18F-FDG PET/CT for detection of the primary tumor in adults with extracervical metastases from cancer of unknown primary

A systematic review and meta-analysis

Burglin, Synne Alexandra; Hess, Søren; Høilund-Carlsen, Poul Flemming; Gerke, Oke

Published in:
Medicine

DOI:
10.1097/MD.0000000000006713

Publication date:
2017

Document version
Final published version

Document license
CC BY-ND

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk

Download date: 12. Oct. 2019
Abstract

Background: Cancer of unknown primary (CUP) is a heterogeneous group of cancers, so called when a biopsy from a patient reveals malignancy without giving a clue to where in the body the primary tumor is located. Whole-body 18-fluorine-fluorodeoxyglucose positron-emission-tomography/computed tomography (18F-FDG PET/CT) is widely used for diagnosis and staging of most cancers. We hypothesized that 18F-FDG PET/CT—especially if used early—is suitable for the detection of the primary tumor in patients with CUP.

Objective: To assess the ability of 18F-FDG PET/CT to detect the primary tumor in adult CUP patients.

Data Sources: PubMed/Medline, Embase, and Web of Science.

Study Eligibility Criteria: Studies on CUP from extracervical metastases in which every patient had received an 18F-FDG PET/CT scan and at least one 18F-FDG PET/CT-positive finding was confirmed by biopsy or clinical follow-up.

Study Appraisal: PRISMA and QUADAS-2 were applied.

Synthesis Methods: The pooled detection rate (DR) of 18F-FDG PET/CT was assessed with a fixed-effects model. Heterogeneity among studies was assessed with the I-squared statistic.

Results: A total of 2953 articles were identified from which N=82 were assessed by full text and N=20, comprising 1942 adult patients, were included in the study. Median (range) number of patients and DR was N=72 (21–316) and 36.3% (9.8%–75.3%), respectively. Two-thirds of included studies were retrospective, and the pooled DR was 40.93% (95% confidence interval: 38.99%–42.87%). There was large heterogeneity between studies (I-squared = 95.9%), randomization was not applied, CUP diagnosis was not standardized, and workup (if described) was characterized by multiple testing procedures resulting in a highly selected, challenging patient group.

Conclusions: Despite great heterogeneity in diagnostic workup and in studies in general, an overall DR of 40.93% suggests that upfront application of 18F-FDG PET/CT may have a role in CUP by obviating a great many futile diagnostic procedures. To what degree 18F-FDG PET/CT used early in the course of disease may improve the detection rate could not be deducted from selected articles. A large, prospective, preferably randomized, study on the potential benefit of using 18F-FDG PET/CT up front in CUP patients is warranted to judge if and when 18F-FDG PET/CT should be applied in these patients.

Abbreviations: 18F-FDG = 18-fluorine-fluorodeoxyglucose, CI = confidence interval, CUP = cancer of unknown primary, DR = detection rate, ES = estimate, FP = false positive, MeSH = medical subheadings, MRI = magnetic resonance imaging, PET/CT = positron-emission-tomography/computed tomography, PICO = population intervention comparison outcome, PRISMA = preferred reporting items for systematic reviews and meta-analysis, QUADAS = quality assessment of diagnostic accuracy studies, s.e. = standard error.

Keywords: cancer, cancer detection, cancer of unknown primary, diagnosis, diagnostic imaging, meta-analysis, positron emission tomography and computed tomography.
1. Introduction

Cancer of unknown primary (CUP) is a heterogeneous clinical entity comprising patients presenting metastases of unknown origin following conventional diagnostic workup procedures. The prevalence of CUP (defined as “unknown and ill-defined cancers”) in Denmark in 2009 to 2013 was 1042 patients which accounts for 2.5% (men) and 3.2% (women) of all cancer diagnoses in this period.[1] In the UK, the proportion of CUP patients among all cancer cases was 3% in 2014,[2] and the American Cancer Society estimated that about 33,770 cases of CUP will be diagnosed in 2017 in the United States which represents about 2% of all cancers.[3] Diagnostic investigations are expensive and time consuming, and they may cause discomfort to the patients. In 40% to 50% of CUP cases, primary tumor is not found.[4] Patients have by definition advanced illness at the time of diagnosis; generic (and usually less effective) treatment regimen rather than specific ones have to be employed when the origin of the cancer cannot be established which in turn may contribute to an overall poor prognosis. The relative 1-year survival (95% confidence interval [95% CI]) in Denmark was 38% (36%–40%) for men and 43% (41%–45%) for women and the relative 5-year survival was 19% (17%–20%) and 19% (18%–20%), respectively.[1]

Whole-body detection of tumor can be achieved with 18-fluorine-fluorodeoxyglucose positron-emission-tomography/computed tomography (¹⁸F-FDG PET/CT).[5] The high metabolic turnover in cancer cells is exploited in the PET scan with the use of a glucose analogue labeled with the radioactive isotope fluorine-18 as a tracer. Previously, studies have shown that ¹⁸F-FDG PET is effective in the diagnosis of many different tumor types.[6] The relative nonspecificity of FDG may pose a challenge although the addition of CT in combined PET/CT scanners has greatly enhanced the assessment of positive metabolic accumulation by adding the anatomical dimension. PET/CT is recommended as an additional diagnostic tool if primary tumor is not detected after conventional workup in patients with CUP and cervical lymph node metastases.[7] Evidence for patients with CUP and extracervical metastases remains to be established.[8] The objective of this systematic literature review was to assess the potential role of ¹⁸F-FDG PET/CT in the detection of the primary tumor in adult CUP patients with extracervical metastases as judged from the available literature.

2. Methods

The systematic review was done in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement.[9] A review protocol does not exist. An ethical review was not necessary due to the nature of this study.

2.1. Systematic literature search

The population, intervention, comparison, outcome (PICO) framework[10] was applied to the research question “Do adult patients with extracervical metastases from cancer of unknown primary experience improved primary tumor detection by means of ¹⁸F-FDG PET/CT compared with other diagnostic investigations?” The target population consisted of adult patients with extracervical metastases from CUP. The intervention of the diagnosing examination was molecular imaging with ¹⁸F-FDG PET/CT which was compared with conventional diagnostic investigations comprising a wide range of examinations like laboratory tests, contrast enhanced CT, magnetic resonance imaging (MRI), and pan-endoscopies for head and neck cancers. Outcome measures were DR (defined as positive PET/CT findings of primary tumors confirmed by biopsy or composite reference standard including follow-up and other imaging modalities on a per-patient basis) and FP proportion (defined as proportion of positive PET/CT findings that could not be confirmed as primary tumors by biopsy or other diagnostic procedures when a biopsy could not be applied).

Three databases were searched: PubMed/Medline, Embase, and Web of Science. For PubMed/Medline, the search strategy comprised both free text search and usage of Medical SubHeadings (MeSH). For Embase, free text search and the Emtree Thesaurus were used. The search was conducted on April 4, 2016 using the following search terms and their derivatives: cancer of unknown primary, positron emission tomography/computed tomography, and ¹⁸F-FDG. Neither date nor language limits were applied. A full search strategy is provided as Supplemental Digital Content 1, http://links.lww.com/MD/B669. All search results were collected, merged, and filtered with EndNote X7 (Thomson Reuters, Philadelphia, PA).

2.2. Selection of literature

One author (SAB) screened titles and abstracts and consulted a second reader (SH) in case of equivocal papers. Another reader (OG) independently assessed full-text articles for eligibility; disagreement was dissolved by consensus between all 3 readers. Inclusion criteria were as follows:

- Original studies on cancer of unknown primary from extracervical metastases.
- Every patient in the population had to have received an ¹⁸F-FDG PET/CT scan.
- At least 1 positive ¹⁸F-FDG PET/CT finding (i.e., lesion) had to be confirmed by biopsy to validate it as being primary tumor; alternatively, a composite reference standard including clinical follow-up and other imaging procedures was applied.
- Access to numbers of: the total number of patients who had an ¹⁸F-FDG PET/CT scan, the number of positive findings that were confirmed as well as the total number of positive findings which were not confirmed by pathology to calculate DR and FP if not stated by the authors.

Exclusion criteria were the following:

- Studies that included patients with a previous cancer.
- Studies that included patients who already received treatment for their current cancer.
- Studies with a population of both known and unknown primaries.

2.3. Data extraction

The definition of DR was the number of patients with positive PET/CT findings which were confirmed by biopsy or composite reference standard including follow-up and other imaging modalities, divided by the total number of patients included. The reported DR was taken as published if the above-mentioned definition was met; if other definitions of DR were used, we recalculated the DR according to the above-mentioned definition. The number of false-positive findings was extracted and used to assess its relative frequency in the study (FP proportion). All data
were extracted by 1 author (SAB). Data on the primary endpoint DR was independently extracted for validation purposes by another author (OG). Details on included studies were listed.

2.4. Statistical analysis
Descriptive statistics were used according to datatype; continuous variables were analyzed by medians and ranges, whereas categorical variables were displayed by means of absolute and relative frequencies by study. When comparing DR and FP proportion across studies, median values and ranges were reported. DR was also meta-analyzed with a fixed-effects model using the inverse variance method. A Forest plot[11] and a Funnel plot[12–15] were derived to graphically display the point estimate and a respective 95% CI for DR on a per-study basis, graphically display the summary estimate for DR (incl. its 95% CI) across studies, derive a measure for the heterogeneity of the studies (incl. an I² value[16]), and assess publication bias visually. All analyses were done by using STATA/MP 14.2 (StataCorp, College Station, TX).

2.5. Quality assessment
Possible sources for bias were assessed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool.[17] Accordingly, 4 domains were evaluated for each study (patient selection; index test; reference standard; flow and timing), with 2 to 3 signaling questions in each domain. If all signaling questions within a domain could be answered with “yes,” a low risk of bias was indicated for that domain. If any signaling question for a domain could be answered with “no” or “unclear” due to lack of details, bias potentially existed.

3. Results
3.1. Literature search
The sum of results from 3 databases was 2953 (Fig. 1). Doublets and foreign language articles other than English, Danish, or Norwegian were excluded (N=933). Thus, 2020 were available for screening by title/abstract. This screening led to exclusion of 1938 records: case reports (N=284), congress/meeting abstracts

![Figure 1. PRISMA flow chart](9)
3.2. Study and patient characteristics

The included studies were reported between 2005 and 2016 and comprised 1942 patients. Thirteen studies (65%) were retrospective, 4 (20%) prospective, and 3 (15%) unclear (Table 1). Median number of patients was N = 72 (range: 21–316). Mean or median age of patients was around 60 years for most studies and respective age ranges within the studies covered often close to whole adulthood. Seventeen of 20 studies included more men than women (median proportion of males: 57.9%).

Diagnostic workup prior to 18F-FDG PET/CT comprised a variety of diagnostic imaging procedures that included, for instance, CT alone, MRI, mammography, ultrasound, and physical examination. Tumor mass was often spread widely across the body in most studies (Supplemental Digital Content 2, http://links.lww.com/MD/B669). Details on 18F-FDG PET/CT imaging in the included studies (dosage, use of contrast enhancement, and area of scan) can be found in Supplemental Digital Content 3, http://links.lww.com/MD/B669.

3.3. Diagnostic performance on PET/CT

The median DR was 36.3% (range: 9.8%–75.3%; Table 1). The pooled DR with (95% CI) was 40.93% (38.99%–42.87%) with an $I^2$ of 95.9% (Fig. 2), the latter indicating large heterogeneity between studies. The respective Funnel plot was roughly symmetrical, but the 5 studies with the lowest detection rates$^{[19,28]}$ as well as the 2 studies with the highest detection rates$^{[19,28]}$ were outside the funnel, possibly indicating publication bias (Supplemental Digital Content 4, http://links.lww.com/MD/B669). The median FP proportion was 7.5% (range 2.3%–22.2%). Results from 3 studies$^{[23,31,36]}$ were restricted to a subgroup of patients to match our inclusion criteria.

3.4. Quality assessment

The quality assessment with QUADAS$^{[17]}$ indicated an overall low risk of bias across studies with respect to patient selection and reference standard, whereas risk of bias was assessed high in 6 of 20 studies (30%) regarding flow and timing and unclear in 12 of 20 studies (60%) regarding the index test (Fig. 3, based on Supplemental Digital Content 5, http://links.lww.com/MD/B669). With respect to the applicability of patient selection and reference standard, around 2 of 3 of the included studies suggested low concerns, but 11 of 20 studies (55%) suggested an unclear applicability of the index test.

4. Discussion

In this study, a comprehensive literature search was conducted to establish the current evidence for the use of PET/CT in adult...
patients with CUP with special reference to DR and false-positive findings in patients with extracervical metastases. We identified 20 studies with an overall DR of 40.93% (95% CI: 38.99%–42.87%); median DR and median FP proportion were 36.3% and 7.5%, respectively.

This study’s strength is the extensive literature search that identified 2953 potential articles, and its findings are based on 1942 patients. Moreover, the PICO framework, the PRISMA statement, and the QUADAS tool were applied. The selection of full-text articles (N = 82) to be assessed was done by only 1 author (SAB), and this screening process could have been strengthened by having 2 readers independently assess all potential articles. Concerning the selection of relevant literature, only original full papers were included, excluding, for instance, conference abstracts with the potential risk of missing recent unpublished work.

Table 2
Diagnostic performance of PET/CT.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Primary tumor detection rate (DR)</th>
<th>False-positive proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yaylali[18]</td>
<td>2016</td>
<td>16/50 32%</td>
<td>6/50 12%</td>
</tr>
<tr>
<td>Jain[19]</td>
<td>2015</td>
<td>13/100 69.3%</td>
<td>3/100 9.2%</td>
</tr>
<tr>
<td>Buer[20]</td>
<td>2014</td>
<td>17/70 24.3%</td>
<td>4/70 8.5%</td>
</tr>
<tr>
<td>Elboga[21]</td>
<td>2014</td>
<td>37/112 33%</td>
<td>5/112 5.7%</td>
</tr>
<tr>
<td>Barbosa[22]</td>
<td>2013</td>
<td>29/89 32.6%</td>
<td>6/89 7.6%</td>
</tr>
<tr>
<td>Daniel[23]</td>
<td>2013</td>
<td>19/51 37.3%</td>
<td>6/51 12.5%</td>
</tr>
<tr>
<td>Wang[24]</td>
<td>2013</td>
<td>13/34 38.2%</td>
<td>5/34 15.1%</td>
</tr>
<tr>
<td>Han[25]</td>
<td>2012</td>
<td>67/142 47.2%</td>
<td>6/142 4.2%</td>
</tr>
<tr>
<td>Hall[26]</td>
<td>2012</td>
<td>54/120 45%</td>
<td>9/120 7.5%</td>
</tr>
<tr>
<td>Moller[27]</td>
<td>2012</td>
<td>38/135 28.1%</td>
<td>10/135 7.5%</td>
</tr>
<tr>
<td>Tamam[28]</td>
<td>2011</td>
<td>37/149 24.8%</td>
<td>11/149 7.5%</td>
</tr>
<tr>
<td>Hu[29]</td>
<td>2011</td>
<td>5/51 9.8%</td>
<td>5/51 9.8%</td>
</tr>
<tr>
<td>Pak[30]</td>
<td>2010</td>
<td>32/74 43.2%</td>
<td>5/74 6.8%</td>
</tr>
<tr>
<td>Yaper[31]</td>
<td>2009</td>
<td>24/43 55.8%</td>
<td>6/43 13.7%</td>
</tr>
<tr>
<td>Fenc[32]</td>
<td>2007</td>
<td>31/190 16.3%</td>
<td>26/190 13.7%</td>
</tr>
<tr>
<td>Ambrosini[33]</td>
<td>2007</td>
<td>50/250 20%</td>
<td>12/250 4.8%</td>
</tr>
<tr>
<td>Pelosi[34]</td>
<td>2006</td>
<td>24/68 35.3%</td>
<td>5/68 7.4%</td>
</tr>
<tr>
<td>Gutzeit[35]</td>
<td>2005</td>
<td>9/27 33.3%</td>
<td>2/27 7.4%</td>
</tr>
<tr>
<td>Nanni[36]</td>
<td>2005</td>
<td>12/21 57.1%</td>
<td>NR</td>
</tr>
</tbody>
</table>

NR = not registered, PET/CT = positron-emission-tomography/computed tomography.

Figure 2. Forest plot on DR. CI = confidence interval, DR = detection rate, ES = estimate.
It is clear from the study characteristics (Table 1, Supplemental Digital Content 2 and 3, http://links.lww.com/MD/B669) that the evidence of PET/CT as a diagnostic tool in CUP is challenged by a significant heterogeneity among included studies (see also $I^2 = 95.9\%$ and Fig. 2 and Supplemental Digital Content 4, http://links.lww.com/MD/B669). First, a majority of studies were retrospective (65%), only 20% were prospective, and none of the included studies employed randomization. Second, the patient populations varied widely ranging from 20 to 316 with a median of 72. Third, the definition in CUP as a diagnosis was not standardized; most studies had biopsy-verified metastases as an inclusion criterion, but some studies also included patients with “clinical suspicion of malignancy.”

Along the same lines, the definition of “standardized diagnostic workup” prior to PET/CT scans was also not standardized. Available details of diagnostic strategy varied from no registration$^{[19,31,33]}$ to presentation of specific investigations and the number of patients receiving each examination.$^{[26,27,30]}$ Based on studies where diagnostic work-up regimens were presented, many patients underwent a significant array of diagnostic tests, both invasive and noninvasive. This has 2-fold consequences: patients included are highly selected toward the difficult ones since obvious diagnosis will be made in many instances from the standard workup; and the number of unnecessary or futile procedures may be overwhelming to individual patients. This may in turn lead to delayed diagnosis, adverse effects from invasive procedures, and increased health care costs.

Another matter is the reference standard. Applying a relevant and accurate reference standard is pivotal in diagnostic accuracy studies, preferably histopathology. Due to our inclusion criteria, all included studies used histology as reference standard to some extent, but when inconclusive or not available, several studies also accepted other diagnostic procedures and investigations, e.g., other imaging, endoscopies, or clinical follow-up. Only in a few of the studies, the information provided allowed us to determine the separate reference standard on a per-patient basis.$^{[25–27,29,30]}$ In these studies, we accepted the reference standard including clinical follow-up (or other imaging during follow-up) as reference standard to provide an improved standard of comparison.

Prior studies have shown similar detection rates in CUP patients. Dong et al$^{[38]}$ extracted 28 studies (comprising 910 patients) published between 1990 and 2007. In 8 of these studies (5 were retrospective and 3 were prospective), 430 patients with CUP were diagnosed by 18F-FDG PET/CT, and 31.4% (N = 135) of the primary tumors were detected. Pooled sensitivity and specificity were 0.81 (95% CI: 0.74–0.87) and 0.83 (95% CI: 0.78–0.87), respectively. Kwee and Kwee$^{[39]}$ included 11 studies (8 retrospective, 3 unclear), comprising a total sample size of 433 patients with CUP, and found an overall primary tumor detection rate, a pooled sensitivity and specificity of 18F-FDG PET/CT of 37%, 84% (95% CI: 78–88%) and 84% (95% CI 78–89%), respectively. Similarly, the systematic review by Møller et al$^{[8]}$ on 4 retrospective studies found that 18F-FDG PET/CT detected the primary tumor in 39.5% of patients with extracervical CUP; the pooled sensitivity, specificity, and accuracy of 18F-FDG PET/CT in the detection of the primary tumor site were 87%, 88%, and 87.5%, respectively. All studies included by Møller et al$^{[8]}$ were also included here,$^{[31,34–36]}$ whereas we included only 5 of 11 studies used by Kwee and Kwee$^{[39]}$ and 5 of 8 studies used by Dong et al,$^{[38]}$ namely.$^{[33–37]}$ All 3 meta-analyses$^{[8,38,39]}$ were affected by large heterogeneity between studies. The authors concluded that prospective studies are warranted to investigate the assumed advantage of 18F-FDG PET/CT over 18F-FDG PET alone and to explore causes of heterogeneity$^{[39]}$ and should employ more uniform inclusion criteria to evaluate the exact value of 18F-FDG PET/CT as a diagnostic tool in CUP patients with extracervical metastases.$^{[8]}$

Finally, there is a special issue when assessing the usefulness of PET in CUP, an issue that most existing literature did not consider, namely the fact that by the very nature of disease PET is useful for the detection of the primary only when applied early in
the course of disease. As soon as more than a few suspicious foci have developed, it becomes increasingly difficult for PET to point to a potential primary tumor the more lesions have become visible. These cases will paradoxically count as “false negatives” despite the fact that PET does actually “see” something and would probably have had a fair chance to detect the primary tumor had the patient come earlier to PET imaging. From the existing literature it was not possible to estimate how often this situation handicapped the performance of PET/CT.

CUP patients have per definition advanced illness and will typically receive a generic rather than a specific (and possibly more effective) treatment if the origin cannot be established. This leads to a poorer prognosis and, thus, establishing the underlying primary is pivotal to ensure timely and effective treatment. Standard diagnostic workup is heterogeneous and often comprises a multitude of more or less invasive procedures. An accurate single modality is therefore in high demand and 18F-FDG PET/CT has shown promising results in this regard; it is another important, but much larger, group, namely patients with serious nonspecific symptoms or signs without any biopsy evidence of cancer.

Acknowledgments
The authors thank Tove Faber Frandsen, PhD (Medical Research Library, Odense University Hospital) for her input to the literature search strategies and Lars Jelsstrup Petersen, MD DMSc CBA (Aalborg University Hospital) for his comments on an earlier version of the manuscript.

References


