Diagnostic role of PET or PET/CT for prosthetic joint infection: A systematic review and Meta-analysis

Keunyoung Kim[1] MD, PhD, Seong-Jang Kim[2,3,4,5] MD, PhD

1. Department of Nuclear Medicine, Pusan National University Hospital, Busan, Korea
2. Department of Nuclear Medicine, Pusan National University Yangsan Hospital, Yangsan, 50612, Korea
3. BioMedical Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, Yangsan, 50612, Korea
4. Department of Nuclear Medicine, College of Medicine, Pusan National University, Yangsan, Korea, 50612
5. Department of Nuclear Medicine and Biomedical Research Institute, Pusan National University Hospital, Busan, Republic of Korea, 49241

Keywords: Positron emission tomography - PET - PET/CT - Arthroplasty - Infection

Corresponding author:
Seong-Jang Kim MD, PhD
Department of Nuclear Medicine, Pusan National University Yangsan Hospital, Yangsan, 50612, Korea
BioMedical Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, Yangsan, 50612, Korea
Department of Nuclear Medicine, College of Medicine, Pusan National University, Yangsan, Korea, 50612
Tel: 82-55-360-1900,
Fax: 82-55-360-1915
growthkim@daum.net &
growthkim@pusan.ac.kr

Received:
9 January 2021
Accepted revised:
8 March 2021

Abstract

Objective: We aimed to explore the role of the diagnostic accuracy of positron emission tomography/computed tomography (PET/CT) or PET for the detection of periprosthetic infection (PPI) of lower limb arthroplasty through a systematic review and meta-analysis. **Methods:** The MEDLINE and EMBASE from the earliest available date of indexing through October 31, 2020, were searched for studies evaluating the diagnostic performance of PET or PET/CT for the detection of PPI of lower limb arthroplasty. We determined the sensitivities and specificities across studies, calculated positive and negative likelihood ratios (LR+ and LR-), and constructed summary receiver operating characteristic curves. **Results:** Across 25 results of 19 studies (826 patients), the pooled sensitivity for PET or PET/CT was 0.88 (95% CI; 0.80-0.93) with heterogeneity (χ^2=119, P=0.00) and a pooled specificity of 0.89 (95% CI; 0.83-0.93) with heterogeneity (χ^2=170, P=0.00). Likelihood ratio syntheses gave an overall positive LR+ of 7.9 (95% CI; 5.1-12.2) and negative LR- of 0.14 (95% CI; 0.08-0.23). The pooled DOR was 57 (95% CI; 31-106). **Conclusion:** The PET or PET/CT demonstrated good sensitivity and specificity for the detection of PPI of lower limb arthroplasty. At present, the literature regarding the use of PET or PET/CT for the detection of PPI of lower limb arthroplasty remains still limited; thus, further large multicenter studies would be necessary to substantiate the diagnostic accuracy of PET or PET/CT for the detection of PPI of lower limb arthroplasty.

Hell J Nucl Med 2021; 24(1): 83-93 *Epub ahead of print: 20 April 2021* *Published online: 30 April 2021*

Introduction

The total hip and knee replacements are one of the common surgical approaches which provide a significant relief of pain and improvement in the functional ability of the advanced degenerative joint diseases [1]. A major complication accompanying these procedures is loosening of the prosthesis. Approximately 10% of lower limb arthroplasties need surgical revision, of which 70% are due to loosening [2, 3]. The accurate diagnosis to distinguish between periprosthetic infection (PPI) and aseptic loosening is crucial to provide optimal treatment. However, the accurate differentiation between PPI and aseptic loosening in the painful lower limb prosthesis is a major clinical challenge. In PPI, it is essential to cure the infection before revision surgery and this condition require extended antibiotic treatment and revision arthroplasty in one or more stages [4].

Before revision arthroplasty, serological testing or microbiologic culture of joint fluid aspirations are probably the most widely used methods to determine the PPI with limited diagnostic accuracy [5]. However, the diagnosis of PPI is difficult because of lack of single test that is highly accurate, sensitive, and cost effective. Recently, various imaging techniques including radiographs, ultrasonography, computed tomography (CT), magnetic resonance imaging (MRI), bone scan, positron emission tomography (PET) could be used in the assessment of suspected PPI [6-9].

Fluorine-18-fluorodeoxyglucose (^{18}F-FDG) PET has been proposed to be a useful imaging modality in the diagnosis and management of patients with suspected infectious diseases [10, 11]. Several previous studies reported the role of ^{18}F-FDG PET or PET/CT for the evaluation of PPI in patients with total hip or knee replacements [12-17]. Apart from the ^{18}F-FDG, other radiopharmaceuticals such as ^{18}F-sodium fluoride or ^{18}F-FDG-leukocyte were also tried to evaluate the diagnostic usefulness in patients with total hip or knee arthroplasty [18-20].

The purpose of our study is to meta-analyze published data on the diagnostic performance of PET or PET/CT using various positron emitting radiopharmaceutical for the detection of PPI of lower limb arthroplasty, in order to provide more evidence-based data

and to address further studies in the evaluation of PPI of lower limb arthroplasty.

Methods

Data sources and search strategy

We conducted electronic English-language literature searches of MEDLINE via PubMed and Embase from the earliest available date of indexing through October 31, 2020. We also hand-searched the reference lists of identified publications for additional studies. We used a search algorithm based on a combination of terms: (1) "PET" OR "positron emission tomography" OR "positron emission tomography/computed tomography" OR "PET/CT" "positron emission tomography-computed tomography" OR "PET-CT"; and (2) "Arthroplasty" OR "Arthroplas*" OR "Prosthesis" OR "Prosthe*"; and (3) "Infection" OR "Septic*".

Study selection

The inclusion criteria for relevant studies were as follows: whole-body PET or PET/CT had been used to diagnose the PPI in lower limb arthroplasty patients; sufficient data to reassess sensitivity and specificity of PET or PET/CT in predicting PPI of total hip and knee replacements or absolute numbers of true-positive, true-negative, false-positive, and false-negative data had been presented; and no data overlap. Studies were excluded if fewer than 10 patients had been included. In addition, duplicate publications were excluded, as were publications such as review articles, case reports, conference papers, and letters, which do not contain the original data. Two researchers independently reviewed titles and abstracts of the retrieved articles, applying the above mentioned selection criteria. Articles were rejected if clearly ineligible. The same two researchers then independently evaluated the full-text version of the included articles to determine their eligibility for inclusion.

Data extraction and quality assessment

Information about basic study (authors, year of publication, and country of origin), study design (prospective or retrospective), patients' characteristics and technical aspects were collected. Each study was analyzed to retrieve the number of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) findings of PET or PET/CT for predicting PPI of lower limb arthroplasty, according to the reference standard. Only studies providing such complete information were finally included in the meta-analysis.

Quality of the included studies was assessed based on 15-item modified Quality Assessment of Diagnostic Accuracy Studies (QUADAS2) [21]. Two reviewers independently assessed each potentially eligible study and assigned them as a quality rating of "good," "fair," or "poor". Quality assessment was conducted based on following criteria: study design and presence of bias including selection, performance, recording, and reporting bias. Studies with high risk of bias were defined as poor quality, presence of moderate risk (did not affect the results) as fair quality, and those with minimal risk as good quality. Disagreements were settled with consensus decision. Disagreement between the 2 authors was resolved by discussion.

Data synthesis and analysis

All data from each eligible study were extracted. Categorical variables are presented as frequencies or percentages, and continuous variables are presented as mean values unless stated otherwise. Measures of the diagnostic performance, including sensitivity, specificity, and diagnostic odds ratios (DOR), are reported as point estimates with 95% confidence intervals (CI). A DOR can be calculated as the ratio of the odds of positivity in a disease state relative to the odds of positivity in the non-disease state, with higher values indicating better discriminatory test performance [22]. Between-study statistical heterogeneity was assessed using I2 and the Cochrane Q test on the basis of the random-effects analysis [23]. Publication bias was examined using the effective sample size funnel plot and associated regression test of asymmetry described by Deeks and colleagues [24]. The bivariate random-effects model was used for analysis and pooling of the diagnostic performance measures across studies, as well as comparisons between different index tests [25, 26]. We also used the model to create hierarchical summary receiver operating characteristic curves and to estimate the area under the curve [27]. When statistical heterogeneity was substantial, we performed meta-regression analysis to identify potential sources of bias [28]. Pooled estimates were also calculated for subgroups of studies that were defined according to specific study designs. Two-sided $P \leq 0.05$ was considered statistically significant. Statistical analyses were performed with commercial software programs (STATA, version 13.1; StataCorp LP).

Results

Literature search and selection of studies

After the comprehensive computerized search was performed and references lists were extensively cross-checked, our research yielded 346 records after removing 163 duplicated studies, of which 321 records (non-relevant studies 177, Case report 85, Review article 59) were excluded after reviewing the title and abstract. Remaining 25 full text articles were assessed for eligibility and 6 articles were excluded due to insufficient data for the calculation of sensitivity and specificity of PET or PET/CT for the prediction of PPI of lower limb arthroplasty. Finally, 19 studies were selected and were eligible for the systematic review and meta-analysis and no additional studies were found screening the references of these articles [12-20, 29-38]. The characteristics of the included studies are presented in Table 1. The detailed procedure of study selection in the meta-analysis is shown in Figure 1.

Study description, quality, publication bias

We conducted all analyses based on per-patient data and/or per-lesion data analysis. Among those 19 studies included in

Table 1. *Characteristics of the included studies.*

Authors	Year	RP	Device	Analysis	Patient number	Prosthesis Number	Age (range)	M/F	Dose	Diagnosis of Infection	Study design	Location
van Acker F	2001	FDG	PET	PB	21		66 (33-78)	8/13	4.6 MBq/kg	Op, C, C-FU	P	Knee
Zhuang H [a]	2001	FDG	PET	LB	62	74	27-81	NA	4.5 MBq/kg	Op, C-FU	R	K & H
Zhuang H [b]				LB		38						Hip
Zhuang H [c]				LB		36						Knee
Chacko TK	2002	FDG	PET	LB	32	41	63.7 (27-89)	20/12	2.5 MBq/kg	Op, C, C-FU	R	Hip
Vanquickenborne B	2003	FDG	PET	PB	17		63.6 (42-77)	8/9	370 MBq	Op, C-FU, A	P	Hip
Stumpe KD	2004	FDG	PET	PB	35		53 (46-89)	12/23	400 MBq	Op, C-FU, A	R	Hip
Love C [a]	2004	FDG	PET	PB	19		35-89	22/37	220 MBq	Op, C-FU	P	Knee
Love C [b]				PB	40						P	Hip
Love C [c]				PB	59						P	K & H
Reinartz P	2005	FDG	PET	LB	63	92	68 (43-88)	32/31	283 MBq	Op, C-FU	P	Hip
Mumme T	2005	FDG	PET	LB	50	70	68.7 (42-86)	19/31	250 MBq	Op, C-FU	R	Hip

(continued)

Pill SG	2006	FDG	PET	LB	89	92	NA	NA	370 MBq	Op, C-FU	P	Hip
Delank K	2006	FDG	PET	LB	27	36	45-82	NA	370 MBq	Op, C-FU	P	K & H
Chryssikos T	2008	FDG	PET	LB	113	127	59 (31-87)	54/59	370 MBq	Op, C-FU	P	Hip
Mayer-Wagner S	2010	FDG	PET	LB	32	74	70.5 (45-90)	13/19	180 MBq	Op, C	P	K & H
Basu S [a]	2014	FDG	PET	LB		87	57 (32-83)	35/52	0.14 mCi/kg	Op, C	P	Knee
Basu S [b]	2014	FDG	PET	LB		134	57 (18-84)	77/57	0.14 mCi/kg	Op, C	P	Hip
Aksoy SY	2014	FDG-WBC	PET	LB	46	54	61 (32-89)	10/36	703 MBq	Op, C-FU	P	K & H
Kumar R [a]	2016	NaF	PET/CT	PB	42		51.4 (23-76)	25/17	0.14 mCi/kg	Op, C	P	Hip
Kumar R [b]	2016	FDG	PET/CT	PB					0.14 mCi/kg		P	Hip
Kumar R, 2016	2016	NaF	PET/CT	PB	45	57	54 (23-76)	27/18	180 MBq	Op, C	P	Hip
Kobayashi N	2011	NaF	PET/CT	LB	49	65	69	NA	185 MBq	Op, C	R	Hip
Chen SH	2008	FDG	PET/CT	PB	24		63 (40-86)	13/11	370 MBq	Op, C, C-FU	R	Hip
Gravius S	2010	FDG	PET	PB	20		63.4 (41-80)	7/13	250 MBq	Op, C, C-FU	P	Knee

Op; operation, C; culture, C-FU; clinical follow-up, A; aspiration, Analysis: LB, Lesion bases; PB, Patient based, NA; Not available, Study design; R, Retrospective: P, Prospective

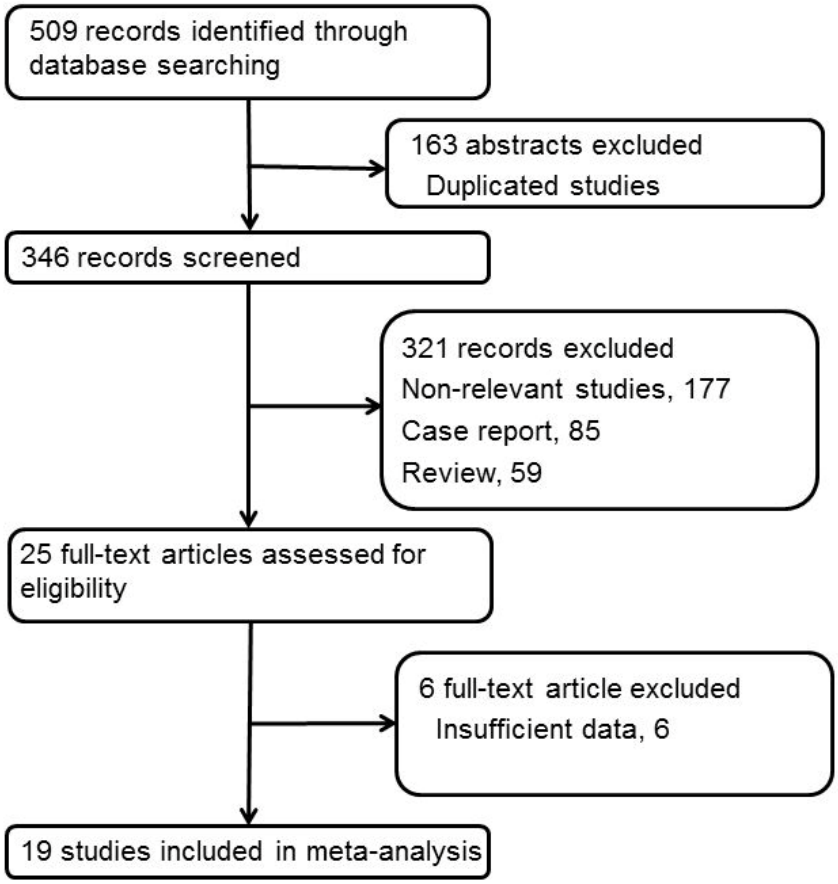

Figure 1. Flow chart of the search for eligible studies on the diagnostic performance of PET or ^{18}F-FDG PET/CT for detection of PPI of lower limb arthroplasty.

the current review, 7 studies conducted patient based analysis of PET or PET/CT [14, 19, 29-31, 37, 38]. Remaining 12 studies conducted lesion-based analysis [12, 13, 15-18, 20, 32-36]. There were a total of 826 patients in the included studies, and the age ranged from 18 to 90 years. A total 382 patients were male, and 438 patients were female. The four studies did not report the number of male and female patients in their population [12, 17, 33, 36]. Among 19 studies, 6 studies enrolled patients retrospectively [12, 13, 20, 30, 32, 37] and 13 studies [14-19, 31, 33-36, 38] were performed prospectively. Fifteen studies used PET [12-18, 29-35, 38] and 4 studies used PET/CT [19, 20, 36, 37] as imaging device in their studies. Fifteen studies used ^{18}F-FDG for PET or PET/CT imaging [12-17, 29-35, 37, 38], 3 studies used ^{18}F-NaF [19, 20, 36], and 1 study used ^{18}F-FDG-leukocyte [18] for the detection of PPI of lower limb arthroplasty. Two studies evaluated the diagnostic role of PET or PET/CT for the detection of PPI in total knee replacement [29, 38], 5 studies conducted in both of total hip and knee replacements [12, 16-18, 31], and 12 studies evaluated in total hip replacement [13-15, 19, 20, 30, 32-37]. The principal characteristics of the 8 studies included in the meta-analysis are included in Table 1. To assess a possible publication bias, Deeks's funnel plot asymmetry tests were designed. The non-significant slope indicates that no significant bias was found. The P value was 0.20 (Figure 2).

Methodological quality assessment

Figure 3 shows the risk of bias and applicability concerns summary and overall, the quality of the studies was deemed satisfactory.

Diagnostic accuracy of PET or PET/CT

The diagnostic performance results of PET or PET/CT for detection of PPI of lower limb arthroplasty of the 19 included studies in the meta-analysis are presented in Figure 4. The pooled sensitivity was 0.88 (95% CI; 0.80-0.93) with heterogeneity (χ^2=119, P=0.00) and a pooled specificity of 0.89 (95% CI; 0.83-0.93) with heterogeneity (χ^2=170, P=0.00). Likelihood ratio syntheses gave an overall positive LR+ of 7.9 (95% CI; 5.1-12.2) and negative LR- of 0.14 (95% CI; 0.08-0.23). The pooled DOR was 57 (95% CI; 31-106). Forest plots of the sensitivity and specificity of PET or PET/CT for the detection of PPI of lower limb arthroplasty are shown in Figure 4. The Figure 5 shows hierarchical summary receiver operating characteristic (ROC) curve and indicates that the areas under the curve was 0.94 (95% CI; 0.92-0.96), indicating good diagnostic accuracy.

Heterogeneity evaluation and Meta-regression analysis

Between-study heterogeneity was present for sensitivity and specificity among studies of PET or PET/CT for the detection

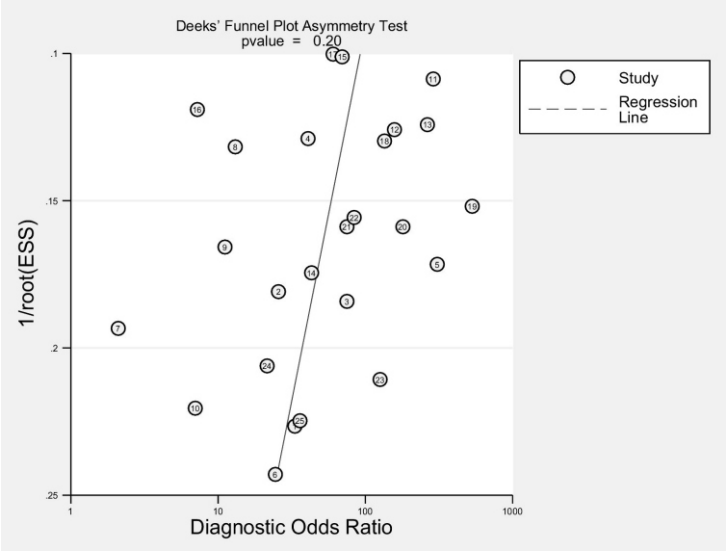

Figure 2. Results of Deeks`s funnel plot of asymmetry test for publication bias. Non-significant slope indicates that no significant bias was found. ESS; Effective sample size.

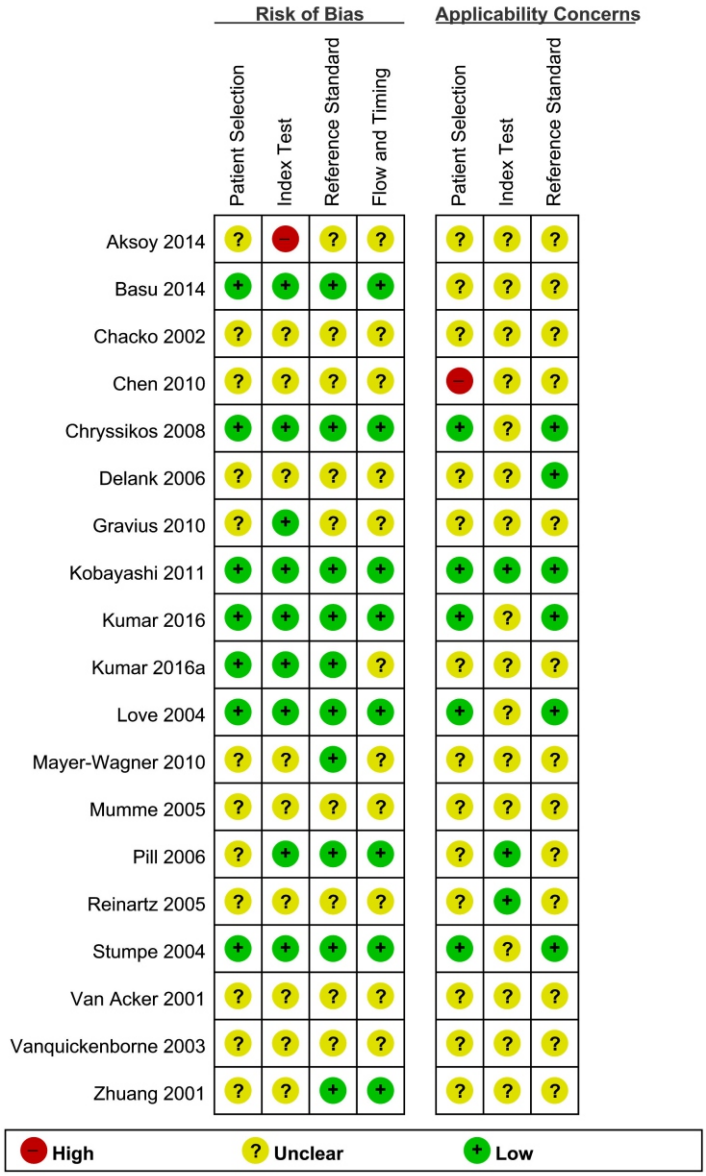

Figure 3. Risk of bias and applicability concerns summary.

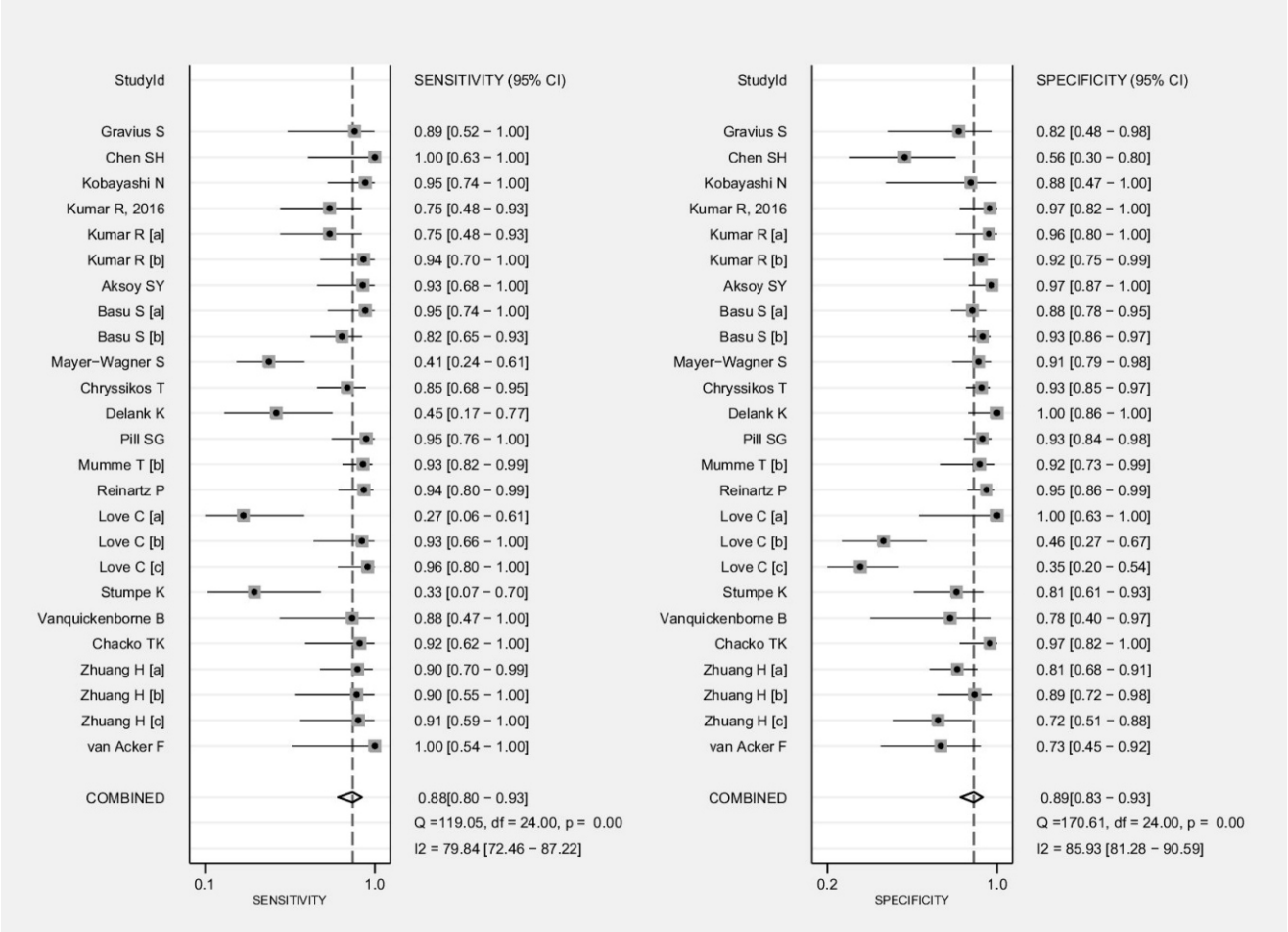

Figure 4. Forest plot of pooled sensitivity and specificity of PET or ^{18}F-FDG PET/CT for detection of PPI of lower limb arthroplasty. Summary of sensitivity and specificity of PET or PET/CT was 0.88 (95% confidence interval [CI]; 0.80-0.93) and 0.89 (95% CI; 0.83-0.93), respectively.

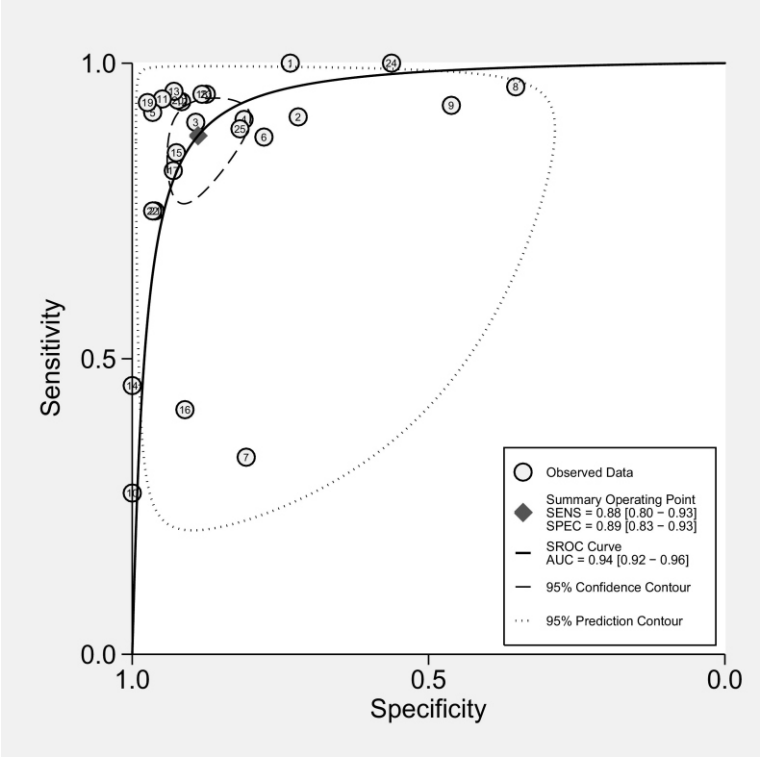

Figure 5. Hierarchical summary receiver operating characteristic (HSROC) curves of PET or ^{18}F-FDG PET/CT for detection of PPI of lower limb arthroplasty.

of PPI of lower limb arthroplasty. A meta-regression analysis was performed to explore other sources of heterogeneity in the studies. Table 2 lists the results of meta-regression analysis for identifying potential sources of heterogeneity. In univariate meta-regression analysis, the study design (prospective vs retrospective) was the potent source of heterogeneity of sensitivity of PET or PET/CT. The potent sources of heterogeneity of specificity of PET or PET/CT were used radiopharmaceuticals, analysis method (patient based vs lesion based), study origin (western vs other countries), and number of included patients. In multivariate meta-regression, the analysis method was the source of the study heterogeneity.

Table 2. *Meta-regression analysis for identifying potential sources of heterogeneity in the diagnostic performance of PET or PET/CT for the detection of periprosthetic infection of lower limb arthroplasty from 25 results of 19 studies.*

		Univariate Analysis				Multivariate Analysis	
Variables	**No of studies**	**Sensitivity**	**P value**	**Specificity**	**P value**	**LRT χ^2**	**P value**
Radiopharmaceutical							
^{18}F-FDG	21	0.88 (0.81-0.95)	0.63	0.87 (0.81-0.93)	0.02	3.69	0.16
Non-^{18}F-FDG	4	0.88 (0.72-1.00)		0.96 (0.91-1.00)			
Analysis							
PB	11	0.85 (0.74-0.97)	0.10	0.80 (0.69-0.91)	0.00	7.45	0.02
LB	14	0.89 (0.81-0.96)		0.92 (0.88-0.96)			
Study origin							
Western	19	0.86 (0.78-0.94)	0.13	0.88 (0.82-0.94)	0.05	2.42	0.30
Other countries	6	0.92 (0.83-1.00)		0.93 (0.85-1.00)			
Prosthesis							
Hip joint	20	0.89 (0.81-0.96)	0.91	0.89 (0.84-0.95)	0.30	1.53	0.47
Non-hip joint	5	0.80 (0.83-0.96)		0.88 (0.76-1.00)			
Study year							
After 2010	9	0.86 (0.74-0.97)	0.11	0.93 (0.88-0.98)	0.40	2.65	0.27
Before 2010	16	0.89 (0.81-0.96)		0.85 (0.78-0.93)			
Study design							
Prospective	19	0.86 (0.78-0.94)	0.05	0.91 (0.85-0.96)	0.39	1.19	0.55
Retrospective	6	0.91 (0.82-1.00)		0.85 (0.73-0.96)			

(continued)

Number of patient							
>49	8	0.93 (0.87-0.99)	0.93	0.88 (0.82-0.94)	0.05	2.42	0.30
≤49	17	0.83 (0.74-0.92)		0.93 (0.85-1.00)			
Image device							
PET/CT	5	0.91 (0.81-1.00)	0.94	0.91 (0.80-1.00)	0.47	0.77	0.68
PET	20	0.87 (0.79-0.94)		0.89 (0.83-0.94)			

PB; Patient-based, LB; Lesion-based

Discussion

Periprosthetic infection following total hip or knee replacements is associated with significant morbidity and economic burdens [39, 40]. After total hip or knee replacements, the PPI rates and subsequent revision operation are reported to be approximately 1% and 3% for hip arthroplasty and 2% and 5% for knee arthroplasty [41]. The major 2 complications requiring revision arthroplasty are aseptic loosening and PPI [42]. The fundamental differences in management of these two conditions emphasize the importance of correct diagnosis of the underlying causes preoperatively.

However, reliable diagnostic methods to differentiate PPI from aseptic loosening are not available. Frequently used diagnostic methods to differentiate between PPI from aseptic loosening involve measurement of C-reactive protein and erythrocyte sedimentation rate (ESR) which is sensitive with limited specificity. Traditional anatomical imaging modalities such as conventional radiography, arthrogram, ultrasonography and radionuclide studies have been proven to be effective in diagnosing PPI. Among these imaging modalities, nuclear medicine imaging represent the current imaging method of choice in most patients. For example, bone marrow scintigraphy showed a sensitivity of 33%, a specificity of 86%, a positive predictive value of 30%, and a negative predictive value of 88% in 72 arthroplasty patients [43].

In our meta-analysis, the results showed that PET or PET/CT with various radiopharmaceuticals has good diagnostic accuracy for diagnosis of the PPI in lower limb arthroplasty patients, with an area under the ROC curve of 0.94 (95% CI; 0.92-0.96). The PET or PET/CT demonstrated a sensitivity of 0.88 (95% CI; 0.80-0.93) and a specificity of 0.89 (95% CI; 0.83-0.93). This result was also consistent with the results of other well-designed studies. Mumme et al. (2005) investigated the diagnostic values of ^{18}F-FDG PET vs triple phase bone scan (TPBS) in hip arthroplasty loosening. In their study, sensitivity, specificity, and accuracy of ^{18}F-FDG PET were 91%, 92%, and 91% compared with 78%, 70%, and 74% for TPBS and they concluded that ^{18}F-FDG PET has a significant higher sensitivity and specificity than TPBS for differentiating between aseptic loosening and PPI [10]. Pill et al. (2006) also demonstrated 95.2% of sensitivity, 93% of specificity, 80% of positive predictive value (PPV), and 98.5% of negative predictive value (NPV) of ^{18}F-FDG PET for the diagnosis of PPI in 89 patients with 92 painful hip prostheses [34]. In knee arthroplasty patients, Basu et al. (2014) showed that ^{18}F-FDG PET showed 94.7% sensitivity, 88.2% specificity, 69.2% PPV, and 98.4% NPV. They concluded that the diagnostic performance of ^{18}F-FDG PET in detecting PPI in painful knee prosthesis is optimal for routine clinical application [16].

Recently, other positron emitting radiopharmaceuticals such as ^{18}F-NaF and ^{18}F-FDG-leucocyte were used for the assessment of the clinical usefulness for the detection of PPI after total hip or knee replacements [18-20, 36]. Aksoy et al. (2014) compared ^{18}F-FDG and ^{18}F-FDG-labelled leucocyte PET/CT in the imaging of PPI infection [18]. They concluded that since ^{18}F-FDG is not specific to infection, the specificity of ^{18}F-FDG PET/CT is low and ^{18}F-FDG-labelled PET/CT with its high specificity may be a useful method [18]. Using ^{18}F-NaF, Kumar et al. (2016) demonstrated fluoride PET/CT had shown optimal in routine clinical practice in detecting PPI in total hip arthroplasty patients [18, 19]. Kobayashi et al. (2011) also reported 94.7% of sensitivity and 87.5% specificity of ^{18}F-NaF PET to differentiate septic from aseptic loosening in total hip arthroplasty patients [20].

However, some studies included in the current review reported intermediate results of sensitivity and specificity of ^{18}F-FDG PET for diagnosis of the PPI in lower limb arthroplasty patients. Mayer-Wagner et al. (2010) showed 75% sensitivity and 71% specificity when differentiating septic from aseptic loosening in the total hip arthroplasty [35].

Previous well conducted meta-analysis included 14 studies comprising 838 prosthesis with suspicious of prosthetic infection after arthroplasty [44]. Jin H et al. (2014) showed that the pooled sensitivity of PET or PET/CT in detecting prosthetic infection was 86% (95% confidence interval [CI] 82%-90%) on a per prosthesis-based analysis. The pooled specificity of PET or PET/CT in detecting prosthetic infection was 86% (95% CI 83-89%) on a per- prosthesis-based analysis and the AUC was 0.93 on a per-prosthesis-based analysis [44]. They concluded that ^{18}F-FDG PET or PET/CT is accurate examination methods in the diagnosis of prosthetic infection based on high sensitivity and specificity [44].

The most important drawback of the current study is heterogeneity. The heterogeneity between studies may represent a potential source of bias. The included studies were

statistically heterogeneous in their estimates of sensitivity and specificity. This heterogeneity is likely to arise through diversity in the used standard reference (Table 1). The baseline differences among the patients in the included studies (Table 1) may have contributed to the observed heterogeneity of the results too. Also, according to the multi-variate meta-regression analysis of the current study, the analysis method (patient-based vs lesion-based) was the source of the study heterogeneity. Furthermore, the small sample size and bias were the potential source of limitations of the current review. To minimize bias in the selection of studies and in the data extraction, reviewers who were blinded to the journal, author, institution, and date of publication independently selected articles based on the inclusion criteria, and scores were assigned to study design characteristics and examination results by using a standardized form that was based on the QUADAS2 tool. Also, publication bias is a major concern in all meta-analyses as studies reporting significant findings are more likely to be published than those reporting non-significant results. We assessed the publication bias in our analysis by using funnel plots which showed some asymmetry (P=0.20).

In conclusion, the PET or ^{18}F-FDG PET/CT demonstrated good sensitivity and specificity for the detection of PPI of lower limb arthroplasty. At present, the literature regarding the use of PET or PET/CT for the detection of PPI of lower limb arthroplasty remains still limited; thus, further large multicenter studies would be necessary to substantiate the diagnostic accuracy of PET or PET/CT for the detection of PPI of lower limb arthroplasty.

The authors declare that they have no conflicts of interest.

Bibliography

1. Zimmerli W. Infection and musculoskeletal conditions: Prosthetic-joint-associated infections. *Best Pract Res Clin Rheumatol* 2006; 20: 1045-63.
2. Furnes O, Lie SA, Espehaug B et al. Hip disease and the prognosis of total hip replacements. A review of 53,698 primary total hip replacements reported to the Norwegian Arthroplasty Register 1987-99. *J Bone Joint Surg Br* 2001; 83: 579-86.
3. Keogh CF, Munk PL, Gee R et al. Imaging of the painful hip arthroplasty. *Am J Roentgenol* 2003; 180: 115-20.
4. Del Pozo JL, Patel R. Clinical practice. Infection associated with prosthetic joints. *N Engl J Med* 2009; 361: 787-94.
5. Barrack RL, Jennings RW, Wolfe MW, Bertot AJ. The Coventry Award. The value of preoperative aspiration before total knee revision. *Clin Orthop Relat Res* 1997; 345: 8-16.
6. Gemmel F, Van den Wyngaert H, Love C et al. Prosthetic joint infections: radionuclide state-of-the-art imaging. *Eur J Nucl Med Mol Imaging* 2012; 39: 892-909.
7. Osmon DR, Berbari EF, Berendt AR et al. Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America. *Clin Infect Dis* 2013; 56 :e1-e25.
8. Palestro CJ. Nuclear medicine, the painful prosthetic joint, and orthopedic infection. *J Nucl Med* 2003; 44: 927-9.
9. Zmistowski B, Della Valle C, Bauer TW et al. Diagnosis of periprosthetic joint infection. *J Orthop Res* 2014; 32: S98-S107.
10. Boerman OC, Rennen H, Oyen WJ, Corstens FH. Radiopharmaceuticals to image infection and inflammation. *Semin Nucl Med* 2001; 31: 286-95.
11. Zhuang H, Alavi A. ^{18}F-Fluorodeoxyglucose positron emission tomographic imaging in the detection and monitoring of infection and inflammation. *Semin Nucl Med* 2002; 32: 47-59.
12. Zhuang H, Duarte PS, Pourdehnad M et al. The promising role of ^{18}F-FDG PET in detecting infected lower limb prosthesis implants. *J Nucl Med* 2001; 42: 44-8.
13. Chacko TK, Zhuang H, Stevenson K et al. The importance of the location of fluorodeoxyglucose uptake in periprosthetic infection in painful hip prostheses. *Nucl Med Commun* 2002; 23: 851-5.
14. Vanquickenborne B, Maes A, Nuyts J et al. The value of ^{18}F-FDG-PET for the detection of infected hip prosthesis. *Eur J Nucl Med Mol Imaging* 2003; 30: 705-15.
15. Reinartz P, Mumme T, Hermanns B et al. Radionuclide imaging of the painful hip arthroplasty: positron-emission tomography versus triple-phase bone scanning. *J Bone Joint Surg Br* 2005; 87: 465-70.
16. Basu S, Kwee TC, Saboury B et al. FDG PET for diagnosing infection in hip and knee prostheses: prospective study in 221 prostheses and subgroup comparison with combined ^{111}In-labeled leukocyte/^{99m}Tc-sulfur colloid bone marrow imaging in 88 prostheses. *Clin Nucl Med* 2014; 39: 609-15.
17. Delank KS, Schmidt M, Michael JW et al. The implications of ^{18}F-FDG PET for the diagnosis of endoprosthetic loosening and infection in hip and knee arthroplasty: results from a prospective, blinded study. *BMC Musculoskelet Disord* 2006; 7: 20.
18. Aksoy SY, Asa S, Ozhan M et al. FDG and FDG-labelled leucocyte PET/CT in the imaging of prosthetic joint infection. *Eur J Nucl Med Mol Imaging* 2014; 41: 556-64.
19. Kumar R, Kumar R, Kumar V, Malhotra R. Potential clinical implication of ^{18}F-FDG PET/CT in diagnosis of periprosthetic infection and its comparison with ^{18}F-Fluoride PET/CT. *J Med Imaging Radiat Oncol* 2016; 60: 315-22.
20. Kobayashi N, Inaba Y, Choe H et al. Use of F-18 fluoride PET to differentiate septic from aseptic loosening in total hip arthroplasty patients. *Clin Nucl Med* 2011; 36: e156-e161.
21. Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529-36.
22. Glas AS, Lijmer JG, Prins MH et al. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003; 56: 1129-35.
23. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994; 309: 1351-5.
24. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; 58: 882-93.
25. Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58: 982-90.
26. Hamza TH, van Houwelingen HC, Stijnen T. The binomial distribution of meta-analysis was preferred to model within-study variability. *J Clin Epidemiol* 2008; 61: 41-51.
27. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001; 20: 2865-84.
28. Lijmer JG, Mol BW, Heisterkamp S et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282: 1061-6.
29. Van Acker F, Nuyts J, Maes A et al. FDG-PET, ^{99m}Tc-HMPAO white blood cell SPET and bone scintigraphy in the evaluation of painful total knee arthroplasties. *Eur J Nucl Med* 2001; 28: 1496-504.
30. Stumpe KD, Nötzli HP, Zanetti M et al. FDG PET for differentiation of infection and aseptic loosening in total hip replacements: comparison with conventional radiography and three-phase bone scintigraphy. *Radiology* 2004; 231: 333-41.
31. Love C, Marwin SE, Tomas MB et al. Diagnosing infection in the failed joint replacement: a comparison of coincidence detection ^{18}F-FDG and ^{111}In-labeled leukocyte/^{99m}Tc-sulfur colloid marrow imaging. *J Nucl Med* 2004; 45: 1864-71.
32. Mumme T, Reinartz P, Alfer J et al. Diagnostic values of positron emission tomography versus triple-phase bone scan in hip arthroplasty loosening. *Arch Orthop Trauma Surg* 2005; 125: 322-9.
33. Pill SG, Parvizi J, Tang PH et al. Comparison of fluorodeoxyglucose positron emission tomography and (111)indium-white blood cell imaging in the diagnosis of periprosthetic infection of the hip. *J Arthroplasty* 2006; 21: 91-7.
34. Chryssikos T, Parvizi J, Ghanem E et al. FDG-PET imaging can diagnose periprosthetic infection of the hip. *Clin Orthop Relat Res* 2008; 466: 1338-42.
35. Mayer-Wagner S, Mayer W, Maegerlein S et al. Use of ^{18}F-FDG-PET in the

diagnosis of endoprosthetic loosening of knee and hip implants. *Arch Orthop Trauma Surg* 2010; 130: 1231-8.
36. Kumar R, Kumar R, Kumar V, Malhotra R. Comparative analysis of dual-phase ^{18}F-fluoride PET/CT and three phase bone scintigraphy in the evaluation of septic (or painful) hip prostheses: A prospective study. *J Orthop Sci* 2016; 21: 205-10.
37. Chen SH, Ho KC, Hsieh PH et al. Potential clinical role of ^{18}F-FDG-PET/CT in detecting hip prosthesis infection: a study in patients undergoing two-stage revision arthroplasty with an interim spacer. *Q J Nucl Med Mol Imaging* 2010; 54: 429-35.
38. Gravius S, Gebhard M, Ackermann D et al. Analysis of ^{18}F-FDG uptake pattern in PET for diagnosis of aseptic loosening versus prosthesis infection after total knee arthroplasty. A prospective pilot study. *Nuklearmedizin* 2010; 49: 115-23.
39. Toms AD, Davidson D, Masri BA, Duncan CP. The management of periprosthetic infection in total joint arthroplasty. *J Bone Joint Surg Br* 2006; 88: 149-55.
40. Bozic KJ, Ries MD. The impact of infection after total hip arthroplasty on hospital and surgeon resource utilization. *J Bone Joint Surg Am* 2005; 87: 1746-51.
41. Hanssen AD, Rand JA. Evaluation and treatment of infection at the site of a total hip or knee arthroplasty. *Instr Course Lect* 1999; 48: 111-22.
42. Clohisy JC, Calvert G, Tull F et al. Reasons for revision hip surgery: a retrospective review. *Clin Orthop Relat Res* 2004; 429: 188-92.
43. Levitsky KA, Hozack WJ, Balderston RA et al. Evaluation of the painful prosthetic joint. Relative value of bone scan, sedimentation rate, and joint aspiration. *J Arthroplasty* 1991; 6: 237-44.
44. Jin H, Yuan L, Li C et al. Diagnostic performance of FDG PET or PET/CT in prosthetic infection after arthroplasty: a meta-analysis. *Q J Nucl Med Mol Imaging* 2014; 58: 85-93.