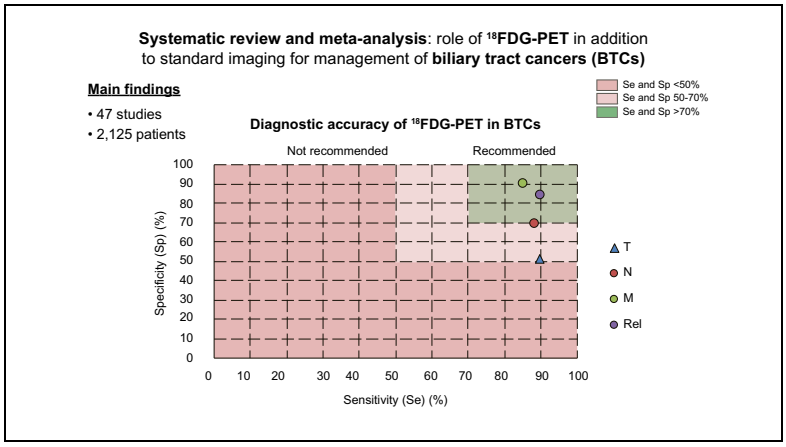
Research Article
Cancer

JOURNAL
OF HEPATOLOGY

# $^{18}$F-fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) for patients with biliary tract cancer: Systematic review and meta-analysis

## Graphical abstract




## Highlights

- Role of $^{18}$FDG-PET in diagnosis (T), staging (N/M) and relapse of BTC was assessed.
- $^{18}$FDG-PET is not recommended for diagnosis (T) in the absence of cytology/histology.
- $^{18}$FDG-PET should be incorporated into current guidelines for staging (N/M) and relapse.
- $^{18}$FDG-PET should be used for staging (N/M) if identification of occult sites of disease will alter management.
- $^{18}$FDG-PET should be used to identify relapse if suspicion remains following standard imaging.

## Authors

Angela Lamarca, Jorge Barriuso, Amarjot Chander, ..., Derek ÓReilly, Prakash Manoharan, Juan W Valle

## Correspondence

angela.lamarca@christie.nhs.uk
(A. Lamarca) juan.valle@christie.nhs.uk
(J.W Valle)

## Lay summary

A positron emission tomography (PET scan), using $^{18}$F-fluorodeoxyglucose ($^{18}$FDG), can help doctors identify areas of cancer in the body by highlighting ‘‘hot spots’’. These hotspots may be cancerous (true positive) but may also be non-cancerous, like inflammation (false positive). We show that PET scans are useful to assess how far advanced the cancer is (by assessing spread to lymph glands and to other organs) and also to identify if the cancer has recurred (for example after surgery), thus helping doctors to make treatment decisions. However, a biopsy is still needed for the initial diagnosis of a biliary tract cancer, because of the high chance of a ‘‘false positive’’ with PET scans.

http://dx.doi.org/10.1016/j.jhep.2019.01.038
 

Research Article
Cancer

JOURNAL OF HEPATOLOGY

# $^{18}$F-fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) for patients with biliary tract cancer: Systematic review and meta-analysis

Angela Lamarca[1,2,*], Jorge Barriuso[1,2], Amarjot Chander[3], Mairéad G McNamara[1,2], Richard A Hubner[1,2], Derek ÓReilly[4], Prakash Manoharan[3], Juan W Valle[1,2,*]

[1]*Medical Oncology Department, The Christie NHS Foundation Trust, Manchester, UK;* [2]*Division of Cancer Sciences, University of Manchester, UK;* [3]*Radiology and Nuclear Medicine Department, The Christie NHS Foundation Trust, Manchester, UK;* [4]*HPB Surgery Department, Manchester Royal Infirmary, Central Manchester University Hospitals, Manchester, UK*

**Background & Aims**: The role of $^{18}$F-fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) in the diagnosis and staging of patients with biliary tract cancers (BTCs) remains controversial, so we aimed to provide robust information on the utility of $^{18}$FDG-PET in the diagnosis and management of BTC.

**Methods**: This systematic review and meta-analysis explored the diagnostic test accuracy of $^{18}$FDG-PET as a diagnostic tool for diagnosis of primary tumour, lymph node invasion, distant metastases and relapsed disease. Subgroup analysis by study quality and BTC subtype were performed. Changes in management based on $^{18}$FDG-PET and impact of maximum standardised uptake values (SUVmax) on prognosis were also assessed. A random effects model was used for meta-analyses.

**Results**: A total of 2,125 patients were included from 47 eligible studies. The sensitivity (Se) and specificity (Sp) of $^{18}$FDG-PET for the diagnosis of primary tumour were 91.7% (95% CI 89.8–93.2) and 51.3% (95% CI 46.4–56.2), respectively, with an area under the curve (AUC) of 0.8668. For lymph node invasion, Se was 88.4% (95% CI 82.6–92.8) and Sp was 69.1% (95% CI 63.8–74.1); AUC 0.8519. For distant metastases, Se was 85.4% (95% CI 79.5–90.2) and Sp was 89.7% (95% CI 86.0–92.7); AUC 0.9253. For relapse, Se was 90.1% (95% CI 84.4–94.3) and Sp was 83.5% (95% CI 74.4–90.4); AUC 0.9592. No diagnostic threshold effect was identified. Meta-regression did not identify significant sources of heterogeneity. Sensitivity analysis revealed no change in results when analyses were limited to studies with low risk of bias/concern. The pooled proportion of change in management was 15% (95% CI 11–20); the majority (78%) due to disease upstaging. Baseline high SUVmax was associated with worse survival (pooled hazard ratio of 1.79; 95% CI 1.37–2.33; $p$ <0.001).

**Conclusions**: There is evidence to support the incorporation of $^{18}$FDG-PET into the current standard of care for the staging (lymph node and distant metastases) and identification of relapse in patients with BTC to guide treatment selection; especially if the identification of occult sites of disease would change management, or if diagnosis of relapse remains unclear following standard of care imaging. The role for diagnosis of the primary tumour remains controversial due to low sensitivity and $^{18}$FDG-PET should not be considered as a replacement for pathological confirmation in this setting.

**Lay summary**: A positron emission tomography (PET scan), using $^{18}$F-fluorodeoxyglucose ($^{18}$FDG), can help doctors identify areas of cancer in the body by highlighting "hot spots". These hotspots may be cancerous (true positive) but may also be non-cancerous, like inflammation (false positive). We show that PET scans are useful to assess how far advanced the cancer is (by assessing spread to lymph glands and to other organs) and also to identify if the cancer has recurred (for example after surgery), thus helping doctors to make treatment decisions. However, a biopsy is still needed for the initial diagnosis of a biliary tract cancer, because of the high chance of a "false positive" with PET scans.



## Introduction

### Introduction to biliary tract cancer

Biliary tract cancers (BTCs) (including cholangiocarcinoma and cancers of the ampulla of Vater and gallbladder) are considered low-incidence malignancies, accounting for approximately 0.7% of all malignant tumours in adults. However, data from the past 25 years suggest that, predominantly due to a rise in diagnosis of intrahepatic cholangiocarcinoma, both incidence and mortality are increasing.[1,2]

A minority of patients (around 20%) are diagnosed with early stage disease when curative resection is possible.[3] Due to the high risk of relapse after curative resection, adjuvant chemotherapy is recommended.[4–6] Unfortunately, more than 65% of patients are diagnosed with unresectable disease, which is associated with a poor prognosis; the 5-year overall survival rate for stage III and IV are 10% and 0%, and palliative chemotherapy is the only available treatment option.[3] In 2010, the results of the phase III randomised NCRN ABC-02 trial, established cisplatin and gemcitabine as the reference regimen for first-line therapy of advanced BTC.[7] The potential role of triple chemotherapy in the first-line setting is also being

Keywords: FDG-PET; Diagnosis; Biliary; SUV; Cholangiocarcinoma.
*Received 5 November 2018; received in revised form 10 January 2019; accepted 30 January 2019; available online 21 February 2019*

* Corresponding authors. Address: Medical Oncology Department, The Christie NHS Foundation Trust, Wilmslow Road, Manchester M20 4BX, UK.
*E-mail addresses:* angela.lamarca@christie.nhs.uk (A. Lamarca), juan.valle@christie.nhs.uk (J.W Valle).

explored.[8] Based on the lack of quality evidence supporting second-line chemotherapy after progression on the cisplatin and gemcitabine combination,[9,10] clinical trials are ongoing (*e.g.* ABC-06; NCT01926236, *etc.*). Development of new-targeted therapies is also awaited.[11–14]

### Current pathway for diagnosis

Patients diagnosed with BTC commonly present with non-specific symptoms; therefore, diagnosis is often challenging. Patients with suspected BTCs are assessed with ultrasound (US), contrast-enhanced computerised tomography (CT), and magnetic resonance imaging (MRI)[15–20] for tumour staging (according to TNM[21]) followed by biopsy or cytology (when feasible) for confirmation of invasive malignancy. Based on the information available, treatment is planned accordingly.

Although the current gold standard for diagnosis of malignancy relies on pathology (histology/cytology), there are two exceptions that would apply only when biopsies are repeatedly non-diagnostic due to challenges of sample acquisition. First, patients with suspected non-pathology-confirmed BTC may be assumed to have a malignant diagnosis if there is evidence of distant metastases. Secondly, the same may apply if benign (*i.e.* no change/growth over time, no distant spread) or malignant (progression, growth in size, distant spread) behaviour is identified on follow-up.

Unfortunately, for some patients, the information provided by the imaging techniques described above is still insufficient for diagnosis and staging of BTC. $^{18}$F-fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) has been postulated as an additional useful tool in selected situations which could potentially provide useful information for patients who have undergone the aforementioned tests.[22,23]

### Potential role of positron emission tomography in BTC

PET is a highly sensitive imaging method used to detect metabolic processes, usually employed for selected patients as an add-on test following a specific diagnostic pathway. $^{18}$F-fluorodeoxyglucose ($^{18}$FDG) is a glucose analogue and a positron tracer. $^{18}$FDG-PET is a method of imaging that utilises glucose metabolism to assess a variety of physiological and disease processes. The use of standardised uptake values (SUVs) is common practice in clinical oncology when reporting results of PET imaging.[24] SUV provides information regarding metabolic activity of the tumour. In oncology, SUVs provide objective and normalised results adjusted to potential variability introduced by factors such as patient size or the amount of injected radiotracer. It is worth mentioning that $^{18}$FDG-PET usually (however not always) incorporates CT ($^{18}$FDG-PET-CT) in order to correct signals according to attenuation and to allow better anatomical localisation.

The role of $^{18}$FDG-PET in the staging and management of a number of malignancies such as lymphoma, head and neck or lung cancer has been previously defined.[25–27]

In contrast, the role of $^{18}$FDG-PET in BTC remains controversial. One of the potential limitations of the use of $^{18}$FDG-PET in patients with BTC may be the false positive results related to biliary stenting, biliary sepsis and local infection, making the interpretation of $^{18}$FDG-PET results challenging.[22,28,29] Multiple small retrospective and prospective series have suggested a potential role for $^{18}$FDG-PET in BTC diagnosis and staging.[30,31] According to the literature, the results of an $^{18}$FDG-PET scan may change the treatment plan in around 20% of patients with resectable BTCs, avoiding unnecessary non-curative resections,[32] with significant implications for individual patients, as well as for the health economy. However, these findings have not been confirmed in larger series. $^{18}$FDG-PET has also been proposed to help with the diagnosis of BTC relapse, but its role is not clear from the available individual studies.[33,34] Based on these small datasets, some countries and centres utilise $^{18}$FDG-PET in the management of BTC,[35] even though the statistical power of individual studies was not sufficient to change practice globally.

Therefore, $^{18}$FDG-PET is currently not used in routine clinical practice. This systematic review and meta-analysis of available studies aims to provide clinicians with more robust information regarding the applicability of $^{18}$FDG-PET in addition to the current standard of care imaging/diagnostic tests utilised (US/CT/MR) in daily clinical practice for patients who have had previous imaging/biopsy for suspected primary or relapsed BTC.

## Material and methods

### Objectives

This systematic review aimed to evaluate the diagnostic test accuracy (DTA) of $^{18}$FDG-PET in addition to the current standard of care imaging/diagnostic tests utilised (US/CT/MR) in the following areas: A) $^{18}$FDG-PET as diagnostic tool for primary tumour (T) malignancy in patients with suspected BTC; this could be of interest when a biopsy is not feasible. B) $^{18}$FDG-PET as a tool for diagnosis of lymph node (N) and distant metastases (M) which could provide information for staging and allow better selection of patients with apparently resectable disease. C) $^{18}$FDG-PET as a tool for identification of relapsed disease.

In addition, the impact of $^{18}$FDG-PET in changing clinical decisions (measured as percentage of participants in whom the management changed based on the $^{18}$FDG-PET results) and prognostic impact of baseline maximum SUV (SUVmax) on survival (overall survival [OS], progression-free survival [PFS] or disease-free survival [DFS]) in studies where this data was available was assessed.

### Search methods for identification of studies

This systematic review was registered in the PROSPERO database, number CRD42018110366.[36] A search (last updated on 19 September 2018) to identify eligible studies was undertaken using the MEDLINE database;[37] search strategy: "fdg pet biliary tract OR ((18f fdg[MeSH Terms] AND Biliary tract[MeSH Terms])) OR ((18f fdg[MeSH Terms] AND Gallbladder Neoplasms [MeSH Terms])) OR ((18f fdg[MeSH Terms] AND Cholangiocarcinoma [MeSH Terms])) OR ((18f fdg[MeSH Terms] AND Ampulla of Vater [MeSH Terms])) ". Abstract from ASCO[38] and ESMO[39] were also screened (search last updated on 20 September 2018); search strategy: "(FDG and cholangiocarcinoma) OR (FDG and gallbladder) OR (FDG and biliary) OR (FDG and ampulla) ".

Potentially-eligible studies were selected from the 2 aforementioned searches by reviewing the abstracts. All studies meeting the inclusion criteria were included irrespective of year or language of publication. If non-English articles were identified, those studies were included if the mandatory data was available in the abstract; alternatively, authors were contacted for further details.

### Study eligibility

This systematic review and meta-analysis focused on BTC (including cholangiocarcinoma, gallbladder and ampullary

malignancies) studies. Prospective and retrospective studies with data available for patients diagnosed with BTC were eligible. Case reports and review publications were excluded. The following data was required for studies to be eligible for each of the main objectives:

- DTA assessment: Eligible studies were required to provide data that allowed the 2 × 2 table to be constructed (true positive, false positive, false negative and true negative). See further details in Supplementary Material 1. "Patient" was the employed unit for the 2 × 2 tables. In studies employing other units (such as for example "lesion") data per patient was extracted from the manuscript (if available); alternatively, authors were contacted, or studies excluded if data per "patient" were unavailable.
- Impact on clinical management: Studies reporting data of change in management measured as percentage of participants in whom the management changed based on the $^{18}$FDG-PET results were eligible.
- Prognostic impact: Eligible studies were those reporting data on impact of baseline SUVmax on survival (OS, PFS or DFS) measured as hazard ratio (HR) or odds ratio (OR) ± 95% confidence interval (CI) or *p* value.

If none of these figures were detailed in the abstract (if full article was not available) or in the full article, studies were deemed ineligible.

In addition to the above, eligible studies were required to meet adequate criteria for the reference standard. For this systematic review, a pathology-proven malignancy (either by cytology or histology) was considered to be the reference standard. Therefore, studies were excluded if the stated reference standard was other than biopsy-proven cancer. See Supplementary Material 2 for further definition and exceptions to the reference standard. For prognostic factors and change in management role, the reference standard definition for T, N or M were applied according to whether patients with localised (T/N) or metastatic (M) disease were included in that study.

The study selection process was summarised in a flow diagram as per PRISMA criteria[40,41] where reasons for exclusion of studies were also recorded.

**Data extraction and assessment of methodological quality**

Data were extracted using predefined data collection tools (Supplementary Material 3) by one of the authors (AL) and queries discussed with a second author (JB); a third author acted as an arbiter in the case of disagreement (JWV).

"Quality" was separated into "risk of bias" and "concerns regarding applicability" following QUADAS-2 guidance.[42] Definitions were predefined and included the 4 main domains: participant selection, index test, reference standard, and flow and timing (Supplementary Material 4). Authors agreed to score the risk of bias as high or unclear if there was at least one 'no' or 'unclear' response to a signalling question for a given domain.

**Statistical analysis and data synthesis**

*Descriptive analysis*

Weighted mean and the 95% CI of age were calculated, weighted according to the number of patients included in the studies (analytic weighting). The same method was employed for comparison of SUVmax between cancer and non-cancer patients in studies with such data reported. Student's *t* test was used to compare SUVmax between groups. Stata v.12 software was employed for this analysis.[43]

*DTA meta-analysis*

Pooled sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR-) and diagnostic odds ratio (DOR) for $^{18}$FDG-PET compared with reference standards were calculated for each one of the DTA endpoints and for each predefined subgroup using Meta-DiSc v.1.4 software.[44] Together with pooled results, 95% CIs and heterogeneity (in the form of the inconsistency ($I^2$) and heterogeneity *p* value) were reported. Data was plotted graphically in forest plots and summary receiver operating characteristic (SROC space); size of dots were proportional to study size. In order to use data from all studies found to be adequate for the DTA meta-analysis (including those with value of "zero" in any cell), a value of "0.5" was added into every 2 × 2 table cell.[44]

Since it was expected that the eligible manuscripts would specify an explicit SUVmax cut-off for definition of "positive or "negative" results and due to the fact that this threshold was expected to vary between studies, analysis of diagnostic threshold for BTC in each DTA endpoint was assessed by calculating the Spearman correlation coefficient (rho) between sensitivity and specificity (using Meta-Disk v.1.4 software[44]). It was considered that the diagnostic threshold effect existed if a strong inverse correlation (defined as rho below −0.4) appeared.[45] The hierarchical summary receiver operating characteristics (HSROC) model was also employed for the analysis with an HSROC curve estimation (using "Metandi" command in Stata v.12 software[43]);[46] mean accuracy (lambda) parameter was employed for interpretation of HSROC results.[47] In addition, if a minimum of 4 studies (minimum number of studies required by the "Metandi" command in Stata v.12 software[43]) were identified to report accuracy data using a common SUVmax cut-off value, a bivariate model was used to estimate pooled DTA measurements of specific SUVmax cut-offs.

The random effects model rather than fixed effects model was employed, since heterogeneity between studies was expected to be present.

Pre-planned subgroup analyses for DTA meta-analysis included analysis by primary tumour site (cholangiocarcinoma, intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, hilar cholangiocarcinoma, gallbladder cancer and ampullary malignancies) and DTA endpoints (identification of primary tumour (T), regional lymph node (N), distant metastases (M) or identification of relapse).

Formal reporting bias analysis for DTA was not performed due to the inadequacy of the existing statistical method for its assessment in such studies.

*Change in patient management*

The proportion of changes with exact 95% CIs was calculated for each study. The Freeman-Tukey double arcsine transformation was chosen for the calculation of pooled estimates and 95% CIs.[48,49] Random effects pooled estimates were calculated in order to take into account heterogeneity between estimates[50] using R-Studio v.8.1 software.[51] Statistical heterogeneity among studies was evaluated.[52] Reporting bias (including publication bias) was assessed using funnel plots and the asymmetry test.[53]

*Prognostic role*

Survival data was meta-analysed using the random effects models in the RevMan v.5 software.[54] Heterogeneity ($I^2$ and *p* value) and presence of reporting bias (including publication bias by exploring funnel-plot asymmetry[54,55]) were assessed.

*Heterogeneity and sensitivity analysis*

The study quality (defined as per predefined QUADAS2 criteria) was considered as a main source of heterogeneity and included in the sensitivity analysis, in which analysis was repeated including studies with low risk/concern (high quality) only. Sensitivity analyses were performed for all predefined subgroup analysis.

Other main sources of heterogeneity were investigated, such as year of publication as a surrogate of PET imaging quality. The year 2005 was used as the cut-off and publications before then were compared with publications from 2005 (inclusive) onwards to investigate this hypothesis. Other relevant variables were included in subgroup analysis (in the case of data availability) such as: adequacy of the reference standard (rate of patients with confirmed cancer or benign disease as per pathology/cytology, radiology or follow-up results), rate of indeterminate/uninterpretable results (situations in which even in the presence of good quality imaging, PET findings are equivocal or situations in which PET findings could not be interpreted due to low quality imaging), and methodological quality (PET *vs.* PET-CT equipment, generation of equipment [PET camera], type of crystal utilised by the PET camera, amount of FDG injected, baseline glucose levels, prevalence of diabetes mellitus in the reported population, protocol utilised for post injection imaging [*e.g.* 60, 75 or 90 min], 2D *vs.* 3D reconstruction, utilisation of iodinated contrast, time of flight capability and the employment of a SUVs threshold [if yes, which]).

In order to formally investigate potential sources of heterogeneity other than the diagnostic threshold effect (assessed by the analysis of diagnostic threshold) and the impact of quality (assessed by sensitivity analyses), meta-regression including the above-mentioned co-variates was used (using Meta-Disk v.1.4 software[44]).[56] Meta-regression was performed for the whole BTC cohort including subgroups by DTA endpoint, but not including subgroups by primary site.

## Results

### Results of the search

A total of 231 abstracts were screened; and 71 selected for full text screening (Fig. 1). Of these 71 abstracts, 25 were excluded due to the following reasons: i) One review paper;[57] ii) Five studies did not report data for patients with BTC (data was provided together with other tumour types such as pancreatic can-

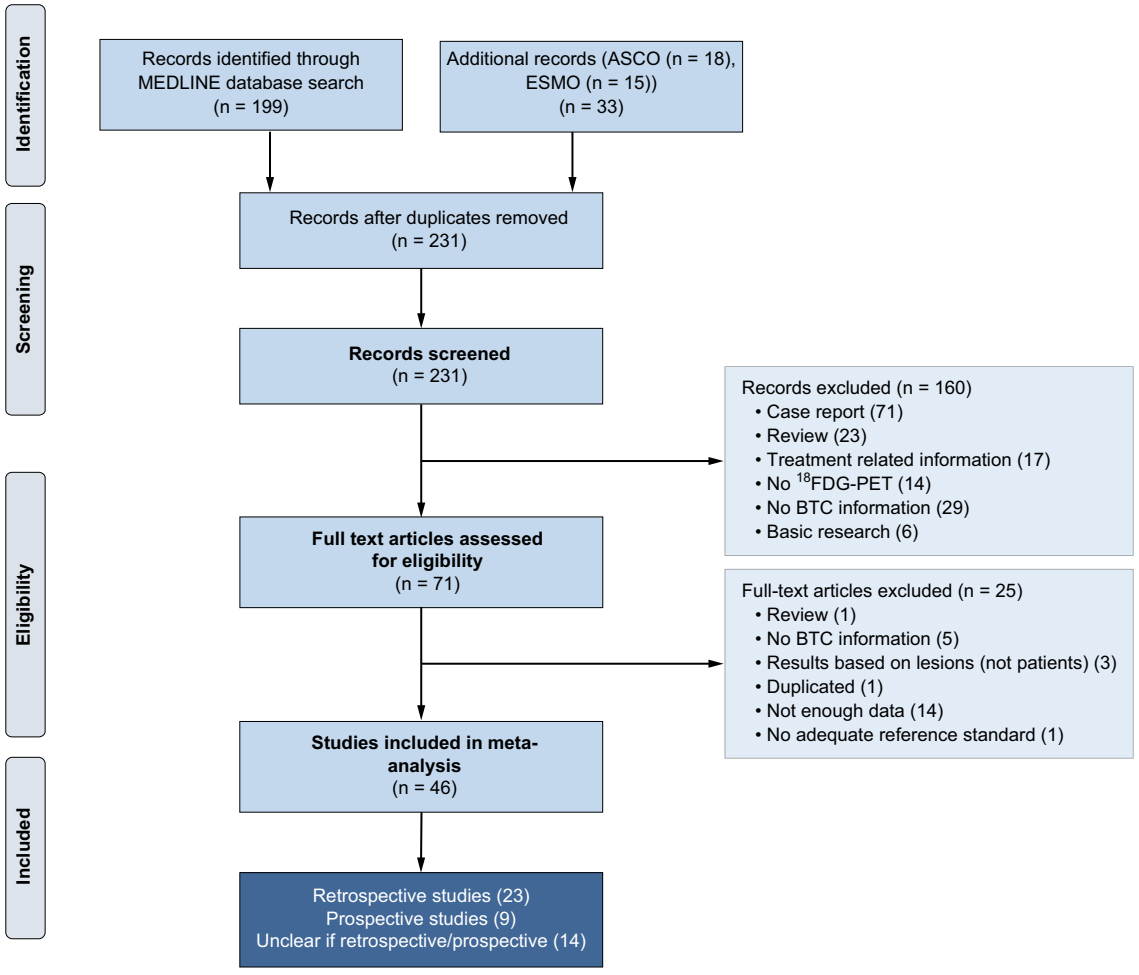



**Fig. 1. PRISMA diagram.** $^{18}$FDG-PET, $^{18}$F-fluorodeoxyglucose positron emission tomography; BTC, biliary tract cancer; n number of studies.

cer[58–61] or liver cancer[62]); iii) Data accuracy and/or 2 × 2 tables were derived from a different unit other than patient (such as number of lymph nodes affected) in 3 studies;[63–65] iv) One paper was excluded as it was a duplicate series[66] already included in this meta-analysis in the form of an updated paper;[67] v) Fourteen papers were excluded due to lack of details for determination of diagnostic accuracy,[34,68–70] extraction of prognostic information[71–73] or other useful information;[74–80] vi) Finally, one paper was excluded due to the reference standard not meeting eligibility criteria.[81]

A total of 47 eligible studies were included;[22,23,32,33,67,82–122] characteristics of each study are summarised in Table 1 and Table S1. Most of these studies were retrospective (23 studies retrospective; 9 prospective; in 14 studies the design was unclear). Six studies included patients with advanced disease due to start palliative treatment, 5 explored the role of $^{18}$FDG-PET following resection, and 16 did so in the pre-operative setting. A total of 19 studies included patients with all stages of BTC. In total, 39 studies reported data on diagnostic accuracy, 13 for changes in management and 11 regarding prognosis.

**Table 1. Eligible studies, study details and characteristics.**

| Study [Ref] | Cancer status | Country | Data extracted from | Study type | DTA info | Management info | Survival info | BTC/ Total patients | F/M | Mean/median age, years ± SD (range) |
|---|---|---|---|---|---|---|---|---|---|---|
| 108 | Before surgery | China | Abstract | Retrospective | Yes | No | Yes | 39/39 | 28/38 | |
| 95 | Advanced stage, before treatment | Korea | Full text | Prospective | No | No | Yes | 75/75 | 32/43 | 64 (46–83) |
| 103 | Before surgery | Korea | Full text | Retrospective | No | No | Yes | 25/25 | 3/22 | 67 ± 6 |
| 121 | Before surgery | China | Full text | Retrospective | Yes | No | No | 30/49 | – | |
| 94 | All stages | China | Abstract | Retrospective | Yes | Yes | No | 65/65 | – | |
| 86 | Advanced stage, before treatment | Korea | Full text | Retrospective | Yes | No | Yes | 106/106 | 31/75 | 61 |
| 82 | All stages | Turkey | Full text | Retrospective | Yes | No | No | 22/22 | – | |
| 117 | After surgery | India | Full text | Retrospective | Yes | No | No | 26/50 | 16/34 | 55 ± 11 |
| 88 | Advanced stage, before treatment | Korea | Full text | Prospective | No | No | Yes | 48/48 | 14/34 | 61 (range 39–75) |
| 97 | Before surgery | Korea | Full text | Unclear | Yes | No | No | 53/53 | 28/25 | 62.4 ± 10.0 (38–82) |
| 128 | Before surgery | Korea | Full text | Unclear | No | No | Yes | 78/78 | 22/42 | 61.3 ± 10.7 |
| 67 | All stages | Spain | Full text | Prospective | Yes | Yes | No | 20/37 | 28/21 | 68.6 ± 11.3 |
| 104 | All stages | Korea | Full text | Retrospective | No | No | Yes | 61/61 | 29/32 | 68.5 ± 8.9 |
| 84 | Before surgery | Taiwan | Abstract | Prospective | Yes | No | No | 36/62 | – | (27–86) |
| 87 | Before surgery | Korea | Abstract | Unclear | Yes | No | No | 34/39 | – | |
| 92 | Before surgery | China | Abstract | Retrospective | Yes | No | No | 32/32 | 14/18 | 56 |
| 102 | After surgery | India | Abstract | Retrospective | Yes | No | No | 42/62 | 34/15 | 52.5 |
| 83 | Advanced stage, before treatment | Germany | Abstract | Retrospective | Yes | No | No | 47/64 | – | |
| 106 | After surgery | Korea | Full text | Retrospective | Yes | No | No | 34/50 | 13/37 | 60 (33–77) |
| 116 | Before surgery | Netherlands | Full text | Retrospective | Yes | Yes | No | 26/30 | – | |
| 93 | Advanced stage, before treatment | Germany | Full text | Unclear | No | No | Yes | 23/23 | 15/11 | |
| 99 | Before surgery | Japan | Full text | Unclear | Yes | No | Yes | 73/73 | 27/46 | 66 ± 8 |
| 105 | All stages | Korea | Full text | Retrospective | Yes | Yes | No | 82/99 | 41/58 | 67 (35–91) |
| 122 | Advanced stage, before treatment | USA | Full text | Prospective | No | No | Yes | 35/35 | 14/21 | 60 (25–82) |
| 33 | After surgery | Japan | Full text | Retrospective | Yes | Yes | No | 29/50 | – | 62 (47–82) |
| 91 | All stages | Japan | Full text | Retrospective | Yes | No | Yes | 36/69 | 29/40 | 69 (46–84) |
| 129 | Before surgery | Germany | Full text | Retrospective | Yes | No | No | 12/17 | 6/11 | 62 |
| 118 | Before surgery | Japan | Full text | Unclear | Yes | Yes | No | 27/27 | 12/15 | 64 (41–78) |
| 23 | Before surgery | Korea | Full text | Prospective | Yes | No | No | 94/123 | 43/80 | 60 (28–78) |
| 89 | All stages | USA | Full text | Retrospective | Yes | Yes | No | 87/93 | – | |
| 111 | All stages | Japan | Full text | Retrospective | Yes | No | No | 29/37 | – | |
| 109 | All stages | Korea | Full text | Retrospective | Yes | Yes | No | 46/54 | 20/34 | 59.2 ± 8.7 |
| 112 | Before surgery | Japan | abstract | Unclear | Yes | No | No | 3/12 | – | |
| 85 | After surgery | Japan | Full text | Unclear | Yes | No | No | 4/5 | 5/0 | (50–69) |
| 119 | All stages | Italy | Full text | Unclear | Yes | Yes | No | 20/25 | – | |
| 110 | All stages | Japan | Full text | Unclear | Yes | No | No | 23/32 | 20/12 | 69.9 (34–83) |
| 32 | All stages | Switzerland | Full text | Prospective | Yes | Yes | No | 61/70 | 30/31 | 64 (35–81) |
| 120 | All stages | Japan | Full text | Prospective | Yes | No | No | 21/30 | – | |
| 114 | All stages | Germany | Full text | Unclear | Yes | No | No | 14/22 | 10/10 | 63 ± 14 |
| 115 | Before surgery | Spain | Full text | Retrospective | Yes | No | No | 5/16 | 11/5 | 67.75 (51–83) |
| 22 | All stages | USA | Full text | Unclear | Yes | Yes | No | 40/50 | 16/20 | 63 (38–84) |
| 130 | All stages | Korea | Full text | Retrospective | Yes | No | No | 21/2 | 10/11 | 57 (34–74) |
| 101 | Before surgery | Japan | Full text | Retrospective | Yes | No | No | 8/16 | – | |
| 96 | All stages | Japan | Full text | Unclear | Yes | No | No | 30/30 | 9/21 | 68 (21–82) |
| 90 | All stages | Germany | Full text | Unclear | Yes | Yes | No | 13/15 | 6/9 | 58 (47–78) |
| 100 | All stages | Germany | Full text | Prospective | Yes | Yes | No | 26/34 | 11/15 | 63 (39–85) |

BTC, biliary tract cancer; Ref, reference; DTA, diagnostic test accuracy; F, female; M, male; SD, standard deviation; info, information.

### Descriptive analysis

Data on a total of 2,125 patients who had $^{18}$FDG-PET imaging were reported. Of these, 1,761 had a confirmed diagnosis of BTC: ampullary cancer (n = 129); cholangiocarcinoma (CCA) (n = 1,130; extrahepatic [eCCA] n = 489, intrahepatic [iCCA] n = 333, hilar [hCCA] n = 163, CCA subtype not specified n = 145); gallbladder cancer (n = 310); BTC subtype not specified (n = 192). Of the studies reporting information on gender distribution, 886 males and 627 females were identified. The male/female ratio for the whole population was 3:2. Information on age was available for 1,113 patients with BTC; weighted mean age was 62.62 years (95% CI 62.38–62.85).

### Methodological quality of included studies

Methodological quality assessed using QUADAS2 is summarised in Table S2. Most studies showed good quality, with low risk/level of concerns. Lack of details available for the index test ($^{18}$FDG-PET) was the most frequent reason for unclear/high risk/concern.

### Diagnostic accuracy

A total of 39 individual studies reported DTA data on 1,416 patients and were included in the DTA meta-analysis.[22,23,32,33,67,82–87,89–92,94,96–102,105–112,114–121]

SUVmax data was available for patients with and without cancer in 49 and 14 studies, respectively. Weighted mean SUVmax in patients with cancer was >5.6 for all subgroups analysed, with standard deviation (SD) varying between 0.2 and 2.1 (Table S3). Contrary to this, in patients without cancer, weighted mean SUVmax varied between 1.8 and 2.9 with SD ranging between 0.2 and 0.4. For the subgroups of patients with BTC in whom the primary tumour (T) was assessed, SUVmax was significantly higher in patients with cancer compared to non-cancer (6.2 [95% CI 5.2–7.2] *vs.* 2.8 [95% CI 2.2–3.4]; $p = 0.0001$). Differences were also significant when the primary tumour (T) of CCA (5.6 [95% CI 3.4–7.8] *vs.* 2.7 [95% CI 1.8–3.6]; $p = 0.0174$) and eCCA (5.9 [95% CI 4.6–7.2] *vs.* 2.9 [95% CI 2.5–3.3]; $p = 0.0108$) patients were analysed. SUVmax within other subgroups could not be compared because of the absence of sufficient observations.

### Diagnostic accuracy: diagnosis of primary tumour

A total of 30 studies reported DTA information in patients with BTC for assessment of the primary tumour (T). Pooled sensitivity and specificity were 91.7% (95% CI 89.9–93.2) and 51.3% (95% CI 46.4–56.2), respectively (Fig. S1.1). The SROC curve showed an AUC of 0.8668 (Fig. S1.1) with Lambda of 2.87 in the HSROC analysis (Table 2). Subgroups analyses are provided in Table 2, which showed the lowest sensitivity in ampullary malignancies (79.9%) and lowest specificity in hCCA (21.9%) and eCCA (27.7%). Sensitivity analysis including low risk/concern studies did not show a significant change in DTA parameters.

### Diagnostic accuracy: diagnosis of lymph node and distant metastases

Data from 12 studies demonstrated that pooled sensitivity and specificity were 88.4% (95% CI 82.6-92.8) and 69.1% (95% CI 63.8-74.1), respectively; AUC 0.8519 (Fig. S1.2). When low risk/-concern studies were analysed in sensitivity analysis, pooled sensitivity did not change significantly (Table 2). The lowest specificity was identified in patients with iCCA (48.9%).

Diagnosis of distant metastases achieved the highest pooled specificity (89.7% (95% CI 86.0-92.7) [pooled sensitivity 85.4% (95% CI 79.5-90.2); AUC 0.9253] in the 9 studies with data available (Fig. S1.3; Table 2).

### Diagnostic accuracy: diagnosis of relapse

Seven studies reported information of diagnostic accuracy of relapse for $^{18}$FDG-PET (Table 2). SROC analysis identified diagnosis of tumour relapse as having the highest AUC (AUC = 0.9592) (Fig. S1.4) compared to other subgroups analysed (Fig. S1.1, Fig. S1.2 and Fig. S1.3). Pooled sensitivity and specificity were 90.1% (95% CI 84.4-94.3) and 83.5% (95% CI 74.4-90.4), respectively.

### Diagnostic threshold analysis

No significant diagnostic threshold effect was identified in any of the groups explored (Table 2). Only 3 studies were identified to use the same SUVmax threshold; thus, bivariate analysis was not performed.

### Meta-regression

Despite multiple DTA parameters showing significant heterogeneity (Table 2), meta-regression did not identify that any of the predefined heterogeneity factors significantly impacted on the DOR (all $p$ values >0.05; Table S4).

### Impact on management

Thirteen individual studies, including data on 591 patients, reported data on management change.[22,32,33,67,89,90,94,100,105,109,116,118,119] Three explored the pre-surgery setting, while 5 studies did so following surgery, when relapsed disease was being investigated. Overall, the pooled proportion of change in management (random effects model) was 15% (95% CI 11–20) (Fig. 2A). Results did not vary significantly when only low risk studies were explored [sensitivity analysis; 11 studies; pooled proportion of change in management (random effects model) was 15% (95% CI 10–20]. Pooled change in management for studies done pre-surgery and after resection (recurrence) was 17% (95% CI 9–25) (Fig. 2B) and 14% (95% CI 8–21) (Fig. 2C), respectively. For the majority (39/50; 78%) of patients in whom details regarding the specific implication on management were provided, the $^{18}$FDG-PET upstaged the disease with identification of previously unknown sites of disease. No significant reporting bias was identified (Fig. S2).

### Prognostic role of $^{18}$FDG-PET imaging

A total of 11 studies (including data on 1,081 patients) explored the prognostic role of $^{18}$FDG-PET.[86,88,91,93,95,99,103,104,108,113,122] Nine studies explored OS.[86,88,91,95,99,103,104,108,122] Baseline high SUVmax was associated with worse survival (pooled HR of 1.79; 95% CI 1.37–2.33; $p$ <0.001 (Fig. 3A)). When a study with unclear bias[108] was removed (sensitivity analysis), results did not vary significantly (pooled HR 2.16; 95% CI 1.49–3.12).

Impact on DFS and PFS was reported in 4 studies. The pooled analysis of the 2 studies reporting on DFS showed no significant impact of high SUVmax on DFS (HR 1.96; 95% CI 0.58–6.65; $p = 0.28$)[108,113] (Fig. 3B). Higher SUVmax at baseline correlated with worse PFS (HR 2.43; 95% CI 1.29–4.56; $p = 0.006$)[95,103] (Fig. 3C).

No significant reporting bias was identified in either OS, DFS or PFS (Fig. S2). Most of the studies explored the prognostic role

**Table 2. Diagnostic accuracy results by subgroups.**

| Primary site | Studies included | Ns | Np | Pooled Se (95% CI); $I^2$ (*p* value) | Pooled Sp (95% CI); $I^2$ (*p* value) | Pooled LR+ (95% CI); $I^2$ (*p* value) | Pooled LR– (95% CI); $I^2$ (*p* value) | Pooled DOR (95% CI); $I^2$ (*p* value) | Diagnostic threshold | HSROC lambda (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Diagnosis of primary tumour** | | | | | | | | | | |
| BTC | All studies | 30 (Fig. S1) | 1,151 | 91.7 (89.8–93.2); 68.9% (<0.001) | 51.3 (46.4–56.2); 82.0% (<0.001) | 1.791 (1.368–2.343); 94.3% (<0.001) | 0.215 (0.159–0.293); 35.6% (0.029) | 11.005 (6.993–17.321); 31.7% (0.051) | Rho: 0.534 *p* value: 0.002 | 2.87 (1.99–3.75) |
| | Low bias/concern | 24 | 986 | 91.9 (89.9–93.6); 65.6% (<0.001) | 50.0 (44.7–55.3); 82.8% (<0.001) | 1.838 (1.341–2.517); 95.1% (<0.001) | 0.198 (0.137–0.287); 45.5% (0.009) | 11.383 (6.609–19.604); 44.4% (0.011) | Rho: 0.362 *p* value: 0.082 | 3.08 (1.98–4.19) |
| CCA | All studies | 21 | 848 | 95.7 (94.0–97.0); 16.1% (0.249) | 38.1 (32.0–44.6); 78.7% (<0.001) | 1.442 (1.138–1.827); 93.2% (<0.001) | 0.139 (0.095–0.203); 0.0% (0.911) | 13.460 (8.039–22.538); 2.9% (0.422) | – | – |
| | Low bias/concern | 18 | 730 | 95.2 (93.2–96.7); 15.3% (0.271) | 40.0 (33.5–46.7); 80.5% (<0.001) | 1.551 (1.163–2.069); 94.7% (<0.001) | 0.139 (0.094–0.205); 0.0% (0.801) | 13.495 (7.443–24.468); 16.0% (0.262) | – | – |
| iCCA | All studies | 10 | 198 | 94.2 (90.1–97.0); 47.6% (0.046) | 68.3 (51.9–81.9); 4.6% (0.398) | 1.890 (1.229–2.905); 22.1% (0.240) | 0.191 (0.106–0.347); 6.5% (0.382) | 23.636 (8.701–64.206); 0.0% (0.949) | – | – |
| | Low bias/concern | 9 | 188 | 94.2 (89.9–97.0); 53.4% (0.028) | 70.5 (53.8–84.0); 0.0% (0.457) | 2.042 (1.288–3.236); 16.4% (0.296) | 0.180 (0.093–0.351); 16.9% (0.292) | 25.499 (9.124–71.262); 0.0% (0.936) | – | – |
| eCCA | All studies | 13 | 425 | 95.3 (92.6–97.3); 0.0% (0.488) | 27.7 (20.5–35.9); 78.6% (<0.001) | 1.271 (0.993–1.628); 90.1% (<0.001) | 0.178 (0.103–0.310); 0.0% (0.674) | 8.512 (3.819–18.973); 11.3% (0.332) | – | – |
| | Low bias/concern | 12 | 379 | 94.8 (91.7–97.0); 0.0% (0.593) | 27.5 (20.3–35.8); 80.3% (<0.001) | 1.262 (0.985–1.615); 90.7% (<0.001) | 0.189 (0.108–0.331); 0.0% (0.733) | 7.964 (3.522–18.010); 12.4% (0.324) | – | – |
| hCCA | All studies | 8 | 130 | 91.9 (85.6–96.0); 0.0% (0.446) | 21.9 (9.30–40.0); 0.0% (0.502) | 1.063 (0.925–1.220); 0.0% (0.616) | 0.319 (0.105–0.973); 0.0% (0.805) | 3.660 (0.983–13.622); 0.0% (0.761) | – | – |
| | Low bias/concern | 5 | 68 | 87.7 (77.5–94.4); 0.0% (0.643) | 30.6 (11.5–56.3); 0.0% (0.510) | 1.085 (0.817–1.441); 9.8% (0.350) | 0.392 (0.105–1.458); 0.0% (0.560) | 3.125 (0.659–14.818); 0.0% (0.476) | – | – |
| GBC | All studies | 11 | 158 | 88.4 (82.6–92.8); 27.4% (0.183) | 72.3 (60.7–82.1); 58.0% (0.008) | 2.084 (1.236–3.513); 58.5% (0.007) | 0.272 (0.172–0.429); 0.0% (0.847) | 10.177 (4.492–23.060); 0.0% (0.672) | – | – |
| | Low bias/concern | 9 | 147 | 89.4 (83.4–93.7); 33.7% (0.148) | 68.2 (54.2–80.1); 60.7% (0.009) | 1.815 (1.069–3.084); 56.6% (0.018) | 0.268 (0.160–0.449); 0.0% (<0.001) | 8.907 (3.599–22.044); 0.0% (0.550) | – | – |
| AMP | All studies | 5 | 89 | 79.9 (69.9–87.7); 58.8% (0.046) | 43.0 (28.0–59.0); 81.2% (<0.001) | 1.397 (0.587–3.324); 86.4% (<0.001) | 0.402 (0.243–0.666); 0.0% (0.482) | 4.672 (1.723–12.671); 1.3% (0.399) | – | – |
| | Low bias/concern | 4 | 53 | 86.4 (72.6–94.8); 59.2% (0.061) | 13.6 (2.90–34.9); 28.8% (0.239) | 0.993 (0.833–1.184); 0.0% (0.727) | 0.651 (0.198–2.145); 0.0% (0.440) | 1.727 (0.312–9.550); 0.0% (0.559) | | |
| **Diagnosis of lymph node metastases** | | | | | | | | | | |
| BTC | All studies | 12 (Fig. S1) | 240 | 88.4 (82.6–92.8); 60.9% (0.003) | 69.1 (63.8–74.1); 81.8% (<0.001) | 2.178 (10.307–3.631); 92.1% (<0.001) | 0.242 (0.107–0.545); 63.4% (0.002) | 11.358 (4.247–30.375); 49.0% (0.028) | Rho: 0.455 *p* value: 0.137 | 2.22 (0.86–3.26) |
| | Low bias/concern | 9 | 156 | 82.6 (73.5–89.6); 58.6% (0.013) | 70.4 (64.8–75.6); 75.3% (<0.001) | 2.295 (1.492–3.530); 77.1% (<0.001) | 0.281 (0.118–0.671); 61.7% (0.008) | 9.485 (3.483–25.830); 43.0% (0.081) | Rho: 0.250 *p* value: 0.516 | 1.82 (0.92–2.73) |
| CCA | All studies | 6 | 83 | 72.7 (57.2–85.0); 52.9% (0.059) | 70.3 (63.3–76.7); 70.5% (0.005) | 2.218 (1.626–3.027); 25.9% (0.240) | 0.397 (0.172–0.913); 30.1% (0.210) | 5.470 (2.318–12.912); 0.0% (0.526) | – | – |
| | Low bias/concern | 6 | 83 | 72.7 (57.2–85.0); 52.9% (0.059) | 70.3 (63.3–76.7); 70.5% (0.005) | 2.218 (1.626–3.027); 25.9% (0.240) | 0.397 (0.172–0.913); 30.1% (0.210) | 5.470 (2.318–12.912); 0.0% (0.526) | – | – |
| iCCA | All studies | 2 | 72 | 98.1 (89.7–100); 37.5% (0.206) | 48.9 (33.7–64.2); 97.0% (<0.001) | 2.114 (0.281–15.899); 94.1% (<0.001) | 0.245 (0.029–2.049); 0.0% (0.776) | 9.964 (0.827–120.04); 0.0% (0.395) | – | – |
| | Low bias/concern | 1 | 6 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| eCCA | All studies | 2 | 36 | 90.9 (58.7–99.8); 0.0% (0.924) | 60.6 (48.3–72.0); 0.0% (0.374) | 2.308 (1.615–3.300); 0.0% (0.414) | 0.148 (0.023–0.967); 0.0% (0.853) | 16.276 (1.958–135.26); 0.0% (0.762) | – | – |
| | Low bias/concern | 2 | 36 | 90.9 (58.7–99.8); 0.0% (0.924) | 60.6 (48.3–72.0); 0.0% (0.374) | 2.308 (1.615–3.300); 0.0% (0.414) | 0.148 (0.023–0.967); 0.0% (0.853) | 16.276 (1.958–135.26); 0.0% (0.762) | – | – |

(*continued on next page*)

**Table 2** (*continued*)

| Primary site | Studies included | Ns | Np | Pooled Se (95% CI); $I^2$ (*p* value) | Pooled Sp (95% CI); $I^2$ (*p* value) | Pooled LR+ (95% CI); $I^2$ (*p* value) | Pooled LR– (95% CI); $I^2$ (*p* value) | Pooled DOR (95% CI); $I^2$ (*p* value) | Diagnostic threshold | HSROC lambda (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Diagnosis of lymph node metastases** | | | | | | | | | | |
| hCCA | All studies | 2 | 18 | 55.6 (30.8–78.5); 61.1% (0.109) | 63.0 (42.4–80.6); 83.9% (0.013) | 1.407 (0.810–2.443); 2.2% (0.312) | 0.693 (0.415–1.158); 0.0% (0.788) | 2.867 (0.717–11.470); 0.0% (0.752) | – | – |
| | Low bias/concern | 2 | 18 | 55.6 (30.8–78.5); 61.1% (0.109) | 63.0 (42.4–80.6); 83.9% (0.013) | 1.407 (0.810–2.443); 2.2% (0.312) | 0.693 (0.415–1.158); 0.0% (0.788) | 2.867 (0.717–11.470); 0.0% (0.752) | – | – |
| GBC | All studies | 2 | 51 | 93.8 (82.8–98.7); 0.0% (0.720) | 70.4 (49.8–86.2); 93.2% (<0.001) | 3.621 (0.033–397.476); 97.2% (<0.001) | 0.087 (0.028–0.267); 0.0% (0.521) | 37.197 (1.372–1008.3); 65.0% (0.091) | – | – |
| | Low bias/concern | 1 | 17 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| AMP | All studies | 1 | 11 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| | Low bias/concern | 0 | 0 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| **Diagnosis of distant metastases** | | | | | | | | | | |
| BTC | All studies | 9 (Fig. S1) | 182 | 85.4 (79.5–90.2); 53.2% (0.029) | 89.7 (86.0–92.7); 89.1% (<0.001) | 8.776 (1.005–76.623); 98.5% (<0.001) | 0.206 (0.117–0.360); 53.4% (0.028) | 44.420 (14.964–131.86); 58.6% (0.013) | Rho: –0.050 *p* value: 0.898 | 3.95 (2.41–5.49) |
| | Low bias/concern | 8 | 148 | 83.4 (76.5–89.0); 48.5% (0.059) | 90.7 (87.0–93.6); 90.0% (<0.001) | 9.903 (0.642–152.807); 98.8% (<0.001) | 0.232 (0.1345–0.400); 48.9% (0.057) | 42.549 (12.243–147.88); 63.2% (0.008) | Rho: –0.238 *p* value: 0.570 | 3.99 (1.93–6.06) |
| CCA | All studies | 5 | 70 | 76.4 (64.8–85.8); 44.0% (0.129) | 86.7 (80.5–91.5); 89.0% (<0.001) | 4.150 (0.606–28.433); 96.8% (<0.001) | 0.357 (0.239–0.532); 0.0% (0.477) | 16.710 (5.645–49.467); 32.7% (0.203) | – | – |
| | Low bias/concern | 5 | 70 | 76.4 (64.8–85.8); 44.0% (0.129) | 86.7 (80.5–91.5); 89.0% (<0.001) | 4.150 (0.606–28.433); 96.8% (<0.001) | 0.357 (0.239–0.532); 0.0% (0.477) | 16.710 (5.645–49.467); 32.7% (0.203) | – | – |
| iCCA | All studies | 1 | 9 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| | Low bias/concern | 1 | 9 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| eCCA | All studies | 1 | 13 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| | Low bias/concern | 1 | 13 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| hCCA | All studies | 2 | 15 | 72.7 (39.0–94.0); 0.0% (0.569) | 77.5 (61.5–89.2); 52.1% (0.148) | 2.595 (1.330–5.064); 0.0% (0.388) | 0.399 (0.152–1.051); 0.0% (0.791) | 7.229 (1.521–34.347); 0.0% (0.823) | – | – |
| | Low bias/concern | 2 | 15 | 72.7 (39.0–94.0); 0.0% (0.569) | 77.5 (61.5–89.2); 52.1% (0.148) | 2.595 (1.330–5.064); 0.0% (0.388) | 0.399 (0.152–1.051); 0.0% (0.791) | 7.229 (1.521–34.347); 0.0% (0.823) | – | – |
| GBC | All studies | 4 | 81 | 91.1 (82.6–96.4); 0.0% (0.709) | 82.4 (71.8–90.3); 85.0% (<0.001) | 4.992 (0.608–41.010); 95.7% (<0.001) | 0.121 (0.059–0.248); 0.0% (0.544) | 48.032 (9.059–254.68); 43.0% (0.153) | – | – |
| | Low bias/concern | 3 | 47 | 88.5 (76.0–65.9); 0.0% (0.869) | 85.0 (72.1–93.5); 89.6% (<0.001) | 6.195 (0.094–409.908); 97.5% (<0.001) | 0.144 (0.064–0.322); 0.0% (0.531) | 36.795 (2.427–557.84); 61.2% (0.076) | – | – |
| AMP | All studies | 0 | 0 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| | Low bias/concern | 0 | 0 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – |
| **Diagnosis of tumour relapse** | | | | | | | | | | |
| BTC | All studies | 7 (Fig. S1) | 150 | 90.1 (84.4–94.3); 0.0% (0.753) | 83.5 (74.4–90.4); 23.8% (0.248) | 4.304 (2.374–7.802); 37.3% (0.144) | 0.131 (0.079–0.216); 0.0% (0.568) | 42.903 (16.449–111.90); 20.4% (0.274) | Rho: –0.450 *p* value: 0.310 | 3.62 (1.40–5.84) |
| | Low bias/concern | 4 | 45 | 89.8 (77.8–86.6); 0.0% (0.716) | 85.3 (68.9–95.0); 39.2% (0.177) | 3.945 (1.135–13.710); 56.1% (0.077) | 0.122 (0.048–0.311); 0.0% (0.647) | 54.654 (11.874–251.56); 0.0% (0.394) | Rho: –0.200 *p* value: 0.800 | 3.49 (–0.83 to 7.81) |
| CCA | All studies | 3 | 42 | 84.0 (70.4–93.1); 0.0% (0.785) | 71.1 (46.1–89.2); 0.0% (0.898) | 2.806 (1.383–5.694); 0.0% (0.798) | 0.233 (0.105–0.474); 0.0% (0.745) | 13.081 (3.611–47.389); 0.0% (0.755) | – | – |
| | Low bias/concern | 2 | 8 | 81.8 (48.2–97.7); 0.0% (0.511) | 66.7 (9.40–99.2); 0.0% (0.668) | 2.147 (0.457–10.088); 0.0% (0.568) | 0.282 (0.046–1.721); 0.0% (0.463) | 9.000 (0.120–192.79); 0.0% (0.482) | – | – |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| iCCA | All studies | 0 | 0 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| | Low bias/concern | 0 | 0 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| eCCA | All studies | 0 | 0 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| | Low bias/concern | 0 | 0 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| hCCA | All studies | 1 | 4 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| | Low bias/concern | 1 | 4 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| GBC | All studies | 2 | 44 | 93.6 (82.5–98.7); 0.0% (0.519) | 90.9 (70.8–98.9); 0.0% (0.480) | 8.638 (2.343–31.848); 0.0% (0.345) | 0.077 (0.025–0.231); 0.0% (0.384) | 116.87 (14.192–962.47); 9.6% (0.293) | – | – | | | | |
| | Low bias/concern | 1 | 2 | n.a. | n.a. | n.a. | n.a. | n.a. | – | – | | | | |
| AMP | All studies | 1 | 26 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |
| | Low bias/concern | 1 | 26 | | n.a. | | n.a. | | n.a. | | n.a. | | – | – |

Interpretation of parameters provided: Se: represents the probability of having a positive $^{18}$FDG-PET if patient has BTC. Sp: represents the probability of having a negative $^{18}$FDG-PET if patient does not have BTC. LR+: represents how many times increases the probability of having diagnosis of BTC if $^{18}$FDG-PET result is positive. Values greater than 1 increase the probability of BTC diagnosis. LR−: represents how many times increases the probability of having a diagnosis of BTC if $^{18}$FDG-PET result is negative. Values between 0 and 1 represent decrease in the probability of BTC diagnosis. DOR: describes the odds of a positive test results in patients diagnosed with BTC compared with the odds of a positive test result in those without the disease.

$^{18}$FDG-PET, $^{18}$F-fluorodeoxyglucose positron emission tomography; AMP, ampullary malignancy; BTC, biliary tract cancer; CCA, cholangiocarcinoma; DOR, diagnostic odds ratio; eCCA, extrahepatic cholangiocarcinoma; GBC, gallbladder cancer; hCCA, hilar cholangiocarcinoma; HSROC, hierarchical summary receiver operating characteristic; $I^2$, inconsistency; iCCA, intrahepatic cholangiocarcinoma; LR+, positive likelihood ratio; LR−, negative likelihood ratio; n.a., not applicable; Np, number of cancer patients included; Ns, number of studies eligible; Se, sensitivity; Sp, specificity. *p* value: heterogeneity *p* value; rho: Spearmen correlation coefficient.

using a previously defined SUVmax cut-off. However, the cut-off employed varied between studies (range 4–9); see Fig. 3A, Fig. 3B and Fig. 3C for further details.

Three studies reported the impact of changes in SUVmax following treatment. The study by Jo and colleagues[95] reported a series of 75 patients with advanced BTC treated with cisplatin and gemcitabine. Jo and colleagues showed that a reduction of 20% or higher in the SUVmax following 2 cycles of chemotherapy was prognostic for better PFS (HR 3.35 for patients with lesser reduction in SUVmax; $p = 0.002$); impact on OS did not reach statistical significance (HR 1.96; $p = 0.082$). The study by Zhu and colleagues showed a worse OS and PFS in patients with increased SUVmax (increase of 20% or higher) after 2 cycles of gemcitabine, oxaliplatin and bevacizumab chemotherapy in a series of 35 patients with advanced BTCs (OS: HR 10.12; 95% CI 2.41–42.53; $p = 0.002$) (PFS: HR 9.71; 95% CI 2.63–35.76; $p < 0.0001$).[122] The impact of SUVmax variations on OS following treatment with liver embolisation was investigated in a series of 23 patients,[93] in this series changes in SUVmax did not correlate with survival in univariate COX regression (HR 0.56; 95% CI 0.15–2.13; $p = 0.39$).

## Discussion

This systematic review and meta-analysis aimed to clarify the role of $^{18}$FDG-PET in addition to the current standard of care imaging/diagnostic tests utilised (US/CT/MR) in the management of patients diagnosed with BTC. The results support the use of $^{18}$FDG-PET, not only for diagnosis of lymph node (N) and distant metastases (M), but also for assessment of relapsed disease. This is supported by the significant change in management identified in patients undergoing $^{18}$FDG-PET, together with the fact that identification of additional sites of disease was the most likely finding impacting management. Based on these findings, there seems to be enough data to incorporate $^{18}$FDG-PET into the patient pathway, especially for staging (N and M) when identification of occult sites of disease would change management (*i.e.* surgery/local therapies) and also for assessment of tumour recurrence if other imaging shows equivocal findings (Fig. 4).

However, the role of $^{18}$FDG-PET for diagnosis of the primary tumour (T) remains controversial, especially due to the limited specificity and LR+ identified. Based on the pooled specificity identified, $^{18}$FDG-PET findings should not be considered as a replacement for pathological confirmation of BTC, since up to half of patients without malignancy could have a false positive result for a primary tumour with $^{18}$FDG-PET imaging (Fig. 4). The only two exceptions for this would be for patients with iCCA and ampullary malignancy, where a higher pooled specificity was seen (even though it is worth highlighting that 95% CIs in these subgroups were wide and still included the 50% within them). The most likely explanations for low specificity in the assessment of primary tumours could include, infection-related false positives within the bile duct and likely co-existence of biliary stents.

Even though the presence of a diagnostic threshold effect could not be proven, it is well known that using a SUVmax for identification of a malignant lesion is of use in clinical practice. Most studies did not provide information of a threshold for SUVmax, and thus different thresholds could not be compared in this study. Despite this, the SUVmax provided for patients with BTC was significantly higher than that reported for patients

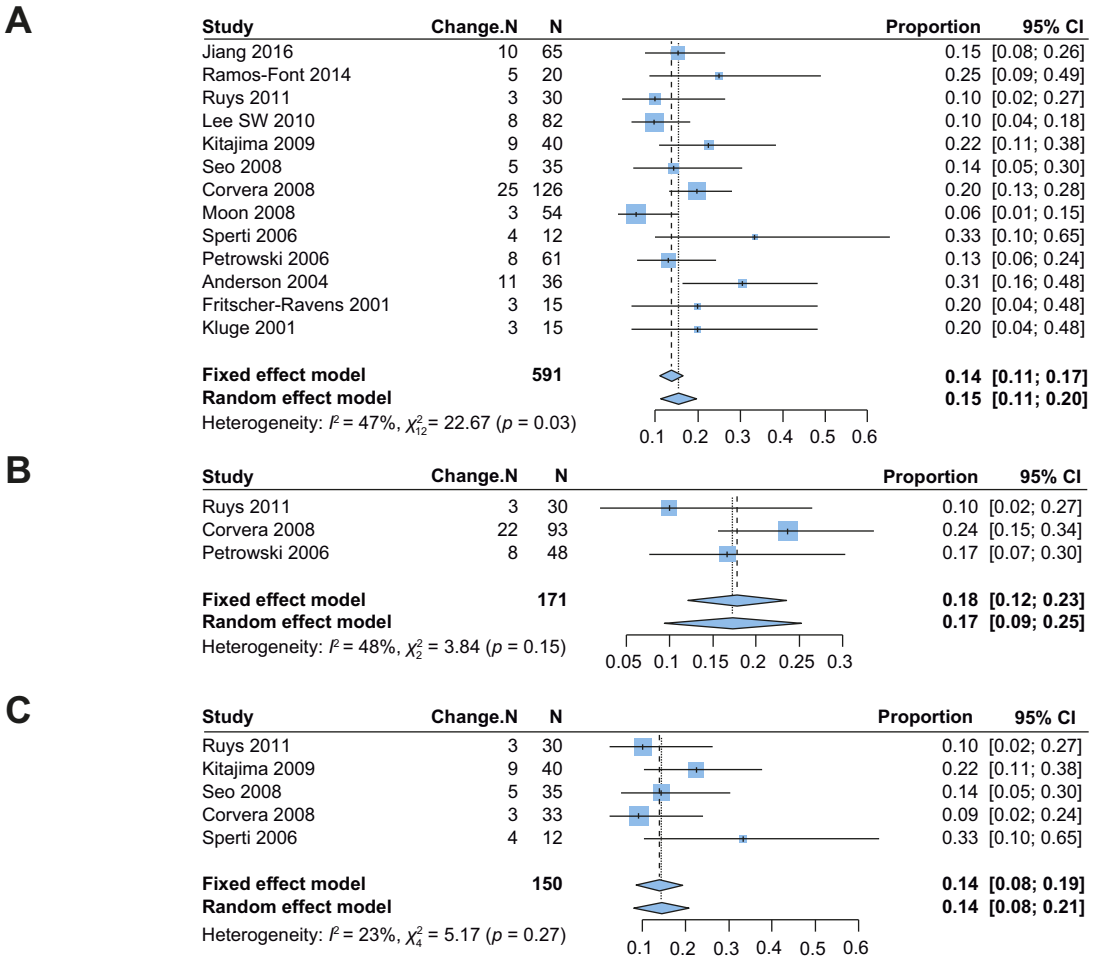



**Fig. 2. Change in management.** (A) Pooled proportion of change in management; (B) Change in management in the pre-surgical setting; (C) Change in management in the post-surgical setting for assessment of relapse.

without cancer. Based on these results, SUVmax cut-offs between 3.4 and 3.6 seemed most adequate for further studies in this scenario, since they represented the lowest and highest end of the 95% CI for delineation of benign from malignant lesions. Significant differences in SUVmax between different scenarios explored were not identified (T, N, M, Relapse), even though this analysis may be limited by the number of observations. Thus, the aforementioned cut-off for SUVmax would hold regardless of the scenario being investigated.

Multiple heterogeneity sources were explored, including PET acquisition characteristics, risk of bias and other factors. No significant impact on the accuracy of $^{18}$FDG-PET was identified for any of the factors explored. Thus, differences in protocol acquisition do not seem to impact performance, which could facilitate immediate incorporation of $^{18}$FDG-PET into the patient management pathway, since currently available protocols in individual institutions could be utilised. Even though no differences in diagnostic accuracy were identified between PET and PET/CT imaging; PET/CT imaging is expected to be superior for lesion delineation and should therefore be employed whenever available. In addition, image acquisition performed after 60 min of tracer injection, which is considered to be current standard of care, should be pursued in clinical practice.

The prognostic role of baseline SUVmax was confirmed in this study. However, currently, clinical implications of high SUVmax may be limited and further prospective randomised studies adjusting treatment (*i.e.* surgery and type of chemotherapy) according to $^{18}$FDG-PET SUVmax are required to assess its real impact, before it is incorporated into the patient pathway for this purpose. Based on these results, baseline $^{18}$FDG-PET could be recommended for patients with this diagnosis until further prospective studies assessing its role as a prognostic/predictive factor become available.

The prognostic role of $^{18}$FDG-PET and its ability to inform change in management have not been explored in any of the other available meta-analyses in this scenario.[123] Other meta-analyses and systematic reviews have explored the DTA of $^{18}$FDG-PET in patients with BTC. This current study concurs with other recent meta-analyses exploring diagnostic accuracy in gallbladder cancer.[124] In addition, other meta-analyses have also suggested a superior diagnostic performance for primary tumour for iCCA compared to other cholangiocarcinoma subtypes, such as eCCA.[125] Diagnosis of the primary tumour remains controversial and some authors have suggested caution in this use, favouring its use for assessment of tumour relapse, for example.[126,127]

Limitations of the current work include significant variability between studies, including, but not limited to, study design. In addition, a significant number of the eligible studies had a retrospective design and selection bias cannot be excluded. Many

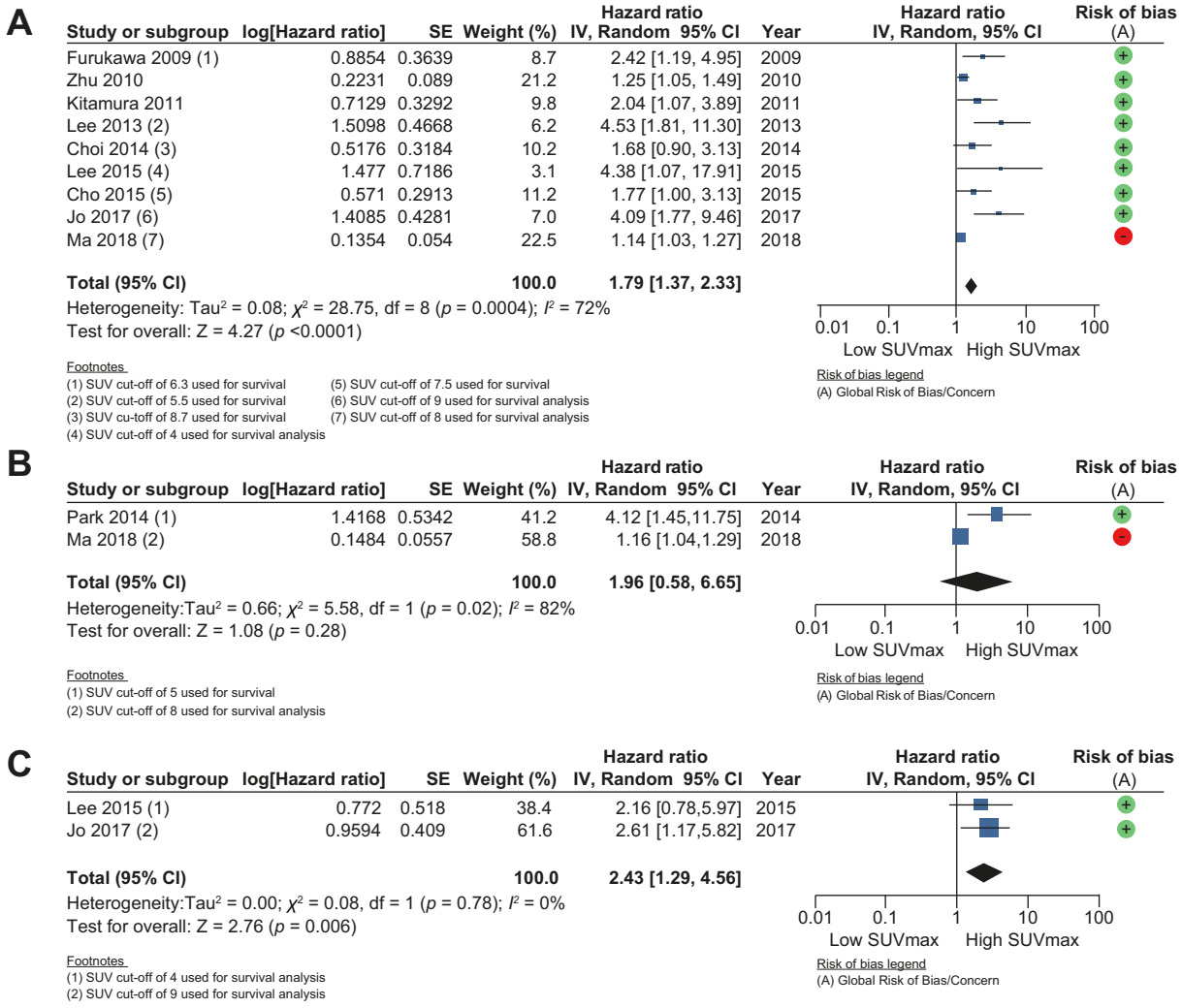



**Fig. 3. Forest plot: survival analysis.** Risk of bias and concern (following predefined definitions, see methods section) is specified for each study. Studies with unclear/high risk are highlighted in red. (A) The effect of SUVmax on overall survival; (B) The effect of SUVmax on disease-free survival; (C) The effect of SUVmax on progression-free survival. SUVmax, maximum standardised uptake values. (This figure appears in colour on the web.)

studies had missing data regarding SUVmax cut-off employed for definition of "positive" findings, which also limited the analysis performed. It would have been of interest to assess whether accuracy of $^{18}$FDG-PET for assessment of primary tumour (T) varied according to the clinical context (*i.e.* suspicion of BTC in patients with jaundice due to biliary obstruction, suspicion of BTC in patients with radiological biliary stricture without jaundice, suspicion of BTC in patients with intrahepatic mass and extrahepatic metastasis); unfortunately, the absence of such detail in the eligible studies made this analysis unfeasible. The lack of individual patient data from each study used for pooled analysis is another limitation worth mentioning. Despite these limitations, this study was performed following a previously defined protocol (designed before study identification and data collection was started) with predefined reference standard criteria, strengthening the findings and reducing any potential inclusion bias. In addition, the quality of studies was carefully addressed and included in the sensitivity analysis.

In conclusion, these findings support the incorporation of $^{18}$FDG-PET imaging in addition to the current standard of care imaging/diagnostic tests utilised (US/CT/MR) in the initial assessment for the presence of lymph node and distant metastases in patients with BTC since it may guide treatment (surgery *vs.* palliative treatment) selection decisions. Its use in this setting would be of particular interest if the identification of occult sites of disease would change management (*i.e.* surgery/local therapies). In addition, $^{18}$FDG-PET is a useful tool for confirmation of disease relapse if diagnosis remains unclear following standard of care imaging. On the contrary, results would not support the use of $^{18}$FDG-PET for diagnosis of the primary tumour in the absence of other disease sites or pathological confirmation, due to low specificity. The prognostic role of $^{18}$FDG-PET and the impact of SUVmax on management requires further investigation in prospective studies.

## Financial support

Dr Angela Lamarca has received funding from European Society for Medical Oncology (ESMO) Fellowship Programme, Pancreatic Cancer Research Fund, Spanish Society of Medical Oncology (SEOM) Fellowship Programme, Cholangiocarcinoma Foundation Research Fellowship Grant, American Society of Clinical Oncology (ASCO) Conquer Cancer Foundation Young Investigator Award and The Christie Charity. Jorge Barriuso received

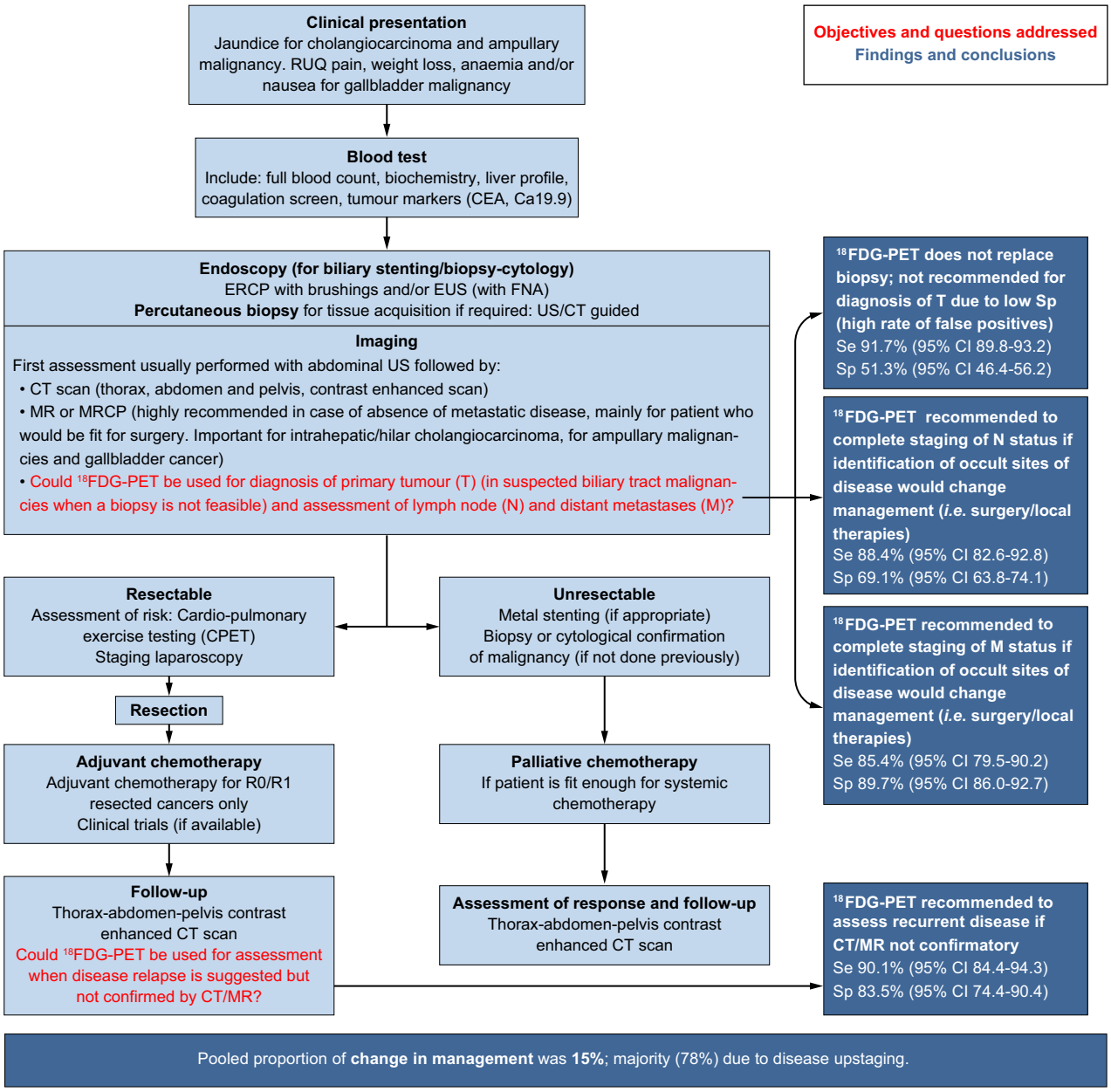



**Fig. 4. Current clinical pathway for the diagnosis of biliary tract malignancies (light blue sections) are shown jointly with the potential roles of $^{18}$FDG-PET explored (red sections) and main study findings and conclusions (dark blue sections).** Blue light sections represent current clinical pathway. Patients diagnosed with BTC commonly present with non-specific symptoms; therefore, diagnosis is challenging. Further diagnostic techniques may not be utilised until patients develop jaundice or severe abdominal pain. The diagnosis is often suspected when a soft tissue mass or narrowing in the biliary tract is identified on imaging studies such as US, CT, MR, or ERCP.[15–17] Biopsy or cytological diagnosis is recommended for confirmation of the suspected malignancy and also for planning of further treatment.[19] The techniques employed for confirmatory pathological diagnosis are EUS-guided-biopsy or FNA, ERCP guided brushings or biopsy and percutaneous CT/MR-guided biopsy or FNA. Red sections represent the potential diagnostic role of $^{18}$FDG-PET. Three potential roles for $^{18}$FDG-PET in biliary tract cancer were explored in this systematic review: 1) diagnosis of primary tumour (T), 2) diagnosis of lymph node metastases (N), 3) diagnosis of distant metastases (M) and 4) diagnosis of relapsed disease). Dark blue sections summarise main findings and conclusions derived from this study: use of $^{18}$FDG-PET for assessment of N/M and relapsed disease is commended for selected patients, while the role for diagnosis of primary tumour (T) is limited. BTC, biliary tract cancer; ERCP, endoscopic retrograde cholangiopancreatography; EUS, endoscopic ultrasound; FNA, fine needle aspiration; R, resection margin; US, ultrasound, RUQ, right upper quadrant; MRCP, magnetic resonance cholangiopancreatography, $^{18}$FDG-PET, $^{18}$F-fluorodeoxyglucose positron emission tomography, CEA, Carcinoembryonic antigen; Ca19.9, carbohydrate antigen 19-9; Se, sensitivity; Sp, specificity. (This figure appears in colour on the web.)

funding from the ENETS Centre of Excellence Fellowship Grant Award.

Please refer to the accompanying ICMJE disclosure forms for further details.

## Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

## Authors' contributions

AL and JWV designed the study. AL performed the literature search and extracted data from selected manuscripts (with

support from JB and JWV in case of discrepancy); AL and JB performed statistical analysis. All authors were involved in interpretation of results and approval of final version of the manuscript.

## Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.01.038.

## References

[1] Patel T. Increasing incidence and mortality of primary intrahepatic cholangiocarcinoma in the United States. Hepatology 2001;33:1353–1357.
[2] Njei B. Changing pattern of epidemiology in intrahepatic cholangiocarcinoma. Hepatology 2014;60:1107–1108.
[3] Valle JW. Advances in the treatment of metastatic or unresectable biliary tract cancer.. Ann Oncol 2010;21 (Suppl 7):vii345–vii348.
[4] Horgan AM, Amir E, Walter T, Knox JJ. Adjuvant therapy in the treatment of biliary tract cancer: a systematic review and meta-analysis. J Clin Oncol 2012;30:1934–1940.
[5] Primrose JNFR, Palmer D, Prasad R, Mirza D, Anthoney DA, et al. Adjuvant capecitabine for biliary tract cancer: the BILCAP randomized study. J Clin Oncol 2017;35 (Suppl):4006.
[6] Neoptolemos JP, Moore MJ, Cox TF, Valle JW, Palmer DH, McDonald AC, et al. Effect of adjuvant chemotherapy with fluorouracil plus folinic acid or gemcitabine vs observation on survival in patients with resected periampullary adenocarcinoma: the ESPAC-3 periampullary cancer randomized trial. JAMA 2012;308:147–156.
[7] Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. N Engl J Med 2010;362:1273–1281.
[8] Shroff RT, Borad MJ, Xiao L, Kaseb AO, Varadhachary GR, Wolff RA, et al. A phase II trial of gemcitabine (G), cisplatin (C), and nab-paclitaxel (N) in advanced biliary tract cancers (aBTCs). J Clin Oncol 2017;35:4018.
[9] Lamarca A, Hubner RA, David Ryder W, Valle JW. Second-line chemotherapy in advanced biliary cancer: a systematic review. Ann Oncol 2014;25:2328–2338.
[10] Lamarca A, Hubner RA, Ryder WD, Valle JW. Reply to the letter to the editor 'second-line chemotherapy in advanced biliary cancer: the present now will later be past' by Vivaldi et al.. Ann Oncol 2014;25:2444–2445.
[11] Valle JW, Lamarca A, Goyal L, Barriuso J, Zhu AX. New horizons for precision medicine in biliary tract cancers. Cancer Discov 2017;7:943–962.
[12] Lamarca A, Barriuso J, McNamara MG, Valle JW. biliary tract cancer: state of the art and potential role of DNA damage repair. Cancer Treat Rev 2018;70:168–177.
[13] Lowery MA, Abou-Alfa GK, Burris HA, Janku F, Shroff RT, Cleary JM, et al. Phase I study of AG-120, an IDH1 mutant enzyme inhibitor: Results from the cholangiocarcinoma dose escalation and expansion cohorts. J Clin Oncol 2017;35 (Suppl 15):4015.
[14] Javle M, Lowery M, Shroff RT, Weiss KH, Springfeld C, Borad MJ, et al. Phase II study of BGJ398 in patients with FGFR-altered advanced cholangiocarcinoma. J Clin Oncol 2018;36:276–282.
[15] Schofl R. Diagnostic endoscopic retrograde cholangiopancreatography. Endoscopy 2001;33:147–157.
[16] Yeh BM, Liu PS, Soto JA, Corvera CA, Hussain HK. MR imaging and CT of the biliary tract. Radiographics 2009;29:1669–1688.
[17] Mittal PK, Moreno CC, Kalb B, Mittal A, Camacho JC, Maddu K, et al. Primary biliary tract malignancies: MRI spectrum and mimics with histopathological correlation. Abdom Imaging 2015;40:1520–1557.
[18] Inui K, Yoshino J, Miyoshi H. Differential diagnosis and treatment of biliary strictures. Clin Gastroenterol Hepatol 2009;7:S79–S83.
[19] Brugge WDJ, Klapman JBFA, Ashfaq RFAU, Shidham VFAU, Chhieng DFAU, et al. Techniques for cytologic sampling of pancreatic and bile duct lesions. Diagn Cytopathol 2014;42:333–337.
[20] Hammoud GM, Almashhrawi A, Ibdah JA. Usefulness of endoscopic ultrasound-guided fine needle aspiration in the diagnosis of hepatic, gallbladder and biliary tract lesions. World J Gastrointest Oncol 2014;6:420–429.
[21] Edge SBD, Compton CC, Fritz AG, Greene FL, Trotti A. AJCC Cancer Staging Manual. New York: Springer-Verlag; 2015.
[22] Anderson CD, Rice MH, Pinson CW, Chapman WC, Chari RS, Delbeke D. Fluorodeoxyglucose PET imaging in the evaluation of gallbladder carcinoma and cholangiocarcinoma. J Gastrointest Surg 2004;8:90–97.
[23] Kim JY, Kim MH, Lee TY, Hwang CY, Kim JS, Yun SC, et al. Clinical role of 18F-FDG PET-CT in suspected and potentially operable cholangiocarcinoma: a prospective study compared with conventional imaging. Am J Gastroenterol 2008;103:1145–1151.
[24] Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. Semin Ultrasound CT MR 2010;31:496–505.
[25] Purandare NC, Pramesh CS, Karimundackal G, Jiwnani S, Agrawal A, Shah S, et al. Incremental value of 18F-FDG PET/CT in therapeutic decision-making of potentially curable esophageal adenocarcinoma. Nucl Med Commun 2014;35:864–869.
[26] Gallamini A, Zwarthoed C, Borra A. Positron Emission Tomography (PET) in oncology. Cancers (Basel) 2014;6:1821–1889.
[27] Bonanni L, de'Liguori Carino N, Deshpande R, Ammori BJ, Sherlock DJ, Valle JW, et al. A comparison of diagnostic imaging modalities for colorectal liver metastases. Eur J Surg Oncol 2014;40:545–550.
[28] Gohy S, Hubert C, Deprez P, Van Beers BE, Annet L, Lhommel R, et al. Benign biliary inflammatory pseudotumor mimicking a Klatskin tumor. Hepatogastroenterology 2007;54:1348–1352.
[29] Nagasaki Y, Yamane H, Ochi N, Honda Y, Takigawa N. High uptake of FDG along a biliary stent. Clin Nucl Med 2016;41:890–891.
[30] Sacks A, Peller PJ, Surasi DS, Chatburn L, Mercier G, Subramaniam RM. Value of PET/CT in the management of primary hepatobiliary tumors, part 2. AJR Am J Roentgenol 2011;197:W260–W265.
[31] De Gaetano AM, Rufini V, Castaldi P, Gatto AM, Filograna L, Giordano A, et al. Clinical applications of (18)F-FDG PET in the management of hepatobiliary and pancreatic tumors. Abdom Imaging 2012;37:983–1003.
[32] Petrowsky H, Wildbrett P, Husarik DB, Hany TF, Tam S, Jochum W, et al. Impact of integrated positron emission tomography and computed tomography on staging and management of gallbladder cancer and cholangiocarcinoma. J Hepatol 2006;45:43–50.
[33] Kitajima K, Murakami K, Kanegae K, Tamaki N, Kaneta T, Fukuda H, et al. Clinical impact of whole body FDG-PET for recurrent biliary cancer: a multicenter study. Ann Nucl Med 2009;23:709–715.
[34] Jadvar H, Henderson RW, Conti PS. [F-18]fluorodeoxyglucose positron emission tomography and positron emission tomography: computed tomography in recurrent and metastatic cholangiocarcinoma. J Comput Assist Tomogr 2007;31:223–228.
[35] Radiologist. Evidence-based indications for the use of PET-CT in the UK 2013. www.rcplondon.ac.uk/sites/default/files/pet-ct_guidelines_2013.pdf 2013 (accessed 21 September 2015).
[36] PROSPERO. https://www.crd.york.ac.uk/PROSPERO/.
[37] MEDLINE. https://www.ncbi.nlm.nih.gov/.
[38] ASCO. https://meetinglibrary.asco.org/.
[39] ESMO. https://academic.oup.com/annonc.
[40] McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, et al. Preferred reporting reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018;319:388–396.
[41] Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006–1012.
[42] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–536.
[43] Stata.v.12. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.
[44] Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
[45] Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. BMC Med Res Methodol 2002;2:9.
[46] Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20:2865–2884.
[47] Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 2007;8:239–251.

[48] Freemana MF, Tukey JW. Transformations related to the angular and the square root. Ann Math Statist 1950;21:607–611.
[49] Jj M. The inverse of the Freeman-Tukey double arcsine transformation. Am Stat 1978;32:138.
[50] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–188.
[51] R-Studio. https://www.r-studio.com/ 2018.
[52] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560.
[53] Egger MDSG, Schneider M, Minder C. Bias in metaanalysis detected by a simple, graphical test. BMJ 1997;315:629–634.
[54] Higgins JPT GSe. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org. In Cohrane website. 2011.
[55] Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343 d4002.
[56] Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med 2002;21:1525–1537.
[57] Hu JH, Tang JH, Lin CH, Chu YY, Liu NJ. Preoperative staging of cholangiocarcinoma and biliary carcinoma using 18F-fluorodeoxyglucose positron emission tomography: a meta-analysis. J Investig Med 2018;66:52–61.
[58] Kurita A, Kodama Y, Nakamoto Y, Isoda H, Minamiguchi S, Yoshimura K, et al. Impact of EUS-FNA for preoperative para-aortic lymph node staging in patients with pancreatobiliary cancer. Gastrointest Endosc 2016;84:467–475.
[59] Vineeth Kumar PM, Verma GR, Mittal BR, Agrawal K, Gupta R, Kochhar R, et al. FLT PET/CT is better than FDG PET/CT in differentiating benign from malignant pancreatobiliary lesions. Clin Nucl Med 2016;41: e244–e250.
[60] Raj P, Kaman L, Singh R, Dahyia D, Bhattacharya A, Bal A. Sensitivity and specificity of FDG PET-CT scan in detecting lymph node metastasis in operable periampullary tumours in correlation with the final histopathology after curative surgery. Updates Surg 2013;65:103–107.
[61] Kalady MF, Clary BM, Clark LA, Gottfried M, Rohren EM, Coleman RE, et al. Clinical utility of positron emission tomography in the diagnosis and management of periampullary neoplasms. Ann Surg Oncol 2002;9:799–806.
[62] Shao H, Cheng W, Yu L, Li Y, Jing H. Use of 18F-FDG PET scan and ultrasound-guided biopsy in the diagnosis of hepatic carcinomas. Hepatogastroenterology 2015;62:978–981.
[63] Park TG, Yu YD, Park BJ, Cheon GJ, Oh SY, Kim DS, et al. Implication of lymph node metastasis detected on 18F-FDG PET/CT for surgical planning in patients with peripheral intrahepatic cholangiocarcinoma. Clin Nucl Med 2014;39:1–7.
[64] Kobayashi S, Nagano H, Hoshino H, Wada H, Marubashi S, Eguchi H, et al. Diagnostic value of FDG-PET for lymph node metastasis and outcome of surgery for biliary cancer. J Surg Oncol 2011;103:223–229.
[65] Keiding S, Hansen SB, Rasmussen HH, Gee A, Kruse A, Roelsgaard K, et al. Detection of cholangiocarcinoma in primary sclerosing cholangitis by positron emission tomography. Hepatology 1998;28:700–706.
[66] Ramos-Font C, Gomez RM, Rodriguez-Fernandez A, Sanchez SR, Llamas Elvira JM. Positron tomography with 18F-fluorodeoxyglucose in the preoperative evaluation of gall bladder lesions suspicious of malignancy. Diagnostic utility and clinical impact. Rev Esp Med Nucl 2011;30:267–275.
[67] Ramos-Font C, Gomez-Rio M, Rodriguez-Fernandez A, Jimenez-Heffernan A, Sanchez SR, et al. Ability of FDG-PET/CT in the detection of gallbladder cancer. J Surg Oncol 2014;109:218–224.
[68] Muller D, Wiedmann M, Kluge R, Berr F, Mossner J, Sabri O, et al. Is 18F-FDG-PET suitable for therapy monitoring after palliative photodynamic therapy of non-resectable hilar cholangiocarcinoma? Z Gastroenterol 2005;43:439–443.
[69] Iwata Y, Shiomi S, Sasaki N, Jomura H, Nishiguchi S, Seki S, et al. Clinical usefulness of positron emission tomography with fluorine-18-fluorodeoxyglucose in the diagnosis of liver tumors. Ann Nucl. Med 2000;14:121–126.
[70] Keiding S, Munk OL, Schiott KM, Hansen SB. Dynamic 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography of liver tumours without blood sampling. Eur J Nucl Med 2000;27:407–412.
[71] Braghiroli MI, Mota JM, Duarte PS, Morita TO, Bariani GM, Nebuloni D, et al. Evaluation of 18F-FDG PET-CT as a prognostic marker in advanced biliary tract cancer. Nucl Med Commun 2018;39:252–259.
[72] Yamada I, Ajiki T, Ueno K, Sawa H, Otsubo I, Yoshida Y, et al. Feasibility of (18)F-fluorodeoxyglucose positron-emission tomography for preoperative evaluation of biliary tract cancer. Anticancer Res 2012;32:5105–5110.
[73] Lee J, Yun M, Kim KS, Lee JD, Kim CK, Alkhawaldeh K, et al. Risk stratification of gallbladder polyps (1–2 cm) for surgical intervention with 18F-FDG PET/CT. J Nucl Med 2012;53:353–358.
[74] Wang ZF, Lan XL, Xiao YR, Liu JP, Ji JS. Correlation between TNM staging of primary cholangiocarcinoma and the maximum standard uptake value of (18)F-2-deoxy-D-glucose positron emission tomography with computerized tomography. Zhonghua Yi Xue Za Zhi 2017;97:3104–3107.
[75] Suzuki H, Komuta M, Bolog A, Yokobori T, Wada S, Araki K, et al. Relationship between 18-F-fluoro-deoxy-D-glucose uptake and expression of glucose transporter 1 and pyruvate kinase M2 in intrahepatic cholangiocarcinoma. Dig Liver Dis 2015;47:590–596.
[76] Sahani DV, Hayano K, Galluzzo A, Zhu AX. Measuring treatment response to systemic therapy and predicting outcome in biliary tract cancer: comparing tumor size, volume, density, and metabolism. AJR Am J Roentgenol 2015;204:776–781.
[77] Song JY, Lee YN, Kim YS, Kim SG, Jin SJ, Park JM, et al. Predictability of preoperative 18F-FDG PET for histopathological differentiation and early recurrence of primary malignant intrahepatic tumors. Nucl Med Commun 2015;36:319–327.
[78] Pak K, Kim SJ, Kim IJ, Kim DU, Kim K, Kim H, et al. Splenic FDG uptake predicts poor prognosis in patients with unresectable cholangiocarcinoma. Nuklearmedizin 2014;53:26–31.
[79] Murata Y, Watanabe H, Kubota K, Toda K, Nakamura S, Okouchi K, et al. PET/CT evaluation of the physiologic accumulation of 18F-FDG within the gallbladder vesicle. Nucl Med Biol 2007;34:961–966.
[80] Lee JD, Yang WI, Park YN, Kim KS, Choi JS, Yun M, et al. Different glucose uptake and glycolytic mechanisms between hepatocellular carcinoma and intrahepatic mass-forming cholangiocarcinoma with increased (18) F-FDG uptake. J Nucl Med 2005;46:1753–1759.
[81] Kornberg A, Kupper B, Thrum K, Wilberg J, Sappler A, Gottschild D. Recurrence-free long-term survival after liver transplantation in patients with 18F-FDG non-avid hilar cholangiocarcinoma on PET. Am J Transplant 2009;9:2631–2636.
[82] Akcam AT, Ulku A, Rencuzogullari A, Parsak CK, Yapar Z, Doran F, et al. 18F-FDG activity PET/CT and CA-19.9 levels for the prediction of histopathological features and localization of peri- ampullary tumors. Turk J Gastroenterol 2015;26:170–175.
[83] Alkhawaldeh K, Faltten S, Biersack HJ, Ezziddin S. The value of F-18 FDG PET in patients with primary sclerosing cholangitis and cholangiocarcinoma using visual and semiquantitative analysis. Clin Nucl Med 2011;36:879–883.
[84] Cheng MF, Wang HP, Tien YW, Liu KL, Yen RF, Tzen KY, et al. Usefulness of PET/CT for the differentiation and characterization of periampullary lesions. Clin Nucl Med 2013;38:703–708.
[85] Chikamoto A, Tsuji T, Takamori H, Kanemitsu K, Uozumi H, Yamashita Y, et al. The diagnostic efficacy of FDG-PET in the local recurrence of hilar bile duct cancer. J Hepatobiliary Pancreat Surg 2006;13:403–408.
[86] Cho KM, Oh DY, Kim TY, Lee KH, Han SW, Im SA, et al. Metabolic characteristics of advanced biliary tract cancer using 18F-fluorodeoxyglucose positron emission tomography and their clinical implications. Oncologist 2015;20:926–933.
[87] Choi EK, Yoo I, Kim SH, O JH, Choi WH, Na S, et al. The clinical value of dual-time point 18F-FDG PET/CT for differentiating extrahepatic cholangiocarcinoma from benign disease. Clin Nucl Med 2013;38: e106–e111.
[88] Choi MK, Choi JY, Lee J, Heo JS, Choi SH, Choi DW, et al. Prognostic and predictive value of metabolic tumor volume on (18)F-FDG PET/CT in advanced biliary tract cancer treated with gemcitabine/oxaliplatin with or without erlotinib. Med Oncol 2014;31:23.
[89] Corvera CU, Blumgart LH, Akhurst T, DeMatteo RP, D'Angelica M, Fong Y, et al. 18F-fluorodeoxyglucose positron emission tomography influences management decisions in patients with biliary cancer. J Am Coll Surg 2008;206:57–65.
[90] Fritscher-Ravens A, Bohuslavizki KH, Broering DC, Jenicke L, Schafer H, Buchert R, et al. FDG PET in the diagnosis of hilar cholangiocarcinoma. Nucl Med Commun 2001;22:1277–1285.

[91] Furukawa H, Ikuma H, Asakura K, Uesaka K. Prognostic importance of standardized uptake value on F-18 fluorodeoxyglucose-positron emission tomography in biliary tract carcinoma. J Surg. Oncol 2009;100:494–499.
[92] Gu XJ, Wang BF, Liu R. Application of (1)(8)F-fluorodeoxyglucose positron emission tomography/computed tomography in preoperative assessment of hilar cholangiocarcinoma. Zhonghua Yi Xue Za Zhi 2012;92:1409–1412.
[93] Haug AR, Heinemann V, Bruns CJ, Hoffmann R, Jakobs T, Bartenstein P, et al. 18F-FDG PET independently predicts survival in patients with cholangiocellular carcinoma treated with 90Y microspheres. Eur J Nucl Med Mol Imaging 2011;38:1037–1045.
[94] Jiang L, Tan H, Panje CM, Yu H, Xiu Y, Shi H. Role of 18F-FDG PET/CT imaging in intrahepatic cholangiocarcinoma. Clin Nucl Med 2016;41:1–7.
[95] Jo J, Kwon HW, Park S, Oh DY, Cheon GJ, Bang YJ. Prospective evaluation of the clinical implications of the tumor metabolism and chemotherapy-related changes in advanced biliary tract cancer. J Nucl. Med 2017;58:1255–1261.
[96] Kato T, Tsukamoto E, Kuge Y, Katoh C, Nambu T, Nobuta A, et al. Clinical role of (18)F-FDG PET for initial staging of patients with extrahepatic bile duct cancer. Eur J Nucl Med Mol Imaging 2002;29:1047–1054.
[97] Kim J, Ryu JK, Kim C, Paeng JC, Kim YT. Is there any role of positron emission tomography computed tomography for predicting resectability of gallbladder cancer? J Korean Med Sci 2014;29:680–684.
[98] Kim YJ, Yun M, Lee WJ, Kim KS, Lee JD. Usefulness of 18F-FDG PET in intrahepatic cholangiocarcinoma. Eur J Nucl Med Mol Imaging 2003;30:1467–1472.
[99] Kitamura K, Hatano E, Higashi T, Seo S, Nakamoto Y, Narita M, et al. Prognostic value of (18)F-fluorodeoxyglucose positron emission tomography in patients with extrahepatic bile duct cancer. J Hepatobiliary Pancreat Sci 2011;18:39–46.
[100] Kluge R, Schmidt F, Caca K, Barthel H, Hesse S, Georgi P, et al. Positron emission tomography with [(18)F]fluoro-2-deoxy-D-glucose for diagnosis and staging of bile duct cancer. Hepatology 2001;33:1029–1035.
[101] Koh T, Taniguchi H, Yamaguchi A, Kunishima S, Yamagishi H. Differential diagnosis of gallbladder cancer using positron emission tomography with fluorine-18-labeled fluoro-deoxyglucose (FDG-PET). J Surg Oncol 2003;84:74–81.
[102] Kumar R, Sharma P, Kumari A, Halanaik D, Malhotra A. Role of 18F-FDG PET/CT in detecting recurrent gallbladder carcinoma. Clin Nucl Med 2012;37:431–435.
[103] Lee EJ, Chang SH, Lee TY, Yoon SY, Cheon YK, Shim CS, et al. Prognostic value of FDG-PET/CT total lesion glycolysis for patients with resectable distal bile duct adenocarcinoma. Anticancer Res 2015;35:6985–6991.
[104] Lee JY, Kim HJ, Yim SH, Shin DS, Yu JH, Ju DY, et al. Primary tumor maximum standardized uptake value measured on 18F-fluorodeoxyglucose positron emission tomography-computed tomography is a prognostic value for survival in bile duct and gallbladder cancer. Korean J Gastroenterol 2013;62:227–233.
[105] Lee SW, Kim HJ, Park JH, Park DI, Cho YK, Sohn CI, et al. Clinical usefulness of 18F-FDG PET-CT for patients with gallbladder cancer and cholangiocarcinoma. J Gastroenterol 2010;45:560–566.
[106] Lee YG, Han SW, Oh DY, Chie EK, Jang JY, Im SA, et al. Diagnostic performance of contrast enhanced CT and 18F-FDG PET/CT in suspicious recurrence of biliary tract cancer after curative resection. BMC Cancer 2011;11:188. https://doi.org/10.1186/1471-2407-11-188.: 188-111.
[107] Li J, Kuehl H, Grabellus F, Muller SP, Radunz S, Antoch G, et al. Preoperative assessment of hilar cholangiocarcinoma by dual-modality PET/CT. J Surg Oncol 2008;98:438–443.
[108] Ma KW, Cheung TT, She WH, Chok KSH, Chan ACY, Dai WC, et al. Diagnostic and prognostic role of 18-FDG PET/CT in the management of resectable biliary tract cancer. World J Surg 2018;42:823–834.
[109] Moon CM, Bang S, Chung JB, Park SW, Song SY, Yun M, et al. Usefulness of 18F-fluorodeoxyglucose positron emission tomography in differential diagnosis and staging of cholangiocarcinomas. J Gastroenterol Hepatol 2008;23:759–765.
[110] Nishiyama Y, Yamamoto Y, Fukunaga K, Kimura N, Miki A, Sasakawa Y, et al. Dual-time-point 18F-FDG PET for the evaluation of gallbladder carcinoma. J Nucl Med 2006;47:633–638.
[111] Nishiyama Y, Yamamoto Y, Kimura N, Miki A, Sasakawa Y, Wakabayashi H, et al. Comparison of early and delayed FDG PET for evaluation of biliary stricture. Nucl Med Commun 2007;28:914–919.
[112] Oe A, Kawabe J, Torii K, Kawamura E, Higashiyama S, Kotani J, et al. Distinguishing benign from malignant gallbladder wall thickening using FDG-PET. Ann Nucl Med 2006;20:699–703.
[113] Park MS, Lee SM. Preoperative 18F-FDG PET-CT maximum standardized uptake value predicts recurrence of biliary tract cancer. Anticancer Res 2014;34:2551–2554.
[114] Reinhardt MJ, Strunk H, Gerhardt T, Roedel R, Jaeger U, Bucerius J, et al. Detection of Klatskin's tumor in extrahepatic bile duct strictures using delayed 18F-FDG PET/CT: preliminary results for 22 patient studies. J Nucl Med 2005;46:1158–1163.
[115] Rodriguez-Fernandez A, Gomez-Rio M, Llamas-Elvira JM, Ortega-Lozano S, Ferron-Orihuela JA, Ramia-Angel JM, et al. Positron-emission tomography with fluorine-18-fluoro-2-deoxy-D-glucose for gallbladder cancer diagnosis. Am J Surg 2004;188:171–175.
[116] Ruys AT, Bennink RJ, van Westreenen HL, Engelbrecht MR, Busch OR, Gouma DJ, et al. FDG-positron emission tomography/computed tomography and standardized uptake value in the primary diagnosis and staging of hilar cholangiocarcinoma. HPB (Oxford) 2011;13:256–262.
[117] Santhosh S, Mittal BR, Kang M, Kapoor R, Gupta R, Rana S, et al. Diagnostic accuracy of F-18 FDG PET/CECT vs. CECT for detecting recurrence of periampullary carcinoma and its prognostic significance. Abdom Imaging 2015;40:1131–1137.
[118] Seo S, Hatano E, Higashi T, Nakajima A, Nakamoto Y, Tada M, et al. Fluorine-18 fluorodeoxyglucose positron emission tomography predicts lymph node metastasis, P-glycoprotein expression, and recurrence after resection in mass-forming intrahepatic cholangiocarcinoma. Surgery 2008;143:769–777.
[119] Sperti C, Pasquali C, Fiore V, Bissoli S, Chierichetti F, Liessi G, et al. Clinical usefulness of 18-fluorodeoxyglucose positron emission tomography in the management of patients with nonpancreatic periampullary neoplasms. Am J Surg 2006;191:743–748.
[120] Wakabayashi H, Akamoto S, Yachida S, Okano K, Izuishi K, Nishiyama Y, et al. Significance of fluorodeoxyglucose PET imaging in the diagnosis of malignancies in patients with biliary stricture. Eur J Surg Oncol 2005;31:1175–1179.
[121] Wang SB, Wu HB, Wang QS, Zhou WL, Tian Y, Ji YH, et al. 18F-FDG PET/CT in differentiating malignant from benign origins of obstructive jaundice. Hepatobiliary Pancreat Dis Int 2015;14:516–522.
[122] Zhu AX, Meyerhardt JA, Blaszkowsky LS, Kambadakone AR, Muzikansky A, Zheng H, et al. Efficacy and safety of gemcitabine, oxaliplatin, and bevacizumab in advanced biliary-tract cancers and correlation of changes in 18-fluorodeoxyglucose PET with clinical outcome: a phase 2 study. Lancet Oncol 2010;11:48–54.
[123] Parikh U, Marcus C, Sarangi R, Taghipour M, Subramaniam RM. FDG PET/CT in pancreatic and hepatobiliary carcinomas: value to patient management and patient outcomes. PET Clin 2015;10:327–343.
[124] Annunziata S, Pizzuto DA, Caldarella C, Galiandro F, Sadeghi R, Treglia G. Diagnostic accuracy of fluorine-18-fluorodeoxyglucose positron emission tomography in gallbladder cancer: a meta-analysis. World J Gastroenterol 2015;21:11481–11488.
[125] Annunziata S, Caldarella C, Pizzuto DA, Galiandro F, Sadeghi R, Giovanella L, et al. Diagnostic accuracy of fluorine-18-fluorodeoxyglucose positron emission tomography in the evaluation of the primary tumor in patients with cholangiocarcinoma: a meta-analysis. Biomed Res Int 2014;2014 247693.
[126] Lan BY, Kwee SA, Wong LL. Positron emission tomography in hepatobiliary and pancreatic malignancies: a review. Am J Surg 2012;204:232–241.
[127] Moon CM, Bang S, Chung JB. The role of (18)F-fluorodeoxyglucose positron emission tomography in the diagnosis, staging, and follow-up of cholangiocarcinoma. Surg Oncol 2011;20:e10–e17.
[128] Park JS, Park HC, Choi DH, Park W, Yu JI, Park YS, et al. Prognostic and predictive value of liver volume in colorectal cancer patients with unresectable liver metastases. Radiat Oncol J 2014;32:77–83.
[129] Li DB, Ye F, Wu XR, Wu LP, Chen JX, Li B, et al. Preoperative administration of bevacizumab is safe for patients with colorectal liver metastases. World J Gastroenterol 2013;19:761–768.
[130] Kim HJ, Lee KT, Kim SH, Lee JK, Lim JH, Paik SW, et al. Differential diagnosis of intrahepatic bile duct dilatation without demonstrable mass on ultrasonography or CT: benign versus malignancy. J Gastroenterol Hepatol 2003;18:1287–1292.