



# Systematic Review on Reproducibility of Nuclear Imaging in the Assessment of Painful Hip and Knee Replacements

S. J. Verberne<sup>1</sup> · J. W. P. Kuiper<sup>1</sup> · W. A. M. Broos<sup>2</sup> · O. P. P. Temmerman<sup>1</sup>

Accepted: 10 August 2020 / Published online: 21 August 2020  
© The Author(s) 2020

## Abstract

Nuclear imaging is used in the evaluation of painful arthroplasties for diagnosing loosening or periprosthetic joint infection (PJI). The purpose of this systematic review is to depict the reproducibility of the various nuclear imaging modalities used in the assessment of painful hip and knee arthroplasties. A systematic review of the literature was conducted with a comprehensive search of MEDLINE to identify clinical studies investigating the intra- and inter-observer agreement of nuclear imaging modalities in diagnosing PJI and mechanical loosening. A total of 3,000 studies, published between 2004 and 2020, were reviewed and 16 studies met the inclusion criteria. Quality assessment identified concerns with regard to the internal validity in approximately 40% of the included studies. A meta-analysis could not be performed because of insufficient available clinical data. The intra-observer agreement was poorly investigated. The included studies reported an inter-observer agreement of “slight” to “excellent” for bone scintigraphy, “moderate” to “substantial” for leukocyte scintigraphy, “substantial” to “almost perfect” for combined leukocyte and bone marrow scintigraphy, and “fair” to “substantial” for anti-granulocyte scintigraphy. Hybrid imaging with SPECT/CT and FDG-PET/CT demonstrated “substantial” and “almost perfect” inter-observer agreement for symptomatic hip prostheses, respectively. The reproducibility of nuclear imaging is underreported in clinical studies investigating painful knee and hip arthroplasties. Moreover, the included studies demonstrated methodological concerns with a high risk of bias. The available evidence demonstrated a wide range of inter-observer agreement using scintigraphy. Hybrid imaging with SPECT/CT and FDG-PET/CT may improve the accuracy of interpretation and reproducibility. However, literature provides limited data to support this assumption.

**Keywords** Nuclear imaging · Prosthetic joint infection · Reproducibility · Review · Painful arthroplasties

## Introduction

The incidence of complications after total hip and knee replacements is low, but due to the increased frequency of implantations, complications are commonly encountered and require an appropriate diagnostic algorithm. Among the variety of complications (including heterotopic ossification, pseudotumor, inflammatory disorders, hardware failure, and foreign body granulomatosis causing osteolysis), the two major complications are mechanical (aseptic) loosening and

periprosthetic joint infection (PJI). PJI is considered one of the most devastating prostheses-related complications. However, reliable differentiation between PJI and aseptic loosening is often difficult since both may be accompanied by similar symptoms.

Currently, there is no diagnostic algorithm and despite multiple diagnostic tests, diagnosing PJI and differentiation from mechanical loosening remains challenging. Although current guidelines do not recommend the routine use of nuclear imaging [1, 2], a variety of radionuclide imaging modalities can contribute to the evaluation of suspected PJI or loosening. These modalities include bone scintigraphy (BS), leukocyte scintigraphy (LS), anti-granulocyte scintigraphy (AGS), fluorodeoxyglucose positron emission tomography (FDG-PET), single-photon emission computed tomography (SPECT), and combined techniques [3–5]. In the assessment of PJI of the hip, a meta-analysis reported pooled sensitivity and specificity of 80–69%, respectively, for BS, 88–92% for

✉ S. J. Verberne  
stevenverberne@hotmail.com

<sup>1</sup> Department of Orthopaedics, Centre for Orthopaedic Research Alkmaar (CORAL), NWZ, Alkmaar, the Netherlands

<sup>2</sup> Department of Radiology & Nuclear Medicine, NWZ, Alkmaar, the Netherlands

LS, 69–96% for combined LS-bone marrow scintigraphy (LS-BMS), 84–75% for AGS, and 86–93% for FDG-PET/CT [4]. For PJI of the knee, the reported pooled sensitivity and specificity were 93–56%, respectively, for BS, 88–77% for LS, 80–93% for LS-BMS, 90–95% for AGS, and 70–84% for FDG-PET/CT [5]. For aseptic loosening of hip arthroplasties, the reported pooled sensitivity and specificity were 85–72%, respectively, for BS [6]. In one clinical study which investigated symptomatic hip arthroplasties with SPECT/CT, a sensitivity of 94% and a specificity of 100% were reported [7]. In the evaluation of aseptic loosening of knee arthroplasties, the reported range of sensitivity and specificity was 76–100% and 33–100%, respectively, for BS, 56–100% and 56–100% for FDG-PET/CT, and 86–100% and 63–93% for SPECT/CT [8].

The reported diagnostic accuracies of the nuclear imaging modalities in the assessment of symptomatic arthroplasties reflect the agreement of the interpretation of the scan (index test) with the reference standard. However, understanding the variables which are of influence to the interpretation of a scan is important for the outcome of (false) positive or (false) negative results. These variables are related to several domains such as the index test, the reference standard, and observer-related factors. Important variables related to the subject (the painful arthroplasty) are the time interval between imaging and surgery, location and type of implant, and severity of the infection. For example, the interpretation of a low grade PJI could be more challenging for the observer and lead to more false negative results compared with diagnosing an acute PJI. The type of imaging modality (index test), the performed imaging protocol, and the used diagnostic criteria could also affect the interpretation and outcome. For example, a variety of diagnostic criteria was reported for various nuclear imaging modalities which could lead to different diagnostic accuracy outcomes regarding the investigated imaging modality [4, 5]. Finally, in the interpretation of the scans, the background of the observer such as the level of expertise and experience may also affect the diagnostic outcome. These factors not only could affect the reported diagnostic accuracy but also may introduce variations in intra- and inter-observer agreement. These variations are important and potential compromising factors for the reproducibility and therefore the clinical application of the imaging modalities in the assessment of symptomatic arthroplasties.

The (statistical) reliability of an imaging modality is defined by the degree of agreement when different observers use the same imaging technique, classification, and procedure to assess the same subject. Each of the nuclear imaging modalities is characterized by specific benefits and drawbacks in diagnosing PJI or aseptic loosening. For example, scintigraphy is potentially hampered by non-specific tracer uptake which could lead to unsatisfactory results of inter-observer reliability. In literature, it has been suggested that the interpretation of nuclear imaging could be improved by more recently

introduced hybrid techniques as SPECT/CT and PET/CT [9, 10]. For other pathologies, including infectious bone diseases, the improved diagnostic accuracy and reproducibility of hybrid imaging have been well established [11]. However, in the assessment of symptomatic arthroplasties, this improvement in terms of reproducibility has not been well defined in literature.

The level of reproducibility and variables that could influence the individual interpretation of images is not always clear and therefore a suitable subject for systematically literature investigation. To the best of our knowledge, no systematic review that investigated this topic has been published before. The objective of this study is to systematically review the reproducibility of currently used nuclear imaging modalities in the evaluation of symptomatic hip and knee arthroplasties.

## Materials and Methods

### Search Criteria

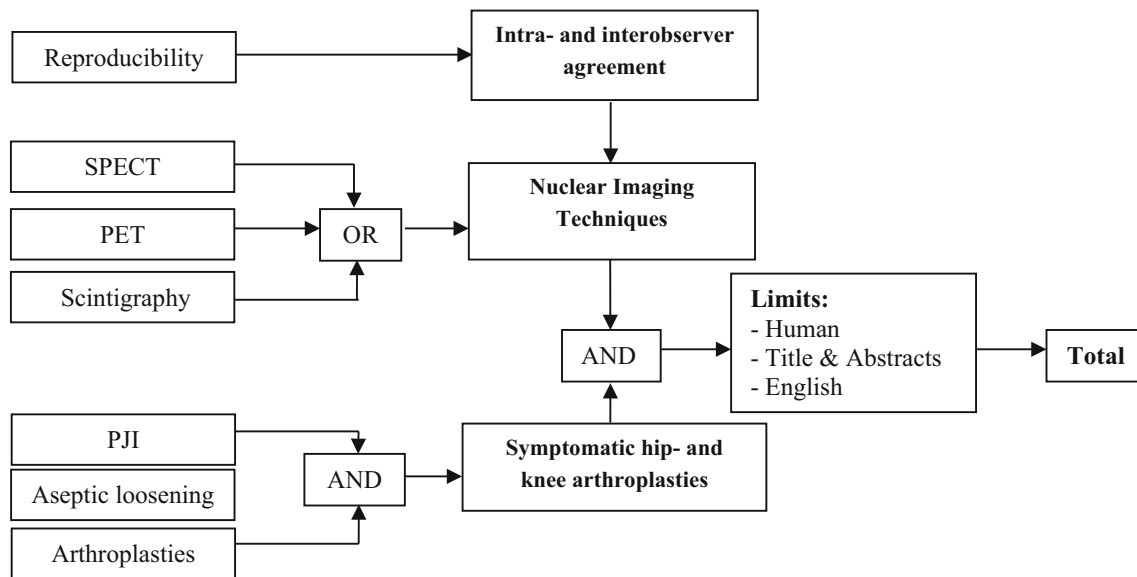
The imaging modalities that were reviewed for the assessment of painful arthroplasties were planar scintigraphy, PET, SPECT, and combined techniques.

### Search Strategy

A computer-aided search of the PubMed databases was conducted in June 2019 and updated in April 2020. The search was restricted to primary studies that were written in English. For each database, a specific search strategy was developed (Fig. 1) with an informatics specialist. Reference lists of the identified studies and relevant reviews were hand-searched for supplementary eligible studies. The search was performed according to the PRISMA guideline (Preferred Reporting Items for Systematic reviews and Meta-Analyses) Statement [12].

### Study Selection

The following inclusion criteria were used to identify eligible, clinical studies: (1) planar scintigraphy, SPECT, or PET had been used to identify suspected PJI or loosening of hip and/or knee arthroplasties. (2) The study used a valid reference standard for PJI (criteria by the Musculoskeletal Infection Society and/or the criteria of Infectious Diseases Society of America) [1, 13] or aseptic loosening with preoperative findings (3), a clinical follow-up interval of at least 6 months when surgery was not performed. (3) The study reported the calculated intra- and inter-observer agreement with multiple observers. Exclusion criteria were (1) nonhuman studies, (2) non-English language studies, or (3) studies in which differentiation between osteomyelitis without arthroplasties and PJI was



**Fig. 1** Flowchart of the conducted search strategy. CT, computed tomography; PET, positron emission tomography; SPECT, single-photon emission computed tomography; PJI, prosthetic joint infection

not possible. There was no restriction regarding the date of publication.

The titles were screened for eligibility by 1 reviewer (S.V.) and were then processed for abstract assessment. The titles and abstracts were independently screened and assessed for eligibility in an unblinded standardized manner by 2 reviewers (S.V. and J.K.) Studies considered to be of dubious eligibility were rejected. The final decision on inclusion was made on the basis of the full text of the article.

### Methodological Quality Assessment

The criteria list of QUADAS2 for evaluating internal and external validity of diagnostic studies recommended by the Cochrane Methods Group on Systematic Review of Screening and Diagnostic Tests was used for grading the methodological quality of the selected studies [14]. Evaluation was performed by the two reviewers (S.V. and J.K.) independently. The internal and external validity were evaluated and the criteria of QUADAS2 were used for the determination of the methodological limitations for descriptive purposes. All studies that met the inclusion criteria were included and, a priori, no studies were excluded from the systematic review on the basis of quality. A quantitative analysis could not be performed due to insufficient available data.

## Results

### Included Studies

The search strategy identified 2,994 studies from MEDLINE. There was no overlap found between the retrieved studies. No

additional studies were identified through other sources. Of the initial 2,994 studies, 2,761 were excluded after analyzing the information provided in the title and abstract. The full articles of the remaining 233 studies were reviewed for eligibility. Six studies were extracted from the reference lists of these studies ( $n = 239$ ). There was no disagreement between the reviewers regarding the definitive inclusion of the studies. The main reason for exclusion was the absence of intra- and inter-observer agreement calculation in the use of imaging for diagnosing PJI or aseptic loosening ( $n = 179$ ). Other reasons for exclusion were the absence of a valid reference standard ( $n = 5$ ), no differentiation between osteomyelitis and periprosthetic complications ( $n = 3$ ), or absence of symptomatic arthroplasties inclusion ( $n = 36$ ). Meta-analysis was not performed due to insufficient available clinical data. In general, no differentiation a priori was made for the type of implant, the interpretation criteria used for the index test, or the time period between surgery and imaging (Table 1).

### Description of Study Characteristics

A total of 16 studies, published between 2004 and 2020, were included for systematic review. Two studies reported the inter-observer agreement for nuclear imaging in the assessment of aseptic loosening of hip and knee arthroplasties [17, 18]. In 4 studies, the results of reproducibility were reported without differentiation between PJI and aseptic loosening using bone scintigraphy [7, 19–21]. Of those studies, one reported the reproducibility of bone scintigraphy combined with SPECT/CT [7]. The intra- or inter-observer agreement of nuclear imaging in the assessment of PJI was investigated by 10 studies. Of those studies, 2 studies used leukocyte scintigraphy [22, 23], 1 study used combined leukocyte and bone scintigraphy

**Table 1** The measurements of observer agreement categorical data

$k < 0$ reflects “poor”	
0 to 0.20 “slight”	
0.21 to 0.4 “fair”	
0.41 to 0.60 “moderate”	
0.61 to 0.8 “substantial”	
above 0.81 “almost perfect”	
The measurement of observer agreement for categorical data [15] Icc correlation [16]	
< 0.40	Poor
0.40 to 0.59	Fair
0.60 to 0.74	Good
0.75 to 1.00	Excellent

[24], 3 studies used combined leukocyte and bone marrow scintigraphy [24–26], 3 studies used anti-granulocyte scintigraphy [15, 16, 27], 1 study used combined anti-granulocyte and bone marrow scintigraphy [27], and 3 studies used FDG-PET/CT [19, 28, 29]. The characteristics of the studies are detailed in Table 2.

### Qualitative Analysis

The external validity showed high concerns with regard to applicability to clinical practice in approximately 40% of the included studies for the domain “patient selection.” The internal validity of the included studies showed more concerns with regard to the risk of bias. Approximately 40% of the included studies did not provide sufficient information regarding the domains “index test” and “patient selection” (Fig. 2) More than 70% of the studies did not provide detailed data regarding the domain “flow and timing.” The details of the internal and external validity are shown in Fig. 3.

### Bone Scintigraphy and Combined Techniques

Of the 6 studies that investigated the use of bone scintigraphy, one study reported “excellent” (ICC = 0.89) intra-observer agreement of two observers in the assessment of the symptomatic hip [21]. In the assessment of aseptic loosening of THA, a “fair” (ICC = 0.43) and “good” (ICC = 0.67) inter-observer agreement was reported for the acetabular and femoral component, respectively [17, 18]. No significant differences between the cemented and uncemented prostheses was found. In 4 studies, the reproducibility was reported without differentiation between PJI and aseptic loosening. For the symptomatic hip, two studies reported a kappa value for inter-observer agreement of 0.11 (slight) [19] and an ICC of 0.81 (excellent) [21]. One study did not differentiate between hip and knee replacements and reported a kappa of 0.92, 0.81, and 0.80 for the blood flow, blood pool, and bone phase,

respectively (overall kappa was not reported) [20]. One study combined bone scintigraphy with SPECT and reported a “fair” ( $\kappa = 0.27$ ) and “moderate” ( $\kappa = 0.60$ ) inter-observer agreement in the assessment of THA and TKA respectively [7]. The same study found a “moderate” ( $\kappa = 0.58$ ) and “substantial” ( $\kappa = 0.80$ ) inter-observer agreement when bone scintigraphy was combined with SPECT/CT in the assessment of THA and TKA respectively.

### Leukocyte Scintigraphy and Combined Techniques

One study [23] reported the inter-observer agreement of leukocyte scintigraphy in diagnosing PJI and reported “moderate” to “substantial” as a result of interpretation by three observers ( $\kappa = 0.55, 0.60, \text{ and } 0.74$ ). There was no significant difference of concordance in 24% ( $n = 55$ ) and 33% ( $n = 40$ ) of the hip and knee group, respectively. Combined leukocyte and bone marrow scintigraphy demonstrated in one study “almost perfect” inter-observer agreement of three pairs of observers in the assessment of TKA and THA together [24]. For the assessment of hip arthroplasty, a kappa value of 0.74 (substantial) was reported by another study [26]. None of the 5 studies that investigated the use of leukocyte scintigraphy reported the intra-observer agreement.

### Anti-granulocyte Scintigraphy and Combined Techniques

The 3 studies [15, 16, 27] that investigated the use of anti-granulocyte scintigraphy in the evaluation of PJI reported “excellent” inter-observer agreement for TKA/UKA ( $\kappa = > 0.75$ ), “fair to moderate” for THA/TKA together ( $\kappa = 0.31\text{--}0.52$ ), and “fair to excellent” agreement depending on the grade of infection (mild to severe). One study [27] reported the anti-granulocyte scintigraphy combined with bone-marrow scintigraphy and reported kappa values of 0.55–0.72 without differentiation between THA and TKA. None of the 3 studies reported the intra-observer agreement of anti-granulocyte scintigraphy in the assessment of PJI.

### FDG-Pet/Ct

Of the 3 studies that investigated the use of PET/CT in the evaluation of PJI, two studies reported the inter-observer agreement of the tracer  $^{18}\text{F}$ -FDG. Stumpe et al. reported “moderate” inter-observer agreement with ( $\kappa = 0.47$ ) [19]. Verberne et al. reported for the inter-observer agreement for 4 different criteria for PJI with kappa values ranging from 0.77–0.85 [29]. One study reported “substantial” ( $\kappa = 0.67$ ) inter-observer agreement when  $^{18}\text{F}$ -PET/CT was combined with  $^{18}\text{F}$ Fluoride. None of the studies reported the intra-observer agreement for the use of PET/CT in the assessment of PJI.

**Table 2** Characteristics of the included studies

Study	Year	Design	No. of patients (male/female)	Patients demographics (male/female)	Age (range)	Age (mean)	No. of prostheses (resection arthroplasty/revision)	Prostheses	C/U	TAS	CFU
El Esper et al.	2004	NR	60	NR	NR	NR	43 (30res)	26H, 17 K	NR	Mean 5.4 years	NR
Pelosi et al.	2004	RS	78	36/42	30–87 years	70 years	95	47H, 35 K	NR	NR	> 1 year
Stumpe et al.	2004	PS	35	12/23	47–89 years	71 years	35(9rev)	35H	18C, 17 U	71 months (12–260months)	> 6 months
Temmerman et al.*	2004	RS	73	19/54	34–54 years	72 years	83	73H	51C, 22 U	NR	NR
Temmerman et al.*	2006	RS	78	20/58	29–88 years	70 years	78	78H	55C, 23 U	3 months (1–10 months)	> 1 year
Fuster et al.	2008	PS	70	28/42	55–81 years	68 years	70	70H	NR	46 months	> 1 year
Gratz et al.	2009	PS	26	9/17	49–85 years	71 years	26	26 K	18C, 8 U	> 1 year	> 6 months
Fuster et al.	2011	PS	40	14/26	56–76 years	66 years	40	21H, 16 K, 3S	NR	Mean 1.5 years	> 1 year
Sousa et al.	2011	PS	27	3/24	NR	65 years (a)	27	2H, 25 K	NR	Mean 29.7 years	> 1 year
Gratz et al.	2012	PS	20	NR	60–76 years	68 years	20	6UK, 14 K	14C	31 months (2–148 months)	> 1 year
Erba et al.	2014	RS	235	104/131	16–93 years	61 years	132	H63, K66, S3	NR	NR	> 6 months
Arican et al.*	2015	RS	50	15/35	41–84 years	63 years	50	H20, 30 K	NR	62 months (3–120 months)	NR
Granados et al.	2015	PS	120	50/70	60–82 years	71 years	120	H63, K57	NR	78 m (2–154 months)	> 1 year
Yoldas et al.	2016	RS	52	17/35	57–77 years	66 years/69 years	52	52H	52 U	> 2 years	NR
Kumar et al.	2016	PS	42	25/17	23–76 years	53 years	42	42H	NR	1–7.6 years	NR
Verberne et al.	2018	RS	43	18/25	54–92 years	76 years	43(12rev)	43H	11C, 22 U	47 months (1–415 months)	> 1 year
Total											

\*Study investigated aseptic loosening, NR not reported, RS retrospective study design, PS prospective study design, (res) resection arthroplasty, (rev) revision, H hip arthroplasty, K knee arthroplasty, U/KP unicompartment knee replacement, C cemented, U uncemented, S shoulder arthroplasty, CFU clinical follow-up

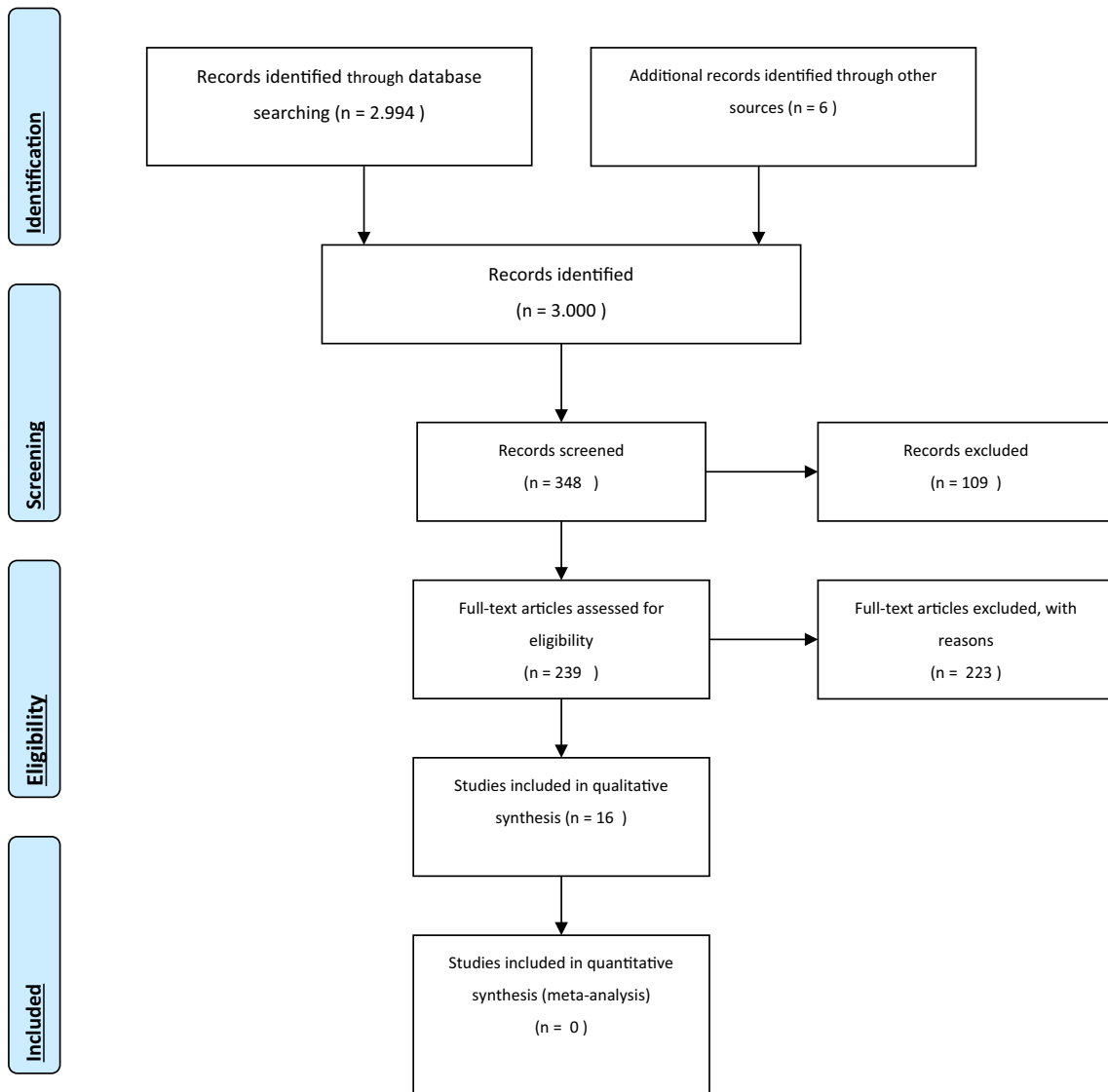


Fig. 2 PRISMA 2009 flow diagram

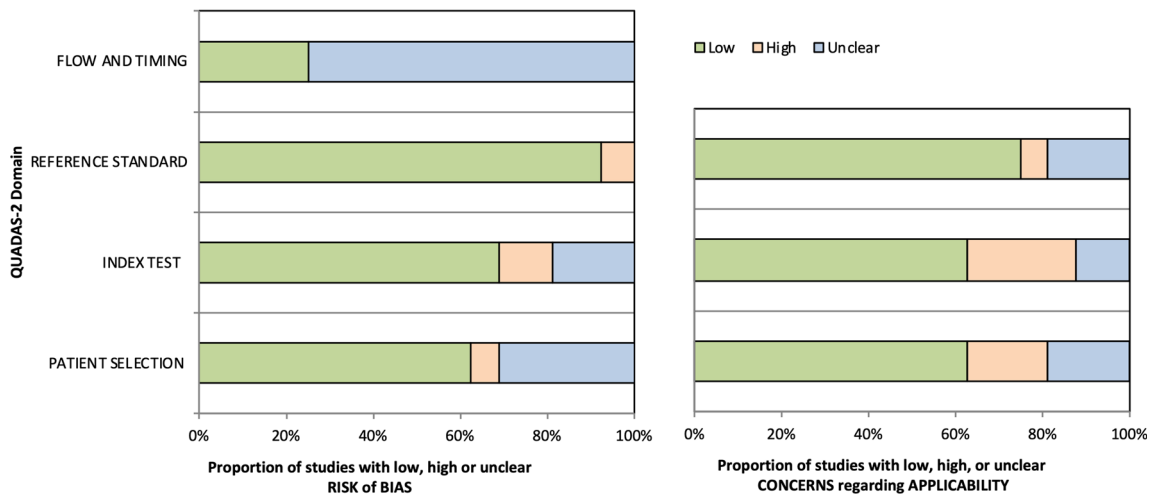


Fig. 3 QUADAS-2, risk bias and applicability

**Table 3** Intra- and inter-observer agreement of bone scintigraphy in diagnosing symptomatic hip and knee replacements

Study	Year	Imaging	Tracer	Doses	Prostheses	Observers	Intra-OA	Inter-OA
Temmerman et al.	2004	BS	<sup>99m</sup> Tc-HDP	370 MBq	THA(a)	2*	-	ICC 0.43 (cemented 0.39, uncemented 0.53)
Stumpe et al.	2004	BS	<sup>99m</sup> Tc-DCP	700 MBq	THA	2(1)	-	$\kappa$ 0.11
Temmerman et al.	2006	BS	<sup>99m</sup> Tc-HDP	370 MBq	THA(f)	2*	-	ICC 0.67
Yoldas et al.	2016	BS	<sup>99m</sup> Tc-HMDP	NR	THA	2*	ICC 0.89	ICC 0.81
Granados et al.	2015	BS	<sup>99m</sup> Tc-HMDP	925 MBq	THA/TKA	3*	-	$\kappa$ 0.80–0.92
Arıcan et al.	2015	BS-SPECT	<sup>99m</sup> Tc-MDP	740 MBq	THA	2(2)	-	$\kappa$ 0.27 (0.43 <sup>1</sup> )
		BS-SPECT/CT	<sup>99m</sup> Tc-MDP	740 MBq	THA	2(2)	-	$\kappa$ 0.58 (0.72 <sup>1</sup> )
		BS-SPECT	<sup>99m</sup> Tc-MDP	740 MBq	TKA	2(2)	-	$\kappa$ 0.60 (0.43 <sup>1</sup> )
		BS-SPECT/CT	<sup>99m</sup> Tc-MDP	740 MBq	TKA			$\kappa$ 0.80 (0.72 <sup>1</sup> )

BS bone scintigraphy, <sup>99m</sup>Tc-DCP, <sup>99m</sup>Tc dicarboxidiphosphonate. <sup>99m</sup>Tc-HDP, <sup>99m</sup>Tc hydroxydiphosphonate. <sup>99m</sup>Tc-HMDP, <sup>99m</sup>Tc hydroxymethylene diphosphonate. <sup>99m</sup>Tc-MDP, <sup>99m</sup>Tc methylene diphosphonate. THA total hip arthroplasty, TKA total knee arthroplasty. (a), acetabular component. (f), femoral component. ICC intra-class correlation coefficients (ICC).  $\kappa$  Kappa

(<sup>1</sup>)(<sup>2</sup>)Two nuclear medicine physicians, > 5-year-experience. Two nuclear medicine physicians, level of experience not reported

\*Level of experience and expertise was not defined. OA observer agreement

## Discussion

Various nuclear imaging modalities can contribute in the challenging diagnosis and differentiation between PJI and mechanical loosening of hip and knee replacements. The reproducibility of these imaging modalities is an important factor for the clinical usability and could concern a major limitation in the evaluation of symptomatic joint replacements and will be discussed in the section below.

### Bone Scintigraphy and SPECT/CT

Bone scintigraphy is known as a highly sensitive imaging modality for diagnosing aseptic loosening. This modality is widely available and can exclude any form of loosening. <sup>99m</sup>Tc-labeled bisphosphonates are the most commonly used radiopharmaceuticals to depict osteoblastic activity. In patients with mechanical loosening of the hip or knee replacement, negative perfusion and blood pool but positive osseous phases are often interpreted as loosening. Throughout literature, one investigation group [17, 18] reported the diagnostic accuracy of bone scintigraphy for the acetabulum and femoral component of THA and reported “fair” and “good” inter-observer agreement, respectively. No substantial differences between cemented and uncemented prostheses were found. However, the results of these studies have not yet been confirmed by other studies. None of the included studies reported the reproducibility of bone scintigraphy for diagnosing mechanical loosening of TKA.

In the assessment of PJI, the increased osteoblastic activity on bone scintigraphy results in a highly sensitive modality. However, in the first years after implantation, specificity is

dramatically lower and differentiation between infection and mechanical loosening (1–2 years for hip prostheses, 4–5 years after knee prostheses) is not accurate [4, 5]. A bone scan is defined as positive for infection when there is uptake in all three phases (positive perfusion, blood pool, and osseous phase), and this criterion was used by all four studies that investigated bone scintigraphy in the assessment of painful prostheses. However, a wide variety in the level of reproducibility was reported (“slight” to “almost perfect” inter-observer agreement) [19–21] This wide range of observer agreement may be explained by the low-resolution images of scintigraphy which could lead to unsatisfactorily results of inter-observer reliability.

The introduction of hybrid imaging with SPECT/CT could improve diagnostic accuracy and reproducibility of the low-resolution scintigraphy. Adding only SPECT may not improve the reproducibility of bone scintigraphy alone [7]. This may be the result of the missing CT portion, which results in less anatomical relation and therefore less accurate interpretation of the images. For other pathologies, it has been reported that adding SPECT/CT to planar images results in higher diagnostic accuracy because of better resolution, direct (anatomic) correlation of functional and morphological abnormalities, and better distinguishing of bone infection from soft tissue infections through improved assessment of the extent of the infection [11]. A previous study showed that when SPECT/CT was added, inter-observer agreement regarding the location of infectious foci increased from  $\kappa = 0.68$  to  $\kappa = 1.0$  in the assessment of osteomyelitis [11]. In agreement with the first clinical studies, 4 studies demonstrated “excellent” intra- and inter-observer agreement when SPECT/CT was added in the interpretation of symptomatic hip and knee

**Table 4** Bone scintigraphy in diagnosing symptomatic hip and knee replacements, index test, and reference standard

Study	Year	Imaging	Prostheses	Reference standard	Criteria for PJI or aseptic loosening
Temmerman et al.	2004	BS	THA(a)	S	Acetabulum: a moderate increased uptake in > two zones or intense uptake in at least one zone (scale 0–5)
Stumpe et al. [1]	2004	BS	THA	M, S	PJI: focally or diffusely increased periprosthetic uptake in all three phases Loosening: increased radionuclide uptake limited to the third phase
Temmerman et al.	2006	BS	THA(f)	S	Femoral: a mild or moderate increase in uptake of contrast in more than two zones or intense uptake in at least one zone
Granados et al. [2]	2015	BS	THA/TKA	M, H	Analyzing images of each phase (blood flow, blood pool, and bone phase) of the BS independently, 2 categories (either positive or negative for periprosthetic infection) were considered. The GL was considered positive for infection when there was any extra-medullary periprosthetic uptake
Yoldas et al. [3]	2016	BS	THA	MSIS	Increased uptake in all three phases
Arican et al. [4]	2015	BS-SPECT	THA	M, S	Focal or minimal linear increased periprosthetic and tip activity of prosthesis on all three phases of the planar bone scan and SPECT with a corresponding irregular periprosthetic radiolucency greater than 2 mm on CT
		BS-SPECT BS-SPECT/CT	TKA TKA	M, S M, S	Focal or minimal linear increased periprosthetic and tip activity of prosthesis on all three phases of the planar bone scan and SPECT with a corresponding irregular periprosthetic radiolucency greater than 2 mm on CT

*MSIS* Musculoskeletal Infection Society for definitive diagnosing of periprosthetic infection. *S* surgery/intra-operative findings. *M* = 2 or more positive peroperative samples that grew the same bacteria. *H* histopathology. 1 = 9/26 patients with PJI, 2 = 18/120 patients with PJI, 3 = 20/52 patients with PJI, 4 = number of PJI not reported

arthroplasties [30–33]. However, these studies did not correlate the nuclear imaging results with surgical findings or clinical follow-up whether the prosthesis was loose or not. Because of this major methodological concern, no conclusions can be drawn regarding the clinical applicability and were therefore excluded from this review. In contrary to these studies, the one study that did correlate the SPECT/CT findings with a valid reference standard reported “moderate” to “substantial” agreement for the hip ( $\kappa = 0.58$ ) and knee ( $\kappa = 0.80$ ) [7]. Unfortunately, these findings are not yet confirmed by other studies. Although it is plausible that adding SPECT/CT improves the reliability of the interpretation of scintigraphy in the assessment of symptomatic joint replacement, there is limited clinical data in literature to support this assumption. The clinical impact of SPECT/CT and its improvement in terms of reproducibility would be a suitable subject for further investigation in the field of symptomatic arthroplasties.

### Leukocyte Scintigraphy and Combined Techniques

Leukocyte scintigraphy has a long history in the evaluation of PJI as it can be used as a more specific imaging modality in order to differentiate between aseptic loosening and PJI. When the appropriate criteria are used, leukocyte scintigraphy may be a specific and accurate imaging modality for diagnosing PJI [4, 5]. Only one study reported “moderate” to “substantial” inter-observer agreement of three different pairs of readers ( $\kappa = 0.55$ – $0.74$ ). In 26 (27.4%) of the 95 investigated knee and hip arthroplasties, there was no concordance between the three observers. There were differences in

concordance between the hip (76%) and knee (67%) group, however, not significant. The diagnostic accuracy of leukocyte scintigraphy is potentially hampered by the fact that labeled leukocytes not only accumulate in infections but also accumulate physiologically in the bone marrow. In order to reduce the number of false-positive results, leukocyte scintigraphy is often combined with bone marrow scintigraphy, which has been proposed as the imaging modality of choice for diagnosing PJI [34]. Besides the improved sensitivity and specificity compared with leukocyte scintigraphy alone, this combined technique demonstrated “substantial” to “almost perfect” inter-observer agreement in two studies [24, 26]. An explanation for the good reproducibility of LS-BMS scans may lie in the fact that comparison of images improve interpretation, which was also found by Sousa et al. [27] It is important to note that different diagnostic accuracies were found between knee and hip arthroplasties in previous studies [4, 5]. The one study that investigated the reproducibility of THA alone reported “moderate” inter-observer agreement [26], while “almost perfect” concordance was found in another study with a combined hip and knee group [24]. For further studies, it is important to separately analyze the reproducibility of both arthroplasties.

### Anti-granulocyte Scintigraphy and Combined Techniques

Anti-granulocyte scintigraphy was introduced as an alternative for leukocyte scintigraphy with the advantage of in vivo labeling of leukocytes and was proposed as a promising



**Table 5** The intra- and inter-observer agreement of leukocyte scintigraphy and combined techniques in diagnosing PJI

Study	Year	Imaging	Tracer	Doses	Prostheses	Observers	Intra-OA	Inter-OA
Pelosi et al.	2004	LS	<sup>99m</sup> Tc-HMPAO	740 MBq	THA / TKA	3(1)	NR	$\kappa$ 0.55, 0.60, 0.74
Erba et al.	2014	LS(SPECT)	<sup>99m</sup> Tc-HMPAO	370–555 MBq	THA/TKA/TSA	2(1)	NR	DTC 11% and FT 19% disagreement <sup>2</sup>
El Esper et al.	2004	LS-BS	<sup>111</sup> In-oxine/ <sup>99m</sup> Tc-HDP	185 MBq/ 30 MBq	THA / TKA	3(1)	NR	$\kappa$ 0.40, 0.49, 0.55
El Esper et al.	2004	LS-BMS	<sup>111</sup> In-oxine/ <sup>99m</sup> Tc-SC <sup>1</sup>	185 MBq/ 650 MBq	THA / TKA	3(1)	NR	$\kappa$ 0.95, 0.95, 1.00
Fuster et al.	2008	LS-BMS	<sup>99m</sup> Tc-HMPAO/ <sup>99m</sup> Tc-SC <sup>1</sup>	185 MBq / nr	THA	3*	NR	$\kappa$ 0.74
Fuster et al.	2011	LS-BMS/CFS	<sup>99m</sup> Tc-HMPAO/ <sup>99m</sup> Tc-SC <sup>1</sup> /CFS	370 MBq	THA TKA	3* 3*	NR NR	$\kappa$ 0.40 $\kappa$ 0.20

1 = <sup>99</sup>Tc-sulfur colloid, 2 = with lower agreement for hip and knee prosthesis (also OM and TSP)

OA observer agreement, LS leukocyte scintigraphy, BMS bone marrow scintigraphy, CFS ciprofloxacin. <sup>99m</sup>Tc-DCP, <sup>99m</sup>Tc-dicarboxi-diphosphonate; <sup>99m</sup>Tc-HMPAO, <sup>99m</sup>Tc-hexamethylpropyleneamine oxime

(1)\*Three experienced nuclear medicine physicians, level of experience not reported. Level of experience and expertise was not defined

diagnostic tool for the detection of PJI. Image acquisition and interpretation are the same as for leukocyte scintigraphy. In recent meta-analyses, less satisfying diagnostic accuracies of anti-granulocyte scintigraphy in the assessment of THA compared with TKA were found [4, 5]. The included studies in this review did not report inter-observer agreement in the assessment of THA alone. For TKA, a wide variety of kappa values between “fair” to “substantial” agreement ( $\kappa = 0.31$  to  $> 0.75$ ) was reported. Interestingly, for the combined knee and hip group [27], only “fair” to “moderate” inter-observer agreement was found, while another study reported “excellent” reproducibility ( $\kappa = > 0.75$ ) for TKA alone [16]. Hence, one may suggest that variation in diagnostic accuracy outcomes between the knee and hip arthroplasties also applies to the level of reproducibility since

interpretation is less accurate for THA. However, there is no available evidence to support this assumption because symptomatic THA has not been separately analyzed. Another important result was found in a study that reported important differences between the inter-observer agreement of mild PJI ( $\kappa = 0.25$ ) versus severe PJI ( $\kappa = > 0.75$ ) [15]. This could implicate that the degree of infection directly influences the interpretation and reproducibility of the anti-granulocyte scintigraphy scan in the assessment of PJI of the knee. Unfortunately, none of the other included studies differentiated in the degree of infection. In accordance with combined leukocyte and bone marrow scintigraphy, one study found higher levels of inter-observer agreement for combined anti-granulocyte scintigraphy and bone marrow scintigraphy compared with anti-granulocyte scintigraphy alone

**Table 6** Leukocyte scintigraphy and combined techniques in diagnosing symptomatic hip and knee replacements, index test, and reference standard

Study	Year	Imaging	Prostheses	Reference standard (PJI)	Criteria
El Esper et al.	2004	LS-BS	THA/TKA	M	Positive; when LS activity at the ROI exceeded that in the adjacent and contralateral region and corresponded to an area of increased bone tracer uptake. Negative; when no focal hyperactivity was identified on LS at the ROI
		LS-BMS	THA/TKA	M	Positive; when increased activity was observed on LS at the ROI without corresponding uptake on BMS (incongruent patterns). Negative; no increased activity on LS at the ROI or when the increased activity on LS was spatially congruent with an increased activity on BMS
Pelosi et al.	2004	LS	THA/TKA	M, S	QA positive; positive when abnormal localized activity in the ROI increased in intensity or in extension in comparison with the contralateral region or with the ipsilateral adjacent bone segment. SQA positive; when K increased with time (Klate > Kearly by at least 10%)
Fuster et al.	2008	LS-BMS	THA	M, H	Any extramedullary periprosthetic focal uptake
Fuster et al.	2011	LS-BMS/CFS	THA/TKA/TSA	M, P	NR
Erba et al.	2014	LS(SPECT)	THA/TKA/TSA	M, H	At least one focus of abnormal uptake characterized by a time-dependent increase in radioactivity or an increase in size from the delayed to the late planar images (FT and DTC)
Total					

M = 2 or more positive peroperative samples that grew the same bacteria, ROI region of interest, QA qualitative analysis, SQA semiquantitative analysis, S surgery/peroperative findings, P macroscopic evidence of infection, H histopathology

**Table 7** The intra- and inter-observer agreement of anti-granulocyte scintigraphy and combined techniques in diagnosing PJI

Study	Year	Imaging	Tracer	Doses	Prostheses	Observers	Intra-observer agreement	Inter-observer agreement
Gratz et al.	2009	AGS	<sup>99m</sup> Tc-monoclonal Fab <sup>1</sup>	1110 ± 185 MBq	TKA	2(1)	NR	$\kappa > 0.75$ , 0.49, 0.25 (severe-mild) <sup>2</sup>
Sousa et al.	2011	AGS	<sup>99m</sup> Tc-sulesomab	20 mCi	THA/TKA	3(2)	NR	$\kappa$ 0.31, 0.50, 0.52 <sup>3</sup>
		AGS-BMS	<sup>99m</sup> Tc-rhenium sulfide colloid	20 mCi		3(2)	NR	$\kappa$ 0.50, 0.55, 0.72 <sup>3</sup>
Gratz et al.	2012	AGS	<sup>99m</sup> Tc-monoclonal Fab <sup>1</sup> / <sup>99m</sup> Tc-besilesomab <sup>2</sup>	555–925 MBq/700–900 MBq	TKA/UKA	2(1)	NR	$\kappa > 0.75$

1 = <sup>99m</sup>Tc-sulesomab. 2 = <sup>99m</sup>Tc-besilesomab, 2 = qualitative/visual evaluation. 3 = inter-observer agreement between the three different pairs of readers (1) = two nuclear medicine physicians, > 5-year-experience. (2) = two nuclear medicine physicians, one orthopedic surgeon; level of experiences was not reported

[27]. This supports the conclusion that the combination of these imaging modalities may improve interpretation and reproducibility.

### FDG-PET/CT

FDG-PET has assumed considerable potential in the assessment of periprosthetic infection. Important advantages of FDG-PET/CT, compared with planar scintigraphy, are its increased resolution, anatomic correlation with coregistered CT, and availability of simple visual interpretation criteria of images. In the assessment of PJI of the hip, in contrary to PJI of the knee [5], FDG-PET/CT is considered a promising imaging modality with assumed simple visual interpretation of uptake patterns. However, the visual interpretation of uptake patterns could potentially be hampered by inter-observer variation in the interpretation of imaging. One of the first studies on clinical application of FDG-PET/CT stated that the simplicity of the diagnostic criteria should reduce potential inter-observer variation and increase the accuracy of the scan [10]. Throughout literature, however, this has not been investigated in literature. Moreover, a variety of diagnostic criteria has

been used. Based on the best available evidence, it is assumed that extended uptake at the middle portion of the femoral bone-prosthesis interface is the most accurate criterion for PJI [29]. This criterion demonstrated “almost perfect” inter-observer agreement ( $\kappa = 0.85$ ) in one study [29]. An alternative criterion, increased uptake along the bone-prosthesis interface compared with the physiological uptake in the bladder, demonstrated only “moderate” inter-observer agreement [19]. Unfortunately, the literature search revealed no additional clinical data to support these findings and more clinical investigation is needed in order to validate these results.

### Limitations and Concerns

This systematic review depicts an overview of the currently available literature on the reproducibility of the various nuclear imaging modalities in the assessment of painful arthroplasties. This review demonstrated that the level of evidence is hampered by serious limitations.

There are several concerns regarding the methodological quality of the included studies. Most studies investigated the diagnostic accuracy and intra- and inter-observer agreement

**Table 8** Anti-granulocyte scintigraphy and combined techniques in diagnosing symptomatic hip and knee replacements, index test, and reference standard

Study	Year	Imaging	Prostheses	Reference standard (PJI)	Criteria
Gratz et al.	2009	AGS	TKP	M, S, H	Four-point scale as follows: 1, sulesomab uptake was similar to that in the bone marrow; 2, sulesomab uptake was increased minimally over that in bone marrow; 3, sulesomab uptake was distinctly higher than uptake in bone marrow; 4, sulesomab uptake was two or more times greater than uptake in bone marrow
Sousa et al.	2011	AGS-BMS	THP/TKP	M, H	Activity on the sulesomab scan without corresponding activity on the bone marrow scan (the images are spatially incongruent). Any other pattern is negative for infection
Gratz et al.	2012	AGS	TKP/UKP	M, S, H	1, uptake was similar to that in the bone marrow; 2, uptake was increased minimally over that in bone marrow; 3, uptake was distinctly higher than uptake in bone marrow; 4, uptake was two or more times greater than uptake in bone marrow
Total					

*M* intra-operative cultures, *H* intraoperative histological findings, *S* surgery/intra-operative findings

**Table 9** Intra- and inter-observer agreement of FDG-PET/CT for diagnosing PJI

Study	Year	Imaging	Tracer	Doses	Prostheses	Observers	Intra-observer agreement	Inter-observer agreement
Stumpe et al.	2004	FDG-PET/CT	<sup>18</sup> F-FDG	300–400 MBq	THP	2(1)	NR	$\kappa$ 0.47
Kumar et al.	2016	FDG-PET/CT Fluoride-PET/CT	<sup>18</sup> F-FDG <sup>18</sup> F-fluoride	0.14 mCi/Kg 0.07–0.14 mCi/Kg	THP	2*	NR	$\kappa$ 0.67
Verberne et al.	2018	FDG-PET/CT	<sup>18</sup> F-FDG	156–351 MBq	THP	2(2) 2(2) 2(2) 2(2)	NR NR NR NR	$\kappa$ 0.85 <sup>1</sup> $\kappa$ 0.81 <sup>1</sup> $\kappa$ 0.77 <sup>1</sup> $\kappa$ 0.79 <sup>1</sup>

1 = 4 different diagnostic criteria were investigated for inter-observer agreement, see Table 10

\*Level of experience and expertise was not defined. (1) = two nuclear medicine physicians, >4-year-experience. (2) = one senior nuclear medicine physician, one nuclear medicine trainee, level of experience not reported

was not the main research question. Ideally, the degree of agreement is investigated between multiple and different observers using the same imaging technique, classification, and procedure to assess the same subject. An important factor of influence to the reproducibility of the images is the level of experience of the observers. Interpretation of images by nuclear medicine physicians with or without experience in musculoskeletal imaging and the level of experience may directly influence the diagnostic accuracy. However, the level of experience and expertise was poorly or not described by 12 of the 16 included studies. Moreover, the number of observers is important to calculate the inter-observer agreement properly and avoid the possibility of agreement occurring by chance. Only 5 of the 16 included studies reported the results of more than 2 readers (Tables 3, 4, 5, 6, 7, 8, 9, and 10).

Recent meta-analyses demonstrated different diagnostic accuracies of nuclear imaging between hip and knee

arthroplasties. It is assumed that the inter-observer agreement could vary between the location of the prostheses, due to different uptake patterns and physiology [4, 5, 21]. In the interpretation of nuclear images, the type of implant (knee or hip) could be of influence to reproducibility, as demonstrated by several studies [7, 16, 27]. Therefore, the reported reproducibility concerning both arthroplasties without differentiation of type and location should be interpreted with caution. Unfortunately, 5 of the included studies did not differentiate between hip and knee arthroplasties. The degree of infection (chronic versus acute PJI) and the time between the implantation (last surgery) and performed imaging are also important factors in the interpretation of scans and were not described in most studies.

In the field of nuclear orthopedic imaging for painful prosthesis, concerns regarding the methodological quality and level of evidence have been previously been described [35]. This

**Table 10** FDG-PET/CT in diagnosing PJI, index test, and reference standard

Study	Year	Imaging	Prostheses	Reference standard (PJI)	Criteria
Stumpe et al.	2004	FDG-PET/CT	THP	M	Scale 0–5. uptake was low and comparable with that in inactive muscles and fat; a score of 2 that FDG uptake was moderate, clearly noticeable, and distinctly higher than the uptake in inactive muscles and fat; a score of 3 that FDG uptake was strong but was distinctly less than the physiologic uptake in the bladder; and a score of 4 that FDG uptake was very strong and was comparable with physiologic urinary uptake in the bladder
Kumar et al.	2016	FDG/Fluoride-PET/CT	THP	H, M, S	One or more of any three criteria: (1) periprosthetic FDG uptake around the femoral/ acetabular component along with FDG avid- associated periprosthetic soft tissue changes on CT images, (2) presence of sinus tract (3), and presence of regional lymph nodes or collection of pus
Verberne et al.	2018	FDG-PET/CT	THP	M, S, (MSIS)	(1) Increased FDG uptake along the femoral bone-prosthesis interface (2) Acetabulum uptake zone I–III (3) Periprosthetic uptake (4) Head zone (physiological)

M = 2 or more positive peroperative samples that grew the same bacteria, H histopathology, S surgery/peroperative findings, MSIS Musculoskeletal Infection Society for PJI

**Table 11** QUADAS-2, risk of bias

	Patient selection	Index test	Reference standard	Flow and timing
El Esper et al.	Unclear	Low	High	Unclear
Pelosi et al.	Low	Low	Low	Unclear
Stumpe et al.	Low	Low	Low	Low
Temmerman et al.*	Low	Low	Low	Low
Temmerman et al.*	Low	Low	Low	Unclear
Fuster et al.	Low	Low	Low	Unclear
Gratz et al.	Unclear	High	Low	Unclear
Fuster et al.	Unclear	High	Low	Unclear
Sousa et al.	Low	Low	Low	Unclear
Gratz et al.	Low	Unclear	Low	Unclear
Erba et al.	Unclear	Unclear	Low	Unclear
Arican et al.*	Low	Low	Low	Unclear
Granados et al.	Low	Low	Low	Low
Yoldas et al.	Low	Unclear	Low	Unclear
Kumar et al.	Low	Low	Low	Low
Verberne et al.	Unclear	Low	Low	Unclear

\*Study which investigated aseptic loosening

systematic review depicts that the interpretation and reproducibility of nuclear imaging in the assessment of symptomatic joint replacements, an important aspect of clinical usability, have thus far been underreported. Moreover, it is important to note in general that most studies reported the consensus results between two readers, and therefore, analyzing the intra- and inter-observer agreement was not possible. Further investigation should report implant characteristics, time interval

between imaging and surgery, and degree of infection, and separately analyze hip and knee arthroplasties.

### Key Points of the Best Available Evidence According to This Systematic Review

Literature provides insufficient data regarding for intra-observer agreement when using nuclear imaging in the assessment of symptomatic joint replacements. Bone scintigraphy demonstrated a wide range of inter-observer agreement. The comparison of images, as for combined leukocyte and bone marrow scintigraphy, could improve the level of inter-observer agreement. The infection grade (mild versus severe) could directly influence the accuracy of interpretation. The interpretation and level of reproducibility for nuclear imaging could be different between hip and knee arthroplasties. Although it is plausible that hybrid imaging improves interpretation for symptomatic joint replacements and could improve the level of inter-observer agreement, there is limited data to support this assumption. The results of this systematic review should be interpreted with caution because of methodological concerns of the included studies with high risk of bias (Tables 11 and 12). Important variables that may influence the level of reproducibility as type of implant, time interval between imaging and surgery, level of expertise of the observers, and degree of infection were underreported. Further studies investigating the diagnostic performances of nuclear imaging in symptomatic orthopedic implants should report intra- and inter-observer agreement and separately analyze hip and knee arthroplasties.

**Table 12** QUADAS-2, applicability

	Patient selection	Index test	Reference standard
El Esper et al.	Unclear	Low	High
Pelosi et al.	Low	Low	Low
Stumpe et al.	Low	Low	Low
Temmerman et al.*	Low	Low	Low
Temmerman et al.*	Low	Low	Low
Fuster et al.	Low	Low	Low
Gratz et al.	High	High	Low
Fuster et al.	Low	High	Low
Sousa et al.	Unclear	High	Low
Gratz et al.	High	Unclear	Unclear
Erba et al.	High	Unclear	Unclear
Arican et al.*	Low	Low	Unclear
Granados et al.	Low	Low	Low
Yoldas et al.	Unclear	Low	Low
Kumar et al.	Low	High	Low
Verberne et al.	Low	Low	Low

\*Study which investigated aseptic loosening

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

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

## References

- 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(1):e1–e25. <https://doi.org/10.1093/cid/cis803>.
- Parvizi J, Tan TL, Goswami K, et al. The 2018 definition of periprosthetic hip and knee infection: an evidence-based and validated criteria. *J Arthroplast*. 2018;33(5):1309–1314.e2. (In eng). <https://doi.org/10.1016/j.arth.2018.02.078>.
- Verberne S, Temmerman O. Imaging of prosthetic joint infections. *Management of Periprosthetic Joint Infections (PJIs) 2016*:259.
- Verberne SJ, Raijmakers PG, Temmerman OP. The accuracy of imaging techniques in the assessment of periprosthetic hip infection: a systematic review and meta-analysis. *J Bone Joint Surg Am*. 2016;98(19):1638–45. (In eng). <https://doi.org/10.2106/jbjs.15.00898>.
- Verberne SJ, Sonnega RJ, Temmerman OP, Raijmakers PG. What is the accuracy of nuclear imaging in the assessment of periprosthetic knee infection? A Meta-analysis *Clin Orthop Relat Res*. 2017;475(5):1395–410. (In eng). <https://doi.org/10.1007/s11999-016-5218-0>.
- Temmerman OP, Raijmakers PG, Berkhof J, Hoekstra OS, Teule GJ, Heyligers IC. Accuracy of diagnostic imaging techniques in the diagnosis of aseptic loosening of the femoral component of a hip prosthesis: a meta-analysis. *J Bone Joint Surg (Br)*. 2005;87(6):781–5. <https://doi.org/10.1302/0301-620X.87B6.15625>.
- Arıcan P, Okudan Tekin B, Sefizade R, Naldoken S, Bastug A, Ozkurt B. The role of bone SPECT/CT in the evaluation of painful joint prostheses. *Nucl Med Commun*. 2015;36(9):931–40. (In eng). <https://doi.org/10.1097/mnm.0000000000000348>.
- Barnsley L, Barnsley L. Detection of aseptic loosening in total knee replacements: a systematic review and meta-analysis. *Skelet Radiol*. 2019;48(10):1565–72. <https://doi.org/10.1007/s00256-019-03215-y>.
- Basu S, Kwee TC, Saboury B, Garino JP, Nelson CL, Zhuang H, 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(7):609–15. <https://doi.org/10.1097/RLU.0000000000000464>.
- Zhuang H, Duarte PS, Pourdehnad M, Maes A, van Acker F, Shnier D, et al. The promising role of 18F-FDG PET in detecting infected lower limb prosthesis implants. *J Nucl Med*. 2001;42(1):44–8 (In eng).
- Horger M, Eschmann SM, Pfannenber C, Storek D, Dammann F, Vonthein R, et al. The value of SPET/CT in chronic osteomyelitis. *Eur J Nucl Med Mol Imaging*. 2003;30(12):1665–73. (In eng). <https://doi.org/10.1007/s00259-003-1321-z>.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Parvizi J, Zmistowski B, Berbari EF, Bauer TW, Springer BD, Della Valle CJ, et al. New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society. *Clin Orthop Relat Res*. 2011;469(11):2992–4. (In Eng). <https://doi.org/10.1007/s11999-011-2102-9>.
- 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(8):529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.
- Gratz S, Behr TM, Reize P, Pfestroff A, Kampen WU, Hoffken H. (99m)Tc-Fab' fragments (sulesomab) for imaging septicallly loosened total knee arthroplasty. *The Journal of international medical research*. 2009;37(1):54–67. (In eng). <https://doi.org/10.1177/147323000903700107>.
- Gratz S, Reize P, Pfestroff A, Hoffken H. Intact versus fragmented 99mTc-monoclonal antibody imaging of infection in patients with septicallly loosened total knee arthroplasty. *The Journal of international medical research*. 2012;40(4):1335–42. (In eng). <https://doi.org/10.1177/147323001204000412>.
- Temmerman OP, Raijmakers PG, David EF, et al. A comparison of radiographic and scintigraphic techniques to assess aseptic loosening of the acetabular component in a total hip replacement. *J Bone Joint Surg Am*. 2004;86-a(11):2456–63 (In eng).
- Temmerman OP, Raijmakers PG, Berkhof J, et al. Diagnostic accuracy and interobserver variability of plain radiography, subtraction arthrography, nuclear arthrography, and bone scintigraphy in the assessment of aseptic femoral component loosening. *Arch Orthop Trauma Surg*. 2006;126(5):316–23. (In eng). <https://doi.org/10.1007/s00402-006-0120-y>.
- Stumpe KD, Notzli 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(2):333–41. <https://doi.org/10.1148/radiol.2312021596>.
- Granados U, Fuster D, Soriano A, et al. Screening with angiographic images prior to (99m)Tc-HMPAO labelled leukocyte scintigraphy in the diagnosis of periprosthetic infection. *Rev Esp Med Nucl Imagen Mol*. 2015;34(4):219–24. <https://doi.org/10.1016/j.remnm.2014.10.003>.
- Yoldas B, Cankaya D, Andic K, et al. Higher reliability of triple-phase bone scintigraphy in cementless total hip arthroplasty compared to cementless bipolar hemiarthroplasty. *Ann Med Surg* (2012). 2016;10:27–31. (In eng). <https://doi.org/10.1016/j.amsu.2016.07.022>.
- Erba PA, Glaudemans AW, Veltman NC, et al. Image acquisition and interpretation criteria for 99mTc-HMPAO-labelled white blood cell scintigraphy: results of a multicentre study. *Eur J Nucl Med Mol Imaging*. 2014;41(4):615–23. (In eng). <https://doi.org/10.1007/s00259-013-2631-4>.
- Pelosi E, Baiocco C, Pennone M, Migliaretti G, Varetto T, Maiello A, et al. 99mTc-HMPAO-leukocyte scintigraphy in patients with symptomatic total hip or knee arthroplasty: improved diagnostic accuracy by means of semiquantitative evaluation. *J Nucl Med*. 2004;45(3):438–44 (In eng).
- El Espera I, Blondet C, Moullart V, et al. The usefulness of 99mTc sulfur colloid bone marrow scintigraphy combined with 111In

- leucocyte scintigraphy in prosthetic joint infection. *Nucl Med Commun.* 2004;25(2):171–5 (In eng).
25. Fuster D, Soriano A, Garcia S, Piera C, Suades J, Rodríguez D, et al. Usefulness of <sup>99m</sup>Tc-ciprofloxacin scintigraphy in the diagnosis of prosthetic joint infections. *Nucl Med Commun.* 2011;32(1):44–51. (In eng). <https://doi.org/10.1097/MNM.0b013e328340e6fb>.
  26. Fuster D, Duch J, Soriano A, García S, Setoain X, Bori G, et al. Potential use of bone marrow scintigraphy in suspected prosthetic hip infection evaluated with <sup>99m</sup>Tc-HMPAO-leukocytes. *Revista española de medicina nuclear.* 2008;27(6):430–5 (In spa).
  27. Sousa R, Massada M, Pereira A, Fontes F, Amorim I, Oliveira A. Diagnostic accuracy of combined <sup>99m</sup>Tc-sulesomab and <sup>99m</sup>Tc-nanocolloid bone marrow imaging in detecting prosthetic joint infection. *Nucl Med Commun.* 2011;32(9):834–9. (In eng). <https://doi.org/10.1097/MNM.0b013e3283496695>.
  28. 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. *Journal of medical imaging and radiation oncology.* 2016;60(3):315–22. (In eng). <https://doi.org/10.1111/1754-9485.12444>.
  29. Verberne SJ, Temmerman OPP, Vuong BH, Raijmakers PG. Fluorodeoxyglucose positron emission tomography imaging for diagnosing periprosthetic hip infection: the importance of diagnostic criteria. *Int Orthop.* 2018;42(9):2025–34. (In eng). <https://doi.org/10.1007/s00264-018-3931-x>.
  30. Barthassat E, Afifi F, Konala P, Rasch H, Hirschmann MT. Evaluation of patients with painful total hip arthroplasty using combined single photon emission tomography and conventional computerized tomography (SPECT/CT) - a comparison of semi-quantitative versus 3D volumetric quantitative measurements. *BMC Med Imaging.* 2017;17(1):31. (In eng). <https://doi.org/10.1186/s12880-017-0204-x>.
  31. Hirschmann MT, Iranpour F, Konala P, Kerner A, Rasch H, Cobb JP, et al. A novel standardized algorithm for evaluating patients with painful total knee arthroplasty using combined single photon emission tomography and conventional computerized tomography. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2010;18(7):939–44. (In eng). <https://doi.org/10.1007/s00167-010-1070-z>.
  32. Rasch H, Falkowski AL, Forrer F, Henckel J, Hirschmann MT. 4D-SPECT/CT in orthopaedics: a new method of combined quantitative volumetric 3D analysis of SPECT/CT tracer uptake and component position measurements in patients after total knee arthroplasty. *Skelet Radiol.* 2013;42(9):1215–23. (In eng). <https://doi.org/10.1007/s00256-013-1643-2>.
  33. Suter B, Testa E, Stampfli P, et al. A novel standardized algorithm using SPECT/CT evaluating unhappy patients after unicondylar knee arthroplasty—a combined analysis of tracer uptake distribution and component position. *BMC Med Imaging.* 2015;15:11. (In eng). <https://doi.org/10.1186/s12880-015-0053-4>.
  34. Palestro CJ, Love C, Tronco GG, Tomas MB, Rini JN. Combined labeled leukocyte and technetium <sup>99m</sup> sulfur colloid bone marrow imaging for diagnosing musculoskeletal infection. *Radiographics.* 2006;26(3):859–70. <https://doi.org/10.1148/rg.263055139>.
  35. Glaudemans AW, Jutte PC, Petrosillo N, Erba PA, Lazzeri E, Signore A. Comment on: “diagnosis of periprosthetic joint infection: the role of nuclear medicine may be overestimated” by Claudio Diaz-Ledezma, Courtney Lamberton, Paul Lichtstein and Javad Parvizi. *J Arthroplast.* 2016;31(2):551–2. (In eng). <https://doi.org/10.1016/j.arth.2015.07.002>.

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