

Colorectal Liver Metastases: CT, MR Imaging, and PET for Diagnosis—Meta-analysis¹

Shandra Bipat, MSc
Maarten S. van Leeuwen,
MD, PhD
Emile F. I. Comans, MD,
PhD
Milan E. J. Pijl, MD, PhD
Patrick M. M. Bossuyt, PhD
Aeilko H. Zwinderman, PhD
Jaap Stoker, MD, PhD

Published online before print
10.1148/radiol.2371042060
Radiology 2005; 237:123–131

Abbreviations:

CI = confidence interval
FDG = fluorine 18
fluorodeoxyglucose
SPIO = superparamagnetic iron
oxide

¹ From the Department of Radiology (S.B., J.S.) and Department of Epidemiology and Biostatistics (P.M.M.B., A.H.Z.), Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands; Department of Radiology, University Medical Center Utrecht, Utrecht, the Netherlands (M.S.v.L.); Departments of Nuclear Medicine and PET Research, VU University Medical Center, Amsterdam, the Netherlands (E.F.I.C.); and Department of Radiology, Martini Hospital, Groningen, the Netherlands (M.E.J.P.). Received December 6, 2004; revision requested January 10, 2005; revision received January 24; accepted February 1. Supported by a grant from the Dutch Order of Medical Specialists. **Address correspondence** to S.B. (e-mail: s.bipat@amc.uva.nl).

Authors stated no financial relationship to disclose.

Author contributions:

Guarantor of integrity of entire study, S.B.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, all authors; statistical analysis, S.B., P.M.M.B., A.H.Z., J.S.; manuscript editing, all authors

© RSNA, 2005

PURPOSE: To perform a meta-analysis to obtain sensitivity estimates of computed tomography (CT), magnetic resonance (MR) imaging, and fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET) for detection of colorectal liver metastases on per-patient and per-lesion bases.

MATERIALS AND METHODS: MEDLINE, EMBASE, Web of Science, and CANCELRLIT databases and Cochrane Database of Systematic Reviews were searched for relevant original articles published from January 1990 to December 2003. Criteria for inclusion of articles were as follows: Articles were reported in the English, German, or French language; CT, MR imaging, or FDG PET was performed to identify and characterize colorectal liver metastases; histopathologic analysis (surgery, biopsy, or autopsy), intraoperative observation (manual palpation, intraoperative ultrasonography [US]), and/or follow-up US was the reference standard; and data were sufficient for calculation of true-positive or false-negative values. A random-effects linear regression model was used to obtain sensitivity estimates in assessment of liver metastases.

RESULTS: Of 165 identified relevant articles, 61 fulfilled all inclusion criteria. Sensitivity estimates on a per-patient basis for nonhelical CT, helical CT, 1.5-T MR imaging, and FDG PET were 60.2%, 64.7%, 75.8%, and 94.6%, respectively; FDG PET was the most accurate modality. On a per-lesion basis, sensitivity estimates for nonhelical CT, helical CT, 1.0-T MR imaging, 1.5-T MR imaging, and FDG PET were 52.3%, 63.8%, 66.1%, 64.4%, and 75.9%, respectively; nonhelical CT had lowest sensitivity. Estimates of gadolinium-enhanced MR imaging and superparamagnetic iron oxide (SPIO)-enhanced MR imaging were significantly better, compared with nonenhanced MR imaging ($P = .019$ and $P < .001$, respectively) and with helical CT with 45 g of iodine or less ($P = .02$ and $P < .001$, respectively). For lesions of 1 cm or larger, SPIO-enhanced MR imaging was the most accurate modality ($P < .001$).

CONCLUSION: FDG PET had significantly higher sensitivity on a per-patient basis, compared with that of the other modalities, but not on a per-lesion basis. Sensitivity estimates for MR imaging with contrast agent were significantly superior to those for helical CT with 45 g of iodine or less.

© RSNA, 2005

Colorectal cancer is the second leading cause of cancer-related deaths in the United States. According to the National Program of Cancer Registries, 146 940 new patients received a diagnosis of the disease in 2004, with an estimated 56 730 deaths due to colorectal cancer in that year. Liver metastasis is a common consequence of colorectal carcinoma; up to 70% of patients with colorectal cancer eventually develop liver metastases. In 30%–40% of those patients, the metastases are still confined to the liver at the time of detection, and only a limited number of patients with colorectal metastases confined to the liver are surgical candidates because of the larger size of the lesions, the broad distribution of the lesions, or the difficulty in assessing the tumors or because the volume of the remaining liver is inadequate (1–5).

Preoperative selection of patients with colorectal metastases who are most likely to benefit from surgery is necessary and challenging. The armamentarium for imaging-based

preoperative selection comprises transabdominal ultrasonography (US), computed tomography (CT), fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET), and magnetic resonance (MR) imaging (6–12). During the past 10 years, improvements in these imaging modalities were either introduced or great progress has been made in their application (6,8,13–16). Although extensive research has been performed in regard to the diagnostic performance of CT, MR imaging, and FDG PET for the detection of colorectal liver metastases, the optimal imaging staging strategy has not been defined.

Kinkel et al (17) performed a meta-analysis to compare current noninvasive imaging methods such as US, CT, MR imaging, and FDG PET for the detection of hepatic metastases from colorectal, gastric, and esophageal cancers. Treatment approaches for liver metastases from various cancerous origins (pancreatic or colorectal cancer), however, are different, and, therefore, the importance of certain findings with respect to these origins differs.

Because of its noninvasive character, low cost, and widespread availability, US is a valuable screening tool for the imaging of liver metastases. US, however, has two relative disadvantages: US is more operator independent than are CT and MR imaging, and parts of the liver remain nonvisible in certain patients at US. In daily practice, though, US is highly efficient in helping to distinguish between two groups of patients with liver metastases: the group of patients with diffuse metastases who are no longer eligible for curative treatment and the group with no metastases or a very limited number of them. The patients in the latter group require CT, MR imaging, or FDG PET for the selection of appropriate therapeutic approaches. Thus, the aim of our study was to perform a meta-analysis to obtain the estimates of sensitivity of CT, MR imaging, and FDG PET for the detection of colorectal liver metastases on a per-patient and a per-lesion basis.

MATERIALS AND METHODS

Literature Search

A comprehensive computer literature search (18) of abstracts about studies in human subjects was performed to identify articles about the diagnostic performance of CT, MR imaging, and FDG PET for the detection of liver metastases in patients with colorectal cancer compared

with the diagnostic performance of intraoperative US, surgery, follow-up US, and histopathologic analysis as the reference standard. The MEDLINE and EMBASE databases, from January 1990 to December 2003, were used with the following keywords: (“Colorectal Neoplasms” [MeSH]) AND (“Liver neoplasms” [MeSH]) AND (“Laparoscopy” [MeSH] OR “Tomography, Emission-Computed” [MeSH] OR “magnetic resonance imaging” [MeSH] OR “Tomography, X-Ray Computed” [MeSH] AND (sensitivity and specificity [MeSH] OR sensitivity [WORD] OR specificity [WORD] OR false negative [WORD] OR false positive [WORD] OR diagnosis [MeSH] OR diagnostic use [MeSH] OR detection [WORD] OR accuracy [WORD])).

Other databases, such as CINAHL and SUMSEARCH, were also checked for relevant articles with the following keywords: Colorectal Neoplasm [MeSH] AND (Liver Neoplasms [MeSH] OR Neoplasm Metastasis [MeSH]). The databases of Web of Science and CANCELIT and the Cochrane Database of Systematic Review were checked with the following words: Colorectal cancer AND (liver metastases OR hepatic metastasis). Review articles, letters, comments, case reports, unpublished articles, and articles that did not include raw data were not selected. The list of articles was supplemented with extensive cross-checking of the reference lists of all retrieved articles.

Selection of Studies

Four observers independently checked all retrieved articles for inclusion criteria. One observer (S.B.) checked all articles. Three observers checked a subset of articles: One observer (M.S.v.L.) checked studies that predominantly focused on evaluation of CT, another (M.E.J.P.) checked studies that predominantly focused on evaluation of MR imaging, and another (E.F.I.C.) checked studies that predominantly focused on evaluation of FDG PET. Disagreements were resolved in consensus. The inclusion criteria were as follows: (a) Articles were reported in the English, German, or French language. (b) CT, MR imaging, or FDG PET was used to identify and characterize colorectal liver metastases. (c) Histopathologic analysis (performed at surgery, biopsy, and autopsy), intraoperative observation (eg, manual palpation, intraoperative US) and/or follow-up US were used as the reference standard. (d) For per-patient or per-lesion statistics, sufficient data were presented to calculate the true-positive and false-negative values for imaging

techniques. (e) When data or subsets of data were presented in more than one article, the article with the most details or the most recent article was chosen. Studies were excluded if results of different imaging modalities were presented in combination and could not be differentiated for performance assessment of tests on an individual modality.

Data Extraction

The same observers independently extracted relevant data about study (design) characteristics and examination results, which will be discussed later, from each article by using a standardized form. One observer (S.B.) extracted data from all articles. Three observers extracted data from a subset of articles: One observer extracted only data from studies that predominantly focused on evaluation of CT (M.S.v.L.), another extracted only data from studies that predominantly focused on evaluation of MR imaging (M.E.J.P.), and still another extracted only data from studies that predominantly focused on evaluation of FDG PET (E.F.I.C.). Observers were not blinded with regard to information about the authors, the authors' affiliation, or the journal name, since this has been shown to be unnecessary (19). To resolve disagreement between reviewers, a fifth reviewer (J.S.) assessed all discrepant items, and the majority opinion was used for analysis.

Study design characteristics.—The QUADAS quality assessment tool was used to extract relevant study design characteristics of each study. This tool and the definitions of the characteristics are fully described elsewhere (20).

Other study characteristics.—In addition, the following characteristics were recorded: (a) year of publication; (b) sample size (number of patients with colorectal liver metastases); (c) description of study population, which included disease severity (tumor stage), age, and male-female distribution; (d) description of interpretation of diagnostic tests, which included the reporting of the characterization of lesions as benign versus malignant or the detailed subcharacterization of lesions as cysts, hemangiomas, or metastases and the confidence rating used for identification of lesions; (e) description of reference tests, which included intraoperative findings (at inspection and/or palpation), intraoperative US features (probe frequency), pathologic features (staining, lamination, thickness of slices), or follow-up characteristics (interval between examinations, frequen-

cies [how many times follow-up examinations were performed], and modality). The following imaging features were extracted: For CT, these features included type of scanner (nonhelical, single-section helical, or multisection helical), section thickness, amount of contrast agent, and number of phases. For MR imaging, these features included magnetic field strength, type of contrast agent (nonspecific or liver-specific agents), sequences, type of coil (body coil or phased-array coil), and section thickness. For FDG PET, these features included system type (dedicated full ring or other), amount of tracer, type of analysis (qualitative or quantitative), and data acquisition characteristics (timing of scanning and time of scanning per table position).

Examination results.—The numbers of true-positive, false-negative, and false-positive results in the detection of liver metastases were extracted on a per-lesion basis. The numbers of true-positive, false-negative, false-positive, and true-negative results were also extracted on a per-patient basis. All tabulated results for different readers (interobserver), for multiple observations per reader (intraobserver), and for multiple CT and MR imaging systems and/or techniques were counted as separate data sets.

Data and Statistical Analysis

Data were separately analyzed for nonhelical CT, helical CT, MR imaging at 1.0 T, MR imaging at 1.5 T, and FDG PET. For each data set, we calculated sensitivity of the imaging techniques as the proportion p of patients with liver metastases (per patient) or as the proportion of liver metastases (per lesion) correctly recognized by the imaging modality.

In our statistical analysis, we used logit-transformed sensitivity $\ln \{p/(1 - p)\}$, where \ln is the natural logarithm. Because of the transformation, these values were approximately normal distributed, with a variance of $1/[n \cdot p \cdot (1 - p)]$, where n is the total number of patients with liver metastases (per patient) or the number of liver metastases (per lesion). Mean logit sensitivity and the standard error were obtained by means of a random-effects linear regression model with a mixed-effects approach that was a procedure that was within the software (SAS; SAS Institute, Cary, NC) (21,22). After antilogit transformation of the mean logit sensitivity, sensitivity estimates with the 95% confidence intervals (CIs) were obtained.

All analyses were performed with sta-

tistical software (SPSS 11.5 for Windows, SPSS, Chicago, Ill; SAS, version 8.02, SAS Institute).

Estimates of sensitivity.—For comparison of the sensitivity estimates of the different imaging techniques, we first determined whether the logit sensitivity values depended on year of publication (1995 or earlier vs later than 1995), sample size (≤ 50 vs > 50 patients), and the study design characteristics (“yes” vs “no” and “unclear” responses). In this analysis, we considered variables as explanatory if the regression coefficient of the variables was significant ($P < .05$).

Subsequently, we developed a multivariable regression model with which we used a backward stepwise algorithm to identify only the most important characteristics. Characteristics were retained in the regression model when the P value for them was less than .10.

Afterward, logit sensitivity values of the imaging techniques of nonhelical CT, helical CT, MR imaging at 1.0 T, MR imaging at 1.5 T, and FDG PET were compared with each other in this random-effects regression model, which included all variables that significantly affected the logit sensitivity of the imaging modalities (set to one, indicating the ideal design, vs zero) as appropriate; in this final model, a factor that indicated the type of diagnostic modality was included, and a P value of less than .05 of the regression coefficient of this factor was considered to indicate a significant difference. Fit of the final regression model was inspected graphically with evaluation of the histograms of the residuals and of the random-effects estimates.

When studies contributed two or more sensitivity values, for instance, when results of multiple readers (interobserver) or multiple observations per readers (intraobserver) were available or when multiple CT or MR imaging systems or multiple MR imaging sequences were evaluated, each sensitivity value was counted as a separate data set. We accounted for the likely correlation between such sensitivity values by calculating so-called robust standard errors, which are provided with the random-effects linear regression model with the mixed-effects approach (23). This approach was also used for intramodality inpatient correlation (in some studies, different modalities were compared in the same patient population).

Subgroup analysis 1.—Enough data sets were available to perform subgroup analyses for helical CT and MR imaging at 1.5 T. For helical CT, subgroup analyses were

used to compare section thicknesses (5 mm or smaller vs larger than 5 mm), the amounts of administered iodine in the contrast agent (≤ 45 g vs > 45 g), and the number of phases (ie, one phase [portal phase] vs two phases [arterial and portal phases]). For MR imaging at 1.5 T, non-enhanced MR imaging, MR imaging enhanced with gadolinium-based contrast agents, and MR imaging enhanced with superparamagnetic iron oxide (SPIO) were compared. The subgroup analyses were performed only for data on a per-lesion basis, as data on a per-patient basis were limited.

Subgroup analysis 2.—In addition, data sets also were analyzed for different lesion sizes (lesions of < 1 cm vs lesions ≥ 1 cm). With subgroup analyses, lesion size was compared for helical CT, nonenhanced MR imaging, gadolinium-enhanced MR imaging, and SPIO-enhanced MR imaging for 1.5-T imagers. This analysis was performed only on a per-lesion basis, as data on a per-patient basis were not available.

RESULTS

Literature Search and Selection of Studies

With the computer search and after extensive cross-checking of reference lists, 315 abstracts were retrieved. After reading of the abstracts was performed, 165 articles were found to be eligible. One hundred four of the 165 relevant articles were excluded because (a) researchers in the articles did not report data about the use of CT, MR imaging, or FDG PET for identification and characterization of colorectal liver metastases ($n = 18$); (b) researchers in the articles did not use histopathologic analysis, intraoperative observation, including manual palpation and intraoperative US, and/or follow-up US as the reference standard ($n = 25$); (c) researchers in the articles did not report data that could be used to construct or calculate true-positive, false-positive, true-negative, and/or false-negative results ($n = 54$); or (d) researchers in the articles presented results from a combination of different imaging modalities that could not be differentiated for assessment of single tests ($n = 7$). Sixty-one articles fulfilled all inclusion criteria and were selected for data extraction and data analysis.

Study Design Characteristics

Most studies (Table 1) had a suboptimal design in regard to the period between the time when the reference stan-

standard was performed and the time when the index test was performed (67.2% for “no” responses to question 4), the description of the execution of the reference standard (63.9% for “no” responses to question 8b), the interpretation of the reference standard results without knowledge of the index test results (91.2% for “no” responses to question 9b), the availability of clinical data when test results were interpreted (70.5% for “no” responses to question 10), reporting of uninterpretable and/or intermediate test results (98.4% for “no” responses to question 11), and explanation of withdrawals from the study (75.4% for “no” responses to question 12). It is impossible to perform an ideal study, however, as the choice for a treatment strategy strongly depends on the outcome of the diagnosis. The description of the execution of the reference standard remains a problem in studies of diagnostic performance of modalities.

Other Study Design Characteristics

The age of the patients included in the selected studies ranged from 12 to 93 years, with a mean age of 61.0 years and a total of 3187 patients. In 57 studies, the sex distribution was described: 1733 patients were male and 1128 patients were female. In all studies, imaging data were presented about the identification of lesions; in only nine studies was a confidence rating scale presented. In 31 of 61 studies, lesions were characterized (benign lesions were distinguished from malignant lesions in 28 studies and detailed subcharacterization of lesions was included in three studies). The reference standard was intraoperative observation (palpation) in 43 studies, intraoperative US in 37, pathologic analysis in 54, and follow-up US in 31. The frequency of the transducer used for intraoperative US varied from 5.0 to 7.5 MHz. In 11 of 54 studies in which pathologic analysis was used as the reference standard, the method of analysis (eg, staining or lamination) and the thickness of slices were described.

Examination Results

Table 2 presents the included data sets (per-patient basis and per-lesion basis), with corresponding numbers of patients and reference numbers. A full list of all included articles with all relevant study characteristics and complete examination results is available on request from the authors of this article. The range in

TABLE 1
Results of Distribution of Study Design Characteristics in 61 Studies

Question about Study Design Characteristic	Response*	
	Yes	No
1. Was the spectrum of patients representative of the patients who receive the test in practice?	50	11
2. Were selection criteria clearly described?	36	25
3. Is the reference standard likely to help correctly classify the target condition?	55	6
4. Is the time between performance of reference standard and index test short enough?	20	41
5. Did the whole sample or a random selection of the sample receive verification by using a reference standard?	52	9
6. Did patients undergo examination with the same reference standard regardless of the index test result?	36	25
7. Was the reference standard performed independently of the index test?	52	9
8a. Was the execution of the index test described in sufficient detail to permit replication of the test?	49	12
8b. Was the execution of the reference standard described in sufficient detail to permit replication of the test?	22	39
9a. Were the index test results interpreted without knowledge of the results of the reference standard?	36	25
9b. Were the reference standard results interpreted without knowledge of the results of the index test?	6	55
10. Were the same clinical data available when test results were interpreted as would be available in practice?	18	43
11. Were uninterpretable and/or intermediate test results reported?	1	60
12. Were withdrawals from the study explained?	15	46
13. Were the data collected after the research question was defined?	36	25

* Data are the numbers of responses from the QUADAS tool. The numbers indicate how many articles were assigned a score of “yes” (for the QUADAS tool) and how many articles were assigned a score of “no.” The responses of “no” and “unclear” were summarized together.

TABLE 2
Study Characteristics of Included Data Sets for Each Imaging Modality

Modality	No. of Data Sets and Articles	No. of Patients in Study	Reference Nos.
Nonhelical CT	58, 28	1915	10, 24–50
Helical CT	53, 15	621	51–65
1.0-T MR imaging	34, 5	173	57, 66–69
1.5-T MR imaging	102, 12	391	27, 51, 53, 54, 70–77
FDG PET	26, 21	1058	41–50, 52, 59, 64, 77–83

section thickness at CT was 5–12 mm (median, 10 mm), and that for MR imaging, 5–10 mm (median, 10 mm). The range in the amount of iodine in the administered contrast agent during CT (reported in 23 studies) was 30–60 g. In 15 studies, either nonspecific gadolinium chelates or liver-specific MR imaging contrast agents, such as SPIO and gadobenate dimeglumine, were used. In most of the studies (15 of 21) about FDG PET, the images were qualitatively analyzed (uptake of FDG).

Sensitivity Estimates on Per-Patient Basis

Following the backward stepwise regression analysis, several variables were

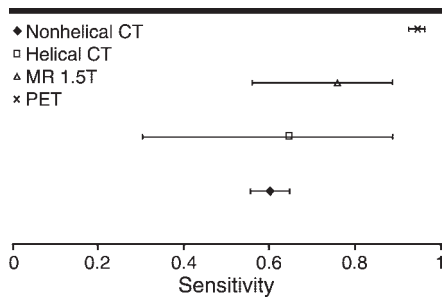
identified as significant predictors of the diagnostic performance of nonhelical CT and FDG PET for assessment of colorectal liver metastases on a per-patient basis (Table 3). No predictors were identified for helical CT and MR imaging at 1.5 T. No data sets were available for MR imaging at 1.0 T. In the final models, all significant variables were included as covariates.

The sensitivity estimates for nonhelical CT, helical CT, MR imaging at 1.5 T, and FDG PET were 60.2% (95% CI: 55.7%, 64.6%), 64.7% (95% CI: 30.4%, 88.5%), 75.8% (95% CI: 55.9%, 88.6%), and 94.6% (95% CI: 92.5%, 96.1%), respectively (Figure). FDG PET had a significantly higher sensitivity estimate compared with that of nonhelical CT ($P <$

TABLE 3
Predictors Identified with Backward Regression Analysis for Each Imaging Modality

Modality	No. of Data Sets	Covariates	Regression Coefficient*	P Value
Per patient				
CT				
Nonhelical	19	Reference standard helped to correctly classify the target condition	-0.85 (-1.37, -0.32)	<.002
Nonhelical	19	Reference standard results were interpreted without knowledge of index test results	-0.55 (-0.88, -0.21)	<.002
Helical	2	No predictors found	...	
1.5-T MR imaging				
1.5 T	6	No predictors found	...	
FDG PET				
FDG PET	15	Index test results were interpreted without knowledge of reference standard results	1.07 (0.40, 1.75)	<.002
Per lesion				
CT				
Nonhelical	22	Execution of the index test was described in sufficient detail	0.87 (0.38, 1.36)	<.001
Nonhelical	22	Reference standard results were interpreted without knowledge of index test results	-0.61 (-0.94, -0.28)	<.001
Helical	39	Reference standard helped to correctly classify the target condition	-0.78 (-1.40, -0.17)	<.012
MR imaging				
1.0 T	22	Spectrum of patients was representative of patients in practice	0.08 (0.07, 0.09)	<.001
1.0 T	22	Was the reference standard performed independently of the index test	-0.77 (0.89, 0.65)	<.001
1.5 T	57	Reference standard helped to correctly classify the target condition	-1.85 (-2.24, -1.46)	<.001
FDG PET	9	No predictors found	...	

* Numbers in parentheses are 95% CIs. A positive regression coefficient indicates better discriminatory power of the modality in studies with that characteristic compared with that in studies without the corresponding characteristic, and a negative regression coefficient indicates reduced diagnostic performance in studies with that characteristic.



Graph shows sensitivity estimates of 60.2%, 64.7%, 75.8%, and 94.6%, with 95% CIs, for nonhelical CT, helical CT, MR imaging at 1.5 T, and FDG PET, respectively, on a per-patient basis. FDG PET was the most accurate modality. A comparison of FDG PET with nonhelical CT, helical CT, and MR imaging at 1.5 T yielded three *P* values (*P* < .001, *P* = .003, *P* < .001).

.001), helical CT (*P* = .003), and MR imaging at 1.5 T (*P* < .001).

Sensitivity Estimates on Per-Lesion Basis

Several variables were identified as significant predictors of the diagnostic performance of nonhelical CT, helical CT, MR imaging at 1.0 T, and MR imaging at 1.5 T for assessment of colorectal liver metastases (Table 3) on a per-lesion basis. No predictors were found for FDG PET. Overall sensitivity estimates for nonhelical CT, helical CT, MR imaging at 1.0 T,

MR imaging at 1.5 T, and FDG PET were 52.3%, 63.8%, 66.1%, 64.4%, and 75.9% (Table 4). Nonhelical CT had the lowest sensitivity estimate compared with helical CT (*P* < .017), MR imaging at 1.0 T (*P* < .001), MR imaging at 1.5 T (*P* < .001), and FDG PET (*P* < .003).

Subgroup Analysis 1

For helical CT, subgroup analyses included a comparison of a section thickness of 5 mm (no data about section thickness of <5 mm were available) with a section thickness of larger than 5 mm, a comparison of the amount of iodine in the contrast agent of 45 g or less with an amount of more than 45 g, and a comparison of the number of phases (one phase [portal phase] vs two phases [arterial and portal phases]). For MR imaging at 1.5 T, nonenhanced MR imaging, gadolinium-enhanced MR imaging, and SPIO-enhanced MR imaging were compared. Sensitivity estimates for a section thickness of 5 mm and a section thickness of larger than 5 mm were comparable: 68.2% and 69.1%, respectively. For the amount of iodine of 45 g or less and that of more than 45 g, the estimates were 61.4% and 64.0%, respectively. Although the sensitivity estimate for the portal phase was higher (71.4%) compared with that of the portal and arterial phases (65.7%), this difference was not significant.

Estimates of sensitivity for nonenhanced MR imaging, gadolinium-enhanced MR imaging, and SPIO-enhanced MR imaging were 59.8%, 78.2%, and 73.2%, respectively. Sensitivity estimates for gadolinium-enhanced MR imaging (*P* = .019) and SPIO-enhanced MR (*P* < .001) were significantly higher compared with the estimate for nonenhanced MR imaging. In addition, sensitivity estimates for gadolinium-enhanced MR imaging (*P* = .02) and SPIO-enhanced MR imaging (*P* < .001) were significantly higher compared with the estimate for helical CT with 45 g or less of iodine.

Subgroup Analysis 2

Sensitivity estimates for nonhelical CT, helical CT, nonenhanced MR imaging, gadolinium-enhanced MR imaging, and SPIO-enhanced MR imaging for lesions smaller than 1 cm were 25.3% (95% CI: 15.9%, 37.6%), 23.1% (95% CI: 7.0%, 54.7%), 12.6% (95% CI: 8.0%, 17.5%), 11.6% (95% CI: 9.5%, 14.2%), and 29.3% (95% CI: 18.2%, 43.6%), respectively. No differences were found between the imaging modalities. Sensitivity estimates for nonhelical CT, helical CT, nonenhanced MR imaging, gadolinium-enhanced MR imaging, and SPIO-enhanced MR imaging for lesions of 1 cm or larger were 74.3% (95% CI: 66.5%, 80.9%), 73.5% (95% CI: 62.2%, 82.4%), 65.7% (95% CI: 56.4%, 73.9%), 68.8% (95% CI: 61.9%,

75.0%), and 90.2% (95% CI: 87.5%, 92.4%), respectively. The sensitivity estimate for SPIO-enhanced MR imaging was significantly higher ($P < .001$).

DISCUSSION

In this meta-analysis, we found that on a per-patient basis, FDG PET was most accurate for detection of colorectal liver metastases. On a per-lesion basis, helical CT, MR imaging at 1.0 T, MR imaging at 1.5 T, and FDG PET were comparable and significantly more accurate than was nonhelical CT.

Data about subgroup analyses indicated no differences between section thicknesses, amounts of administered iodine, and numbers of phases for helical CT. Gadolinium-enhanced MR imaging and SPIO-enhanced MR imaging, however, were significantly better compared with nonenhanced MR imaging and helical CT with an amount of iodine of 45 g or less.

As treatment policies differ for liver metastases of various cancerous origins (pancreatic cancer or colorectal cancer), only data about colorectal cancer were extracted and analyzed.

To avoid selection bias, not only the MEDLINE database but also the CINAHL, SUMSEARCH, Web of Science, and CANCERLIT databases and the Cochrane Database of Systematic Review were searched for relevant articles. In addition, all reference lists were checked manually.

To minimize bias in the selection of studies and in data extraction, reviewers independently selected articles on the basis of inclusion criteria, and scores were assigned to study design characteristics and examination results by using a standardized form that was based on the QUADAS tool. The QUADAS tool is an evidence-based quality assessment tool, which was developed for use in systematic reviews of studies of diagnostic accuracy (20).

Data were analyzed by means of a random-effects approach, which accounts for the heterogeneity between studies caused by different threshold settings (as in regular summary receiver operating characteristic curves) (84,85), for the error of estimation of the sensitivity values in each study that represents the size of the population, and finally for the residual heterogeneity that may remain even after adjustment for study design characteristics (86,87). Lijmer et al (88) showed that studies of diagnostic performance of modalities with methodological short-

TABLE 4
Sensitivity Estimates for Nonhelical CT, Helical CT, 1.0-T MR Imaging, 1.5-T MR Imaging, and FDG PET on a Per-Lesion Basis

Modality and Subgroup	Sensitivity Estimate (%)*
Nonhelical CT, overall	52.3 (52.1, 52.5)
Helical CT	
Overall	63.8 (54.4, 72.2) [†]
Section thickness of 5 mm	68.2 (50.5, 81.9)
Section thickness of >5 mm	69.1 (59.8, 77.1)
Amount of iodine of ≤45 g	61.4 (43.5, 76.6)
Amount of iodine of >45 g	64.0 (55.1, 72.0)
Two phases (arterial and portal phases)	65.7 (56.8, 73.7)
One phase (portal phase only)	71.4 (57.7, 82.1)
1.0-T MR imaging, overall	66.1 (65.9, 66.3) [†]
1.5-T MR imaging	
Overall	64.4 (57.8, 70.5) [†]
Nonenhanced MR imaging	59.8 (49.0, 69.7)
Gadolinium-enhanced MR imaging	78.2 (63.0, 88.3) [‡]
SPIO-enhanced MR imaging	73.2 (62.3, 81.9) [‡]
FDG PET, overall	75.9 (61.1, 86.3) [†]

* Sensitivity estimates were obtained by means of a logit-transformed data analysis, and percentages were not calculated with raw numbers. Numbers in parentheses are 95% CIs expressed as percentages.

[†] Significantly higher compared with nonhelical CT.

[‡] Significantly higher compared with nonenhanced MR imaging and amount of iodine of 45 g or less.

comings may cause overestimation of the accuracy of a diagnostic test; we, therefore, evaluated the effect of these characteristics on diagnostic performance and made adjustments when appropriate.

During the years, substantial improvements in CT (eg, introduction of spiral CT, multisection CT) and MR imaging (eg, liver-specific agents and more widespread use of higher magnetic field strength) have been introduced (13–16). To account for these improvements, data for techniques were analyzed separately, and, if possible, subgroup analyses were performed. In addition, data were extracted and analyzed on a per-patient, as well as on a per-lesion, basis. This is important, since the treatment policy depends not only on distinguishing patients with or without liver metastases but also on the number, size, location, and surgical margin of the liver metastases in the first group. Thereby, the prevalence of liver metastases in patients with primary colorectal cancer is high.

A potential limitation of any meta-analysis is the possibility of publication bias. We did not perform any analysis for detection of and adjustment for publication bias. A recent systematic review conducted by the Cochrane Collaboration showed that a large number of methods have been developed, and when the methods are compared with one another, results can provide different estimates in terms of direction and magnitude of pub-

lication bias (89). In addition, studies about publication bias focus mostly on randomized trials, and these studies are registered; the registration of studies about diagnostic studies is either limited or difficult to achieve. We attempted to examine publication bias by using an evaluation of whether the size of studies was associated with the results for diagnostic accuracy. In particular, small studies with optimistic results may be published more easily than small studies with unfavorable results. Larger studies with optimistic results may also be published more easily than larger studies with unfavorable results, but this difference usually is smaller. There was no association between sample size and diagnostic performance.

Characteristics of the patients, such as the stage of disease, differentiation between synchronous and metachronous liver metastases, presence of extrahepatic disease, and age or sex distribution, are also important for diagnostic accuracy, but because of variation in data presentation or incomplete reporting of data, the effect of these variables could not be examined.

The reference standard used in this systematic review ranged from histopathologic analysis to follow-up US. It was impossible to examine the effect of each reference test on diagnostic accuracy. As stated in the Standards for Reporting of Diagnostic Accuracy initiative, a refer-

ence standard can be either a single method or a combination of methods to establish the presence of the target condition (90). The major problem, however, was the absence of critical information, such as data about the execution of the reference test, the confidence rating, or the characterization of lesions, and these data were insufficiently described or not described in a large subset of articles. This has also been described by authors of other meta-analyses (91,92). Therefore, the Standards for Reporting of Diagnostic Accuracy initiative was developed to improve the quality of the reporting of diagnostic studies. The items in the checklist and the flowchart can help authors in describing essential elements of the design and conduct of the study, the execution of tests, and the results.

Another limitation is the consideration of 2×2 tables for different readers, for multiple observations per reader, and for multiple CT and MR imaging techniques as separate data sets. This has been performed to avoid selection bias. We are aware of the dependency in data sets from the same patient population. Analysis of this dependency is not possible with our software, as the random-effects linear regression model with mixed-effects approach is able to adjust for this potential dependency only if the same numbers of data sets in each study are available. We examined this correlation by using the empirical standard error calculated by using the "sandwich estimator," which is possible with the software for the regression model with mixed-effects approach (25). We also used this approach to adjust for correlation between imaging modalities performed in the same patient population.

A final limitation of this study was the absence of information on specificity values. On a per-patient basis, specificity can be calculated. To minimize selection bias, we included all studies in which data were presented about colorectal liver metastases (also other conditions) and not studies in which only colorectal liver metastases were presented. We summarized and analyzed only data on colorectal liver metastases. The specificity can be underestimated in studies in which only colorectal liver metastases are evaluated and overestimated in studies in which other conditions are also evaluated.

Kinkel et al (17) also performed a meta-analysis to compare US, CT, MR imaging, and FDG PET in the detection of liver metastases. In studies with specificity higher than 85%, the sensitivity for US was 55%, and that for CT was 72%, that

for MR imaging was 76%, and that for FDG PET was 90%. They, however, analyzed hepatic metastases that originated from colorectal, gastric, and esophageal cancers. Treatment policies and reference standards differ for liver metastases of different origins, and, therefore, the importance of certain findings differs. In addition, the literature search was performed only in the MEDLINE database, and some study design characteristics were used to include studies, thereby introducing selection bias. Moreover, they combined results on a per-patient basis and on a per-lesion basis, thereby causing overestimation of the diagnostic accuracy of FDG PET. In our per-patient analysis, FDG PET had a significantly higher sensitivity estimate (94.6%) compared with that of helical CT, nonhelical CT, and MR imaging, although the sensitivity estimate was comparable on a per-lesion basis. Analysis of combined data would, therefore, lead to overestimation of the diagnostic accuracy of FDG PET.

Although on a per-patient basis, FDG PET was found to be most accurate, the treatment policy depends not only on distinguishing patients with liver metastases from patients without liver metastases but also mainly on the number, size, location, and surgical margin of the liver metastases. In addition, FDG PET mostly was performed in selected patients (10,49) or with a long time between CT and FDG PET (>4 weeks) (42,46,79), thereby increasing the detection of liver metastases by using FDG PET compared with the detection by using CT. In general, the time between diagnostic tests should be short to avoid differences in disease status. In several studies, scans were not corrected for attenuation with CT (46,64,78).

Because of its noninvasive character, low cost, and widespread availability, US can be used to help distinguish patients with diffuse disease who are not eligible for curative treatment from the group of patients with no liver metastases or the group with a limited number of them. The patients in the latter group should undergo CT, MR imaging, or FDG PET. On a per-lesion basis, helical CT, MR imaging at 1.0 T, MR imaging at 1.5 T, and FDG PET were comparable and significantly more accurate compared with nonhelical CT. In the subgroup analyses, however, SPIO-enhanced MR imaging and gadolinium-enhanced MR imaging were significantly more accurate compared with nonenhanced MR imaging and helical CT performed with a contrast agent that has 45 g or less of iodine.

The choice between portal phase helical CT performed with more than 45 g of iodine and MR imaging with a gadolinium-based contrast agent or SPIO should, therefore, also depend on availability and expertise, and not on diagnostic accuracy only. The role of FDG PET at this moment is limited and, therefore, it will be used mainly as an additional imaging modality for detection of extrahepatic disease.

References

- Jatzko GR, Lisborg PH, Stettner HM, Klimpfinger MH. Hepatic resection for metastases from colorectal carcinoma: a survival analysis. *Eur J Cancer* 1995;31A:41-46.
- Castells A, Bessa X, Daniels M, et al. Value of postoperative surveillance after radical surgery for colorectal cancer: results of a cohort study. *Dis Colon Rectum* 1998;41:714-723.
- Komborozos VA, Skrekas GJ, Pissiotis CA. The contribution of follow-up programs in the reduction of mortality of rectal cancer recurrences. *Dig Surg* 2001;18:403-408.
- Safi F, Link KH, Beger HG. Is follow-up of colorectal cancer patients worthwhile? *Dis Colon Rectum* 1993;36:636-643.
- Ketteniss M, Schutz G, Ulrich B. Costs and efficiency of a tumor follow-up program for the detection of colorectal liver metastases. *Int J Colorectal Dis* 2001;16:28-31.
- Albrecht T, Blomley MJ, Burns PN, et al. Improved detection of hepatic metastases with pulse-inversion US during the liver-specific phase of SHU 508A: multicenter study. *Radiology* 2003;227:361-370.
- Balfe DM. Hepatic metastases from colorectal cancer: radiologic strategies for improved selection. *Radiology* 1992;185:18-19.
- Bellin MF, Zaim S, Auberton E, et al. Liver metastases: safety and efficacy of detection with superparamagnetic iron oxide in MR imaging. *Radiology* 1994;193:657-663.
- Bernatik T, Strobel D, Hahn EG, Becker D. Detection of liver metastases: comparison of contrast-enhanced wide-band harmonic imaging with conventional ultrasonography. *J Ultrasound Med* 2001;20:509-515.
- Boykin KN, Zibari GB, Lillien DL, McMillan RW, Aultman DF, McDonald JC. The use of FDG-positron emission tomography for the evaluation of colorectal metastases of the liver. *Am Surg* 1999;65:1183-1185.
- Caudana R, Morana G, Pirovano GP, et al. Focal malignant hepatic lesions: MR imaging enhanced with gadolinium benzylxypropionictetra-acetate (BOPTA)—preliminary results of phase II clinical application. *Radiology* 1996;199:513-520.
- Ch'en IY, Katz DS, Jeffrey RB Jr, et al. Do arterial phase helical CT images improve detection or characterization of colorectal liver metastases? *J Comput Assist Tomogr* 1997;21:391-397.
- Kamel IR, Choti MA, Horton KM, et al. Surgically staged focal liver lesions: accuracy and reproducibility of dual-phase helical CT for detection and characterization. *Radiology* 2003;227:752-757.
- Kim MJ, Kim JH, Chung JJ, Park MS, Lim JS, Oh YT. Focal hepatic lesions: detection

- and characterization with combination gadolinium- and superparamagnetic iron oxide-enhanced MR imaging. *Radiology* 2003;228:719-726.
15. Dill-Macky MJ, Burns PN, Khalili K, Wilsson SR. Focal hepatic masses: enhancement patterns with SH U 508A and pulse-inversion US. *Radiology* 2002;222:95-102.
 16. Even-Sapir E, Lerman H, Figer A, et al. Role of (18)F-FDG dual-head gamma-camera coincidence imaging in recurrent or metastatic colorectal carcinoma. *J Nucl Med* 2002;43:603-609.
 17. Kinkel K, Lu Y, Both M, Warren RS, Thoeni RF. Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): a meta-analysis. *Radiology* 2002;224:748-756.
 18. Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53:65-69.
 19. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 1997;350:185-186.
 20. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25. <http://www.biomedcentral.com/1471-2288/3/25>. Published November 10, 2003.
 21. Van Houwelingen HC, Zwinderman KH, Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993;12:2273-2284.
 22. Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;21:589-624.
 23. Liang KY, Zeger SL. Regression analysis for correlated data. *Annu Rev Public Health* 1993;14:43-68.
 24. Acunas B, Rozanes I, Acunas G, et al. Preoperative CT staging of colon carcinoma (excluding the recto-sigmoid region). *Eur J Radiol* 1990;11:150-153.
 25. Lundstedt C, Ekberg H, Hederstrom E, Tranberg KG. The accuracy of computed tomography of the liver in colorectal carcinoma. *Clin Radiol* 1990;42:335-339.
 26. Machi J, Isomoto H, Kurohiji T, et al. Accuracy of intraoperative ultrasonography in diagnosing liver metastasis from colorectal cancer: evaluation with postoperative follow-up results. *World J Surg* 1991;15:551-556.
 27. Wernecke K, Rummeny E, Bongartz G, et al. Detection of hepatic masses in patients with carcinoma: comparative sensitivities of sonography, CT, and MR imaging. *AJR Am J Roentgenol* 1991;157:731-739.
 28. Yamaguchi A, Ishida T, Nishimura G, et al. Detection by CT during arterial portography of colorectal cancer metastases to liver. *Dis Colon Rectum* 1991;34:37-40.
 29. Oudkerk M, van Ooijen B, Mali SP, et al. Liver metastases from colorectal carcinoma: detection with continuous CT angiography. *Radiology* 1992;185:157-161.
 30. Soyer P, Levesque M, Elias D, Zeitoun G, Roche A. Preoperative assessment of resectability of hepatic metastases from colonic carcinoma: CT portography vs sonography and dynamic CT. *AJR Am J Roentgenol* 1992;159:741-744.
 31. Soyer P, Levesque M, Elias D, Zeitoun G, Roche A. Detection of liver metastases from colorectal cancer: comparison of intraoperative US and CT during arterial portography. *Radiology* 1992;183:541-544.
 32. Kerner BA, Oliver GC, Eisenstat TE, Rubin RJ, Salvati EP. Is preoperative computerized tomography useful in assessing patients with colorectal carcinoma? *Dis Colon Rectum* 1993;36:1050-1053.
 33. Knol JA, Marn CS, Francis IR, et al. Comparisons of dynamic infusion and delayed computed tomography, intraoperative ultrasound, and palpation in the diagnosis of liver metastases. *Am J Surg* 1993;165:81-87.
 34. Ohlsson B, Tranberg KG, Lundstedt C, Ekberg H, Hederstrom E. Detection of hepatic metastases in colorectal cancer: a prospective study of laboratory and imaging methods. *Eur J Surg* 1993;159:275-281.
 35. Soyer P, Elias D, Zeitoun G, Roche A, Levesque M. Surgical treatment of hepatic metastases: impact of intraoperative sonography. *AJR Am J Roentgenol* 1993;160:511-514.
 36. Stewart PJ, Chu JM, Kos SC, Chapuis PH, Bokey EL. Intra-operative ultrasound for the detection of hepatic metastases from colorectal cancer. *Aust N Z J Surg* 1993;63:530-534.
 37. Paul MA, Mulder LS, Cuesta MA, et al. Impact of intraoperative ultrasonography on treatment strategy for colorectal cancer. *Br J Surg* 1994;81:1660-1663.
 38. Leen E, Angerson WJ, Wotherspoon H, et al. Detection of colorectal liver metastases: comparison of laparotomy, CT, US, and Doppler perfusion index and evaluation of postoperative follow-up results. *Radiology* 1995;195:113-116.
 39. Carter R, Hemingway D, Cooke TG, et al. A prospective study of six methods for detection of hepatic colorectal metastases. *Ann R Coll Surg Engl* 1996;78:27-30.
 40. van Ooijen B, Oudkerk M, Schmitz PI, Wiggers T. Detection of liver metastases from colorectal carcinoma: is there a place for routine computed tomography arteriography? *Surgery* 1996;119:511-516.
 41. Lai DT, Fulham M, Stephen MS, et al. The role of whole-body positron emission tomography with [18F]fluorodeoxyglucose in identifying operable colorectal cancer metastases to the liver. *Arch Surg* 1996;131:703-707.
 42. Delbeke D, Vitola JV, Sandler MP, et al. Staging recurrent metastatic colorectal carcinoma with PET. *J Nucl Med* 1997;38:1196-1201.
 43. Ogunbiyi OA, Flanagan FL, Dehdashti F, et al. Detection of recurrent and metastatic colorectal cancer: comparison of positron emission tomography and computed tomography. *Ann Surg Oncol* 1997;4:613-620.
 44. Abdel-Nabi H, Doerr RJ, Lamonica DM, et al. Staging of primary colorectal carcinomas with fluorine-18 fluorodeoxyglucose whole-body PET: correlation with histopathologic and CT findings. *Radiology* 1998;206:755-760.
 45. Yasuda S, Makuuchi Y, Sadahiro S, et al. Colorectal cancer recurrence in the liver: detection by PET. *Tokai J Exp Clin Med* 1998;23:167-171.
 46. Valk PE, Abella-Columna E, Haseman MK, et al. Whole-body PET imaging with [18F]fluorodeoxyglucose in management of recurrent colorectal cancer. *Arch Surg* 1999;134:503-511.
 47. Zhuang H, Sinha P, Pourdehnad M, et al. The role of positron emission tomography with fluorine-18-deoxyglucose in identifying colorectal cancer metastases to liver. *Nucl Med Commun* 2000;21:793-798.
 48. Whiteford MH, Whiteford HM, Yee LF, et al. Usefulness of FDG-PET scan in the assessment of suspected metastatic or recurrent adenocarcinoma of the colon and rectum. *Dis Colon Rectum* 2000;43:759-767.
 49. Hung GU, Shiau YC, Tsai SC, et al. Value of 18F-fluoro-2-deoxyglucose positron emission tomography in the evaluation of recurrent colorectal cancer. *Anticancer Res* 2001;21:1375-1378.
 50. Johnson K, Bakhsh A, Young D, Martin TE Jr, Arnold M. Correlating computed tomography and positron emission tomography scan with operative findings in metastatic colorectal cancer. *Dis Colon Rectum* 2001;44:354-357.
 51. Hagspiel KD, Neidl KF, Eichenberger AC, Weder W, Marinck B. Detection of liver metastases: comparison of superparamagnetic iron oxide-enhanced and unenhanced MR imaging at 1.5 T with dynamic CT, intraoperative US, and percutaneous US. *Radiology* 1995;196:471-478.
 52. Vitola JV, Delbeke D, Sandler MP, et al. Positron emission tomography to stage suspected metastatic colorectal carcinoma to the liver. *Am J Surg* 1996;171:21-26.
 53. Strotzer M, Gmeinwieser J, Schmidt J, et al. Diagnosis of liver metastases from colorectal adenocarcinoma: comparison of spiral-CTAP combined with intravenous contrast-enhanced spiral-CT and SPIO-enhanced MR combined with plain MR imaging. *Acta Radiol* 1997;38:986-992.
 54. Lencioni R, Donati F, Cioni D, et al. Detection of colorectal liver metastases: prospective comparison of unenhanced and ferumoxides-enhanced magnetic resonance imaging at 1.5 T, dual-phase spiral CT, and spiral CT during arterial portography. *MAGMA* 1998;7:76-87.
 55. Valls C, Lopez E, Guma A, et al. Helical CT versus CT arterial portography in the detection of hepatic metastasis of colorectal carcinoma. *AJR Am J Roentgenol* 1998;170:1341-1347.
 56. Rahusen FD, Cuesta MA, Borgstein PJ, et al. Selection of patients for resection of colorectal metastases to the liver using diagnostic laparoscopy and laparoscopic ultrasonography. *Ann Surg* 1999;230:31-37.
 57. Ward J, Naik KS, Guthrie JA, Wilson D, Robinson PJ. Hepatic lesion detection: comparison of MR imaging after the administration of superparamagnetic iron oxide with dual-phase CT by using alternative-free response receiver operating characteristic analysis. *Radiology* 1999;210:459-466.
 58. Schmidt J, Strotzer M, Fraunhofer S, Boedeker H, Ziringibl H. Intraoperative ultrasonography versus helical computed tomography and computed tomography with arteriography in diagnosing colorectal liver metastases: lesion-by-lesion analysis. *World J Surg* 2000;24:43-47.

59. Arulampalam T, Costa D, Visvikis D, et al. The impact of FDG-PET on the management algorithm for recurrent colorectal cancer. *Eur J Nucl Med* 2001;28:1758-1765.
60. Baulieu F, Bourlier P, Scotto B, et al. The value of immunoscintigraphy in the detection of recurrent colorectal cancer. *Nucl Med Commun* 2001;22:1295-1304.
61. Scott DJ, Guthrie JA, Arnold P, et al. Dual phase helical CT versus portal venous phase CT for the detection of colorectal liver metastases: correlation with intra-operative sonography, surgical and pathological findings. *Clin Radiol* 2001;56:235-242.
62. Valls C, Andia E, Sanchez A, et al. Hepatic metastases from colorectal cancer: preoperative detection and assessment of resectability with helical CT. *Radiology* 2001;218:55-60.
63. Haider MA, Amitai MM, Rappaport DC, et al. Multi-detector row helical CT in preoperative assessment of small (≤ 1.5 cm) liver metastases: is thinner collimation better? *Radiology* 2002;225:137-142.
64. Ruers TJ, Langenhoff BS, Neeleman N, et al. Value of positron emission tomography with [F-18]fluorodeoxyglucose in patients with colorectal liver metastases: a prospective study. *J Clin Oncol* 2002;20:388-395.
65. Pijl ME, Wasser MN, Joeke EC, van De Velde CJ, Bloem JL. Metastases of colorectal carcinoma: comparison of soft and hard copy helical CT interpretation. *Radiology* 2003;227:747-751.
66. Soyer P, Laissy JP, Sibert A, et al. Hepatic metastases: detection with multisection FLASH MR imaging during gadolinium chelate-enhanced arterial portography. *Radiology* 1993;189:401-405.
67. Soyer P, Laissy JP, Sibert A, et al. Focal hepatic masses: comparison of detection during arterial portography with MR imaging and CT. *Radiology* 1994;190:737-740.
68. Ward J, Chen F, Guthrie JA, et al. Hepatic lesion detection after superparamagnetic iron oxide enhancement: comparison of five T2-weighted sequences at 1.0 T by using alternative-free response receiver operating characteristic analysis. *Radiology* 2000;214:159-166.
69. del Frate C, Bazzocchi M, Morteale KJ, et al. Detection of liver metastases: comparison of gadobenate dimeglumine-enhanced and ferumoxides-enhanced MR imaging examinations. *Radiology* 2002;225:766-772.
70. Soyer P, Levesque M, Caudron C, et al. MRI of liver metastases from colorectal cancer vs. CT during arterial portography. *J Comput Assist Tomogr* 1993;17:67-74.
71. Peterson MS, Baron RL, Murakami T. Hepatic malignancies: usefulness of acquisition of multiple arterial and portal venous phase images at dynamic gadolinium-enhanced MR imaging. *Radiology* 1996;201:337-345.
72. Oudkerk M, van den Heuvel AG, Wielopolski PA, et al. Hepatic lesions: detection with ferumoxide-enhanced T1-weighted MR imaging. *Radiology* 1997;203:449-456.
73. Pijl MEJ, Wasser MNJM, van Persijn van Meerten EL, et al. Comparison of inversion-recovery gradient- and spin-echo and fast spin-echo techniques in the detection and characterization of liver lesions. *Radiology* 1998;209:427-434.
74. Semelka RC, Cance WG, Marcos HB, Mauro MA. Liver metastases: comparison of current MR techniques and spiral CT during arterial portography for detection in 20 surgically staged cases. *Radiology* 1999;213:86-91.
75. Matsuo M, Kanematsu M, Itoh K, et al. Detection of malignant hepatic tumors: comparison of gadolinium- and ferumoxide-enhanced MR imaging. *AJR Am J Roentgenol* 2001;177:637-643.
76. Ward J, Guthrie JA, Wilson D, et al. Colorectal hepatic metastases: detection with SPIO-enhanced breath-hold MR imaging—comparison of optimized sequences. *Radiology* 2003;228(3):709-718.
77. Yang M, Martin DR, Karabulut N, Frick MP. Comparison of MR and PET imaging for the evaluation of liver metastases. *J Magn Reson Imaging* 2003;17:343-349.
78. Schiepers C, Penninckx F, De Vadder N, et al. Contribution of PET in the diagnosis of recurrent colorectal cancer: comparison with conventional imaging. *Eur J Surg Oncol* 1995;21:517-522.
79. Flanagan FL, Dehdashti F, Ogunbiyi OA, Kodner IJ, Siegel BA. Utility of FDG-PET for investigating unexplained plasma CEA elevation in patients with colorectal cancer. *Ann Surg* 1998;227:319-323.
80. Staib L, Schirrmeister H, Reske SN, Beger HG. Is (18)F-fluorodeoxyglucose positron emission tomography in recurrent colorectal cancer a contribution to surgical decision making? *Am J Surg* 2000;180:1-5.
81. Willkomm P, Bender H, Bangard M, et al. FDG PET and immunoscintigraphy with 99mTc-labeled antibody fragments for detection of the recurrence of colorectal carcinoma. *J Nucl Med* 2000;41:1657-1663.
82. Rohren EM, Paulson EK, Hagge R, et al. The role of F-18 FDG positron emission tomography in preoperative assessment of the liver in patients being considered for curative resection of hepatic metastases from colorectal cancer. *Clin Nucl Med* 2002;27:550-555.
83. Rydzewski B, Dehdashti F, Gordon BA, et al. Usefulness of intraoperative sonography for revealing hepatic metastases from colorectal cancer in patients selected for surgery after undergoing FDG PET. *AJR Am J Roentgenol* 2002;178:353-358.
84. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;13:313-321.
85. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;12:1293-1316.
86. Rutter CM, Gatsonis CA. Regression methods for meta-analysis of diagnostic test data. *Acad Radiol* 1995;2(suppl 1):S48-S56.
87. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865-2884.
88. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-1066.
89. Pham B, Platt R, McAuley L, Klassen TP, Moher D. Is there a "best" way to detect and minimize publication bias? an empirical evaluation. *Eval Health Prof* 2001;24:109-125.
90. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Radiology* 2003;226:24-28.
91. de Vries SO, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. *Acad Radiol* 1996;3:361-369.
92. Nelemans PJ, Leiner T, de Vet HC, van Engelshoven JM. Peripheral arterial disease: meta-analysis of the diagnostic performance of MR angiography. *Radiology* 2000;217:105-114.