;56:77–84.
8. Martin E, Coert JH, Flucke UE, et al. A nationwide cohort study on treatment and survival in patients with malignant peripheral nerve sheath tumours. *Eur J Cancer*. 2020;124:77–87.
9. Stucky CC, Johnson KN, Gray RJ, et al. Malignant peripheral nerve sheath tumors (MPNST): the Mayo Clinic experience. *Ann Surg Oncol*. 2012;19(3):878–885.
10. Nelson CN, Dombi E, Rosenblum JS, et al. Safe marginal resection of atypical neurofibromas in neurofibromatosis type 1. *J Neurosurg*. 2019:1–11. [published online ahead of print]. doi:10.3171/2019.7.JNS191353.
11. Hajiabadi MM, Campos B, Sedlaczek O, et al. Interdisciplinary approach allows minimally invasive, nerve-sparing removal of retroperitoneal peripheral nerve sheath tumors. *Langenbecks Arch Surg*. 2020;405(2):199–205.
12. Wasa J, Nishida Y, Tsukushi S, et al. MRI features in the differentiation of malignant peripheral nerve sheath tumors and neurofibromas. *AJR Am J Roentgenol*. 2010;194(6):1568–1574.
13. Ferner RE. Neurofibromatosis 1. *Eur J Hum Genet*. 2007;15(2):131–138.
14. Derlin T, Tornquist K, Munster S, et al. Comparative effectiveness of ^{18}F-FDG PET/CT versus whole-body MRI for detection of malignant peripheral nerve sheath tumors in neurofibromatosis type 1. *Clin Nucl Med*. 2013;38(1):e19–e25.
15. Reinert CP, Schuhmann MU, Bender B, et al. Comprehensive anatomical and functional imaging in patients with type I neurofibromatosis using simultaneous FDG-PET/MRI. *Eur J Nucl Med Mol Imaging*. 2019;46(3):776–787.
16. Perez-Roman RJ, Shelby Burks S, Debs L, Cajigas I, Levi AD. The risk of peripheral nerve tumor biopsy in suspected benign etiologies. *Neurosurgery*. 2020;86(3):E326–E332.
17. Graham DS, Russell TA, Eckardt MA, et al. Oncologic accuracy of image-guided percutaneous core-needle biopsy of peripheral nerve sheath tumors at a high-volume sarcoma center. *Am J Clin Oncol*. 2019;42(10):739–743.
18. Spurlock G, Knight SJ, Thomas N, Kiehl TR, Guha A, Upadhyaya M. Molecular evolution of a neurofibroma to malignant peripheral nerve sheath tumor (MPNST) in an NF1 patient: correlation between histopathological, clinical and molecular findings. *J Cancer Res Clin Oncol*. 2010;136(12):1869–1880.
19. Benz MR, Czernin J, Dry SM, et al. Quantitative F18-fluorodeoxyglucose positron emission tomography accurately characterizes peripheral nerve sheath tumors as malignant or benign. *Cancer*. 2010;116(2):451–458.
20. Ahlawat S, Blakeley JO, Rodriguez FJ, Fayad LM. Imaging biomarkers for malignant peripheral nerve sheath tumors in neurofibromatosis type 1. *Neurology*. 2019;93(11):e1076–e1084.
21. Broski SM, Johnson GB, Howe BM, et al. Evaluation of ^{18}F-FDG PET and MRI in differentiating benign and malignant peripheral nerve sheath tumors. *Skeletal Radiol*. 2016;45(8):1097–1105.
22. Frankel S, Smith GD, Donovan J, Neal D. Screening for prostate cancer. *Lancet*. 2003;361(9363):1122–1128.
23. Sauter ER. Reliable Biomarkers to Identify New and Recurrent Cancer. *Eur J Breast Health*. 2017;13(4):162–167.

24. Chu GCW, Lazare K, Sullivan F. Serum and blood based biomarkers for lung cancer screening: a systematic review. *BMC Cancer.* 2018;18(1):181.
25. Guo J, Riebler A, Rue H. Bayesian bivariate meta-analysis of diagnostic test studies with interpretable priors. *Stat Med.* 2017;36(19):3039–3058.
26. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: a principled, practical approach to constructing priors. *Stat Sci.* 2017;32(1):1–28.
27. Ahlawat S, Fayad LM. Imaging cellularity in benign and malignant peripheral nerve sheath tumors: utility of the "target sign" by diffusion weighted imaging. *Eur J Radiol.* 2018;102:195–201.
28. Chhabra A, Soldatos T, Durand DJ, Carrino JA, McCarthy EF, Belzberg AJ. The role of magnetic resonance imaging in the diagnostic evaluation of malignant peripheral nerve sheath tumors. *Indian J Cancer.* 2011;48(3):328–334.
29. Demehri S, Belzberg A, Blakeley J, Fayad LM. Conventional and functional MR imaging of peripheral nerve sheath tumors: initial experience. *AJNR Am J Neuroradiol.* 2014;35(8):1615–1620.
30. Karsy M, Guan J, Ravindra VM, Stilwill S, Mahan MA. Diagnostic quality of magnetic resonance imaging interpretation for peripheral nerve sheath tumors: can malignancy be determined? *J Neurol Surg A Cent Eur Neurosurg.* 2016;77(6):495–504.
31. Li CS, Huang GS, Wu HD, et al. Differentiation of soft tissue benign and malignant peripheral nerve sheath tumors with magnetic resonance imaging. *Clin Imaging.* 2008;32(2):121–127.
32. Matsumine A, Kusuzaki K, Nakamura T, et al. Differentiation between neurofibromas and malignant peripheral nerve sheath tumors in neurofibromatosis 1 evaluated by MRI. *J Cancer Res Clin Oncol.* 2009;135(7):891–900.
33. Matsumoto Y, Endo M, Harimaya K, et al. Malignant peripheral nerve sheath tumors presenting as spinal dumbbell tumors: clinical outcomes and characteristic imaging features. *Eur Spine J.* 2015;24(10):2119–2125.
34. Schwabe M, Spiridonov S, Yanik EL, et al. How effective are noninvasive tests for diagnosing malignant peripheral nerve sheath tumors in patients with neurofibromatosis type 1? Diagnosing MPNST in NF1 patients. *Sarcoma.* 2019;2019:4627521.
35. Well L, Salamon J, Kaul MG, et al. Differentiation of peripheral nerve sheath tumors in patients with neurofibromatosis type 1 using diffusion-weighted magnetic resonance imaging. *Neuro Oncol.* 2019;21(4):508–516.
36. Yu YH, Wu JT, Ye J, Chen MX. Radiological findings of malignant peripheral nerve sheath tumor: reports of six cases and review of literature. *World J Surg Oncol.* 2016;14:142.
37. Furniss D, Swan MC, Morritt DG, et al. A 10-year review of benign and malignant peripheral nerve sheath tumors in a single center: clinical and radiographic features can help to differentiate benign from malignant lesions. *Plast Reconstr Surg.* 2008;121(2):529–533.
38. Razek AAKA, Ashmalla GA. Assessment of paraspinal neurogenic tumors with diffusion-weighted MR imaging. *Eur Spine J.* 2018;27(4):841–846.
39. Fayad LM, Wang X, Blakeley JO, et al. Characterization of peripheral nerve sheath tumors with 3T proton MR spectroscopy. *Am J Neuroradiol.* 2014;39(5):1035–1041.
40. Azizi AA, Slavc I, Theisen BE, et al. Monitoring of plexiform neurofibroma in children and adolescents with neurofibromatosis type 1 by [^{18}F]FDG-PET imaging. Is it of value in asymptomatic patients? *Pediatr Blood Cancer.* 2018;65(1). doi:10.1002/pbc.26733.
41. Cook GJR, Lovat E, Siddique M, Goh V, Ferner R, Warbey VS. Characterisation of malignant peripheral nerve sheath tumours in neurofibromatosis-1 using heterogeneity analysis of ^{18}F-FDG PET. *Eur J Nucl Med Mol Imaging.* 2017;44(11):1845–1852.
42. Lerman L, Zehou O, Ortonne N, et al. Interest of ^{18}F-FDG PET/CT in neurofibromatosis type 1, 10-year experience from the national reference centre Henri-Mondor. *Med Nucl.* 2019;43(5–6):370–380.
43. Moharir M, London K, Howman-Giles R, North K. Utility of positron emission tomography for tumour surveillance in children with neurofibromatosis type 1. *Eur J Nucl Med Mol Imaging.* 2010;37(7):1309–1317.
44. Nose H, Otsuka H, Otomi Y, et al. Correlations between F-18 FDG PET/CT and pathological findings in soft tissue lesions. *J Med Invest.* 2013;60(3–4):184–190.
45. Salamon J, Derlin T, Bannas P, et al. Evaluation of intratumoural heterogeneity on ^{18}F-FDG PET/CT for characterization of peripheral nerve sheath tumours in neurofibromatosis type 1. *Eur J Nucl Med Mol Imaging.* 2013;40(5):685–692.
46. Tsai LL, Drubach L, Fahey F, Irons M, Voss S, Ullrich NJ. [^{18}F]-Fluorodeoxyglucose positron emission tomography in children with neurofibromatosis type 1 and plexiform neurofibromas: correlation with malignant transformation. *J Neurooncol.* 2012;108(3):469–475.
47. Warbey VS, Ferner RE, Dunn JT, Calonje E, O'Doherty MJ. [^{18}F]FDG PET/CT in the diagnosis of malignant peripheral nerve sheath tumours in neurofibromatosis type-1. *Eur J Nucl Med Mol Imaging.* 2009;36(5):751–757.
48. Salamon J, Veldhoen S, Apostolova I, et al. ^{18}F-FDG PET/CT for detection of malignant peripheral nerve sheath tumours in neurofibromatosis type 1: tumour-to-liver ratio is superior to an SUV_{max} cut-off. *Eur Radiol.* 2014;24(2):405–412.
49. Combemale P, Valeyrie-Allanore L, Giammarile F, et al. Utility of ^{18}F-FDG PET with a semi-quantitative index in the detection of sarcomatous transformation in patients with neurofibromatosis type 1. *PLoS One.* 2014;9(2):e85954.
50. Bredella MA, Torriani M, Hornicek F, et al. Value of PET in the assessment of patients with neurofibromatosis type 1. *AJR Am J Roentgenol.* 2007;189(4):928–935.
51. Cardona S, Schwarzbach M, Hinz U, et al. Evaluation of F18-deoxyglucose positron emission tomography (FDG-PET) to assess the nature of neurogenic tumours. *Eur J Surg Oncol.* 2003;29(6):536–541.
52. Chirindel A, Chaudhry M, Blakeley JO, Wahl R. ^{18}F-FDG PET/CT qualitative and quantitative evaluation in neurofibromatosis type 1 patients for detection of malignant transformation: comparison of early to delayed imaging with and without liver activity normalization. *J Nucl Med.* 2015;56(3):379–385.
53. Ferner RE, Golding JF, Smith M, et al. [^{18}F]2-fluoro-2-deoxy-D-glucose positron emission tomography (FDG PET) as a diagnostic tool for neurofibromatosis 1 (NF1) associated malignant peripheral nerve sheath tumours (MPNSTs): a long-term clinical study. *Ann Oncol.* 2008;19(2):390–394.
54. Ferner RE, Lucas JD, O'Doherty MJ, et al. Evaluation of ^{18}fluorodeoxyglucose positron emission tomography (^{18}FDG PET) in the detection of malignant peripheral nerve sheath tumours arising from within plexiform neurofibromas in neurofibromatosis 1. *J Neurol Neurosurg Psychiatry.* 2000;68(3):353–357.
55. Karabatsou K, Kiehl TR, Wilson DM, Hendler A, Guha A. Potential role of ^{18}fluorodeoxyglucose-positron emission tomography/computed tomography in differentiating benign neurofibroma from malignant peripheral nerve sheath tumor associated with neurofibromatosis 1. *Neurosurgery.* 2009;65(4 Suppl):A160–A170.
56. Mautner VF, Brenner W, Fünsterer C, Hagel C, Gawad K, Friedrich RE. Clinical relevance of positron emission tomography and magnetic resonance imaging in the progression of internal plexiform neurofibroma in NF1. *Anticancer Res.* 2007;27(4A):1819–1822.
57. Meany H, Dombi E, Reynolds J, et al. 18-fluorodeoxyglucose-positron emission tomography (FDG-PET) evaluation of nodular lesions in patients with neurofibromatosis type 1 and plexiform neurofibromas (PN) or malignant peripheral nerve sheath tumors (MPNST). *Pediatr Blood Cancer.* 2013;60(1):59–64.

58. Hummel TR, Jessen WJ, Miller SJ, et al. Gene expression analysis identifies potential biomarkers of neurofibromatosis type 1 including adrenomedullin. *Clin Cancer Res.* 2010;16(20):5048–5057.
59. Johansson G, Peng PC, Huang PY, et al. Soluble AXL: a possible circulating biomarker for neurofibromatosis type 1 related tumor burden. *PLoS One.* 2014;9(12):e115916.
60. Park SJ, Sawitzki B, Kluwe L, Mautner VF, Holtkamp N, Kurtz A. Serum biomarkers for neurofibromatosis type 1 and early detection of malignant peripheral nerve-sheath tumors. *BMC Med.* 2013;11:109.
61. EUnetHTA. Meta-analysis of diagnostic test accuracy guideline. https://eunethta.eu/wp-content/uploads/2018/01/Meta-analysis-of-Diagnostic-Test-Accuracy-Studies_Guideline_Final-Nov-2014.pdf. Published 2014. Accessed May 4, 2020.
62. Kaalep A, Sera T, Oyen W, et al. EANM/EARL FDG-PET/CT accreditation - summary results from the first 200 accredited imaging systems. *Eur J Nucl Med Mol Imaging.* 2018;45(3):412–422.
63. Brahmi M, Thiesse P, Ranchere D, et al. Diagnostic accuracy of PET/CT-guided percutaneous biopsies for malignant peripheral nerve sheath tumors in neurofibromatosis type 1 patients. *PLoS One.* 2015;10(10):e0138386.
64. Gatidis S, Schmidt H, Gücke B, et al. Comprehensive oncologic imaging in infants and preschool children with substantially reduced radiation exposure using combined simultaneous ^{18}F-fluorodeoxyglucose positron emission tomography/magnetic resonance imaging: a direct comparison to ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography. *Invest Radiol.* 2016;51(1):7–14.
65. Chawla SC, Federman N, Zhang D, et al. Estimated cumulative radiation dose from PET/CT in children with malignancies: a 5-year retrospective review. *Pediatr Radiol.* 2010;40(5):681–686.
66. Beaulieu S, Rubin B, Djang D, Conrad E, Turcotte E, Eary JF. Positron emission tomography of schwannomas: emphasizing its potential in preoperative planning. *AJR Am J Roentgenol.* 2004;182(4):971–974.
67. Peeken JC, Spraker MB, Knebel C, et al. Tumor grading of soft tissue sarcomas using MRI-based radiomics. *EBioMedicine.* 2019;48:332–340.
68. Wang H, Chen H, Duan S, Hao D, Liu J. Radiomics and machine learning with multiparametric preoperative MRI may accurately predict the histopathological grades of soft tissue sarcomas. *J Magn Reson Imaging.* 2019;51(3):791–797. doi:10.1002/jmri.26901.
69. Vos M, Starmans MPA, Timbergen MJM, et al. Radiomics approach to distinguish between well differentiated liposarcomas and lipomas on MRI. *Br J Surg.* 2019;106(13):1800–1809.