



Diagnostic performance of [¹⁸F]FDG-PET/MRI for liver metastasis in patients with primary malignancy: a systematic review and meta-analysis

Seung Baek Hong^{1,2} · Sang Hyun Choi¹ · Kyung Won Kim¹ · Seong Ho Park¹ · So Yeon Kim¹ · So Jung Lee¹ · Seung Soo Lee¹ · Jae Ho Byun¹ · Moon-Gyu Lee¹

Received: 11 July 2018 / Revised: 18 October 2018 / Accepted: 22 November 2018 / Published online: 4 February 2019
© European Society of Radiology 2019

Abstract

Objectives To systematically determine the diagnostic accuracy of 18-fluorodeoxyglucose positron emission tomography/magnetic resonance imaging ([¹⁸F]FDG-PET/MRI) for the detection of liver metastases and evaluate the sources of heterogeneity in the reported results.

Methods PubMed and EMBASE databases were searched up until December 31, 2017, to identify original research studies reporting the diagnostic performance (Se and Sp) of PET/MRI for liver metastases, in comparison with PET/CT. Study quality was assessed using QUADAS-2. The summary Se and Sp of the studies were estimated using hierarchical modeling methods. To determine causes of study heterogeneity, the presence of a threshold effect was analyzed, and meta-regression analysis was performed.

Results Of 546 articles screened, eight suitable articles were identified, with seven for per-lesion analysis, and four for per-patient analysis. The meta-analytic summary Se and Sp for per-patient-based analysis were 99.2% (95% CI, 31.4–100.0%, $I^2 = 89.4%$) and 98.6% (95% CI, 84.0–99.9%, $I^2 = 0.0%$), respectively, while for per-lesion-based analysis they were 95.4% (95% CI, 78.3–99.2%, $I^2 = 99.7%$) and 99.3% (95% CI, 93.8–99.9%, $I^2 = 96.5%$). PET/MRI showed higher Se (95.4% vs. 68.3%) and Sp (99.3% vs. 95.8%) than PET/CT. Meta-regression analysis showed five significant factors affecting study heterogeneity: study subject characteristics, study design, MRI technique (DWI, HBP after injection of liver-specific contrast media), imaging review method, and reference standard.

Conclusion The diagnostic accuracy of [¹⁸F]FDG-PET/MRI for liver metastasis was high overall, but substantial heterogeneity was found. Further randomized controlled studies or prospective studies are needed to investigate the role of PET/MRI in liver metastasis in comparison with PET/CT.

Key Points

- [¹⁸F]FDG-PET/MRI has high meta-analytic Se and Sp for the diagnosis of liver metastasis.
- PET/MRI using DWI and HBP images significantly increased diagnostic accuracy.
- Study heterogeneity was associated with subject characteristics, study design, MRI technique, image review method, and reference standard.

Keywords Fluorodeoxyglucose F18 · Positron emission tomography-computed tomography · Magnetic resonance imaging · Neoplasms · Liver

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-018-5909-x>) contains supplementary material, which is available to authorized users.

✉ Sang Hyun Choi
edwardchoi83@gmail.com

² Department of Radiology, Pusan National University of Hospital, Busan, South Korea

¹ Department of Radiology and the Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, 88 Olympic-Ro 43-Gil, Songpa-Gu, Seoul 05505, South Korea

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
CI	Confidence intervals
DWI	Diffusion-weighted imaging
FN	False negatives
FP	False positives
HBP	Hepatobiliary phase
HSROC	Hierarchical summary receiver operating characteristic
MRI	Magnetic resonance imaging
PET	Positron emission tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
Se	Sensitivity
Sp	Specificity
TN	True negatives
TP	True positives

Introduction

The liver is one of the most common sites for metastatic disease, accounting for 25% of all metastases to solid organs [1]. As modern chemotherapy, interventional therapy and resection of liver metastases may increase survival [2, 3], accurate diagnosis, and staging of liver metastases is important for treatment planning and outcome improvement.

Many imaging tests have been used for the diagnosis of liver metastasis, including computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). Although CT is the mainstay for the staging of liver metastasis because of its easy availability and speed, it has several limitations, such as ionizing radiation, nephrotoxicity of the contrast media, and inability to characterize sub-centimeter lesions [4, 5]. By contrast, MRI can detect and characterize sub-centimeter lesions using hepatocyte-specific contrast media without subjecting the patient to ionizing radiation [6–8]. PET can give additional information for determining management strategy when CT or MRI is equivocal for the presence of metastasis, and can also play an important role in detecting postoperative recurrence [9]. However, the metabolic activity measured in PET is not specific to metastasis and may result in false-positive studies [5].

Recently, combined 18-fluorodeoxyglucose PET (^{18}F FDG-PET)/MRI has been introduced to overcome these limitations of the current imaging tests. As PET/MRI combines the advantages of PET (metabolic information) and MRI (precise anatomy, multiparametric signal) [10], its diagnostic sensitivity and specificity for liver metastasis might be enhanced compared with either modality alone. The first study

on PET/MRI used a fusion of PET and MRI data acquired separately [11]; however, fully integrated PET/MRI systems allow simultaneous data acquisition. As such systems are now available for clinical practice, several studies have reported on the diagnostic accuracy of PET/MRI for the diagnosis of liver metastasis [11–18]. These studies generally agree on the potential of PET/MRI as a promising imaging method for the diagnosis of liver metastasis; however, the reported accuracies are variable, with Se ranging from 48 to 100% and Sp from 71 to 100% [11–18]. In addition, as some studies included only a small number of subjects, their results may have limited generalizability [11, 15, 18]. Therefore, we considered it timely and important to systematically determine the diagnostic performance of ^{18}F FDG-PET/MRI in the diagnosis of liver metastasis in comparison with ^{18}F FDG-PET/CT. We used hierarchical modeling methods to understand the sources of study heterogeneity between studies.

Materials and methods

This study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for conducting and reporting meta-analyses, and the analysis was performed using methods advocated by the Diagnostic Test Accuracy Working Group of the Cochrane Collaboration [19] and the Agency for Healthcare Research and Quality (AHRQ) [20].

Literature search strategy

A comprehensive search of PubMed and EMBASE databases was conducted. The search query was designed with maximum Se, with the following MeSH terms being used as keywords: (“Positron-emission tomography” OR “PET”) AND (“MR” OR “MRI” OR “magnetic resonance imaging” OR “magnetic resonance”) AND (“metastases” OR “metastasis”) AND (“liver” OR “hepatic”) AND (accuracy OR Se OR Sp OR “diagnostic performance”). The search included articles published up until December 31, 2017, with no start date limits. The studies were confined to those published in the English language and concerned with human patients. The identified articles were managed using Endnote version X7 (Thomson Reuters).

Eligibility criteria

After removal of duplicate articles, one author reviewed the identified articles to determine their eligibility according to the following criteria: (a) patients who had suspected liver metastasis from any primary malignancy; (b) index test, ^{18}F FDG-PET/MRI with or without contrast-enhanced study; (c) comparisons with ^{18}F FDG-PET/CT; (d) outcomes, diagnostic

accuracy including Se and Sp for the dichotomous diagnosis of liver metastasis; (e) reference standard, pathological diagnosis (surgical or biopsy specimens), or imaging follow-up results (interval growth or stability); and (f) study design, both observational studies (retrospective or prospective), and clinical trials. All study subjects presented malignancies that primarily metastasized to the liver. All studies that performed [18F]FDG-PET/MRI were included without any restrictions. Studies with a retrospective fusion of PET and MRI were also included as well as studies with combined PET/MRI devices. PET/MRI studies with no comparison with PET/CT were included but were separated from those with comparative data. If data were from various organ metastases, the data on liver metastasis were separated from those of other metastases. Any articles with the inseparable article were not included. Studies were excluded according to the following criteria: (a) studies with insufficient information to allow construction of a diagnostic 2×2 table for clinical and imaging results; (b) review articles, letters, comments, editorials, conference abstracts, and case reports; (c) studies not in the field of interest; and (d) studies with duplicated patients and data. If duplicate publications used the same cohort of patients, the data from the most recent article were included.

The retrieved articles were first screened according to their abstracts and titles, and then full-text of potentially eligible articles was reviewed. Both review sessions (abstracts and full-text) were performed by two reviewers (with 6 and 2 years of experience in the meta-analysis) working independently. Throughout the review sessions, only articles with definite ineligibility were excluded. In the screening of abstract, a disagreement rate of 4.0% was noted as to whether articles were within the field of interest, while in the full-text screening a disagreement rate of 9.5% was noted over whether data could be extracted for 2×2 table to determine diagnostic accuracy. Disagreements between the two reviewers were resolved at a consensus meeting in the presence of a third reviewer.

Data extraction

The following data were extracted to a predefined form: (a) study characteristics (authors, year of publication, and study design); (b) study population characteristics (patient number, age, and sex); (c) lesion characteristics (lesion number, lesion size, and primary cancer origin); (d) imaging techniques (the use of contrast-enhanced images, DWI, and hepatobiliary phase [HBP] images and the amount of FDG); (e) methods to interpret images including the presence of predefined imaging criteria for metastasis on each imaging test and the method of image review, i.e., consensus review or independent review; (f) details of the reference standard; (g) image fusion methods (retrospective or simultaneous); (h) study outcomes, i.e., diagnostic performance of the dichotomous

diagnosis of liver metastasis. To determine the diagnostic accuracy, the exact numbers of true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) for liver metastasis on PET/MRI and PET/CT were extracted. When insufficient data were present, the authors were emailed and the supplementary information was requested. Two reviewers independently performed data extraction, and all discrepancies were resolved at a consensus meeting in the presence of a third reviewer.

Assessment of study quality

Using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) criteria, two reviewers independently assessed the quality of the eligible articles, including the risk of bias and the applicability of each study. All discrepancies were resolved by arriving at a consensus between the two reviewers and a third reviewer.

Data synthesis and statistical analysis

Accuracy of liver metastasis diagnosis

Descriptive analysis The relevant data were recorded and analyzed to perform per-patient- and per-lesion-based analysis of the diagnostic performance of PET/MRI and PET/CT for liver metastasis. The Se and Sp for the diagnosis of liver metastasis and 95% confidence intervals (CIs) were calculated using the generic inverse variance method for calculating weights. The Higgins I^2 statistic was used to evaluate the presence of heterogeneity between studies with respect to the Se and Sp ($I^2 > 50\%$: substantial heterogeneity). When substantial heterogeneity was noted, the presence of a threshold effect was analyzed by visual assessment of coupled forest plots of Se and Sp, as well as by calculating the Spearman correlation coefficient between the Se and FP rate. A correlation coefficient > 0.6 was considered to indicate a considerable threshold effect.

Meta-analytic summary A meta-analytic summary was calculated using hierarchical modeling methods. The summary Se and Sp and their 95% CIs were acquired using a bivariate random effects model. The summary receiver-operating characteristic curve was obtained using a hierarchical summary receiver-operating characteristic (HSROC) model. For available articles, we performed subgroup analyses of the diagnostic accuracy of PET/MRI for diagnosing liver metastasis according to primary malignancy tumor type and the lesion size (≤ 1 cm versus > 1 cm).

Comparative analysis with PET/CT Those PET/MRI studies making a comparison with PET/CT were included in a comparative analysis. The meta-analytic summary diagnostic

accuracies of PET/MRI and PET/CT were compared using 95% CIs and meta-regression analytical methods. For available articles, we performed subgroup comparative analyses between PET/MRI and PET/CT according to the primary malignancy tumor type and the lesion size (≤ 1 cm versus > 1 cm).

Meta-regression analysis Meta-regression analysis was performed to identify further causes of study heterogeneity, with the following covariates being considered: (a) study subjects (one type of primary cancer vs. various types of primary cancers); (b) study design (prospective vs. retrospective); (c) PET/MRI data acquisition method (simultaneous fusion vs. retrospective fusion); (d) MRI technique (acquisition with DWI vs. acquisition without DWI; acquisition with HBP image vs. acquisition without HBP image); (e) imaging review method (consensus review or independent review); and (g) reference standard (reference standard determined using pathology in $\geq 50\%$ subjects vs. in $< 50\%$ subjects).

Analysis of publication bias A Deeks' funnel plot was visually evaluated to determine any publication bias, with its statistical significance being examined using Deeks' asymmetry test.

Statistical analysis was performed with Stata version 15.0 (StataCorp LP). A value of $p < 0.05$ was considered statistically significant.

Results

Literature search

A total of 546 articles were screened after the exclusion of duplicate articles. Of these, we removed 504 articles according to their titles and abstracts. In addition, a further 34 articles were excluded after a full-text review (Fig. 1). Of the remaining eight articles [1, 2, 8, 10, 12, 14, 17, 21], three reported the diagnostic performance of PET/MRI for diagnosing liver metastasis on both a per-patient and per-lesion basis [11, 12, 14], and one reported it only on a per-patient basis [13]. Therefore, four articles were available for per-patient analysis [11–14]. In addition, of the eight articles, after excluding one article that only reported per-patient results, seven articles were available for per-lesion analysis [11, 12, 14–18]. Of the seven articles for the per-lesion analysis, five reported diagnostic results for both PET/CT and PET/MRI [11, 12, 14, 15, 18], but two only covered PET/MRI [16, 17].

Of the eight articles, seven clearly defined test positivity criteria on PET/MRI (Supplementary Table 1) [11–17], with six using a 3-point or 5-point diagnostic confidence scale for test positivity [11–15, 17]. Seven articles qualitatively assessed FDG uptake of the focal lesion [11, 12, 14–18] and only one article quantitatively assessed FDG uptake [13] to determine test positivity.

Article characteristics and quality according to QUADAS-2

Table 1 summarizes the characteristics of the individual studies. Of the eight articles, the study subjects in three articles had only one type of primary cancer [13, 15, 18]. Four articles were prospective studies [12, 14, 17, 18], and four were retrospective studies [11, 13, 15, 16]. Most articles used DWI-MRI with PET [12–15, 17, 18], obtaining simultaneous imaging fusion [12, 13, 15, 17, 18]. Pathology and imaging follow-up were used together as a reference standard in five articles [12–15, 17], with two articles using only imaging follow-up [16, 18], and one article using only pathology [11].

The quality of the included articles is summarized in Supplementary Fig. 1. The lack of external reference standard and the unclearness of blinding between the index test and reference standard were notable areas of quality concern. All eight studies were unclear as to whether interpret the reference standard results were interpreted without knowledge of the results of the index test. In addition, five studies were unclear as to whether interpret the index test results were interpreted without knowledge of the results of the reference standard.

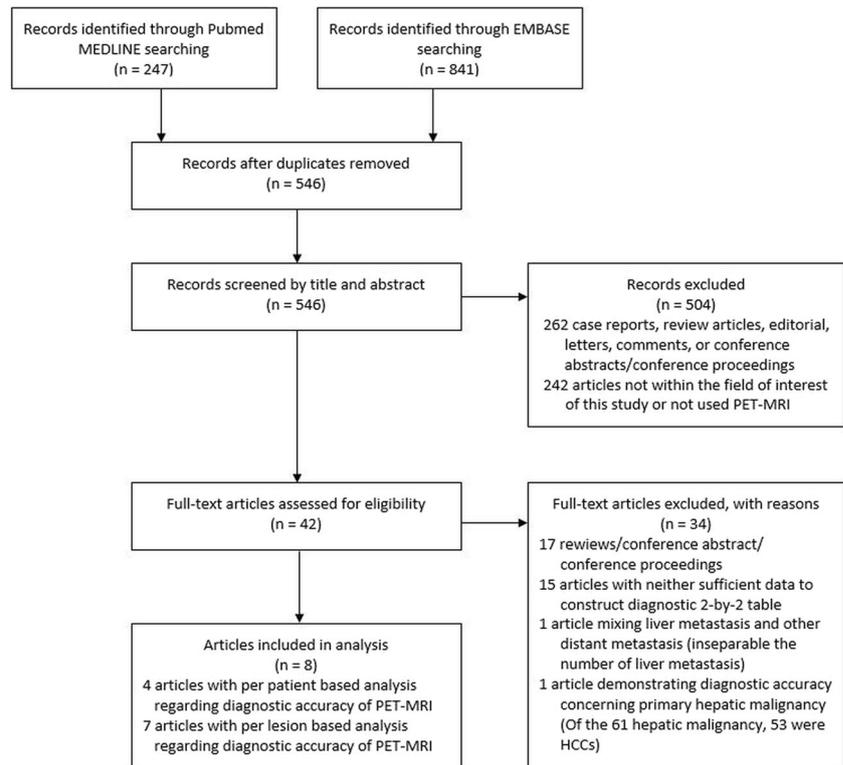
Diagnostic accuracy of PET-MRI for liver metastasis

As the per-patient-based and per-lesion-based results were considered as separate sets of results for the analyses, we analyzed a total of 11 sets of results from eight articles (four studies for the per-patient analysis and seven studies for the per-lesion analysis) in the meta-analysis of the diagnostic performance of PET/MRI for liver metastasis. The results of the meta-analysis for the per-patient and per-lesion analyses are summarized in Table 2.

In the per-patient-based analysis, the sensitivities and specificities of individual studies were 85–100% and 96–100%, respectively. Substantial study heterogeneity was noted with regard to Se ($I^2 = 89.4\%$), but there was no study heterogeneity with regard to Sp ($I^2 = 0.0\%$). The coupled forest plot of Se and Sp did not show an apparent threshold effect (Fig. 2a). The Spearman correlation between Se and FP rate was 0.333 ($p = 0.667$), also indicating the lack of a threshold effect. The meta-analytic summary Se and Sp were 99.2% (95% CI, 31.4–100.0%) and 98.6% (95% CI, 84.0–99.9%), respectively.

In the per-lesion-based analysis, the sensitivities and specificities of individual studies were 48–100% and 80–100%, respectively. Substantial study heterogeneity was noted with regard to both Se ($I^2 = 99.7\%$) and Sp ($I^2 = 96.5\%$). The coupled forest plot of Se and Sp did not reveal an apparent threshold effect (Fig. 2b), and the Spearman correlation between Se and FP rate was -0.249 ($p = 0.591$), indicating the lack of a threshold effect. The meta-analytic summary Se and Sp were 95.4% (95% CI, 78.3–99.2%) and 99.3% (95% CI,

Fig. 1 Flow diagram of study selection for the systematic review and meta-analysis



93.8–99.9%), respectively. The HSROC curve showed a large difference between the 95% confidence and prediction regions, thus indicating considerable heterogeneity between the studies (Fig. 3).

Subgroup analysis

Regarding the tumor type of primary malignancy, three articles reported the diagnostic accuracy of PET/MRI in each primary malignancy [13, 15, 18], but five other articles did not report such information [11, 12, 14, 16, 17]. In patients with colorectal cancer, Se and Sp of PET/MRI were 85.4 and 100.0% on per-patient analysis, and 72.7 and 80.0% on per-lesion analysis. In patients with breast cancer, Se and Sp of PET/MRI were 100.0 and 97.1% on per-lesion analysis.

Regarding the lesion size, three articles reported a subgroup analysis of the diagnostic accuracy of PET/MRI [11, 13, 16]. Lesions 1 cm in diameter or smaller showed lower Se and Sp on PET/MRI than lesions larger than 1 cm in diameter (44–71 vs. 50–98% for Se; 71–100 vs. 94–100% for Sp).

Comparative analysis with PET/CT

We included two studies with per-patient analysis and five studies with the per-lesion analysis in the meta-analysis of the diagnostic performance of PET/CT for liver metastasis. In the per-patient-based analysis, the sensitivities and specificities of individual studies were 71–86 and 96–100%,

respectively. Due to the small number of included studies for per-patient analysis, meta-analytic summary Se and Sp could not be calculated. In the per-lesion-based analysis, the sensitivities and specificities of individual studies were 32–76 and 90–98%, respectively. Substantial study heterogeneity was noted in Se ($I^2 = 77.3\%$), but no study heterogeneity was noted in Sp ($I^2 = 0.0\%$).

The meta-analytic summary Se and Sp of PET/CT were 68.3 (95% CI, 54.7–79.4%) and 95.7% (95% CI, 91.4–97.9%), respectively (Supplementary Table 2). PET/CT showed a significantly lower Se than PET/MRI (68.3% vs. 95.4%, $p = 0.02$) and a lower Sp than PET/MRI without a significant difference (95.8 vs. 99.4%, $p = 0.10$).

In patients with colorectal cancer, Se of PET/CT was lower than that of PET/MRI (70.8 vs. 85.4% on per-patient analysis and 31.8 vs. 72.7% on per-lesion analysis), while Sp of PET/CT was the same or higher than that of PET/MRI (100.0% for both tests on per-patient analysis and 93.3 vs. 80.0% on per-lesion analysis). In patients with breast cancer, Se of PET/CT was lower than that of PET/MRI (75.0 vs. 100.0%), while Sp of PET/CT was the same as that of PET/MRI (97.1% for both tests).

In lesions with 1 cm in diameter or smaller, PET/CT showed a lower sensitivity than PET/MRI (29–30 vs. 44–71%), but a similar specificity (86 vs. 71–100%). In lesions larger than 1 cm in diameter, PET/CT showed similar sensitivity (87–89 vs. 50–98%) and specificity (94 vs. 94–100%) to PET/MRI.

Table 1 Characteristics of the articles included

Author (year of publication)	Study design	Study type	Patient age, year (range)	No. of patients	No. of tumor	Tumor size (range, cm)	Primary tumor origin	Comparison with PET/CT	MRI magnet	MRI technique-CE	MRI DWI (b, s/mm ²)	FDG amount	PET/MRI fusion	Imaging review method	Reference standards
Beidenwellen et al (2015) [12]	Prospective Cohort	Cohort	18 (N.A.)	32	45	14 ± 8 mm*	Melanoma (n = 7), breast (n = 7), colon (n = 4), others (n = 14)	Per-lesion analysis	1.5 T	CE without HBP: 0.1 mmol/kg gadobutrol	b = 0, 500, 1000	286 ± 45 MBq	Simultaneous	Independent review	Pathology and imaging follow-up
Brendle et al (2016) [15]	Retrospective Cohort	Cohort	N.A. (10–62)	N.S.	22	N.A.	Colon	Per-lesion analysis	3.0 T	No CE	b = 50, 800	337 ± 59 MBq	Simultaneous	Independent review	Pathology and imaging follow-up
Donati et al (2010) [11]	Retrospective Cohort	Cohort	60.2 (35–82)	37	55	< 2 cm (n = 32) ≥ 2 cm (n = 23)	Colon (n = 20), pancreas (n = 11), others (n = 6)	Per-lesion analysis	1.5 T	CE with HBP: 0.025 mmol/kg gadoxetic acid	No DWI	370–400 MBq	Retrospective fusion	Independent review	Pathology only
Heusner et al (2012) [16]	Retrospective Cohort	Cohort	56.1 (35–71)	43	483	N.A.	Colon (n = 14), lung (n = 6), breast (n = 9), genitourinary (n = 4), others (n = 10)	N.A.	1.5 T	CE without HBP: 0.1 mmol/kg gadobutrol or gadoteric acid	No DWI	< 150 mg/dl	Retrospective fusion	Consensus review	Imaging follow-up only
Kirchner et al (2017) [17]	Prospective Cohort	Cohort	55.9 (29–79)	41	57	N.A.	Colon (n = 19), liver† (n = 7), melanoma (n = 4), others (n = 11)	N.A.	3.0 T	CE with HBP: 0.05 mmol/kg gadobenate dimeglumine	b = 0, 500, 1000	244 ± 53 MBq	Simultaneous	Consensus review	Pathology and imaging follow-up
Lee et al (2016) [13]	Retrospective Cohort	Cohort	62.9 (38–80)	48	98	1.28 (0.2–9.5)	Colon	Per-patient analysis	3.0 T	CE with HBP: 0.025 mmol/kg gadoxetic acid	b = 0, 100, 500, 1000	290 ± 40 MBq	Simultaneous	Independent review	Pathology and imaging follow-up
Melsaether et al (2016) [18]	Prospective Cohort	Cohort	N.A.	N.S.	40	N.A.	Breast	Per-lesion analysis	3.0 T	CE without HBP: 0.1 mmol/kg gadopentetate dimeglumine	b = 0, 350, 700	547.6 MBq (485–566)	Simultaneous	Consensus review	Imaging follow-up only
Reiner et al (2014) [14]	Prospective Cohort	Cohort	61.0 (32–79)	21	79	N.A. (0–10.0)	Colon (n = 41), pancreas (n = 5), GIST (n = 4), liver (n = 3), stomach and esophagus (n = 2)	Per-lesion analysis, per-patient analysis	1.5 T	CE without HBP: 0.1 mmol/kg gadodiamide	b = 0, 400, 800	300 MBq	Retrospective fusion	Consensus review	Pathology and imaging follow-up

Articles are listed according to year of publication and the alphabetical order of the names of the first authors within the same year of publication

CE contrast-enhanced, DWI diffusion-weighted image, HBP hepatobiliary phase, N.A. not applicable, GIST gastrointestinal stromal tumor

* Data are mean ± standard deviation

† Liver includes four hepatocellular carcinoma and three cholangiocarcinoma

Table 2 Per-patient and per-lesion accuracy of PET/MRI for diagnosing metastasis

Author (year of publication)	Number of study subject	Number of patients				Sensitivity (95% CI)	Specificity (95% CI)
		TP	FP	FN	TN		
Beiderwellen et al (2015) [12]	32	12	0	0	20	100 (74–100)	100 (83–100)
Donati et al (2010) [11]	37	24	0	0	13	100 (86–100)	100 (75–100)
Lee et al (2016) [13]	55	41	0	7	7	85 (72–94)	100 (59–100)
Reiner et al (2014) [14]	55	21	1	0	23	100 (84–100)	96 (79–100)
Higgins I^2 for study heterogeneity: inconsistency (I-square)						89.4%	0.0%
Meta-analytic summary estimate using the bivariate model						99.2% (31.4, 100.0)	98.6% (84.0, 100.0)
Author (year of publication)	Number of study subject	Number of lesions				Sensitivity (95% CI)	Specificity (95% CI)
		TP	FP	FN	TN		
Beiderwellen et al (2015) [12]	113	42	0	3	68	93 (82–99)	100 (95–100)
Brendle et al (2016) [15]	37	16	3	6	12	73 (50–89)	80 (52–96)
Donati et al (2010) [11]	85	51	1	4	29	93 (82–98)	97 (83–100)
Heusner et al (2012) [16]	556	233	0	250	73	48 (44–53)	100 (95–100)
Kirchner et al (2017) [17]	137	56	0	1	80	98 (91–100)	100 (95–100)
Melsaether et al (2016) [18]	75	40	1	0	34	100 (91–100)	97 (85–100)
Reiner et al (2014) [14]	120	79	0	0	41	100 (95–100)	100 (91–100)
Higgins I^2 for study heterogeneity: inconsistency (I-square)						99.7%	96.5%
Meta-analytic summary estimate using the bivariate model						95.4% (78.3, 99.2)	99.3% (93.8, 99.9)

Articles are listed according to year of publication and in alphabetical order of the names of the first authors within the same year of publication
 TP true positive, FP false positive, FN false negative, TN true negative, CI confidence interval

Meta-regression analysis

The results of the meta-regression analysis are summarized in Table 3. The study subject characteristics, study design, MR imaging technique, imaging review method, and type of reference standard were significant factors influencing study heterogeneity. Study subjects with various types of primary cancers

increased the Sp of PET-MRI (100 vs. 92%, $p < 0.001$). Prospective study design, PET/MRI with DWI, PET/MRI with HBP images, consensus image review, and reference standard determined by pathology showed significantly higher Se and Sp on PET/MRI (Se, $p \leq 0.05$; Sp, $p \leq 0.01$).

There was no significant publication bias among the studies (Supplementary Fig. 2; $p = 0.93$).

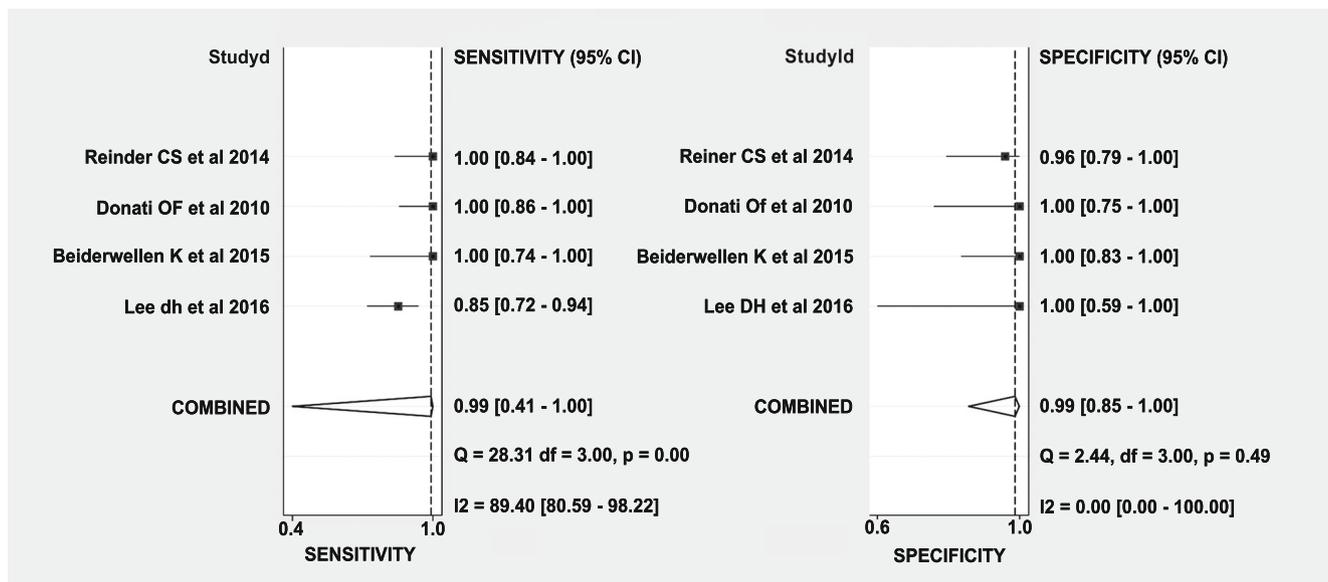


Fig. 2 Coupled forest plots of the Se and Sp for diagnosis of liver metastasis with PET/MRI. **a** The coupled forest plot of Se and Sp for per-patient-based analysis, and **(b)** the coupled forest plot of Se and Sp for per-lesion-based analysis, did not show an inverse correlation, suggesting no threshold effect

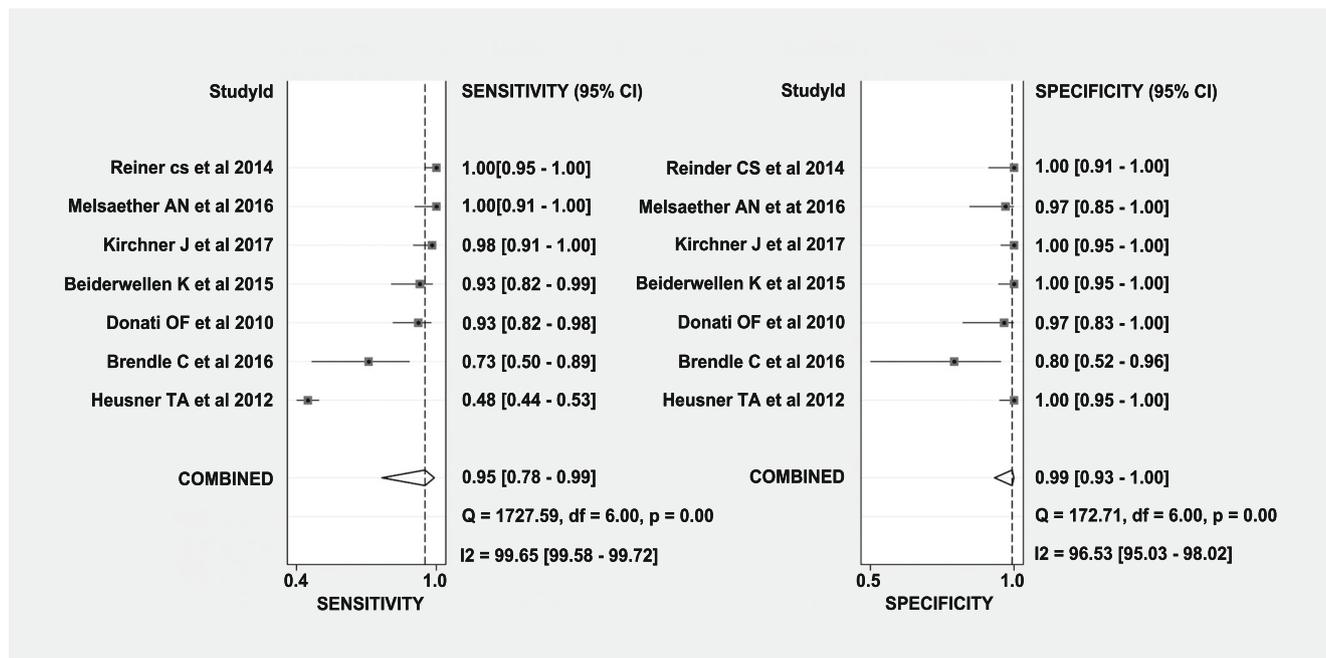


Fig. 2 (continued)

Discussion

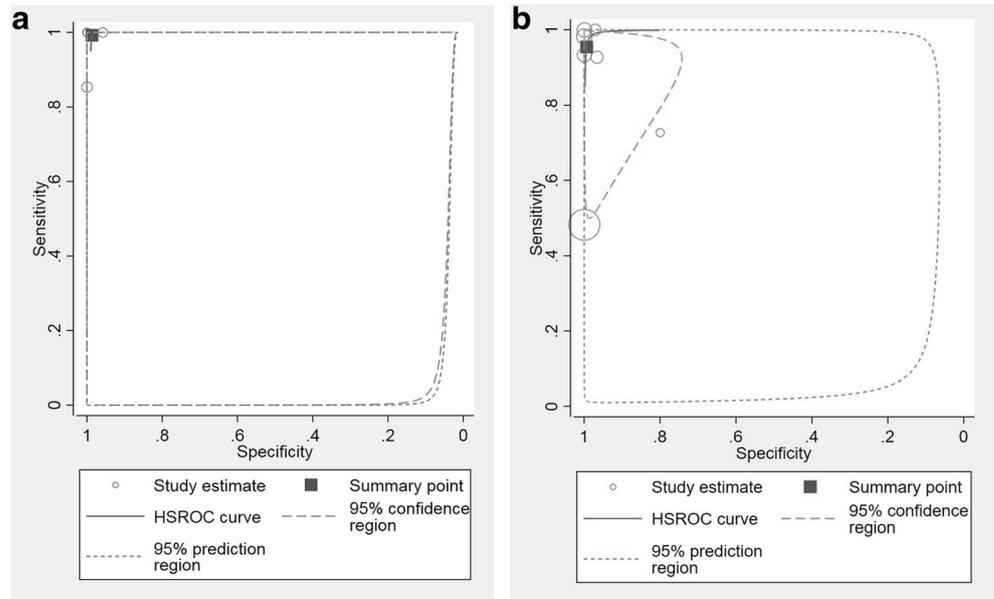
This study demonstrates that [18F]FDG-PET/MRI generally has a very good diagnostic accuracy, with a per-patient meta-analytic summary Se of 99.2% (95% CI, 31.4–100.0%) and Sp of 98.6% (95% CI, 84.0–99.9%), and a per-lesion meta-analytic summary Se of 95.4% (95% CI, 78.3–99.2%) and Sp of 99.3% (95% CI, 96.5–99.9%). These values were higher than those of [18F]FDG-PET/CT. However, there is insufficient robust data to substantiate the value of PET/MRI in routine clinical practice for the evaluation of liver metastasis.

In this study, PET/MRI demonstrated a high Se and Sp for the diagnosis of liver metastasis. This high Se could have resulted from MRI with liver-specific contrast, i.e., the excellent liver-to-lesion conspicuity on HBP images [22, 23], as well as the additional lesion detection on DWI [21]. Notably, our meta-regression analysis showed that studies with DWI or HBP had a higher Se for the diagnosis of liver metastasis than studies without either of these sequences. Considering that DWI can play a critical role in the detection of small liver metastasis [21, 24] and HBP image can provide excellent liver-to-lesion conspicuity for small metastatic lesions [19, 23], both DWI and HBP image might be clinically useful, providing incremental value for the diagnosis of liver metastasis. In addition, PET/MRI showed a high Sp for the diagnosis of liver metastasis, which can be attributed to the fact that PET provides metabolic information that can reduce the possibility of FP by excluding benign lesions, which sometimes mimic liver metastasis on conventional anatomical imaging such as CT or MRI [6].

Meta-regression analysis revealed that the study subject characteristics, study design, MR imaging technique (i.e., the use of DWI or HBP image), imaging review method, and type of reference standard were significant factors affecting study heterogeneity. Of the eight articles, Heusner et al showed the lowest Se for PET/MRI (48.2%) with 250 FN. The large number of FN could be explained by the absence of DWI or HBP and the inclusion of patients with systemic chemotherapy, which can result in a decrease in FDG uptake in liver metastasis. We found that prospective studies showed significantly higher Se and Sp than retrospective studies and those studies with subjects with only one type of primary cancer showed significantly lower Sp. As prospective studies usually have fewer potential sources of bias and confounding than retrospective studies, the reported results from prospective studies would be expected to be more reliable. Although studies including only one type of primary cancer showed lower Sp than those including various primary cancers, the diagnostic accuracy of PET/MRI in studies including only one type of primary cancer varied according to the type of primary cancer, i.e., 72.7 and 70.6% for colorectal cancer; 100.0 and 97.1% for breast cancer. In contrast, the results from studies including various types of primary cancers might be not an individual result but a mixed result from various cancers.

Notably, studies using clinical follow-up as a reference standard showed lower Se and Sp than those using pathology as a reference standard. As it is difficult to obtain a pathological diagnosis for small indeterminate lesions because of the risk of biopsy-related complications, clinicians may have a tendency to perform clinical follow-up without a pathological diagnosis

Fig. 3 HSROC curve for the accuracy of PET/MRI in the diagnosis of liver metastasis. **a** The HSROC curve for per-patient-based analysis, and **(b)** the HSROC curve for per-lesion-based analysis. The sizes of the circles (individual study estimates) indicate the sample sizes of the individual studies. There is a very large difference between the 95% prediction and 95% confidence regions, indicating considerable heterogeneity between studies.



in the case of the small indeterminate lesions. The lower diagnostic accuracy in studies using clinical follow-up as a reference standard might be explained by the presence of such small indeterminate lesions. In addition, consensus image review and image acquisition with DWI or HBP images increased the Se and Sp of PET/MRI. With consensus image review, misinterpretation should be minimized, as discrepancies between the reviewers would be re-evaluated and discussed.

This study has several limitations. First, as mentioned previously, large heterogeneity was noted between the studies, and this could preclude the creation of solid meta-analytic summary estimates. However, we robustly explored the reasons for the heterogeneity using meta-regression. Second, the small number of studies included in the meta-analysis is a major limitation; we could not perform a per-patient meta-regression analysis because of the small number of studies included. To minimize this

Table 3 Results of meta-regression analysis of PET/MRI for the diagnosis of metastasis

Covariates	Subgroup	Number of studies	Meta-analytic summary estimate				
			Sensitivity (95% CI)	<i>p</i> value	Specificity (95% CI)	<i>p</i> value	
Study subject	One primary cancer	2	97 (86–100)	0.14	92 (80–100)	< 0.001	
	Various primary cancer	5	95 (86–100)		100 (99–100)		
Study design	Prospective study	4	99 (97–100)	< 0.001	100 (99–100)	< 0.001	
	Retrospective study	3	76 (52–99)		97 (89–100)		
Data acquisition method	Simultaneous fusion	4	97 (89–100)	0.08	99 (95–100)	0.11	
	Retrospective fusion	3	93 (78–100)		100 (99–100)		
MR imaging technique	Acquisition with DWI	5	98 (94–100)	< 0.001	99 (97–100)	0.01	
	Acquisition without DWI	2	77 (36–100)		99 (97–100)		
	Acquisition with HBP images	2	97 (88–100)		100 (98–100)		< 0.001
	Acquisition without HBP images	5	94 (84–100)		99 (97–100)		
Imaging review method	Consensus review	4	98 (93–100)	0.01	100 (99–100)	< 0.001	
	Independent review	3	90 (67–100)		97 (91–100)		
Reference standard	Clinical follow-up*	5	92 (79–100)	< 0.001	99 (97–100)	< 0.001	
	Pathologic diagnosis [#]	2	99 (95–100)		99 (98–100)		

The results were obtained using meta-regression analysis with the bivariate model

DWI diffusion-weighted image, HBP hepatobiliary phase

* Reference standard determined using clinical follow-up in more than 50% of subjects

[#] Reference standard determined using pathologic diagnosis in more than 50% of subjects

limitation, we performed a per-lesion meta-regression analysis, including three of the four articles that had reported per-patient results. Third, the number of studies using a fully integrated PET/MRI system was small ($n = 4$). Although we found no significant difference in diagnostic performance between the two PET/MRI data acquisition methods (retrospective fusion vs. simultaneous fusion), further studies using a fully integrated PET/MRI system would be needed to validate the diagnostic performance of PET/MRI. Fourth, as we included only articles with [^{18}F]FDG-PET/MRI, we could not discuss PET/MRI with other tracers. In those studies identified according to our inclusion criteria, only one article reported the Se of 91.2% and Sp of 95.6% for PET/MRI using ^{68}Ga -DOPA-peptide for diagnosing liver metastasis of neuroendocrine tumor [25]. Further studies of PET/MRI with different tracers are required.

In conclusion, the overall diagnostic accuracy of [^{18}F]FDG-PET/MRI for the diagnosis of liver metastasis is high. However, large heterogeneities were noted between studies, and five significant causes of study heterogeneity were identified: study subject characteristics, study design, MR imaging technique (i.e., the use of DWI or HBP image), imaging review method, and type of reference standard. Further randomized controlled studies or prospective studies are needed to investigate the role of PET/MRI in liver metastasis in comparison with PET/CT.

Funding This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number HI18C2383).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Sang Hyun Choi.

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

Statistics and biometry One of the authors (Sang Hyun Choi) has significant statistical expertise.

Informed consent Written informed consent was not required for this study because this study was a meta-analysis.

Ethical approval Institutional review board approval was not required because this study was a meta-analysis.

Methodology

- Meta-analysis
- Diagnostic study
- Performed at one institution

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

References

1. Centeno BA (2006) Pathology of liver metastases. *Cancer Control* 13:13–26
2. Kokudo N, Tada K, Seki M et al (2001) Anatomical major resection versus nonanatomical limited resection for liver metastases from colorectal carcinoma. *Am J Surg* 181:153–159
3. Qian J (2011) Interventional therapies of unresectable liver metastases. *J Cancer Res Clin Oncol* 137:1763–1772
4. Frankel TL, Gian RK, Jamagin WR (2012) Preoperative imaging for hepatic resection of colorectal cancer metastasis. *J Gastrointest Oncol* 3:11–18
5. Smith TJ, Korngold E, Orloff SL (2014) Preoperative imaging in colorectal liver metastases: current practices. *Curr Surg Rep* 2:39
6. Schulz A, Viktil E, Godt JC et al (2016) Diagnostic performance of CT, MRI and PET/CT in patients with suspected colorectal liver metastases: the superiority of MRI. *Acta Radiol* 57:1040–1048
7. Choi SH, Kim SY, Park SH et al (2018) Diagnostic performance of CT, gadoxetate disodium-enhanced MRI, and PET/CT for the diagnosis of colorectal liver metastasis: systematic review and meta-analysis. *J Magn Reson Imaging* 47:1237–1250
8. Nickel MC, Bipat S, Stoker J (2010) Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology* 257:674–684
9. Chua SC, Groves AM, Kayani I et al (2007) The impact of 18F-FDG PET/CT in patients with liver metastases. *Eur J Nucl Med Mol Imaging* 34:1906–1914
10. Torigian DA, Zaidi H, Kwee TC et al (2013) PET/MR imaging: technical aspects and potential clinical applications. *Radiology* 267:26–44
11. Donati OF, Hany TF, Reiner CS et al (2010) Value of retrospective fusion of PET and MR images in detection of hepatic metastases: comparison with 18F-FDG PET/CT and Gd-EOB-DTPA-enhanced MRI. *J Nucl Med* 51:692–699
12. Beiderwellen K, Geraldo L, Ruhlmann V et al (2015) Accuracy of [^{18}F]FDG PET/MRI for the detection of liver metastases. *PLoS One* 10:e0137285
13. Lee DH, Lee JM, Hur BY et al (2016) Colorectal cancer liver metastases: diagnostic performance and prognostic value of PET/MR imaging. *Radiology* 280:782–792
14. Reiner CS, Stolzmann P, Husmann L et al (2014) Protocol requirements and diagnostic value of PET/MR imaging for liver metastasis detection. *Eur J Nucl Med Mol Imaging* 41:649–658
15. Brendle C, Schwenzer NF, Rempp H et al (2016) Assessment of metastatic colorectal cancer with hybrid imaging: comparison of reading performance using different combinations of anatomical and functional imaging techniques in PET/MRI and PET/CT in a short case series. *Eur J Nucl Med Mol Imaging* 43:123–132
16. Heusner TA, Mikat C, Hahn S et al (2012) FDG-PET/MRI fused data sets for the detection of liver metastases in patients undergoing systemic anticancer treatment. *Iran J Radiat Res* 9(4):209–219
17. Kirchner J, Sawicki LM, Deuschl C et al (2017) 18 F-FDG PET/MR imaging in patients with suspected liver lesions: value of liver-specific contrast agent Gadobenate dimeglumine. *PLoS One* 12:e0180349
18. Melsaether AN, Raad RA, Pujara AC et al (2016) Comparison of whole-body (18)F FDG PET/MR imaging and whole-body (18)F FDG PET/CT in terms of lesion detection and radiation dose in patients with breast cancer. *Radiology* 281:193–202
19. Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P (2013) Cochrane diagnostic test accuracy reviews. *Syst Rev* 2:82

20. Trikalinos TA, Balion CM, Coleman CI et al (2012) Meta-analysis of test performance when there is a “gold standard”. In: Chang SM, Matchar DB, Smetana GW, Umscheid CA (eds) *Methods Guide for Medical Test Reviews*, Rockville (MD)
21. Bruegel M, Gaa J, Waldt S et al (2008) Diagnosis of hepatic metastasis: comparison of respiration-triggered diffusion-weighted echo-planar MRI and five t2-weighted turbo spin-echo sequences. *AJR Am J Roentgenol* 191:1421–1429
22. Lee KH, Lee JM, Park JH et al (2013) MR imaging in patients with suspected liver metastases: value of liver-specific contrast agent gadoxetic acid. *Korean J Radiol* 14:894–904
23. Yu MH, Lee JM, Hur BY et al (2015) Gadoxetic acid-enhanced MRI and diffusion-weighted imaging for the detection of colorectal liver metastases after neoadjuvant chemotherapy. *Eur Radiol* 25: 2428–2436
24. Coenegrachts K, Delanote J, Ter Beek L et al (2007) Improved focal liver lesion detection: comparison of single-shot diffusion-weighted echoplanar and single-shot T2 weighted turbo spin echo techniques. *Br J Radiol* 80:524–531
25. Schreiter NF, Nogami M, Steffen I et al (2012) Evaluation of the potential of PET-MRI fusion for detection of liver metastases in patients with neuroendocrine tumours. *Eur Radiol* 22:458–467