

## Using fluorodeoxy-D-glucose-positron emission tomography to monitor neoadjuvant chemotherapy response in sarcoma: A meta-analysis

Yu-Ting Wang, Hong Pu, Long-Lin Yin, Jia-Yuan Chen

Yu-Ting Wang, Hong Pu, Long-Lin Yin, Jia-Yuan Chen, Department of Radiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu 610072, Sichuan Province, China

Author contributions: Wang YT and Chen JY designed the research; Wang YT, Pu H and Yin LL performed the research; Wang YT and Pu H analyzed the data; Wang YT wrote the manuscript.

Supported by Scientific Research Subject of Health Department of Sichuan, China, No. 070045

Correspondence to: Yu-Ting Wang, MD, Department of Radiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, No. 32, Section 2, 1<sup>st</sup> Ring Road (West), Chengdu 610072, Sichuan Province, China. [wangyuting\\_330@163.com](mailto:wangyuting_330@163.com)

Telephone: +86-28-87394280

Received: May 13, 2014 Revised: August 11, 2014

Accepted: September 4, 2014

Published online: November 26, 2014

### Abstract

**AIM:** To systematically evaluate the accuracy of 18-fluorodeoxy-D-glucose-positron emission tomography (18-FDG PET) to assess response to neoadjuvant chemotherapy in bone and soft tissue sarcomas.

**METHODS:** Studies published in English language regarding the accuracy of F-18 FDG PET for the indication were retrieved from MEDLINE. The QUADAS tool was utilized for methodological quality appraisal. Relevant data were extracted, and quantitative data synthesis included pooled estimation and subgroup analysis.

**RESULTS:** A total of fifteen studies involving 420 patients with pathologically confirmed sarcoma were collected. Methodological quality was relatively high. The pooled sensitivity and specificity of PET to predict histopathological response were 87% (95%CI: 81%-91%) and 83% (95%CI: 77%-87%), respectively. Ten stud-

ies employed a lower standardized uptake value (SUV) after chemotherapies (mostly 2.5) and/or a higher SUV reduction rate (mostly around 50%) as PET criteria of good response. Subgroup analysis showed that PET exhibited a significantly better specificity in osteosarcoma (OS) and Ewing sarcoma (ES) than in soft-tissue sarcoma (STS) (91% *vs* 75%,  $P < 0.05$ ), and a higher specificity in pediatric patients than in adults (90% *vs* 74%,  $P < 0.01$ ). PET yielded a lower specificity in ifosfamide-contained chemotherapies than in the alternative regimen (70% *vs* 97%,  $P < 0.01$ ).

**CONCLUSION:** F-18 FDG PET is promising to predict neoadjuvant therapy response in sarcoma, especially in pediatric patients with OS or ES. Certain chemotherapeutic agents could potentially cause false positives of PET.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** 18-fluorodeoxy-D-glucose-positron emission tomography; Sarcoma; Therapy monitoring; Meta-analysis

**Core tip:** Histopathological response to neoadjuvant chemotherapy is a significant prognostic factor and could guide the following treatment for patients with sarcoma. Novel molecular imaging such as 18-fluorodeoxy-D-glucose-positron emission tomography (PET) could offer non-invasive and accurate assessment of such response in sarcoma based on our relatively large-sample statistical analysis for this relatively rare tumor. More information that clinical physicians concerned with was generated, such as recommended subgroup of patients, potential candidates for quantitative criteria of response and possible causes of false positives of PET.

Wang YT, Pu H, Yin LL, Chen JY. Using fluorodeoxy-D-glucose-positron emission tomography to monitor neoadjuvant chemo-

therapy response in sarcoma: A meta-analysis. *World J Meta-Anal* 2014; 2(4): 212-220 Available from: URL: <http://www.wjgnet.com/2308-3840/full/v2/i4/212.htm> DOI: <http://dx.doi.org/10.13105/wjma.v2.i4.212>

## INTRODUCTION

Soft tissue and bone sarcomas account for about 1% of malignancies and remain a diagnostic and therapeutic challenge. Osteosarcoma (OS) and Ewing's sarcoma (ES) are the most common malignant primary bone tumors in children and adolescents, while soft tissue sarcomas (STS) are a heterogeneous group of malignancies that exhibit mesenchymal differentiation, and may occur in both children and the elderly<sup>[1]</sup>.

While sarcoma is not a common tumor, it had a high mortality rate<sup>[2]</sup>. The rate of systemic spread is so high that surgical treatment alone was rarely enough for cure. The introduction of neoadjuvant chemotherapy (NCT) brought a significant improvement in the long-term survival of patients and was therefore widely used in clinical practice<sup>[3,4]</sup>. Although histopathological response to neoadjuvant chemotherapy (mostly defined by tumor necrosis of more than 90%) was showed to be associated with a significantly higher survival rate, it is known that only a subcollective of patients benefit from NCT (32%-88% for OS and ES, 19%-48% for STS)<sup>[5]</sup>. Therefore, there is an urgent need for non-invasive methods which allow accurate identification of responses to NCT, and accordingly determine whether to switch to a more intensified chemotherapy regimen, and also to determine the most appropriate surgical approach.

Radiologic imaging has some limitations for this utility; for instance, the OS does not suffer evident shrinkage after chemotherapy<sup>[6]</sup>. In contrast to conventional anatomic imaging such as MRI and CT, [F-18]-fluorodeoxy-D-glucose (FDG)-positron emission tomography (PET) has the unique ability to detect response at the molecular level, and is being increasingly used in patients with sarcoma for therapy monitoring<sup>[7]</sup>.

Nonetheless, the sample sizes in conducted trials were relatively small and the results among them were remarkably discrepant. We performed this meta-analysis to systematically estimate the accuracy of F-18 FDG PET to assess histopathological response to neoadjuvant therapy in patients with bone and soft tissue sarcomas.

## MATERIALS AND METHODS

### Literature search

This meta-analysis was performed following guidelines for the systematic review and meta-analysis of diagnostic studies<sup>[8]</sup>. Ethical committee approval and patients' consents were not required for this work due to its statistic nature. MEDLINE (January 1980 to August 2012) were used for searching English-language articles with the following keywords: ("positron emission tomography OR fluorodeoxyglucose") AND ("prediction OR response

OR response monitoring OR neoadjuvant OR chemotherapy") AND ("carcinoma OR cancer OR tumor OR tumour") AND ("sensitivity OR specificity OR accuracy") AND ("sarcoma"). Two reviewers independently checked the abstracts of retrieved publications and obtained the fulltext of each eligible article. References of the retrieved studies were screened to expand the scope of the computerized literature search. Disagreements were resolved by discussion between the two reviewers.

Studies were eligible for inclusion based on the following criteria: (1) using F-18 FDG PET to assess response to neoadjuvant therapies in patients with primary bone or soft tissue sarcoma; (2) employing pathological outcome as the "gold standard"; and (3) clinical raw data could be drawn from data in the original published study. A minimal sample size for inclusion was not required given the relative rareness of sarcomas in clinical scenarios. Conference abstracts, case reports, comments, and letters to the editors were excluded.

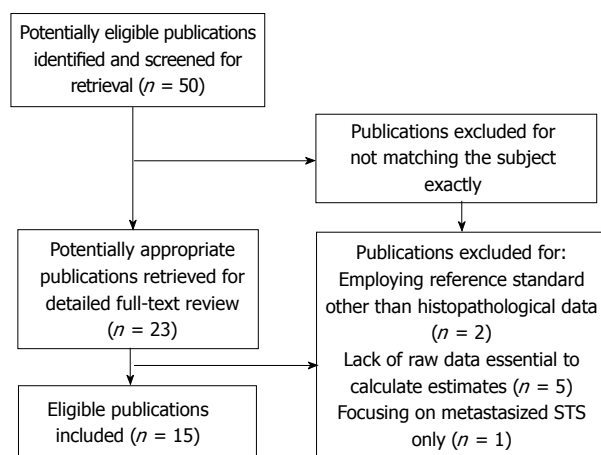
### Quality appraisal and data extraction

To identify the methodological quality of included article studies, two reviewers independently assessed each article using the QUADAS tool, which is specifically designed and widely used in systematic reviews of diagnostic accuracy studies<sup>[9]</sup>. It provides a checklist of items regarding the representativeness and methodological quality of investigated studies and presents an accumulated score as a result. Description of the execution of PET scans and the pathological interpretation were assessed. The optimal time interval between F-18 FDG PET and histopathological reading was defined as less than 30 d. Whether the pathological interpretation was blind to the PET scan results was also recorded. However, the PET readers were naturally blind to pathological information, since these scans were performed before the pathological results from following surgeries.

Two reviewers independently extracted the following data by using a standardized data extraction form: (1) characteristics of the study, including investigated tumor pathology (OS or STS), sample size, patient age, timing for PET scans, specific criteria of PET and pathology, *etc.*; (2) methodological details which needed to be examined with the QUADAS tool; and (3) clinical outcome. A 2 × 2 contingency table was constructed consisting of true-positive, false-positive, true-negative and false-negative for each study. We considered response to therapies as "positive", while therapy failures as "negative".

### Statistical analysis

Based on the 2 × 2 contingency tables, sensitivity was calculated as the ratio of the number of true-positive findings showed in the index test to the number of all positive findings supported by the golden standard. Specificity was calculated as the ratio of the number of true-negative findings reported by the index test to the number of all negative findings showed in the reference standard. The diagnostic odds ratio (DOR) was calculated as  $[\text{True positive (TP)} \times \text{True negative (TN)}] / [\text{False posi-}$



**Figure 1** Flow diagram of publication selection for this meta-analysis. STS: Soft-tissue sarcoma.

tive (FP)  $\times$  False negative (FN)], where TP, TN, FP and FN stand for true-positive, true-negative, false-positive and false-negative, respectively. If the DOR could not be calculated because one of the cells in the  $2 \times 2$  table was empty, 0.5 was added to all cells<sup>[10]</sup>. We used the  $\chi^2$  test to assess the heterogeneity among studies. The  $I^2$  index was employed, with an  $I^2$  value of over 50% indicating statistical heterogeneity (a random-effects model was accordingly utilized), and an  $I^2$  value of less than 50% suggesting that homogeneity was reached (a fixed-effects model was then applied)<sup>[11,12]</sup>.

If statistical heterogeneity was found, subgroup analyses were then performed to explore its possible sources (e.g., different tumor pathology, patient age group and chemotherapy regimen). We employed the Z test to examine the statistical differences between subgroups. Two-sided  $P < 0.05$  was considered statistically significant.

Our summary receiver operating characteristic curves (SROC) was plotted in accordance with Mose's linear model, reflecting the summary trade-off between sensitivity and specificity across the included studies. The areas under the curves (AUC) and  $Q^*$  value were also calculated.  $Q^*$  value is defined by the point where sensitivity and specificity are equal, which is the point closest to the ideal top-left corner of the SROC space, reflecting the diagnostic value<sup>[13]</sup>.

To analyze publication bias, we created inverted funnel plots of individual study log DORs plotted against the effective sample size. Both Begg's rank correlation test and Egger's regression test were utilized to assess the asymmetry of the funnel plot<sup>[14]</sup>. An asymmetrical funnel plot would suggest that additional small-sample studies may have been conducted but not published due to their unfavorable results.

## RESULTS

### Study evaluation and selection

After independent review, our research initially yielded a total of 50 articles in English-language from MEDLINE.

Twenty-three articles were considered candidates after careful review of abstracts. After fulltext reading of these articles, eight were excluded because of employing reference standard other than pathological outcome ( $n = 2$ ); lack of raw data that was essential to built the  $2 \times 2$  contingency tables ( $n = 5$ ); or focusing on metastasized STS only ( $n = 1$ ). Finally, 15 articles fulfilled all eligibility criteria and were included for data extraction and following analysis<sup>[3,6,7,15-26]</sup>. Flow diagram of publication selection is showed in Figure 1.

### Study characteristics and quality appraisal

Among the 15 studies, four were retrospectively designed<sup>[6,17,21,25]</sup>, one had unclear design of whether prospective or retrospective<sup>[24]</sup>, and the rest ten were prospective. All were in English. A total of 420 patients were included. Median sample size was 21 (range, 9 to 50), and mean age ranged from 10 to 54 years old. Three studies focused on pediatric patients only<sup>[3,6,24]</sup>.

Detailed characteristics of the 15 included trials are outlined in Table 1. All articles assessed treatment of neoadjuvant chemotherapy, and in four studies some patients received neoadjuvant radiotherapy as well<sup>[19,20,23,25]</sup>. All trials had a sequential F-18 FDG PET both before and after the neoadjuvant therapies were performed.

All studies included in the meta-analysis fulfilled 9 or more of the 14 standards in the guidelines of the QUADAS tool. Drawbacks were commonly seen in description of how patients were selected and whether histopathological interpretation was blind to PET results. Moreover, the availability of clinical data in the diagnostic test reading was mostly unclear. The median methodological quality score (represented as the number of "Yes" to fourteen questions) was 12.

### Data analysis

Estimation of overall accuracy: Table 2 exhibited the absolute numbers from  $2 \times 2$  contingency tables from each article. When more than one standard of PET was present in one trial, we took the one that generated higher accuracy into statistic calculation. The sensitivity of F-18 FDG PET to assess therapy response in sarcoma ranged from 40%-100%, and specificity from 44%-100%. The forest plot displayed the sensitivity, specificity and DOR with 95%CI for each individual study (Figure 2).

Based on the random-effects model, the pooled sensitivity and specificity were 87% (95%CI: 81%-91%) and 83% (95%CI: 77%-87%), respectively, and the pooled DOR was 40.5 (95%CI: -14.6-112.0). The SROC indicated an AUC of 0.930 and  $Q^*$  value of 0.865 (Figure 3). Heterogeneity of sensitivity and specificity was detected, but no significant heterogeneity of DOR was found, as their  $I^2$  exhibited in Figure 2.

### Subgroup analysis

Several subgroup analyses were then performed to examine the potential sources of heterogeneity displayed among the 15 studies. It was noted that these studies

**Table 1** Characteristics of the included studies

	Year	Tumor pathology	Patient No. (% of males)	Age, mean (range), yr	Therapy, regimen	No. (%) of responders	Pathologic criteria	Adopted PET criteria	Timing of PET	Quality score <sup>b</sup>
Im	2012	OS (resectable high-grade)	14 (71)	15 (10-25)	Neo-chemo: various combinations of D, C, M, I and E	5 (36)	Necrosis fraction $\geq 90\%$ (by Salzer-Kuntschik <i>et al</i> )	Post-chemo SUVmax < 3	Before, between the first and second cycles	12
Kim	2011	OS or ESFT	23 (69)	10 (3-19)	Neo-chemo: various combinations, I contained	15 (65)	Viable tumor $\leq 10\%$	SUV reduction rate $\geq 50\%$ and SUV2 $\leq 2.5$	Before, after completion of chemo	10
Bajpai	2011	OS	31 (81)	17 (5-66)	Neo-chemo: C + D	10 (32)	Necrosis fraction $\geq 90\%$	V1 $\leq 300$ mL and SUV2:SUV1 $\leq 0.48$	Before, after 3 cycle	13
Dimitrako	2010	STS (non-metastasized)	24 (unclear)	Unclear	Neo-chemo: E + D	14 (58)	Necrosis fraction $\geq 90\%$	SUVmean+ Influx, value unclear	Before, after 2 cycles	9
Cheon	2009	OS (high-grade)	70 (68)	14 <sup>a</sup> (5-59)	Neo-chemo: M, A and C	33 (47)	Necrosis fraction $\geq 90\%$	SUV2 < 2 or MVCR > 65% when SUV2 is between 2 to 5	Before, after completion of chemo	12
Hamada	2009	OS	9	51 <sup>a</sup> (20-80)	Neo-chemo: D, C, I and M	5 (56)	Necrosis fraction $\geq 90\%$	SUV2 < 2.5	Before, after completion of chemo	10
Benz	2009	STS (resectable high-grade)	50 (52)	51 <sup>a</sup> (20-80)	Neo-chemo: mostly I contained	8 (16)	Necrosis fraction $\geq 90\%$	SUV reduction rate $\geq 35\%$ after the first cycle	Before, after the first cycle, after completion of chemo	12
Benz	2008	STS (high-grade)	20 (50)	49 (19-86)	Neo-chemo: I + D or G + D; some had radiotherapy	6 (30)	Necrosis fraction $\geq 95\%$	SUVmean changes, value unclear	Before, after completion of chemo	12
Evilevitch	2008	STS (high-grade)	42 (57)	54 (20-86)	Neo-chemo: I in 84%, G in 16% of patients, some had radiotherapy	8 (19)	Necrosis fraction $\geq 95\%$	SUV reduction rate $\geq 60\%$	Before, after completion of chemo	13
Iagaru	2008	OS and STS	14 (57)	36 (18-56)	Neo-Chemo: I based	6 (43)	Necrosis fraction $\geq 90\%$	SUV reduction rate $\geq 27\%$	Before, after completion of chemo	11
Ye	2008	OS	15 (60)	17 <sup>a</sup> (7-33)	Neo-chemo: M, C, and A	8 (53)	Necrosis fraction $\geq 90\%$	TBR2/TBR1 < 0.46	Before, after completion of chemo	13
Hawkins	2005	ESFT	34 (unclear)	19 (6-46)	Neo-chemo: V, D, Cy, I and E; some had radiotherapy	25 (74)	Necrosis fraction $\geq 90\%$	SUV2 < 2.5	Before, after completion of chemo	10
Hawkins	2002	OS and ES	31 (74)	13 (6-19)	Neo-chemo: D, C, M, with or without I	18 (58)	Necrosis fraction $\geq 90\%$	SUV2 < 2	Before, after completion of chemo	10
Franzius	2000	OS and ES	17 (76)	13 <sup>a</sup> (5-36)	Neo-chemo: D, M, C and I; some had radiotherapy	15 (88)	Necrosis fraction $\geq 90\%$	TBR reduction rate > 30%	Before, after completion of chemo	12
Schulte	1999	OS	27 (63)	17 <sup>a</sup> (5-36)	Neo-chemo: D, M, I and C	17 (63)	Necrosis fraction $\geq 90\%$	TBR2/TBR1 < 0.6	Before, after completion of chemo	12

<sup>a</sup>: Median age; <sup>b</sup>: Evaluated by QUADAS tool, presented as the number of “Yes” to the fourteen questions of QUADAS criteria. A: Adriamycin; C: Cisplatin, Cy: Cyclophosphamide; D: Doxorubicin, E: Etoposide, G: Gemcitabine; M: Methotrexate, I: Ifosfamide; V: Vincristine; OS: Osteosarcoma; STS: Soft-tissue sarcomas; ES: Ewing sarcoma; ESFT: Ewing sarcoma family of tumors; Neo-chemo: Neoadjuvant chemotherapy; SUV: Standardized uptake value; SUV1: Pre-chemo SUV; SUV2: Post-chemo SUV; MVCR: Metabolic volume change ratio; TBR: Tumor to background ratio; TBR1: Pre-chemo TBR; TBR 2: post-chemo TBR.

investigated different tumor pathology, patient age group and chemotherapy regimen. Subgroup analysis results are showed in Table 3.

**Tumor pathology:** The  $I^2$  values of sensitivity and specificity were close to 50% within both the group “OS or ES” and “STS”, while within the group “STS” the



**Table 2** Absolute numbers from 2 × 2 contingency tables for the included studies

	TP	FN	FP	TN
Im	5	0	1	8
Kim	14	1	1	7
Bajpai <sup>1</sup>	4	6	0	20
Dimitrako	11	3	1	9
Cheon	31	2	1	36
Hamada	5	0	0	4
Benz-2009	8	0	14	28
Benz-2008	6	0	0	14
Evilevitch	8	0	10	24
Iagaru	3	3	3	5
Ye	8	0	0	7
Hawkins-2005	19	6	5	4
Hawkins-2002	14	4	1	12
Franzius	15	0	0	2
Schulte	17	0	2	8

<sup>1</sup>One patient was not taken into calculation due to the unavailability of the relevant PET data. TP: True positive; FP: False positive; FN: False negative; TN: True negative.

heterogeneity of DOR was eliminated ( $I^2 = 0\%$ ). Statistic suggested that PET had better specificity and DOR when assessing the “OS or ES” group than the “STS” group.

**Patient age group:** Three studies focused on pediatric patients, and four studies investigated adult patients only. The heterogeneity of sensitivity, specificity and DOR was all resolved within the pediatric patient group ( $I^2$ : 0%-37%), and the  $I^2$  of DOR within the adult patient group was 50%. Results showed that PET had significantly better specificity and DOR when used among pediatric patients than adults.

**Chemotherapy regimen:** Ifosfamide was reported to cause inflammatory reactions and consequently increase the chance of mismatch between PET scans and pathology<sup>[21]</sup>. To assess the effect of ifosfamide, we conducted this subgroup analysis. We found that within the group using chemotherapy agents without ifosfamide, the heterogeneity of specificity and DOR was resolved ( $I^2$ : 0% and 15.7%), and significantly better specificity and DOR than the group using ifosfamide were gained.

**PET response criteria:** Ten studies employed a lower standardized uptake value (SUV) after chemotherapies (mostly between 2 and 3) and/or a higher SUV reduction rate (27%-60%, mostly around 50%) as PET criteria of good response. Three studies used tumor to background ratio (TBR) related parameters to define good response. However, subgroup analysis did not resolve the heterogeneity of any parameters (sensitivity, specificity and DOR), hindering a reliable further comparison.

### Publication bias

The inverted funnel plot suggested some asymmetry of the distribution of the points (Figure 4). Begg's test indicated that publication bias was unlikely with a  $P$  value of

0.06. However, Egger's test showed that such bias might exist with a  $P$  value of 0.003.

## DISCUSSION

The diagnosis and response assessment of bone and soft tissue sarcomas by imaging remain a challenge, largely due to its high intratumoral heterogeneity. A previous meta-analysis on the value of F-18 FDG PET in detecting sarcoma has reported a relatively high accuracy (pooled sensitivity of 91%, specificity of 85%)<sup>[1]</sup>. Clinical evidence regarding the value of F-18 FDG PET to monitor response of sarcoma to neoadjuvant therapy is increasing, but a solid and quantitative evaluation is lacking. To our knowledge, this meta-analysis quantitatively summarized increased but different data from individual studies, and explored specific clinical uses of PET, such as the optimal patient groups and potential causes of false results, which could provide vital clues for the clinical applications of F-18 FDG PET in sarcoma. The statistical results supported the important role of F-18 FDG PET in assessing histopathological response to NCT in bone and soft tissue sarcomas, with relatively high pooled sensitivity (87%), specificity (83%), DOR (40.5), and AUC (0.930).<sup>1</sup>

### Optimal patient groups for F-18 FDG PET

The difference of pathology and age predilection between bone and soft tissue sarcomas rationalize identification of optimal patient groups of using F-18 FDG PET to assess response to NCT. It is noted that all three studies that focused on pediatric patients investigated OS or ES. Further discrimination of OS and ES was hardly possible, since many studies did not report such data separately. Almost all studies investigating STS were comprised of merely adult patients (one study did not clearly report the patient age range<sup>[7]</sup>). One study included both bone and soft tissue sarcomas<sup>[21]</sup>. We conducted subgroup analysis between studies that included merely pediatric patients or adult patient to identify the effect of patient age, and so as to the effect of two pathological groups. Statistical results suggested that PET was especially promising among pediatric patients with OS and ES, with relatively high sensitivity and specificity, both around 90%.

### Possible cause of misjudgment

The pooled sensitivity of PET in response assessment of sarcoma was relatively high and stable among all subgroups in this meta-analysis, suggesting a relatively low chance of false negatives in clinical practice. By contrast, the pooled specificity was lower, and almost all the significant difference gained in subgroup analysis was attributed to specificity, calling for clinical caution of false positives. Certain chemotherapy agents have been reported to potentially cause misjudgments in PET scans after treatment of sarcoma, and ifosfamide was among the most widely known ones<sup>[21]</sup>. It is reported in bone tumors to theoretically increase SUVmax of adjacent inflammatory tissues,

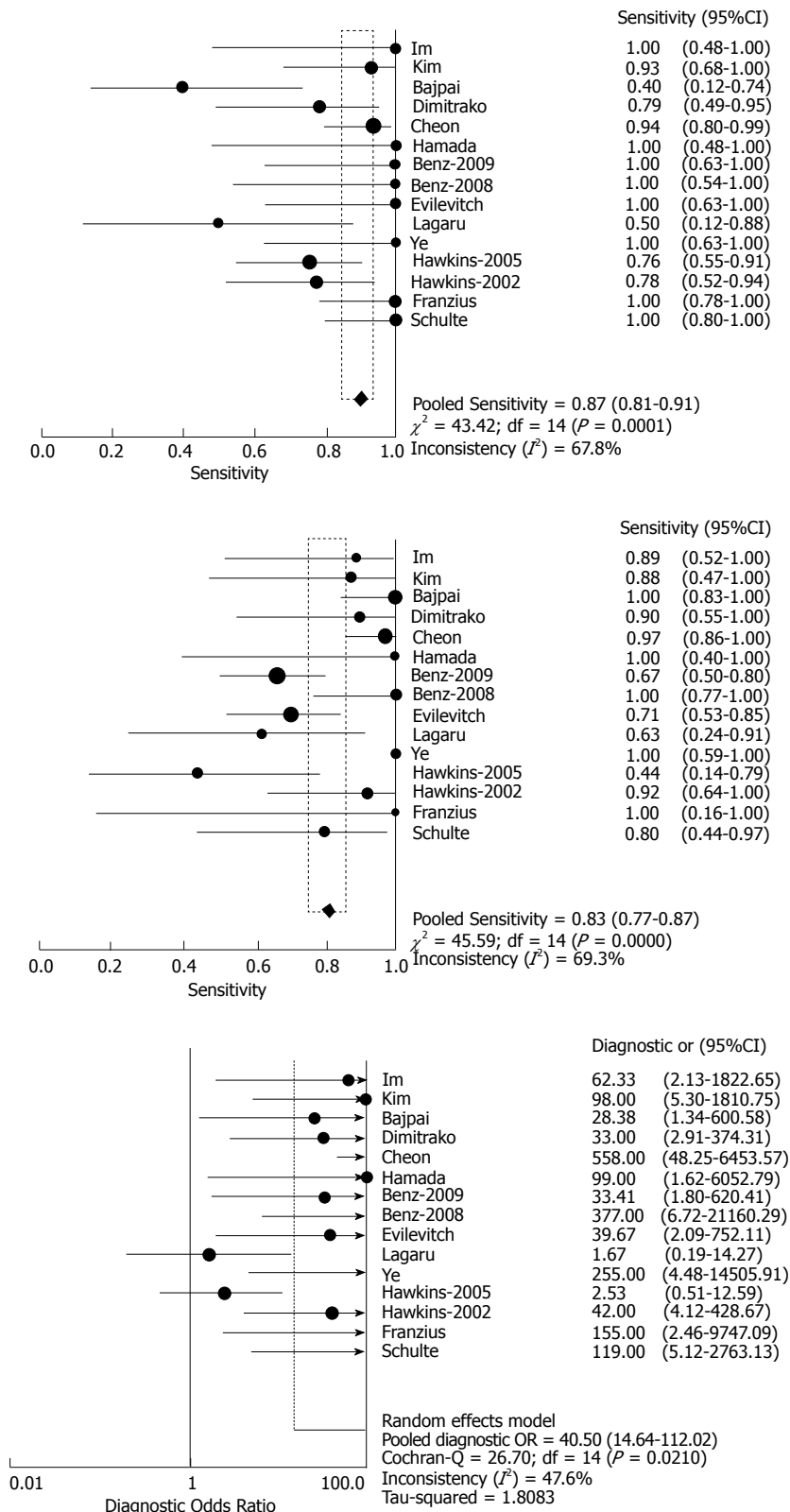


Figure 2 Pooled sensitivity, specificity, and diagnostic odds ratio in forest plots for neoadjuvant therapy response prediction in bone and soft tissue sarcomas by F-18-fluorodeoxy-D-glucose-positron emission tomography.

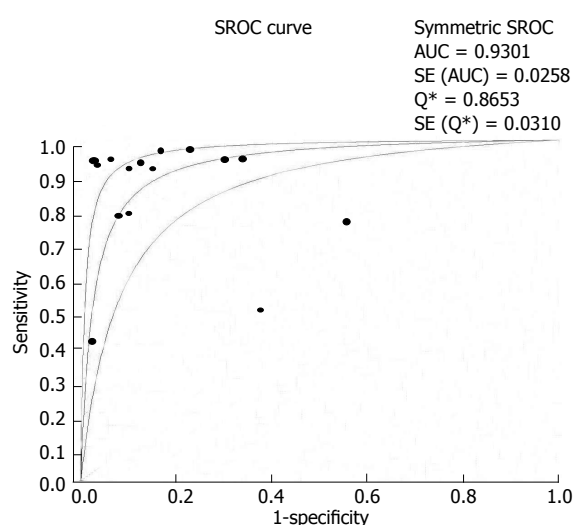
causing false negatives consequently. Differing from such possibility, our subgroup analysis found more chance of false positives. The most distinct and largest numbers of false positives were generated in studies by Benz *et al*<sup>[18]</sup>

and Evilevitch *et al*<sup>[20]</sup>, both investigating STS, suggesting that ifosfamide might have different mechanisms in STS from bone tumors. Since it was reported in clinical practice that addition of ifosfamide could improve patients'

**Table 3** Subgroup analysis based on tumor pathology, patient age groups and chemotherapy regimen

Subgroup criteria	Characteristics	No. of studies	No. of patients	Se %	P of Se	Sp %	P of Sp	DOR
Tumor pathology <sup>a</sup>	OS or ES	10	270	87 (95%CI: 81-92)	> 0.05	91 (95%CI: 84-95)	< 0.05	57.8
	(1-3, 5, 6, 11-15)							
Patient age groups <sup>b</sup>	STS (4, 7, 8, 9)	4	55	92 (95%CI: 78-98)		75 (95%CI: 65-83)		47.9
	Pediatric patients (1, 2, 13)	3	68	87 (95%CI: 72-96)	> 0.05	90 (95%CI: 73-98)	< 0.01	59.2
	Adult patients (6, 7, 8, 9, 10)	5	135	91 (95%CI: 76-98)		74 (95%CI: 64-82)		25.6
Chemotherapy regimen <sup>c</sup>	With ifosfamide (2, 6, 7, 9, 10, 12, 14, 15)	8	216	90 (95%CI: 82-95)	> 0.05	70 (95%CI: 61-78)	< 0.01	20.8
	Without ifosfamide (3, 4, 5, 11)	4	139	83 (95%CI: 72-91)		97 (95%CI: 91-100)		103.8

<sup>a</sup>[22] not included due to covering both tumor pathology; <sup>b</sup>7 studies not included due to age ranges covering both pediatric and adult patients; <sup>c</sup>3 studies not included [3,20,25] for unclear portion of ifosfamide use. Se: Sensitivity; Sp: Specificity; OS: Osteosarcoma; ES: Ewing sarcoma; DOR: Diagnostic odds ratio.

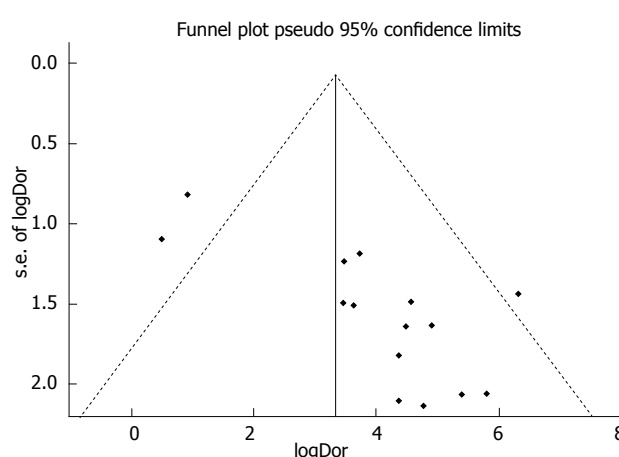


**Figure 3** Summary receiver operating characteristic curves for neoadjuvant therapy response prediction in bone and soft tissue sarcomas by F-18-fluorodeoxy-D-glucose-positron emission tomography.

survival in both OS and ES, and ifosfamide was widely used in STS as well, it is noteworthy that this chemotherapeutic agent has the potential to cause misjudgment on PET scan results of sarcoma.

#### Candidate parameters for quantitative evaluation

Different parameters have been employed to explore a better accuracy of F-18 FDG PET, among which maximum SUVs were considered quantitative and reproducible, therefore widely used in clinical trials<sup>[17]</sup>. Six studies included in this meta-analysis relied on post-treatment SUV (SUV2, mostly maximum SUV) and eight trials relied on SUV reduction rate to identify responders, and overlaps existed. It was pointed out that SUV2 may remain elevated because of inflammation or reactive changes. Moreover, tumors with an initial high pre-treatment SUV might still have a high SUV2, despite a significant decrease in the SUV value. SUV reduction rate is therefore proposed as a more reliable indicator<sup>[6]</sup>. In contrast to such knowledge, three included studies that



**Figure 4** Funnel plot to test publication bias shows some asymmetry.

compared the accuracy of SUV2 and SUV reduction rate all presented that SUV2 had a better correlation with histopathology than the latter<sup>[17,23,24]</sup>. TBR was employed in three studies, but was reported as nonreproducible and less practical due to inter-observer variability<sup>[3]</sup>. Since the subgroup analysis of different criteria still suffered unexplained heterogeneity, the optimal criteria remain to be defined in larger-sample, prospectively designed studies that include and compare directly different parameters.

#### Clinical roles of pre-operative F-18 FDG PET

The insufficient accuracy of current imaging modalities raised reasonable concerns of clinicians regarding the roles of pre-operative PET. Potential clinical applications may include: (1) providing prognostic indicator of disease-free survival after multimodal treatment<sup>[5]</sup>; (2) guiding the selection of resection technique, because surgical oncologists may be more willing to perform limb-sparing surgical resections in the setting of a favorable response to therapy<sup>[27]</sup>; and (3) potentially identifying non-responders early (pooled NPV = 90%), since several studies explored the use of interim PET information and all generated positive results<sup>[3,7,15,18]</sup>. Benz *et al*<sup>[18]</sup> compared the performance of PET after the first cycle and after the

completion of NCT, reporting that the earlier scans had an even better accuracy. Because the currently available treatments are associated with significant toxicities<sup>[28]</sup>, early identification of non-responders would allow clinicians to minimize the toxicity of ineffective chemotherapies and switch the treatment strategy to more aggressive regimen or addition of radiotherapy. However, further large-scale prospective studies are required to adequately position PET in sarcoma treatment management.

### Limitations of the study

Our meta-analysis had several drawbacks. First, the included studies had relatively small sample sizes in general, giving that sarcoma lacks abundant cases in clinical scenarios. Second, some extent of heterogeneity among studies was detected, while subgroup analysis could not eliminate all of the heterogeneity (although some was resolved). Third, some biases might be present such as publication bias, since *P* value yielded by Egger's test was 0.003. Spectrum biases were possible due to the failure of finding out comprehensive relevant articles or unpublished data, caused by the restriction of searching conditions.

In summary, this meta-analysis reveals a beneficial value of F-18 FDG PET to assess histopathological response after NCT in bone and soft tissue sarcomas. It is especially promising to apply to pediatric patients with OS or ES. Post-NCT SUV and SUV reduction rate on PET scans are potential candidates for quantitative criteria of response. It is noteworthy that certain chemotherapeutic agents, specifically ifosfamide, could contribute to false positives of PET results. Given the potential of PET to guide clinical treatment strategies and establish personalized medical plans, large-scale prospective studies are required to adequately position this modality in sarcoma management.

## COMMENTS

### Background

Sarcoma is a rare but highly malignant tumor that frequently requires neoadjuvant chemotherapy, but only a subcollective of patients benefit from such treatment. Therefore, there is an urgent need for non-invasive methods to accurately identify response of sarcoma to chemotherapy, and accordingly decide whether to switch to a more intensified schedule.

### Research frontiers

Radiologic imaging has some limitations for this utility due to the biological characteristics of sarcoma and the disadvantages of conventional anatomic imaging. Positron emission tomography (PET) using [F-18]-fluorodeoxy-D-glucose (F-18 FDG) is a newly applied imaging modality that measures tissue metabolism and has the unique ability to detect response at the molecular level, therefore is being increasingly used in chemotherapy response monitoring among patients with sarcoma. However, the actual accuracy of PET and some specific issues in clinical practice such as optimal patient groups, standards for response and possible causes for misjudgment remain to be explored.

### Innovations and breakthroughs

There have been multiple clinical trials assessing the accuracy of PET to monitor response to chemotherapy in sarcoma. Nonetheless, the sample sizes in conducted trials were relatively small and the results among them were remarkably discrepant. This meta-analysis provided a systematic and quantitative evaluation of the accuracy of PET for this utility. Moreover, it generated further information that clinical physicians concerned with, such as recommended subgroup of patients, potential candidates for quantitative criteria of response and

possible causes of false positives of PET.

### Applications

The study results suggested that F-18 FDG PET is promising to predict neo-adjuvant therapy response in sarcoma, especially in pediatric patients with osteosarcoma or Ewing sarcoma. A lower standardized uptake value (SUV) after chemotherapies and/or a higher SUV reduction rate are potential candidates for quantitative criteria of response. Certain chemotherapeutic agents such as ifosfamide could potentially cause false positives of PET.

### Terminology

Positron emission tomography is a functional imaging technique that produces a three-dimensional image of the body. When fludeoxyglucose (FDG), an analogue of glucose, was chosen as the tracer for PET, its concentrations will indicate tissue metabolic activity and therefore explore the possibility of viable tumor; Standardized uptake value is a frequently used quantitative parameter generated on PET scans, indicating concentrations of the tracer (e.g., FDG) in targeted tissues and therefore reflecting their metabolic activity.

### Peer review

Interesting meta-analysis on the use of PET CT in the evaluation of patients with sarcomas of both soft tissue and bone after chemoradiation with a curative intent. The authors have used all the appropriate scientific methodologies used for the preparation of meta-analyses and their results are extremely interesting and supported by the literature. The authors conclude that PET scan is a promising diagnostic tool in the evaluation of neo-adjuvant treatment response in sarcomas.

## REFERENCES

- 1 **Bastiaannet E**, Groen H, Jager PL, Cobben DC, van der Graaf WT, Vaalburg W, Hoekstra HJ. The value of FDG-PET in the detection, grading and response to therapy of soft tissue and bone sarcomas; a systematic review and meta-analysis. *Cancer Treat Rev* 2004; **30**: 83-101 [PMID: 14766127 DOI: 10.1016/j.ctrv.2003.07.004]
- 2 **Ioannidis JP**, Lau J. 18F-FDG PET for the diagnosis and grading of soft-tissue sarcoma: a meta-analysis. *J Nucl Med* 2003; **44**: 717-724 [PMID: 12732672]
- 3 **Im HJ**, Kim TS, Park SY, Min HS, Kim JH, Kang HG, Park SE, Kwon MM, Yoon JH, Park HJ, Kim SK, Park BK. Prediction of tumour necrosis fractions using metabolic and volumetric 18F-FDG PET/CT indices, after one course and at the completion of neoadjuvant chemotherapy, in children and young adults with osteosarcoma. *Eur J Nucl Med Mol Imaging* 2012; **39**: 39-49 [PMID: 21953008 DOI: 10.1007/s00259-011-1936-4]
- 4 **Meyers PA**, Heller G, Healey J, Huvos A, Lane J, Marcove R, Applewhite A, Vlamis V, Rosen G. Chemotherapy for non-metastatic osteogenic sarcoma: the Memorial Sloan-Kettering experience. *J Clin Oncol* 1992; **10**: 5-15 [PMID: 1370176]
- 5 **Bielack SS**, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jürgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *J Clin Oncol* 2002; **20**: 776-790 [PMID: 11821461 DOI: 10.1200/JCO.20.3.776]
- 6 **Kim DH**, Kim SY, Lee HJ, Song BS, Kim DH, Cho JB, Lim JS, Lee JA. Assessment of Chemotherapy Response Using FDG-PET in Pediatric Bone Tumors: A Single Institution Experience. *Cancer Res Treat* 2011; **43**: 170-175 [PMID: 22022294 DOI: 10.4143/crt.2011.43.3.170]
- 7 **Dimitrakopoulou-Strauss A**, Strauss LG, Egerer G, Vasami-llette J, Mechttersheimer G, Schmitt T, Lehner B, Haberkorn U, Stroebel P, Kasper B. Impact of dynamic 18F-FDG PET on the early prediction of therapy outcome in patients with high-risk soft-tissue sarcomas after neoadjuvant chemotherapy: a feasibility study. *J Nucl Med* 2010; **51**: 551-558 [PMID: 20351350 DOI: 10.2967/jnumed.109.070862]
- 8 **Barker FG**, Carter BS. Synthesizing medical evidence: systematic reviews and metaanalyses. *Neurosurg Focus* 2005; **19**:



- E5 [PMID: 16241107]
- 9 **Whiting P**, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25 [PMID: 14606960 DOI: 10.1186/1471-2288-3-25]
- 10 **Sweeting MJ**, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004; **23**: 1351-1375 [PMID: 15116347 DOI: 10.1002/sim.1761]
- 11 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]
- 12 **Reitsma JB**, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982-990 [PMID: 16168343 DOI: 10.1016/j.jclinepi.2005.02.022]
- 13 **Zamora J**, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006; **6**: 31 [PMID: 16836745 DOI: 10.1186/1471-2288-6-31]
- 14 **Song F**, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002; **31**: 88-95 [PMID: 11914301 DOI: 10.1093/ije/31.1.88]
- 15 **Bajpai J**, Kumar R, Sreenivas V, Sharma MC, Khan SA, Rastogi S, Malhotra A, Gamnagatti S, Kumar R, Safaya R, Bakhshi S. Prediction of chemotherapy response by PET-CT in osteosarcoma: correlation with histologic necrosis. *J Pediatr Hematol Oncol* 2011; **33**: e271-e278 [PMID: 22193290 DOI: 10.1097/MPH.0b013e31820ff78e]
- 16 **Cheon GJ**, Kim MS, Lee JA, Lee SY, Cho WH, Song WS, Koh JS, Yoo JY, Oh DH, Shin DS, Jeon DG. Prediction model of chemotherapy response in osteosarcoma by 18F-FDG PET and MRI. *J Nucl Med* 2009; **50**: 1435-1440 [PMID: 19690035 DOI: 10.2967/jnumed.109.063602]
- 17 **Hamada K**, Tomita Y, Inoue A, Fujimoto T, Hashimoto N, Myoui A, Yoshikawa H, Hatazawa J. Evaluation of chemotherapy response in osteosarcoma with FDG-PET. *Ann Nucl Med* 2009; **23**: 89-95 [PMID: 19205843 DOI: 10.1007/s12149-008-0213-5]
- 18 **Benz MR**, Czernin J, Allen-Auerbach MS, Tap WD, Dry SM, Elashoff D, Chow K, Evilevitch V, Eckardt JJ, Phelps ME, Weber WA, Eilber FC. FDG-PET/CT imaging predicts histopathologic treatment responses after the initial cycle of neoadjuvant chemotherapy in high-grade soft-tissue sarcomas. *Clin Cancer Res* 2009; **15**: 2856-2863 [PMID: 19351756 DOI: 10.1158/1078-0432.CCR-08-2537]
- 19 **Benz MR**, Allen-Auerbach MS, Eilber FC, Chen HJ, Dry S, Phelps ME, Czernin J, Weber WA. Combined assessment of metabolic and volumetric changes for assessment of tumor response in patients with soft-tissue sarcomas. *J Nucl Med* 2008; **49**: 1579-1584 [PMID: 18794268 DOI: 10.2967/jnumed.108.053694]
- 20 **Evilevitch V**, Weber WA, Tap WD, Allen-Auerbach M, Chow K, Nelson SD, Eilber FR, Eckardt JJ, Elashoff RM, Phelps ME, Czernin J, Eilber FC. Reduction of glucose metabolic activity is more accurate than change in size at predicting histopathologic response to neoadjuvant therapy in high-grade soft-tissue sarcomas. *Clin Cancer Res* 2008; **14**: 715-720 [PMID: 18245531 DOI: 10.1158/1078-0432.CCR-07-1762]
- 21 **Iagaru A**, Masamed R, Chawla SP, Menendez LR, Fedenko A, Conti PS. F-18 FDG PET and PET/CT evaluation of response to chemotherapy in bone and soft tissue sarcomas. *Clin Nucl Med* 2008; **33**: 8-13 [PMID: 18097248 DOI: 10.1097/RLU.0b013e31815c4fd4]
- 22 **Ye Z**, Zhu J, Tian M, Zhang H, Zhan H, Zhao C, Yang D, Li W, Lin N. Response of osteogenic sarcoma to neoadjuvant therapy: evaluated by 18F-FDG-PET. *Ann Nucl Med* 2008; **22**: 475-480 [PMID: 18670853 DOI: 10.1007/s12149-008-0147-y]
- 23 **Hawkins DS**, Schuetze SM, Butrynski JE, Rajendran JG, Vernon CB, Conrad EU, Eary JF. [18F]Fluorodeoxyglucose positron emission tomography predicts outcome for Ewing sarcoma family of tumors. *J Clin Oncol* 2005; **23**: 8828-8834 [PMID: 16314643 DOI: 10.1200/JCO.2005.01.7079]
- 24 **Hawkins DS**, Rajendran JG, Conrad EU, Bruckner JD, Eary JF. Evaluation of chemotherapy response in pediatric bone sarcomas by [F-18]-fluorodeoxy-D-glucose positron emission tomography. *Cancer* 2002; **94**: 3277-3284 [PMID: 12115361 DOI: 10.1002/cncr.10599]
- 25 **Franzius C**, Sciuk J, Brinkschmidt C, Jürgens H, Schober O. Evaluation of chemotherapy response in primary bone tumors with F-18 FDG positron emission tomography compared with histologically assessed tumor necrosis. *Clin Nucl Med* 2000; **25**: 874-881 [PMID: 11079583 DOI: 10.1097/00003072-200011000-00004]
- 26 **Schulte M**, Brecht-Krauss D, Werner M, Hartwig E, Sarkar MR, Keppler P, Kotzerke J, Gohlmann A, Delling G, Reske SN. Evaluation of neoadjuvant therapy response of osteogenic sarcoma using FDG PET. *J Nucl Med* 1999; **40**: 1637-1643 [PMID: 10520703]
- 27 **Picci P**, Sangiorgi L, Rougraff BT, Neff JR, Casadei R, Campanacci M. Relationship of chemotherapy-induced necrosis and surgical margins to local recurrence in osteosarcoma. *J Clin Oncol* 1994; **12**: 2699-2705 [PMID: 7989947]
- 28 **Lorigan P**, Verweij J, Papai Z, Rodenhuis S, Le Cesne A, Leahy MG, Radford JA, Van Glabbeke MM, Kirkpatrick A, Hogendoorn PC, Blay JY. Phase III trial of two investigational schedules of ifosfamide compared with standard-dose doxorubicin in advanced or metastatic soft tissue sarcoma: a European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group Study. *J Clin Oncol* 2007; **25**: 3144-3150 [PMID: 17634494 DOI: 10.1200/JCO.2006.09.7717]

**P- Reviewer:** Rapidis AD, Reis JS **S- Editor:** Song XX  
**L- Editor:** Wang TQ **E- Editor:** Wu HL





Published by **Baishideng Publishing Group Inc**

8226 Regency Drive, Pleasanton, CA 94588, USA

Telephone: +1-925-223-8242

Fax: +1-925-223-8243

E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)

Help Desk: <http://www.wjgnet.com/esps/helpdesk.aspx>

<http://www.wjgnet.com>

