
Nuclear Imaging for Classic Fever of Unknown Origin: Meta-Analysis

Motoki Takeuchi¹, Issa J. Dahabreh², Takashi Nihashi³, Mitsunaga Iwata¹, George M. Varghese⁴, and Teruhiko Terasawa¹

¹Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, Japan; ²Center for Evidence-based Medicine, Brown University, Providence, Rhode Island; ³Department of Radiology, Nagoya University Graduate School of Medicine, Nagoya, Japan; and ⁴Department of Infectious Diseases, Christian Medical College, Vellore, India

Several studies have assessed nuclear imaging tests for localizing the source of fever in patients with classic fever of unknown origin (FUO); however, the role of these tests in clinical practice remains unclear. We systematically reviewed the test performance, diagnostic yield, and management decision impact of nuclear imaging tests in patients with classic FUO. **Methods:** We searched PubMed, Scopus, and other databases through October 31, 2015, to identify studies reporting on the diagnostic accuracy or impact on diagnosis and management decisions of ¹⁸F-FDG PET alone or integrated with CT (¹⁸F-FDG PET/CT), gallium scintigraphy, or leukocyte scintigraphy. Two reviewers extracted data. We quantitatively synthesized test performance and diagnostic yield and descriptively analyzed evidence about the impact on management decisions. **Results:** We included 42 studies with 2,058 patients. Studies were heterogeneous and had methodologic limitations. Diagnostic yield was higher in studies with higher prevalence of neoplasms and infections. Nonneoplastic causes, such as adult-onset Still's disease and polymyalgia rheumatica, were less successfully localized. Indirect evidence suggested that ¹⁸F-FDG PET/CT had the best test performance and diagnostic yield among the 4 imaging tests; summary sensitivity was 0.86 (95% confidence interval [CI], 0.81–0.90), specificity 0.52 (95% CI, 0.36–0.67), and diagnostic yield 0.58 (95% CI, 0.51–0.64). Evidence on direct comparisons of alternative imaging modalities or on the impact of tests on management decisions was limited. **Conclusion:** Nuclear imaging tests, particularly ¹⁸F-FDG PET/CT, can be useful in identifying the source of fever in patients with classic FUO. The contribution of nuclear imaging may be limited in clinical settings in which infective and neoplastic causes are less common. Studies using standardized diagnostic algorithms are needed to determine the optimal timing for testing and to assess the impact of tests on management decisions and patient-relevant outcomes.

Key Words: FUO; FDG-PET; FDG-PET/CT; gallium scintigraphy; leukocyte scintigraphy

J Nucl Med 2016; 57:1913–1919

DOI: 10.2967/jnumed.116.174391

Classic fever of unknown origin (FUO) is defined as fever of 38.3°C (101°F) or higher for 3 or more weeks in immunocompetent, otherwise healthy patients with no identified cause of fever after undergoing a set of obligatory investigations (1). Common causes of classic FUO include infections, tumors, and noninfectious inflammatory diseases (2). With recent advances in diagnostic technologies, such as sophisticated imaging tests, improved culture techniques, and molecular diagnostics, the 2 most common causes (infections and neoplasms) have become less common and undiagnosed cases have become more challenging to investigate (3,4).

⁶⁷Ga scintigraphy was the mainstay staging modality in oncology practice (5) until the introduction of ¹⁸F-FDG PET in the 1990s. Because gallium accumulates in both malignant tumors and inflammation, gallium scintigraphy is still used as a component of workup strategies for patients with classic FUO (2,6). Other scintigraphy methods using autologous white blood cells labeled with ¹¹¹In or ^{99m}Tc are also used when infectious causes are suspected (2,6). ¹⁸F-FDG PET is a functional imaging modality that can be used for localizing malignant tumors, as well as infectious and noninfectious inflammatory lesions, because ¹⁸F-FDG accumulates in malignant cells and activated leukocytes (6).

Several studies have assessed the diagnostic usefulness of nuclear imaging tests for localizing a source of fever in such patients. However, the interpretation of these studies is not straightforward. First, classic FUO has a broad differential diagnosis and FUO causes vary across clinical settings (3,7). Such variation, coupled with differences in preimaging workup algorithms across settings, can affect test performance. Second, variability in the reference standards used across studies (e.g., biopsy for malignancies, cultures for infections, and operational diagnostic criteria for autoimmune disorders) affects the cause-specific performance of tests. Third, nuclear imaging tests are often evaluated using routinely collected clinical data, without a standardized postimaging diagnostic algorithm; in such cases, the imaging results influence the selection of further (confirmatory) tests and may introduce differential verification (8). Fourth, studies focus on a single nuclear imaging test and rarely report direct comparisons among alternative modalities. Previous systematic reviews (9–12) have also focused on single tests and have not provided comparative information. Our review attempts to address some of these challenges by synthesizing current evidence on the diagnostic performance and clinical utility of ¹⁸F-FDG PET alone or integrated with CT (¹⁸F-FDG PET/CT), gallium scintigraphy, and leukocyte scintigraphy for classic FUO.

Received Feb. 18, 2016; revision accepted May 31, 2016.

For correspondence or reprints contact either of the following:
Teruhiko Terasawa, Fujita Health University Hospital, 1-98 Dengakugakubo, Kutsukakecho, Toyoake, Aichi, 470-1192 Japan.
E-mail: terasawa@fujita-hu.ac.jp

Motoki Takeuchi, Fujita Health University Hospital, 1-98 Dengakugakubo, Kutsukakecho, Toyoake, Aichi, 470-1192 Japan.

E-mail: motoki-t@fujita-hu.ac.jp

Published online Jun. 23, 2016.

COPYRIGHT © 2016 by the Society of Nuclear Medicine and Molecular Imaging, Inc.

MATERIALS AND METHODS

Data Sources and Search Strategy

We searched PubMed and Scopus (from inception until October 31, 2015) with no language restrictions. We used prespecified search terms for the target condition (e.g., “fever of unknown origin” or “FUO”) and the tests of interest (e.g., “scintigraphy”, “PET”, “PET/CT”, or “SPECT”). We perused the reference lists of eligible primary papers and relevant reviews and meta-analyses. We also tracked citations to eligible papers through Scopus, Web of Science, and Google Scholar. Supplemental Table 1 (supplemental materials are available at <http://jnm.snmjournals.org>) provides the complete search strategy.

Study Selection

Two reviewers independently screened abstracts and examined the full text of potentially eligible papers to identify studies that evaluated ^{18}F -FDG PET or PET/CT using a full-ring scanner, gallium or leukocyte scintigraphy, or SPECT for at least 10 patients with classic FUO. We included only studies that reported sufficient information to calculate sensitivity and specificity, diagnostic yield (the proportion of patients in whom the imaging results were reported to contribute to the diagnosis of FUO causes), or the proportion of patients in whom the imaging results were deemed to have contributed to changes in diagnostic or therapeutic strategies planned before imaging. When multiple studies reported results from potentially overlapping patient groups, we used only information from the largest patient group. We excluded studies that exclusively evaluated patients 17 y or younger or studies that included patients infected with HIV and did not report separate data on non-HIV participants.

Data Extraction and Assessments of Risk of Bias and Applicability

One investigator extracted descriptive information, which was confirmed by a second investigator; discrepancies were resolved by consensus. We extracted information on study design, preimaging tests, characteristics of enrolled patients, FUO causes, index imaging tests and diagnostic criteria (13–16), and the reference standard. We categorized studies into 3 groups (first-, second-, and third-level examinations) based on their preimaging diagnostic workup algorithm. We also classified the reported reference standards into 3 degrees of accuracy (high, moderate, and low accuracy). The supplemental materials provide detailed descriptions.

Two independent reviewers assessed risk of bias and applicability for each eligible study using items based on the QUADAS-2 tool (17). We evaluated the risk of differential verification bias as proposed elsewhere (8). Discrepancies were resolved by consensus.

Data Synthesis

For each study, we constructed a 2×2 contingency table consisting of true-positive, false-positive, false-negative, and true-negative results, whereby patients were categorized according to their nuclear imaging test results (positive or negative) and whether the cause of FUO was correctly identified in the imaging-positive sites or not (i.e., a cause was identified outside the imaging-positive sites or the cause remained unknown).

We estimated summary sensitivity and specificity with their corresponding 95% confidence intervals (95% CIs) using bivariate random-effects meta-analysis with binomial within-study likelihood when 4 or more studies were available for the same imaging test (18,19). When a bivariate model failed to converge, we calculated summary sensitivity and specificity separately by univariate random-effects meta-analysis using mixed-effects logistic regression (20,21). We visually assessed between-study heterogeneity by plotting study estimates in the receiver-operating-characteristic (ROC) space (22).

We also constructed hierarchical summary ROC curves and obtained the confidence regions for the summary sensitivity and specificity (23). We performed meta-analysis of diagnostic yield using univariate random-effects logistic regression (20). We quantified between-study heterogeneity by estimating the between-study variance. Data on diagnostic or therapeutic decision impact were qualitatively synthesized.

To explore heterogeneity, we used subgroup analyses and univariable meta-regressions, when 10 or more studies with pertinent data were available (24). Specifically, we examined study design, geographic area, clinical context, proportion of identified neoplasms and infections, and use of contrast-enhancement or not for ^{18}F -FDG PET/CT.

In stability analyses, we excluded studies that enrolled patients younger than 18 y and did not report separate data on adult participants. To address the heterogeneity in classification and reporting of benign (spontaneously regressing) causes, we recalculated sensitivity and specificity for each study considering benign cases as disease negative along with cases in which no cause was identified (supplemental materials; Supplemental Table 2).

We visually assessed the results of studies directly comparing test modalities by plotting all estimates from the same study in the ROC space. We did not perform meta-analyses of comparative studies because few studies were available for each comparison. We also indirectly compared test performance by visually assessing summary ROC curves and by estimating relative diagnostic odds ratios comparing alternative imaging tests (24,25). We also performed study-level univariable meta-regressions to assess the difference in diagnostic yield between 2 imaging tests. The detailed methods used for indirect comparisons are described in the supplemental materials.

We did not perform tests for funnel plot asymmetry because they do not provide a valid way for assessing the extent and impact of missing data (26). All analyses were conducted using Stata SE 13.1 (Stata Corp.) and WinBUGS 1.4.3 (MRC Biostatistics Unit) (27). *P* values were 2-tailed, and statistical significance was defined as a *P* value of less than 0.05.

RESULTS

Literature Flow and Eligible Studies

We screened 6,351 abstracts and evaluated 83 full-text articles (Fig. 1). The supplemental materials provide a list of excluded studies along with their reasons for exclusion. A total of 43 unique publications (4 comparative and 38 noncomparative studies, including 2,058 unique patients) met our eligibility criteria (Supplemental Table 3). Twenty-two studies (1,137 patients) evaluated ^{18}F -FDG PET/CT, 12 (522 patients) ^{18}F -FDG PET, 6 (397 patients) gallium scintigraphy, and 6 (153 patients) ^{111}In -labeled leukocyte scintigraphy.

Study and Patient Characteristics

Studies of gallium scintigraphy and leukocyte scintigraphy have been published since the 1980s, whereas studies of ^{18}F -FDG PET and ^{18}F -FDG PET/CT are more recent, having been published after 2001 and 2008, respectively (Supplemental Table 3). Studies included a median of 48 patients (minimum–maximum, 10–162). Typically, studies retrospectively assessed nuclear imaging tests as second- or third-level examinations, performed after diagnostic workup that was not standardized in each study. Eight studies (19%) were prospective: 1 study adopted a standardized workup algorithm (28), and 3 studies specified mandatory preimaging tests (29–31) as selection criteria, one of which routinely performed thoracoabdominal contrast-enhanced CT (31). The average

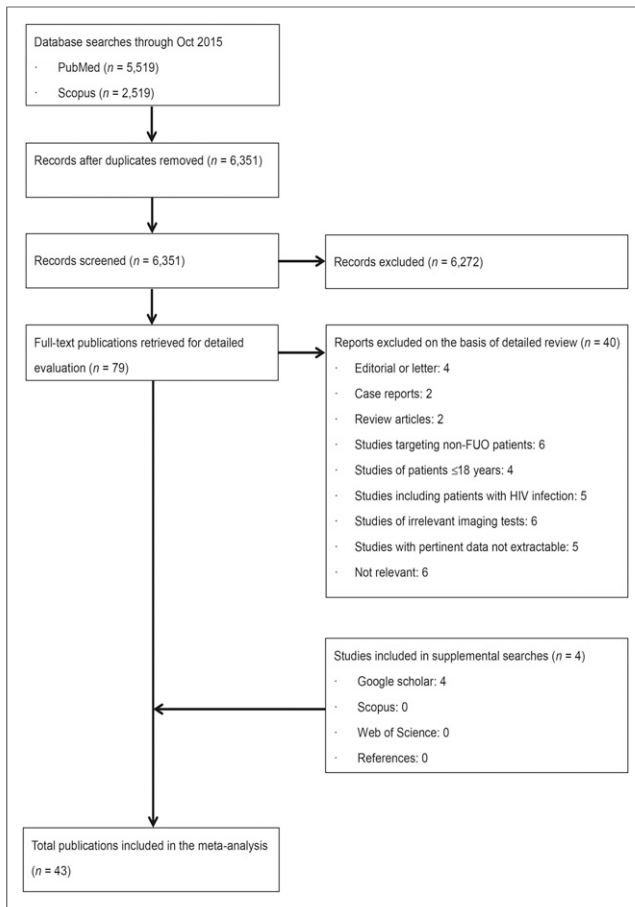


FIGURE 1. Study flow diagram.

preimaging disease duration ranged from 5 to 41 wk. Thirteen studies (31%) reported details on the postimaging diagnostic workup used for each patient. Positive scan results were often verified with invasive but high-accuracy reference standards, such as biopsy (median, 44%; 25th–75th percentile, 38%–64%, of study participants with positive scan vs. median, 13%; 25th–75th percentile, 0%–20%, with negative scan); negative results were typically verified with less-invasive but low-accuracy reference standards, such as clinical follow-up (median, 70%; 25th–75th percentile, 60%–88%). The average postimaging follow-up ranged from 3 to 30 mo (median, 15 mo).

The average age ranged from 42 to 62 y (median, 54 y), and the proportion of patients with specific FUO etiologies varied across studies (Supplemental Table 4). The median proportion of patients with infectious or neoplastic etiologies as a final diagnosis was 48% (minimum–maximum, 23%–80%) for studies of ^{18}F -FDG PET/CT, 31% (minimum–maximum, 10%–80%) for ^{18}F -FDG PET, 42% (minimum–maximum, 26%–58%) for gallium scintigraphy, and 37% (minimum–maximum, 12%–56%) for leukocyte scintigraphy. The proportion of cases deemed to be of infectious or neoplastic etiology was not associated with preimaging workup algorithms (Spearman $\rho = 0.065$; $P = 0.78$) or publication year (Spearman $\rho = 0.22$; $P = 0.15$). The proportion of undiagnosed cases also varied substantially (minimum–maximum, 4%–56%).

Test Characteristics

The supplemental materials describe how imaging was performed and interpreted. Studies generally adopted standard imaging protocols (13–16), and multiple nuclear medicine physicians visually interpreted the results (Supplemental Tables 5 and 6). Few studies reported diagnostic thresholds for quantitative assessment to support the visual assessment.

Assessment of Study Risk of Bias and Applicability

We had concerns about high risk of bias and limited generalizability in most included studies for all nuclear imaging tests (Supplemental Fig. 1). Specifically, risk of bias due to differential verification (i.e., selection of reference standard tests on the basis of index test results) was a concern in all studies. Further, the nuclear imaging results were suggested as the basis for the final diagnosis of at least 1 case of noninfectious inflammatory diseases (e.g., vasculitis) in 7 studies (32–38); such incorporation of index test results into the reference standard can lead to overestimation of test performance.

Test Performance, Diagnostic Yield, and Impact on Management

Studies of ^{18}F -FDG PET/CT produced heterogeneous estimates of sensitivity and specificity (Supplemental Fig. 2A). The summary sensitivity and specificity were 0.86 (95% CI, 0.81–0.90) and 0.52 (95% CI, 0.36–0.67), respectively (Fig. 2). Summary estimates were stable in subgroup analyses, and meta-regression analysis identified no covariates that significantly related with both sensitivity and specificity (data not shown).

The diagnostic yield of ^{18}F -FDG PET/CT also varied across studies (Supplemental Fig. 3A). Studies reporting a higher proportion of neoplasms and infections as the cause of FUO also reported a higher diagnostic yield (Spearman $\rho = 0.44$; $P = 0.038$) (Fig. 3A). ^{18}F -FDG PET/CT successfully localized a cause of FUO in approximately 60% of patients (summary diagnostic yield, 0.58; 95% CI, 0.51–0.64) (Fig. 4). The summary estimates were similar across subgroups and in stability analyses; meta-regression indicated that the proportion of neoplasms and infections was positively associated with diagnostic yield ($P = 0.003$). Nonneoplastic causes were less successfully localized; adult-onset Still's disease, tuberculosis, and polymyalgia rheumatica were the 3 causes for which ^{18}F -FDG PET/CT most often showed no pathologic uptake leading to diagnosis (Supplemental Tables 7 and 8).

Studies of ^{18}F -FDG PET reported variable estimates of sensitivity and specificity (Supplemental Fig. 2B). Summary sensitivity and specificity were 0.76 (95% CI, 0.66–0.83) and 0.50 (95% CI, 0.30–0.70), respectively (Fig. 2). In meta-regression, no covariates were significantly associated with both sensitivity and specificity (data not shown). The diagnostic yield also varied across studies (Supplemental Fig. 3B), with a summary estimate of 0.44 (95% CI, 0.31–0.58) (Fig. 4). Again, studies with a higher diagnostic yield also reported a higher proportion of combined neoplasms and infections (Spearman $\rho = 0.66$; $P = 0.020$) (Fig. 3B). A positive association was also suggested by meta-regression between the prevalence of neoplasms and infections and diagnostic yield ($P = 0.010$). Similarly to ^{18}F -FDG PET/CT, stand-alone ^{18}F -FDG PET frequently failed to localize nonneoplastic causes (Supplemental Tables 7 and 8).

Test performance of gallium scintigraphy was heterogeneous (Fig. 2C). The summary sensitivity and specificity were 0.60 (95% CI, 0.45–0.73) and 0.63 (95% CI, 0.37–0.84), respectively (Fig. 2). Diagnostic yield ranged from 0.21 to 0.54 (Supplemental Fig. 3C), and on average, the location of a source of fever was

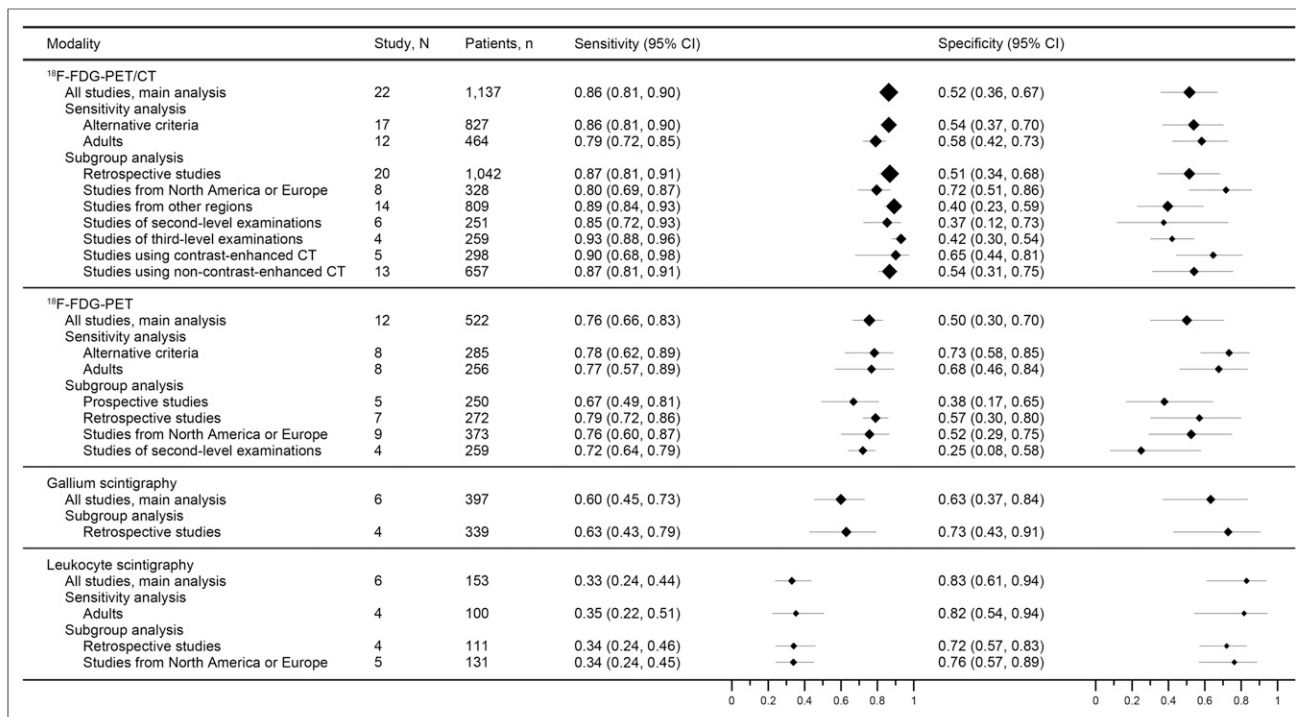


FIGURE 2. Meta-analysis of sensitivity and specificity. Diamonds (proportional to number of patients) represent point estimates; extending lines represent 95% CI of each estimate.

correctly localized in approximately a third of patients (summary diagnostic yield, 0.35; 95% CI, 0.25–0.46) (Fig. 4). Data on gallium scintigraphy were too sparse for reliable subgroup analysis.

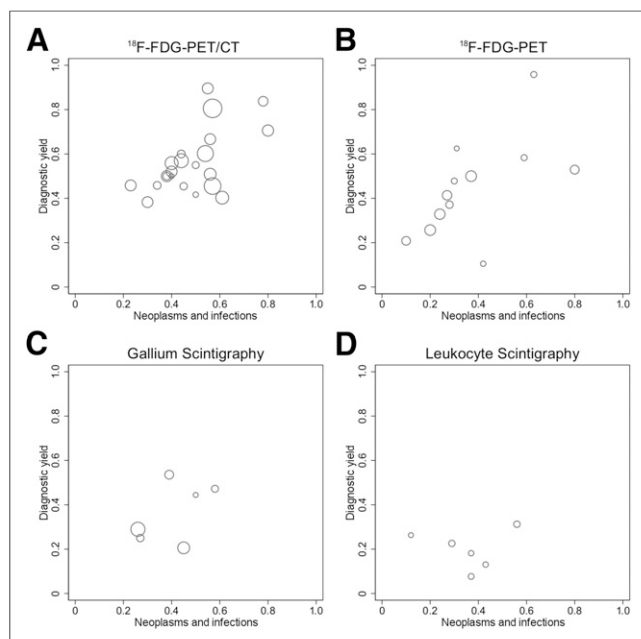


FIGURE 3. Diagnostic yield plotted against prevalence of infections and neoplasms for studies of ¹⁸F-FDG PET/CT (A), ¹⁸F-FDG PET (B), gallium scintigraphy (C), and leukocyte scintigraphy (D). Size of each circle is proportional to sample size for each study.

Studies of leukocyte scintigraphy reported rather homogeneous estimates of specificity; however, estimates of sensitivity were variable (Supplemental Fig. 2D). The summary sensitivity and specificity were 0.33 (95% CI, 0.24–0.44) and 0.83 (95% CI, 0.61–0.94), respectively (Fig. 2). Estimates of diagnostic yield ranged from 0.08 to 0.31 (Supplemental Fig. 3D), and overall, the FUC was correctly identified on the basis of the scan results in only a fifth of patients (summary diagnostic yield, 0.20; 95% CI, 0.14–0.28) (Fig. 4). The summary estimates were stable in subgroup and sensitivity analyses, although data were limited.

Five studies of ¹⁸F-FDG PET/CT (31,39–42), 3 studies of ¹⁸F-FDG PET (28,29,37), and 3 studies of leukocyte scintigraphy (43–45) performed univariable or multivariable analyses to identify predictors of the impact of tests on diagnosis or therapeutic management (Supplemental Table 9). The 3 most commonly assessed predictors were c-reactive protein, leukocyte counts, and erythrocyte sedimentation rate, although studies were highly heterogeneous regarding how candidate predictors were measured and how they were incorporated in the models (e.g., dichotomized or transformed). Overall, no predictive model was validated. Three studies (28,34,46) reported how often scans altered diagnosis, and 2 studies (37,46) reported how often scans altered therapeutic decisions (Supplemental Table 10). For example, ¹⁸F-FDG PET or PET/CT affected diagnostic and therapeutic management in 44% (46) and 36% of cases (37), respectively.

Comparisons of Test Performance and Diagnostic Yield Among Imaging Tests

One study directly compared ¹⁸F-FDG PET with ¹⁸F-FDG PET/CT (47), another compared ¹⁸F-FDG PET with gallium scintigraphy (48), and 2 compared ¹⁸F-FDG PET with leukocyte scintigraphy (49,50). This limited evidence was insufficient to establish any

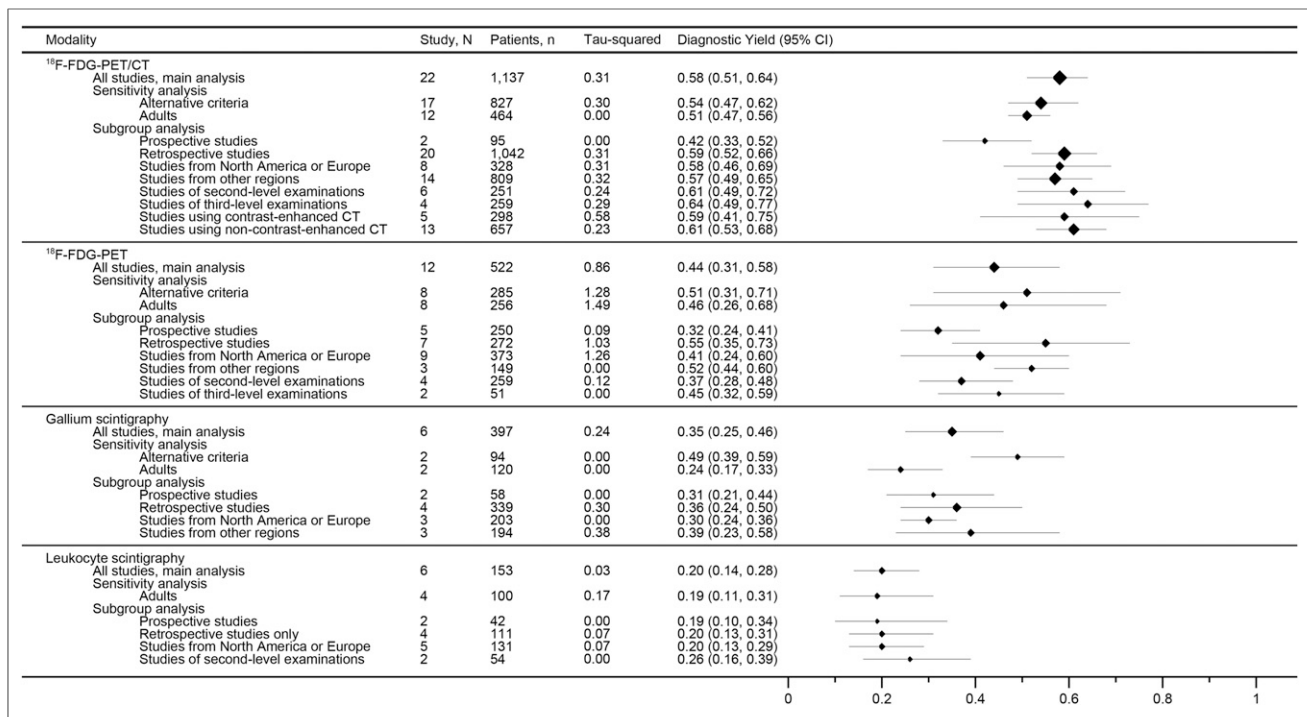


FIGURE 4. Meta-analysis of diagnostic yield. Squares (proportional to numbers of patients) represent point estimates; extending lines represent 95% CI of each estimate.

nuclear imaging modality as superior to any other (Supplemental Fig. 4). Regarding indirect comparisons of test performance, visual assessment of the summary ROC curves and meta-regression suggested that ¹⁸F-FDG PET/CT outperformed standalone ¹⁸F-FDG PET, gallium scintigraphy, and leukocyte scintigraphy (Supplemental Figs. 2 and 5). Similarly, visual and quantitative indirect comparisons of diagnostic yield suggested that ¹⁸F-FDG PET/CT was more likely to correctly identify the cause of FUO than alternative tests (Supplemental Figs. 3 and 6). Detailed descriptions are reported in the supplemental materials.

DISCUSSION

We examined 42 studies involving nearly 2,000 patients with classic FUO to evaluate the performance and clinical validity of nuclear imaging modalities for localizing the source of fever. We found that ¹⁸F-FDG PET/CT had the best test performance among nuclear imaging tests and correctly pinpointed the anatomic location of FUO pathologies in approximately 60% of patients for whom basic laboratory tests and anatomic imaging had failed. Indirect comparisons of test modalities suggested that ¹⁸F-FDG PET/CT had a better performance than alternative nuclear imaging modalities. However, few studies reported direct comparison among tests, and the diagnostic yield of ¹⁸F-FDG PET and PET/CT was lower in clinical contexts in which the prevalence of neoplastic and infectious causes of FUO was low. Furthermore, the data on gallium scintigraphy and leukocyte scintigraphy were derived from older studies and thus may not be applicable to contemporary clinical practice.

Our findings about the test performance of ¹⁸F-FDG PET and ¹⁸F-FDG PET/CT are in general agreement with previous systematic reviews (9–12). However, we extended previous work by examining diagnostic yield and demonstrating a positive associa-

tion between yield and the prevalence of infections and neoplasms. This suggests that ¹⁸F-FDG PET and ¹⁸F-FDG PET/CT may be less useful in clinical settings in which these causes are less common. In addition, we assessed evidence on comparative test performance and other outcomes more thoroughly than previous reviews. We found that ¹⁸F-FDG PET/CT is the best choice and can localize most malignant or infectious lesions. A negative scan, however, does not exclude other causes of fever, such as noninfectious inflammatory diseases. Gallium scintigraphy can be used to localize malignancies when it is the only readily available modality. Leukocyte scintigraphy has a limited role in most clinical contexts, in view of its low diagnostic yield. Finally, we qualitatively summarized data on the impact of ¹⁸F-FDG PET and ¹⁸F-FDG PET/CT on diagnostic and therapeutic management. Lack of data, however, precluded reliable assessment on how ¹⁸F-FDG PET or ¹⁸F-FDG PET/CT affects estimates of the probability of particular FUO cases or how it alters diagnostic or therapeutic management decisions. Arguably, these outcomes are more relevant to clinical practice than test performance, suggesting that more research, ideally comparative studies of alternative testing strategies incorporating nuclear imaging tests, is needed to assess clinical outcomes beyond test performance (51).

Our review has several limitations. First, our summary estimates were based on heterogeneous estimates from studies with varied designs, and we noted substantial between-study heterogeneity. We believe that variability in the prevalence of underlying FUO etiologies across studies and differences in workup algorithms are the main sources of the observed statistical heterogeneity; it is hard to account for these factors without access to primary study data. Second, the included studies used multiple and imperfect reference standards and were deemed likely to have produced

biased results because of differential verification and incorporation of the index test result in the reference standard. Differential verification and incorporation bias are likely to lead to over-estimation of test performance (52), and thus our summary estimates should be interpreted with caution. Third, our comparative results depend largely on indirect comparisons, which may be confounded by differences across studies that cannot be addressed analytically (25). We thus view our indirect comparisons as suggestive of possible differences among modalities that merit further research in properly designed and conducted comparative studies. Fourth, the available literature provides only limited evidence on the impact of nuclear imaging tests on clinical management.

CONCLUSION

¹⁸F-FDG PET/CT, ¹⁸F-FDG PET, gallium scintigraphy, and leucocyte scintigraphy are useful imaging modalities for localizing the source of fever in patients with classic FUO for whom a routine diagnostic workup has been unsuccessful in establishing a diagnosis. ¹⁸F-FDG PET/CT is the most promising nuclear test, with the highest diagnostic yield. However, little is known about the impact of all nuclear imaging tests on diagnostic and therapeutic decisions. Future studies should use designs with a prespecified diagnostic algorithms before nuclear imaging, and standardized protocols for imaging and image interpretation, to clarify which modalities are most useful. Also, given the heterogeneous test performance across studies and the high cost of nuclear imaging, identifying factors that modify test performance and affect management and using methods that combine nuclear imaging results with other information should be a priority.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734. This study was supported by MEXT, Japan (No. 26460755). No other potential conflict of interest relevant to this article was reported.

REFERENCES

- Bleeker-Rovers CP, van der Meer JWM. Fever of unknown origin. In: Fauci A, Hauser S, Longo D, Jameson J, Loscalzo J, Kasper D, eds. *Harrison's Principles of Internal Medicine*. 19th ed. New York, NY: McGraw Hill; 2015:135–142.
- Knockaert DC, Vanderschueren S, Blockmans D. Fever of unknown origin in adults: 40 years on. *J Intern Med*. 2003;253:263–275.
- Varghese GM, Trowbridge P, Doherty T. Investigating and managing pyrexia of unknown origin in adults. *BMJ*. 2010;341:C5470.
- Horowitz HW. Fever of unknown origin or fever of too many origins? *N Engl J Med*. 2013;368:197–199.
- Hicks RJ, Hofman MS. Is there still a role for SPECT-CT in oncology in the PET-CT era? *Nat Rev Clin Oncol*. 2012;9:712–720.
- Rennen HJM, Bleeker-Rovers C, Oyen WG. Imaging infection and inflammation. In: Baert AL, Sartor K, Schiepers C, eds. *Diagnostic Nuclear Medicine*. Berlin Heidelberg, Germany: Springer; 2006:113–126.
- Mourad O, Palda V, Detsky AS. A comprehensive evidence-based approach to fever of unknown origin. *Arch Intern Med*. 2003;163:545–551.
- Naaktgeboren CA, de Groot JA, van Smeden M, Moons KG, Reitsma JB. Evaluating diagnostic accuracy in the face of multiple reference standards. *Ann Intern Med*. 2013;159:195–202.
- Dong MJ, Zhao K, Liu ZF, Wang GL, Yang SY, Zhou GJ. A meta-analysis of the value of fluorodeoxyglucose-PET/PET-CT in the evaluation of fever of unknown origin. *Eur J Radiol*. 2011;80:834–844.
- Hao R, Yuan L, Kan Y, Li C, Yang J. Diagnostic performance of ¹⁸F-FDG PET/CT in patients with fever of unknown origin: a meta-analysis. *Nucl Med Commun*. 2013;34:682–688.
- van der Ende EL, Punwasia RVG, van Daele PLA, Schurinkc KAM. Diagnostic value of FDG-PET/CT in fever of unknown origin. *Erasmus J Med*. 2013;3:32–37.
- Sioka C, Assimakopoulos A, Fotopoulos A. The diagnostic role of ¹⁸F fluorodeoxyglucose positron emission tomography in patients with fever of unknown origin. *Eur J Clin Invest*. 2015;45:601–608.
- Jamar F, Buscombe J, Chiti A, et al. EANM/SNMMI guideline for ¹⁸F-FDG use in inflammation and infection. *J Nucl Med*. 2013;54:647–658.
- Palestro C, Brown M, Forstrom L, et al. Society of Nuclear Medicine procedure guideline for gallium scintigraphy in inflammation 3.0. SNMMI website. http://snmmi.files.cms-plus.com/docs/Gallium_Scintigraphy_in_Inflammation_v3.pdf. Approved June 2, 2004. Accessed July 20, 2016.
- Roca M, de Vries EF, Jamar F, Israel O, Signore A. Guidelines for the labelling of leucocytes with ¹¹¹In-oxine. *Eur J Nucl Med Mol Imaging*. 2010;37:835–841.
- Palestro C, Brown M, Forstrom L, et al. Society of Nuclear Medicine Procedure guideline for ¹¹¹In-leukocyte scintigraphy for suspected infection/inflammation. SNMMI website. http://snmmi.files.cms-plus.com/docs/Leukocyte_v3.pdf. Approved June 2, 2004. Accessed July 20, 2016.
- Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536.
- Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58:982–990.
- Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol*. 2006;59:1331–1332, author reply 1332–1333.
- Trikalinos TA, Trow P, Schmid CH. Simulation-based comparison of methods for meta-analysis of proportions and rates. AHRQ Publication No. 13(14)-EHC084-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013.
- Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res*. June 26, 2015 [Epub ahead of print].
- Phillips B, Stewart LA, Sutton AJ. ‘Cross hairs’ plots for diagnostic meta-analysis. *Res Synth Methods*. 2010;1:308–315.
- 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.
- Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. Cochrane website. <http://dta.cochrane.org/sites/dta.cochrane.org/files/uploads/Chapter%2010%20-%20Version%201.0.pdf>. Accessed July 20, 2016.
- Takwoingi Y, Leeftang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Ann Intern Med*. 2013;158:544–554.
- Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
- Thompson J. *Bayesian Analysis with Stata*. College Station, TX: Stata Press; 2014.
- Bleeker-Rovers CP, Vos FJ, Mudde AH, et al. A prospective multi-centre study of the value of FDG-PET as part of a structured diagnostic protocol in patients with fever of unknown origin. *Eur J Nucl Med Mol Imaging*. 2007;34:694–703.
- Buysschaert I, Vanderschueren S, Blockmans D, Mortelmans L, Knockaert D. Contribution of ¹⁸fluoro-deoxyglucose positron emission tomography to the work-up of patients with fever of unknown origin. *Eur J Intern Med*. 2004;15:151–156.
- Meller J, Altenvoerde G, Munzel U, et al. Fever of unknown origin: prospective comparison of [¹⁸F]FDG imaging with a double-head coincidence camera and gallium-67 citrate SPET. *Eur J Nucl Med*. 2000;27:1617–1625.
- Singh N, Kumar R, Malhotra A, Bhalla AS, Kumar U, Sood R. Diagnostic utility of fluorodeoxyglucose positron emission tomography/computed tomography in pyrexia of unknown origin. *Indian J Nucl Med*. 2015;30:204–212.
- Balink H, Collins J, Bruyn GA, Gemmel F. F-18 FDG PET/CT in the diagnosis of fever of unknown origin. *Clin Nucl Med*. 2009;34:862–868.
- Federici L, Blondet C, Imperiale A, et al. Value of ¹⁸F-FDG-PET/CT in patients with fever of unknown origin and unexplained prolonged inflammatory syndrome: a single centre analysis experience. *Int J Clin Pract*. 2010;64:55–60.
- Manohar K, Mittal BR, Jain S, et al. F-18 FDG-PET/CT in evaluation of patients with fever of unknown origin. *Jpn J Radiol*. 2013;31:320–327.
- Pedersen TI, Roed C, Knudsen LS, Loft A, Skinhoj P, Nielsen SD. Fever of unknown origin: a retrospective study of 52 cases with evaluation of the diagnostic utility of FDG-PET/CT. *Scand J Infect Dis*. 2012;44:18–23.
- Pelosi E, Skanjeti A, Penna D, Arena V. Role of integrated PET/CT with [¹⁸F]-FDG in the management of patients with fever of unknown origin: a single-centre experience. *Radiol Med (Torino)*. 2011;116:809–820.

37. Kubota K, Nakamoto Y, Tamaki N, et al. FDG-PET for the diagnosis of fever of unknown origin: a Japanese multi-center study. *Ann Nucl Med*. 2011;25:355–364.
38. Lorenzen J, Buchert R, Bohuslavizki KH. Value of FDG PET in patients with fever of unknown origin. *Nucl Med Commun*. 2001;22:779–783.
39. Crouzet J, Boudousq V, Lechiche C, et al. Place of ¹⁸F-FDG-PET with computed tomography in the diagnostic algorithm of patients with fever of unknown origin. *Eur J Clin Microbiol Infect Dis*. 2012;31:1727–1733.
40. Kim YJ, Kim SI, Hong KW, Kang MW. Diagnostic value of ¹⁸F-FDG PET/CT in patients with fever of unknown origin. *Intern Med J*. 2012;42:834–837.
41. Zheng Y, Lu W, Pan Q. Effect of ¹⁸F-FDG PET/CT on etiological diagnosis for fever of unknown origin (FUO). *J Clin Oncol*. 2013;19:936–939.
42. Gafter-Gvili A, Raibman S, Grossman A, et al. [¹⁸F]FDG-PET/CT for the diagnosis of patients with fever of unknown origin. *QJM*. 2015;108:289–298.
43. Kelly MJ, Kalff V, Hicks R, Spicer W, Spelman D. ¹¹¹In-oxine labelled leukocyte scintigraphy in the detection and localization of active inflammation and sepsis. *Med J Aust*. 1990;152:352–357.
44. Kjaer A, Lebech A. Diagnostic value of ¹¹¹In-granulocyte scintigraphy in patients with fever of unknown origin. *J Nucl Med*. 2002;43:140–144.
45. Syrjälä MT, Valtonen V, Liewendahl K, Myllylä G. Diagnostic significance of indium-111 granulocyte scintigraphy in febrile patients. *J Nucl Med*. 1987;28:155–160.
46. Tokmak H, Ergonul O, Demirkol O, Cetiner M, Ferhanoglu B. Diagnostic contribution of ¹⁸F-FDG-PET/CT in fever of unknown origin. *Int J Infect Dis*. 2014;19:53–58.
47. Kang L, Xu X, Fan Y, et al. Diagnostic value of fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography in fever of unknown origin [in Chinese]. *Beijing Da Xue Xue Bao*. 2015;47:175–180.
48. Blockmans D, Knockaert D, Maes A, et al. Clinical value of [¹⁸F]fluorodeoxyglucose positron emission tomography for patients with fever of unknown origin. *Clin Infect Dis*. 2001;32:191–196.
49. Kjaer A, Lebech AM, Eigtved A, Hojgaard L. Fever of unknown origin: prospective comparison of diagnostic value of ¹⁸F-FDG PET and ¹¹¹In-granulocyte scintigraphy. *Eur J Nucl Med Mol Imaging*. 2004;31:622–626.
50. Seshadri N, Sonoda LI, Lever AM, Balan K. Superiority of ¹⁸F-FDG PET compared to ¹¹¹In-labelled leucocyte scintigraphy in the evaluation of fever of unknown origin. *J Infect*. 2012;65:71–79.
51. Dahabreh IJ, Gatsonis C. A flexible, multifaceted approach is needed in health technology assessment of PET. *J Nucl Med*. 2014;55:1225–1227.
52. Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clin Epidemiol*. 2013;66:1093–1104.