Technology Assessment

Technology Assessment Program

Positron Emission Tomography for Nine Cancers (Bladder, Brain, Cervical, Kidney, Ovarian, Pancreatic, Prostate, Small Cell Lung, Testicular)

December 1, 2008

Prepared for:

Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, Maryland 20850

Positron Emission Tomography for Nine Cancers (Bladder, Brain, Cervical, Kidney, Ovarian, Pancreatic, Prostate, Small Cell Lung, Testicular)

Technology Assessment Report

Project ID: PETC1207

December 1, 2008

University of Alberta Evidence-based Practice Center

Research team:
Maria B. Ospina, M.Sc.[1] (*Project coordinator)*
Jennifer Horton, M.Sc.[1]
Jennifer Seida, M.P.H.[1]
Ben Vandermeer, M.Sc.[1]
Grace Liang, Ph.D.[1]

Clinical Investigators:
Alexander J. McEwan, F.R.C.P., M.B.B.S.[2]
Karen Gulenchyn, M.D., F.R.C.P.C.[3]

[1] University of Alberta Evidence-based Practice Center, Edmonton (AB), Canada
[2] Chair, Department of Oncology, University of Alberta, Edmonton (AB), Canada
[3] Chief, Department of Nuclear Medicine, Hamilton Health Sciences and St. Joseph's Healthcare, Hamilton (ON), Canada

This report is based on research conducted by the University of Alberta Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (HHSA 290 2007 10021 I, Task Order: 4; Work Assignment: 1). The findings and conclusions in this document are those of the authors who are responsible for its contents. The findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of the Agency for Healthcare Research and Quality or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decision-makers, patients and clinicians, health system leaders, and policymakers make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

None of the investigators has any affiliations or financial involvement related to the material presented in this report.

Acknowledgments

The following members of the University of Alberta Evidence-based Practice Center have contributed to the preparation of this report:

- Liza Bialy, B.Sc.
- Kenneth Bond, B.Ed, M.A.
- Nina Buscemi, Ph.D.
- Donna M Dryden, Ph.D.
- Lisa Hartling, M.Sc.
- Krystal Harvey, B.Sc.
- Mohammad Karkhaneh, M.D., Ph.D (Cand)
- Andrea Milne, B.Sc.N.

Many thanks to Thane Chambers, M.L.I.S. from the John W. Scott Health Sciences Library at the University of Alberta for designing the literature search strategies of this technology report.

Structured Abstract

Objective: To review and synthesize the evidence on the use of ^{18}FDG-PET in the assessment and treatment of nine types of cancer with respect to the following clinical situations: diagnosis, staging, restaging, and monitoring response to treatment.

Data Sources: Comprehensive searches were conducted in four relevant electronic databases for the time period from 2003 to March 2008.

Review Methods: Studies should be published in English, with more than 12 adult participants with primary cancer of the following type: bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular. Restrictions regarding study design were not imposed. Two independent reviewers assessed study relevance, extracted the data and assessed the methodological quality of the studies. A combination of qualitative and quantitative approaches was used to synthesize the data.

Results: One hundred and nine articles were included in this report. The strongest evidence for the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT has been produced for staging of locally advanced cervical cancer and the detection and restaging of recurrent disease, the detection of ovarian cancer recurrences following treatment, and the diagnosing and initial staging of pancreatic cancer. These and other indications require further research to show the impact of^{18}FDG-PET or ^{18}FDG-PET/CT on patient management or added value in the diagnostic and therapeutic process.

Conclusion: For some type of cancers (e.g., cervical, ovarian, and pancreatic cancer), there is some evidence of the utility of ^{18}FDG-PET or ^{18}FDG-PET/CT for diagnosing, staging, or detecting recurrences, but they may still require more studies to augment the evidence base. Further studies are needed to reach firm conclusions about the clinical effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT in terms of the impact on diagnosis and treatment options, patient-centered outcomes, and economic costs. It is still unclear how ^{18}FDG-PET and ^{18}FDG-PET/CT affects patient treatment and ultimately their outcome. For other types of cancer examined in the review (e.g., bladder, kidney, prostate, SCLC, and testicular) the answers are still inconclusive and require more careful study.

Table of Contents

Executive Summary……..1
Chapter 1. Introduction 7
Background 7
Scope of the report 10
Structure of the Report 12
Chapter 2. Methods 15
Overview 15
Literature Search and Retrieval 15
Criteria for Selection of Studies 16
Study Selection Process 17
Evaluating the Methodological Quality of Studies and Grading the Evidence 17
Data Collection 18
Evidence Synthesis 19
Chapter 3. Results 21
Search Results 21
1. Bladder Cancer 23
2. Brain Cancer 32
3. Cervical Cancer 39
4. Kidney Cancer 71
5. Ovarian Cancer 84
6. Pancreatic Cancer 110
7. Prostate Cancer 131
8. Small Cell Lung Cancer 136
9. Testicular Cancer 146
Chapter 4. Discussion 152
1. Bladder Cancer 152
2. Brain Cancer 153
3. Cervical Cancer 153
4. Kidney Cancer 155
5. Ovarian Cancer 156
6. Pancreatic Cancer 157
7. Prostate Cancer 158
8. Small Cell Lung Cancer 159
9. Testicular Cancer 160
Chapter 5. Conclusions 162
Strengths and Limitations 162
Conclusions and Future Areas of Research 164
Abbreviations 165
References and Included Studies 167

Tables and Figures

Tables:

Table 1. Hierarchy of Diagnostic Efficacy 14
Table 2. Databases Searched for Relevant Studies 15
Table 3. Inclusion Criteria 16
Table 4. Grading Scheme for Diagnostic Studies 18
Table 5. Bladder cancer stages 24

Table 6. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for bladder cancer 27
Table 7. Results of meta-analysis of the accuracy of ^{18}FDG-PET for bladder cancer 29
Table 8. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for bladder cancer 31
Table 9. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for brain cancer 36
Table 10. Main findings and types of bias that affected the evidence on ^{18}FDG-PET as part of a management strategy in brain cancer 38
Table 11. FIGO staging of cervical cancer 40
Table 12. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer 43
Table 13. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer 62
Table 14. Main findings and types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer 70
Table 15. Robson stages of renal cancer 72
Table 16. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for kidney cancer 74
Table 17. Results of meta-analyses of the accuracy of ^{18}FDG-PET for kidney cancer 79
Table 18. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for kidney cancer 83
Table 19. FIGO staging of ovarian cancer 85
Table 20. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer 88
Table 21. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer 101
Table 22. Main findings and types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer 109
Table 23. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer 113
Table 24. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer 123
Table 25. Main findings and types of bias that affected the evidence on the diagnostic thinking impact, effect on patient-centered outcomes and economic outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer 130
Table 26. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for prostate cancer 135
Table 27. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for SCLC 140
Table 28. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for SCLC 145
Table 29. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for testicular cancer 149
Table 30. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for testicular cancer 151

Figures:

Figure 1. Flow diagram for study retrieval and selection for the technology report 22
Figure 2. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer 27
Figure 3. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies) 28
Figure 4. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies) 28
Figure 5. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies) 29
Figure 6. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology for the staging of brain cancer 36
Figure 7. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer 45
Figure 8. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site) 47
Figure 9. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site) 48
Figure 10. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site) 49
Figure 11. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer 50
Figure 12. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer (retrospective studies) 50
Figure 13. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer (retrospective studies) 50
Figure 14. Summary ROC plot of ^{18}FDG-PET v. any reference standard for detecting recurrences of cervical cancer (retrospective studies) 51
Figure 15. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer 52
Figure 16. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies) 52
Figure 17. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies) 53
Figure 18. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies) 53
Figure 19. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer 54
Figure 20. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 54
Figure 21. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 55

Figure 22. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 55
Figure 23. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer 56
Figure 24. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies) 56
Figure 25. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies) 56
Figure 26. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies) 57
Figure 27. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer 58
Figure 28. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies) 58
Figure 29. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies) 58
Figure 30. Summary ROC plot of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies) 59
Figure 31. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer 60
Figure 32. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 60
Figure 33. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 60
Figure 34. Summary ROC plot of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies) 61
Figure 35. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer 75
Figure 36. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies) 75
Figure 37. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies) 75
Figure 38. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies) 76
Figure 39. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer 76
Figure 40. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies) 77
Figure 41. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies) 77
Figure 42. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies) 78

Figure 43. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer 89
Figure 44. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies) 89
Figure 45. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies) 90
Figure 46. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)..... 90
Figure 47. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer 91
Figure 48. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies) 91
Figure 49. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies) 91
Figure 50. Summary ROC plot of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies) 92
Figure 51. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (subgroup analysis)..... 93
Figure 52. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (prospective studies)..... 93
Figure 53. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (prospective studies)..... 93
Figure 54. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)..... 94
Figure 55. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer 94
Figure 56. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (retrospective studies) 95
Figure 57. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (retrospective studies) 95
Figure 58. Summary ROC plot of ^{18}FDG-PET v. any reference standard to detect recurrences of ovarian cancer (retrospective studies)..... 96
Figure 59. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (subgroup analysis)..... 97
Figure 60. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies) 97
Figure 61. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies) 97

Figure 62. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies) 98
Figure 63. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (subgroup analysis)..... 99
Figure 64. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies)..... 99
Figure 65. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies)..... 99
Figure 66. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies) 100
Figure 67. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer..... 114
Figure 68. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 114
Figure 69. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 115
Figure 70. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 116
Figure 71. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer 116
Figure 72. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 117
Figure 73. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 117
Figure 74. Summary ROC plot of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies) 118
Figure 75. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer..... 118
Figure 76. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies) 119
Figure 77. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies) 119
Figure 78. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies)..... 120
Figure 79. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer..... 120
Figure 80. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies)..... 121

Figure 81. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies) 121
Figure 82. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies) 122
Figure 83. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of SCLC 140
Figure 84. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. any reference standard for the staging of SCLC (retrospective studies) 141

Executive Summary

Introduction

The University of Alberta Evidence-based Practice Center (UAEPC) reviewed and synthesized the published literature on the use of ^{18}FDG-PET in the assessment and treatment of nine types of cancer with respect to the following clinical situations: diagnosis, staging, restaging, and monitoring response to treatment. The research questions were organized under four key questions:

- **1**^{18}FDG-PET and ^{18}FDG-PET/CT diagnostic test performance;
- Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT;
- ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy; and
- Cost-effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT.

Positron emission tomography (PET) is a form of nuclear medicine imaging that detects and establishes metabolic abnormalities in tissue. Compared to structural imaging techniques (X-ray, computed tomography [CT], and magnetic resonance imaging [MRI]), ^{18}FDG-PET is considered a potentially major advancement in clinical practice because it may provide information about the behavior of tumors in addition to the anatomic extent and thus can provide evidence to guide therapeutic choices. In the United States, the Centers for Medicaid and Medicare Services (CMS) has determined that there is sufficient evidence to show that a ^{18}FDG-PET scan is reasonable and necessary for certain indications in the pretreatment and management phase of non-small cell lung cancer, esophageal cancer, colorectal cancer, lymphoma, melanoma, breast cancer, head and neck cancers and thyroid cancer. There is a need to evaluate the evidence emerging recently on the benefits of ^{18}FDG-PET for other oncologic indications.

Methods

The UAEP established a prospectively designed protocol for this technology report. To accomplish the tasks as directed, a core research team composed of clinical investigators and methodologists was assembled. Comprehensive searches were conducted in four relevant electronic databases for the time period from 2003 to March 2008. Two independent reviewers applied a set of

inclusion and exclusion criteria to determine eligibility of potentially relevant studies. Disagreements about the inclusion or exclusion of studies were resolved by consensus among reviewers. Studies should be published in English, with more than 12 adult participants with primary cancer of the following type: bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular, and providing quantitative data on at least one outcome of interest. Restrictions regarding study design were not imposed.

Trained research assistants extracted the data using a comprehensive and pretested data extraction form. One reviewer verified the accuracy and completeness of the data. The methodological quality of studies on the diagnostic performance and the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT was assessed using the Scottish Intercollegiate Guidelines Network (SIGN) assessment tool for diagnostic studies. An individual components approach that considered important aspects of design, conduct, and reporting of effectiveness studies was used to assess the methodological quality of studies assessing the impact of ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy to improve patient-centered outcomes. Finally, the methodological quality of economic evaluations of ^{18}FDG-PET and ^{18}FDG-PET/CT was assessed using the Consensus on Health Economic Criteria (CHEC). Studies were graded according to a system adopted by the Veterans Affairs Technology Assessment Program (VATAP) to classify the level of evidence regarding the clinical utility of studies on PET. Two reviewers assessed the methodological quality of studies independently. Disagreements were resolved by consensus.

A combination of qualitative and quantitative approaches was used to synthesize the data. Details of individual studies were summarized in evidence tables including information on article source, study design, study population, characteristics of ^{18}FDG-PET and ^{18}FDG-PET/CT and reference tests, and outcomes. Two-by-two tables were constructed for studies on the diagnostic test performance of ^{18}FDG-PET and ^{18}FDG-PET/CT. Sensitivity and specificity were calculated for each study. Individual study results were grouped when two or more studies assessed the same type of PET for similar purposes (e.g., primary diagnosis, staging, restaging, recurrences), had similar study design (i.e., prospective or retrospective), and had usable data for common outcomes of interest. Summary estimates of the likelihood ratio, both positive and negative were meta-analyzed using the DerSimonian and Laird random effects method to support inferences regarding the magnitude and direction of the diagnostic performance of ^{8}FDG-PET and ^{18}FDG-PET/CT.

Results

One hundred and nine articles were included in this report. The results are presented by type of cancer.

Bladder cancer: Evidence from three studies is available on the diagnostic performance of ^{18}FDG-PET and ^{18}FDG-PET/CT for bladder cancer. Data from a meta-analysis of two prospective studies totalling 88 participants showed that ^{18}FDG-PET does not seem to be helpful in identifying the stage of the disease. The diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT has not been evaluated for other clinical situations. Only one study of moderate sample size and moderate quality evaluated the diagnostic thinking impact of ^{18}FDG-PET on the treatment of bladder cancer and reported that 17 percent of the treatment decisions were changed after knowing the results of ^{18}FDG-PET. Since the amount and quality of the evidence is limited, firm conclusions about the diagnostic thinking impact of ^{18}FDG-PET in bladder cancer cannot be drawn.

Brain cancer. Six studies provided evidence on the use of ^{18}FDG-PET for brain cancer. We did not find studies that reported on the use of ^{18}FDG-PET/CT for brain cancer. The majority of the studies evaluated the diagnostic accuracy of ^{18}FDG-PET. ^{18}FDG-PET does not seem to be highly discriminative in identifying the stage of the disease, and in distinguishing between necrosis and recurrences. The sensitivity and specificity values of the studies were modest and had wide confidence intervals, precluding firm conclusions about the diagnostic utility of ^{18}FDG-PET for brain cancer. The effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes continue to be scarcely evaluated. There is limited evidence from low quality studies that the best indication of ^{18}FDG-PET seems to be differentiating between high and low grade gliomas. There is, however, no consensus regarding the utility of ^{18}FDG-PET in predicting histological grading and survival of brain tumors.

Cervical cancer. ^{18}FDG-PET and ^{18}FDG-PET/CT have been evaluated for a variety of clinical indications in the assessment of patients with cervical cancer. These include 1) initial staging, 2) detection of recurrence, and 3) restaging, including planning for salvage therapy. There is evidence around the diagnostic accuracy for each of these indications, and also limited evidence for the

diagnostic thinking impact. Only two studies provided evidence for the effects of ^{18}FDG-PET on patient-centered outcomes and these addressed problems of detection of recurrence and restaging. There is no data on the cost effectiveness of either ^{18}FDG-PET or ^{18}FDG-PET/CT. The strongest evidence for the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT has been produced for staging of locally advanced cervical cancer and the detection and restaging of recurrent disease.

Kidney cancer. There is some evidence to suggest that the accuracy of ^{18}FDG-PET may be sufficient to support its use in the initial staging of renal cancer. When ^{18}FDG-PET was compared against any reference standard in prospective studies, statistically significant results were obtained for both the positive and negative likelihood ratios. Because of the high sensitivities and specificities reported in prospective studies of diagnostic accuracy, the application of ^{18}FDG-PET to at least initial staging of renal cancer seems to be worthy of additional study by well-designed prospective trials. There is insufficient evidence to support its widespread adoption at this time.

Ovarian cancer. ^{18}FDG-PET and ^{18}FDG-PET/CT have been evaluated for a variety of clinical indications in the assessment of patients with ovarian cancer. These include 1) primary diagnosis, 2) initial staging, 3) detection of recurrence and 3) restaging. There is evidence around the diagnostic accuracy for each of these indications, and also limited evidence for the diagnostic thinking impact. Only one study provided evidence for the effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes, and there is no data on the cost effectiveness of either ^{18}FDG-PET or ^{18}FDG-PET/CT for ovarian cancer. Most of the studies have evaluated the use of ^{18}FDG-PET/CT. ^{18}FDG-PET has not been studied in women with suspected ovarian cancer for primary diagnostic purposes. The evidence on the efficacy of ^{18}FDG-PET and ^{18}FDG-PET/CT for the initial staging of ovarian cancer is very limited and no firm conclusions can be made regarding their utility for this indication. The clinical indication for which ^{18}FDG-PET has been evaluated the most is for the detection of recurrences following treatment. Meta-analyses of the diagnostic accuracy of both ^{18}FDG-PET and ^{18}FDG-PET/CT showed a consistent, statistically significant effect in both the positive and negative LRs across a range of reference standards and study designs, providing evidence to support the usefulness of ^{18}FDG-PET and ^{18}FDG-PET/CT in detecting ovarian cancer recurrences.

Pancreatic cancer. Seventeen studies evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT. The majority of studies are on ^{18}FDG-PET but some of them have evaluated ^{18}FDG-PET/CT. The findings were consistently significant, suggesting that both ^{18}FDG-PET and

^{18}FDG-PET/CT are useful in diagnosing and establishing initial stage of the disease. Further, there is no clear evidence for the choice of ^{18}FDG-PET v. ^{18}FDG-PET/CT, since the observed heterogeneity indicates considerable uncertainty in these estimates. We found that when ^{18}FDG-PET was evaluated for primary diagnostic purposes, the positive LR was slightly better for ruling in the disease, but the negative LR remained almost the same. There is some evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT that indicates that the management plan is altered in an important number of patients (up to 69 percent), more often resulting in a conservative course of management thus avoiding unnecessary surgery. Finally, one study conducted a cost-minimization analysis on the use of ^{18}FDG-PET/CT for pancreatic cancer. The analysis, however, is insufficient to demonstrate changes in costs relative to changes in clinical effects due to the implementation of ^{18}FDG-PET/CT in the diagnostic workup of pancreatic cancer.

Prostate cancer. The studies identified for analysis in this technology assessment addressed only the detection of recurrence in patients with an increased PSA following primary therapy for prostate cancer. No study addressed the diagnosis or staging of disease. Due to heterogeneity across studies in terms of study design, type of PET and indications for its use, no pooled estimate of the accuracy of ^{18}FDG-PET could be obtained. The impact of ^{18}FDG-PET on diagnostic thinking or patient-centered outcomes was not assessed in any studies. Furthermore, there is no data available describing the use of ^{18}FDG-PET/CT technology.

Small cell lung cancer. There is no evidence to support the role of ^{18}FDG-PET or ^{18}FDG-PET/CT in the diagnosis of SCLC and no cost-effectiveness data is available for this tumour for any indication. However, there is preliminary evidence from three cohort studies of moderate quality supporting the technique's use in staging and restaging.

Testicular cancer. The recent evidence for ^{18}FDG-PET in testicular cancer is very limited and inconclusive. Four studies of moderate quality evaluated the diagnostic accuracy of ^{18}FDG-PET for testicular cancer. ^{18}FDG-PET/CT has not been evaluated for this condition. The clinical indications that were evaluated included initial staging, restaging and recurrences; however, a pooled analyzes of the data was precluded due to the limited number of studies. One small retrospective study of moderate methodological quality evaluated the physician decision-making impact when ^{18}FDG-PET imaging is used in the assessment of the recurrence of testicular cancer. The management plan was changed in only a small proportion of the patients, suggesting that ^{18}FDG-PET imaging had minimal effect on subsequent treatment decisions. Finally, none of the studies evaluated the effects of ^{18}FDG-

PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT and therefore, no firm conclusions on the clinical and economic impact of ^{18}FDG-PET or ^{18}FDG-PET/CT can be made.

Future Research

Further evaluations of the utility of this technology should be done with developments concentrating on enhancing patient throughput and establishing new and more focused clinical applications in various subpopulations of patients. Because it may be quite challenging to enrol patients into a study for low-incidence cancers, multicenter studies will be required to adequately address these important issues.

Finally, some of the most important roles of ^{18}FDG-PET and ^{18}FDG-PET/CT have not been sufficiently explored (e.g., estimating prognosis, selecting and changing treatment modalities, estimating their role in the evaluation of tumor burden regardless of histology). Evaluations of the procedures or therapies forestalled or cancelled based on ^{8}FDG-PET and ^{18}FDG-PET/CT must be explored further to inform policy decisions, these information gaps need to be filled with new methodological approaches.

Conclusions

For some type of cancers (e.g., cervical, ovarian, and pancreatic cancer), there is some evidence of the utility of ^{18}FDG-PET or ^{18}FDG-PET/CT for diagnosing, staging, or detecting recurrences, but they may still require more studies to augment the evidence base. Further studies are needed to reach firm conclusions about the clinical effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT in terms of the impact on diagnosis and treatment options, patient-centered outcomes, and economic costs. It is still unclear how ^{18}FDG-PET and ^{18}FDG-PET/CT affects patient treatment and ultimately their outcome. For other types of cancer examined in the review (e.g., bladder, kidney, prostate, SCLC, and testicular) the answers are still inconclusive and require more careful study.

Chapter 1. Introduction

Background

Positron emission tomography (PET) is a form of nuclear medicine imaging that detects and establishes metabolic abnormalities in tissue. A PET scanner produces an image of the area of interest through the detection of radiation emitted from a positron-emitting radionuclide that is introduced into the patient and that accumulates in the target tissue.

Different radiotracers allow for various aspects of tumor metabolism to be imaged. The most commonly used radioisotope tracer is ^{18}F-fluorodeoxyglucose (^{18}FDG), a glucose analog with the addition of a radioactive fluorine atom, which has a half-life of 109.8 minutes. The relatively long half-life of this radioisotope allows the operation of imaging sites up to 2 to 4 hours travelling distance from the production site. [1] Like glucose, ^{18}FDG is taken up into cells through glucose transport proteins (GLUT) and then phosphorylated by a hexokinase. At this point glucose is further metabolized while deoxyglucose is not, leaving the ^{18}FDG to accumulate intra-cellularly as ^{18}F-FDG-6-phosphate. Images may be interpreted qualitatively by visual assessment for regions of increased uptake. Quantitative measurement of the glucose metabolism by cells in a region of interest is performed using the standardized uptake value (SUV). The SUV is calculated by measuring the tissue radioactivity concentration (μCi/mL) and dividing by the total injected dose (μCi/kg), normalized to the patient's body weight. Results may be variable depending upon the scanner image resolution, time of image acquisition after radioisotope injection (later images will have higher SUVs as ^{18}FDG accumulates), the presence of hyperglycemia, method of normalization (use of body surface area or lean body mass), and the method of quantitative measurement.

Compared to structural imaging techniques (X-ray, computed tomography [CT], and magnetic resonance imaging [MRI]), ^{18}FDG-PET may be a more accurate technique for diagnosis, staging, and treatment decisions in oncology. It has been suggested that PET imaging may be able to differentiate between benign and malignant lesions (detecting malignancies as small as 6 millimeters, allowing for the early detection of disease before structural changes become apparent), establish the grade of malignancy (the stage of disease, the existence of recurrent or residual disease, the site of the disease and the primary site of a tumor for biopsy), evaluate a patient's response to

therapy, and can be used for radiotherapy planning in certain types of tumors. [2] Thus, ^{18}FDG-PET holds promise for decreasing the utilization of other diagnostic tests and invasive procedures, and providing more accurate knowledge about the extent of the disease. This information may influence patient management decisions, such as the aggressiveness of planned chemotherapy or radiotherapy, which, in turn may significantly impact patient mortality and quality of life. [3]

Several authors have discussed the sequence of evaluations that can be done in a diagnostic test study. [4,5] These include diagnostic test performance, therapeutic impact and clinical outcome.

The diagnostic performance of a test can be evaluated based on its sensitivity, specificity, accuracy or likelihood ratios (LR). Evaluating a test's performance involves comparing test results against a valid reference or "gold" standard, which represents the actual or accepted disease status. Appropriate reference standards can include pathology findings (e.g., histopathological confirmation of the presence or absence of disease) or clinical outcome (e.g., subsequent disease progression or resolution of symptoms and signs).

Therapeutic impact is measured as the change in treatment decision, or decision for additional diagnostic workup, made by clinicians in response to the information provided by the test. The evaluation of outcome assesses if, and to what degree, the patients who had the test have better health outcomes. This can be assessed by randomized controlled trials (RCT) of the test and subsequent management resulting from test information. Changes in outcome may also be reasonably inferred from a combination of evidence of improved diagnostic accuracy, evidence of changes in management and evidence of the effective treatment of a given condition. That is, in conjunction with evidence of improved diagnostic accuracy and changes in management, there should be evidence (ideally from RCTs) that alternative treatment or management results in improved long term health outcomes for patients. For example, if a diagnostic test allowed earlier diagnosis of a condition, evidence that earlier treatment is more effective than delayed treatment is needed to infer that improved outcomes result from the diagnostic test result.

^{18}FDG-PET is considered a potentially major advancement in clinical practice because it can provide information about the behavior of tumors in addition to the anatomic extent and thus may provide evidence to guide therapeutic choices. [6] The use of ^{18}FDG-PET for the diagnosis of several cancers has been evaluated, [6-9] and it is estimated that applications of ^{18}FDG-PET in oncology may soon account for 80 to 90 percent of the technology's utilization. [2] In the United States, the Centers for Medicaid and Medicare Services (CMS) has determined that there is sufficient evidence to show

that a ^{18}FDG-PET scan is reasonable and necessary for certain indications in the pretreatment and management phase of non-small cell lung cancer, esophageal cancer, colorectal cancer, lymphoma, melanoma, breast cancer, head and neck cancers and thyroid cancer. [10] Since the decision by the CMS to cover PET for these cancers, the use of PET scanning has increased anywhere from 80 percent for esophageal and brain cancer to 128 percent for head and neck cancers, based on claims data from 2001-2004. Increases of over 1,000 percent were recorded for the initial coverage period 1999-2001 for lymphoma. [11] The issue of assessing ^{18}FDG-PET for a range of other cancers currently designated as "coverage with evidence development" remains unaddressed. "Coverage with evidence development" refers to the designation of a ^{18}FDG-PET scan being considered "reasonable and necessary" only when the provider is participating in, and patients are enrolled in a prospective clinical study designed to collect additional information to assist in patient management. [10]

With the exception of central nervous system (CNS) neoplasms, PET for oncologic indications has only been in use since 1995 when the first scanners capable of whole body imaging were introduced. Despite the rapidly expanding evidence for the use of PET, [2] researchers have noted the small number of high-quality ^{18}FDG-PET studies and uncertainty surrounding the possibility of publication bias. [3] There remain many unanswered questions with respect to the diagnostic accuracy of ^{18}FDG-PET for other cancers and the role of ^{18}FDG-PET in grading, restaging and monitoring response to treatment. In addition, because of its putative high cost, [12] ($179.4 million was paid by CMS to providers and facilities for 112,729 PET scans in 2002) [13] it would be beneficial to assess the cost-effectiveness of ^{18}FDG-PET in light of the most recent reports of the technology's clinical effectiveness.

Scope of the report

In 2004, the Duke Evidence-based Practice Center (EPC) completed a technology assessment on the evidence produced from 1966 to 2003 regarding the utility of PET for six cancers: brain, cervical, small-cell lung, ovarian, pancreatic and testicular. [14] This technology assessment suggested that PET might be beneficial in helping physicians with clinical questions such as staging and detecting metastatic disease and recurrence. However, the literature had many limitations including the use of older generation technology, inclusion of heterogeneous groups of patients without presentation of results by clinically relevant subgroups, absence of data that would allow the reader to infer the information contributed by PET beyond what was available from conventional studies, and in some cases, lack of a comparator.

The National Oncologic PET Registry (NOPR) was launched in May 2006 in response to the CMS "Coverage with evidence development" policy to collect data through a clinical registry to inform the CMS's ^{18}FDG-PET coverage determination decisions for currently non-covered cancer indications. Since then, NOPR has collected questionnaire data from referring physicians on intended patient management before and after an ^{18}FDG-PET scan. One publication from the NOPR Working Group[15] has reviewed survey data from referring physicians regarding changes in treatment decisions before and after ^{18}FDG-PET. The authors found that clinicians report they often change their intended management based on the ^{18}FDG-PET results. ^{18}FDG-PET was associated with a 36.5 percent change in the treatment or no-treatment decisions. One of the limitations of the NOPR database is that the registry does not document whether the physicians actually completed the planned management changes. Therefore the information is based on an intention-to-treat approach, and the relative impact of ^{18}FDG-PET on the actual management of patients with cancer has not been assessed. Recently, NOPR has formally asked CMS to reconsider the current national coverage decision on ^{18}FDG-PET and to end the data collection requirements for diagnosis, staging and restaging. CMS will review the published data and determine the next steps related to reimbursement for ^{18}FDG-PET scans now only covered through the NOPR.

The Coverage and Analysis Group at CMS requested from The Technology Assessment Program (TAP) at the Agency for Healthcare Research and Quality (AHRQ) that an evaluation be performed on the available scientific evidence on the use of ^{18}FDG-PET for nine different cancers (bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular). AHRQ

assigned this evaluation and report to the following Evidence-based Practice Center: the University of Alberta/Capital Health Evidence-based Practice Center (U of A EPC; Contract Number HHSA 290 2007 10021 I).

Structure of the Report

To provide a framework for the report, we first present the key questions and our analytic approach to address them. Next, a general methods section applicable to all the cancers considered in the report is presented. We describe the literature review methods, outline our inclusion and exclusion criteria, the search strategy for identifying articles relevant to the key questions, and the process for abstracting and synthesizing information from eligible studies. We also describe the methods for assessing the methodological quality of individual studies, the data analysis and synthesis.

The results are reported by type of cancer; each section addressing a particular cancer is organized so that it can be considered a stand-alone report. The bibliography of included studies and appendices, including the search strings, data extraction and quality assessment forms and detailed evidence tables for each cancer, have been placed at the end of the document.

The following four key questions examine the degree to which current evidence supports confident judgments about the use of ^{18}FDG-PET in the assessment and treatment of nine types of cancer in clinical practice. It encompasses both dedicated PET and newer PET/CT technology that integrates PET and CT into one device.

Q1: ^{18}FDG-PET and ^{18}FDG-PET/CT diagnostic test performance

How does the diagnostic test performance of ^{18}FDG-PET and ^{18}FDG-PET/CT compare to conventional imaging modalities (e.g., CT and MRI) or other diagnostic procedures (e.g., biopsy, serum tumor markers) with respect to the following clinical situations:

1. Diagnosis
2. Staging
3. Restaging
4. Monitoring response to treatment

Q2: Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT

What is the magnitude of the impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making regarding approaches to diagnosis and management with respect to the following clinical situations:

1. Diagnosis
2. Staging
3. Restaging
4. Monitoring response to treatment

Q3: ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy

What is the impact of ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy to improve patient-centered outcomes? What is the ability of ^{18}FDG-PET and ^{18}FDG-PET/CT to improve patient-centered outcomes when used as a diagnostic test to identify patients suitable for a particular treatment?

Q4: Cost-effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT

What is the cost-effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT with respect to the following clinical situations:

1. Diagnosis
2. Staging
3. Restaging
4. Monitoring response to treatment

The six-tiered efficacy model of technology assessment introduced by Fryback and Thornbury[16] was used as a framework to quantify the level of evidence available to address the questions of this report (Table 1). This report focuses on all evidence between levels 2 and 6 (excluding technical imaging quality data).

Table 1. Hierarchy of Diagnostic Efficacy

Level of evidence	Description
Level 1: Technical	Resolution of line pairs Modulation transfer function change Gray-scale range Amounts of mottle Sharpness Computerized imaging parameters
Level 2: Diagnostic accuracy efficacy	Yield of abnormal or normal diagnoses in a case series Diagnostic accuracy (percentage of correct diagnoses in case series) Sensitivity, specificity, and positive/negative predictive value in a defined clinical problem setting Measures of area under the receiver operating characteristic (ROC) curve
Level 3: Diagnostic thinking efficacy	Number (percentage) of cases in a series in which image was judged "helpful" for making the diagnosis Entropy change in differential diagnosis probability distribution Difference in clinicians' subjectively estimated diagnosis probabilities before and after test information
Level 4: Therapeutic efficacy	Number (percentage) of times image was judged "helpful" in planning patient care in a case series Percentage of times medical or surgical procedure avoided due to image information Number or percentage of times planned therapy pretest changed after the image information was obtained (retrospectively inferred from clinical records) Number or percentage of times clinicians' prospectively stated therapeutic choices changed after test information
Level 5: Patient outcome efficacy	Percentage of patients improved with test v. without test Morbidity (or procedures) avoided after having image information Change in quality-adjusted life expectancy Expected value of test information in quality-adjusted life years (QALYs) Cost per QALY saved with image information Patient utility assessment (e.g., Markov modeling, time trade-off)
Level 6: Societal efficacy	Benefit-cost analysis from societal viewpoint Cost-effectiveness analysis from societal viewpoint

Chapter 2. Methods

Overview

In this chapter, we document a prospectively designed protocol that the U of A EPC followed for this technology assessment report on the use of ^{18}FDG-PET and ^{18}FDG-PET/CT for nine cancers.

To accomplish the tasks as directed, a core research team composed of clinical investigators and methodologists was assembled. The core research team was trained and experienced in systematic review methodology or critical appraisal of the scientific literature in diagnostic tests.

In this chapter, we describe the technology assessment methods. We outline our inclusion and exclusion criteria, the study selection process for identifying relevant articles, and the process for abstracting information from eligible studies. Finally, we describe the methods for assessing the methodological quality of individual studies, the analysis and synthesis of the results.

Literature Search and Retrieval

Comprehensive searches of four biomedical electronic databases (Table 2) were conducted for the time periods specified. All search strategies were developed by a research librarian with input from the project team. The search strategy comprised both controlled vocabulary and keywords. Separate searches were done for each cancer. The search was not restricted by language and articles were retrieved from 2003 to the present. No study design filters were used since the research questions could be answered by a large variety of study types. See Appendix A for detailed search strings.

Table 2. Databases Searched for Relevant Studies

Database	Years/issues	Date of search
MEDLINE®	2003 - 2008	12 March, 2008
EMBASE	2003 - 2008	12 March, 2008
CENTRAL (Cochrane Central Register of Controlled Trials, NHS Economic Evaluation Database)	1st Quarter 2008	20 March, 2008
Scopus	2003 - 2008	19 March, 2008

Criteria for Selection of Studies

A set of inclusion and exclusion criteria was used to determine eligibility of studies for the technology assessment (Table 3). Briefly, eligible studies were published in English and evaluated the use of ^{18}FDG-PET or ^{18}FDG-PET/CT in a sample of 12 or more adult participants (older than 16 years of age) with primary cancer of the following type: bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular. Restrictions in terms of study design were not imposed and both prospective and retrospective studies were included.

Studies must have reported numeric data on at least one outcome of interest for the key questions of the technology assessment. A study can contribute with data for more than one of the questions addressed in the technology report. For studies on the diagnostic performance of ^{18}FDG-PET and ^{18}FDG-PET/CT (Q1), the outcomes of interest were: sensitivity, specificity, positive and negative predictive values, and LR. Other outcomes that were examined for Q2, Q3 and Q4 included:

- additional diagnostic test work-up;
- treatment decisions and management strategy;
- changes in therapy;
- patient-centered outcomes (e.g., survival, quality of life, prognostic indicators, time until recurrence); and
- economic outcomes.

Table 3. Inclusion Criteria

Category	Criteria
Source	• English language studies reporting original research from 2003 to March 2008; • Study not duplicated or superseded by later study with the same purpose from the same institution
Population	• Studies ≥12 human participants; • The study provides separate data for a population consisting of adults (>16 years) with primary cancer of the following type: bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular
Test	• Studies of PET or PET/CT using ^{18}FDG as radioisotope tracer
Comparator	• ^{18}FDG-PET or ^{18}FDG-PET/CT should be compared to a reference standard, where either: a) all participants undergo both ^{18}FDG-PET or ^{18}FGD-PET/CT and the reference standard; or b) one group of patients undergo ^{18}FDG-PET or ^{18}FGD-PET/CT and the other undergo the reference standard • Examples of possible reference standard: MRI, CT, biopsy/histology, X-rays, ultrasound, PET with other radioisotope tracer, clinical followup
Study design	• Prospective or retrospective studies
Outcomes of interest	• Study should provide numeric data for the outcomes of interest in the review

CT = computed tomography; ^{18}FDG = fluorodeoxyglucose; MRI = magnetic resonance imaging; PET = positron emission tomography

Study Selection Process

Screening of titles and abstracts

One of four reviewers evaluated the title and abstract of each study to select references potentially relevant to the topics of the report (Appendix B). The full-text of studies meeting the criteria was retrieved as was the full-text of those that reported insufficient information to determine eligibility.

Identification of studies eligible for the report

Two independent reviewers appraised the full-text of potentially relevant articles using a standard form (Appendix B). Disagreements about the inclusion or exclusion of studies were resolved by consensus among reviewers.

Evaluating the Methodological Quality of Studies and Grading the Evidence

The methodological quality of studies that assessed the diagnostic performance (Q1) and the diagnostic thinking impact (Q2) of ^{18}FDG-PET and ^{18}FDG-PET/CT was assessed using the Scottish Intercollegiate Guidelines Network (SIGN) assessment tool for diagnostic studies, [17] which is based on the Quality Assessment of Studies of Diagnostic Accuracy (QUADAS) tool. [18] Studies assessing the impact of ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy to improve patient-centered outcomes (Q3) are different from diagnostic performance studies and would be more akin to standard effectiveness studies (e.g., clinical trials, observational analytical cohort studies). Therefore, an individual components approach that considered important aspects of design, conduct, and reporting of effectiveness studies was used to assess the methodological quality of Q3 studies. Finally, the methodological quality of economic evaluations of ^{18}FDG-PET and ^{18}FDG-PET/CT was assessed using the Consensus on Health Economic Criteria (CHEC). [11] See Appendix B for the quality assessment instruments used in this technology report.

Evidence from the selected studies was graded using a system adopted by the Veterans Affairs Technology Assessment Program (VATAP) to classify the level of evidence regarding the clinical utility of studies on PET (Table 4).

Table 4. Grading Scheme for Diagnostic Studies

Grade	Criteria
A	Prospective studies with broad generalizability to a variety of patients and no significant flaws in research methods.
B	Prospective studies with a narrower spectrum of generalizability, and with only a few flaws that are well described (and impact on conclusions can be assessed).
C	Studies with several methods flaws (e.g., small sample size and retrospective)
D	Studies with multiple flaws in methods (e.g., no credible reference standard for diagnosis)

Adapted from Robert et al. 1999[1]

Two reviewers assessed the methodological quality of studies independently. Disagreements were resolved by consensus or, when no consensus could be reached, a senior methodologist adjudicated.

Data Collection

Information regarding the study design and methods, characteristics of participants, PET and comparison tests, and outcomes of interest were extracted using a pretested data extraction form that was adapted to each of the four key questions (Appendix B).

General data relevant to the review were collected on a general data extraction form. This included information on the country, year and type of publication, study design, setting, duration of the study, and number of participating centers. Data on characteristics of study participants included type of primary cancer, how participants were enrolled, inclusion and exclusion criteria, demographic characteristics, and stage or severity of their condition.

Data on characteristics of ^{18}FDG-PET and ^{18}FDG-PET/CT included a description of the purpose of their use within the study, technical details of the devices and administration procedures, and characteristics of the reference tests. Likewise, information on the criteria for interpretation was extracted. Specific forms were used to collect data for each of the four key questions of the report. Finally, information was collected on study conclusions, as reported by the authors of the primary studies. Data from the primary studies were extracted by one reviewer and then independently verified for accuracy and completeness by a second reviewer. Any discrepancies in data extraction were resolved by consensus between the data extractor and the data verifier. Study selection, methodological quality assessment, and data extraction were managed with Microsoft Excel™

(Microsoft Corporation, Redmond, WA). Extraction of data from graphs was performed using Corel Draw®, version 9.0 (Vector Capital, San Francisco, CA).

Evidence Synthesis

Characteristics of the included studies were summarized using descriptive statistics (i.e., proportions and percentages for categorical data; means with standard deviations [SD], or medians with interquartile ranges [IQR], for continuous data).

Data were analyzed qualitatively. Evidence tables were constructed to report information on each article's source, study design, study population, characteristics of ^{18}FDG-PET and ^{18}FDG-PET/CT and reference tests, and outcomes. The evidence tables also included summaries of study quality and comments to help interpret the outcomes. Data were combined by type of cancer to provide summary information across studies for each of the key questions, if appropriate.

For each of the four key questions, the following study characteristics were summarized and discussed:

a. Inclusion criteria of studies (patients and disease characteristics)
b. PET technology used (^{18}FDG-PET alone, ^{18}FDG-PET/CT) and comparator
c. Tests used prior to, concurrent with or after the PET scanning and whether the studies indicated the information contributed by PET beyond that provided by other tests
d. Overall quality of the body of evidence
e. The generalizability of the summarized evidence to the Medicare population (aged >65 years)
f. The generalizability of the summarized evidence to other cancers
g. Homogeneity of SUVs with respect to ^{18}FDG dose, timing of study, and scanner variability

For the question related to ^{18}FDG-PET and ^{18}FDG-PET/CT diagnostic test performance, two-by-two tables (or two-by-one if only reference standard positive or reference standard negative subjects were included) were constructed for each comparison test or combination of tests within the individual studies. Sensitivity and specificity were calculated for each study using standard formulas. Results were graphed in forest plots for visual analysis, but were not pooled statistically due to the different diagnostic thresholds of the various studies. Individual study results were grouped when two or more studies assessed the same type of PET (i.e., ^{18}FDG-PET or ^{18}FDG-

PET/CT) for similar purposes (e.g., primary diagnosis, staging, restaging, recurrences), had similar study design (i.e., prospective or retrospective), and had usable data for common outcomes of interest. Studies using different methods to confirm the final diagnosis were considered for grouping, but results were also presented separately by type of reference standard.

Summary estimates of the LR, both positive and negative, were meta-analyzed using the DerSimonian and Laird random effects method. [19] The LRs are a measure of the performance of diagnostic tests, expressing the magnitude by which the odds of a diagnosis in a given patient is modified by the result of a test. [20] For example, if an individual has probability of disease of 0.1 prior to taking a test (odds = 0.09) and the test has a positive LR of 4 and a negative LR of 0.2, the patients' post test odds of having the disease would be 4*0.09 = 0.36 (probability = 0.27) if the test was positive and 0.2*0.09 = 0.02 (probability = 0.02) if the test was negative. A test with a higher positive LR and lower negative LR is considered a better test. Where studies presented more than one estimate of test performance for the same test (e.g., at different cut-off points or for different patient subgroups) we only included one estimate in the pooled analysis. We aimed to select the data set most similar to the estimates provided by the other studies in terms of patient population.

Data on diagnostic performance were also synthesized using the summary receiver operating characteristic (SROC) approach, which is a measure of test accuracy. [21] The ROC curve is a plot of true positive rate (sensitivity) versus the false positive rate (1 - specificity) for various possible cutpoints of a diagnostic test. The closer the curve follows the left hand and top borders of the ROC curve space, the more accurate the test is, while a test of lower accuracy will come closer to the 45 degree diagonal. Thus, the area under the curve is indicative of the accuracy of a diagnostic test, where area of 1.0 represents a perfect test; an area of 0.5 represents a worthless test.

Homogeneity tests were carried out to evaluate the consistency of findings across the studies. We used the I^2 statistic to determine the percentage of total variation in the LR across the studies due to heterogeneity rather than to chance. [22] A value of 0 percent indicates no observed heterogeneity. I^2 values of 25, 50 and 75 percent were used as anchors in guiding the classification of meta-analyses into low, moderate and high heterogeneity categories, respectively, [22] where larger values indicate increased heterogeneity. Possible reasons for heterogeneity, such as patient characteristics and the nature of the reference method (biopsy/histology, clinical followup or a composite reference standard) were explored. All analyses were performed using RevMan software version 5.0 (Cochrane Collaboration, Oxford, UK).

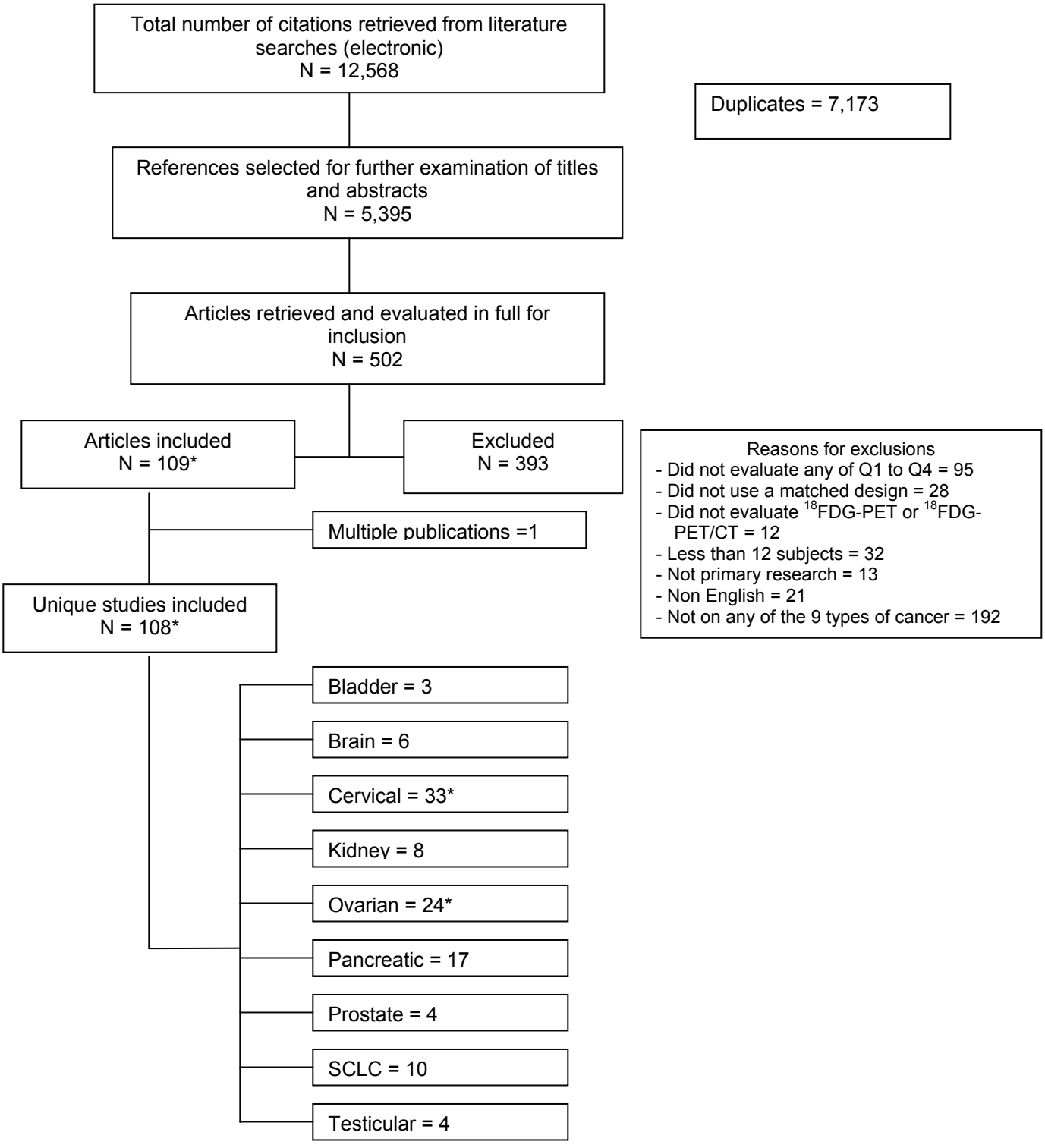
Chapter 3. Results

Search Results

Overall, the literature search (electronic and reference lists) resulted in the identification of 12,568 citations of which 7,173 were duplicates. After screening titles and abstracts (5,395 citations), the full-texts of 502 potentially relevant articles were retrieved and evaluated for inclusion. The application of the selection criteria to the 502 articles resulted in 393 articles being excluded, while 109 articles were relevant to the questions addressed in this review. Figure 1 outlines study retrieval and selection.

The primary reasons for exclusion of studies were as follows: (1) the study did not report on any of the nine types of cancer (n = 192), (2) the study did not evaluate the questions of interest (n = 95), (3) the study reported on less than 12 participants (n = 32), (4) the study did not use a matched design (n = 28), (5) the study did not evaluate ^{18}FDG-PET or ^{18}FDG-PET/CT (n = 12), (6) the study was not primary research (n = 13), and (7) the study was published in a language other than English (n = 21) (Appendix C).

Figure 1. Flow Diagram for Study Retrieval and Selection for the Technology Report



*One study provided data for both cervical and ovarian cancer

1. Bladder Cancer

1.1. Background

Approximately five to ten percent of all malignancies in men are bladder cancer. Throughout the United States and Europe it is the fourth most common cancer diagnosed. [131] It is also the most common cancer of the urinary tract. [132] Men are diagnosed with bladder cancer three to four times more frequently than women. [131] It is estimated that 68,810 new cases of bladder cancer will be diagnosed in the United States in 2008. Of these cases: 51,230 will be men and 17,580 will be women. Furthermore, approximately 14,100 deaths will occur as a result of this malignancy: 9,950 men and 4,150 women. [133] African-Americans are at half the risk of Caucasian-Americans in developing bladder cancer, but African-Americans have a poorer overall survival. [131] Bladder cancer tends to present in an older age group, with the median age of diagnosis at 73 years of age. [133]

Bladder cancer is a heterogeneous disease and its natural history varies. [131] At one end of the spectrum bladder cancer may be of low-grade with slow progression, where on the other end it may be high-grade, highly malignant with significant progression and result in death. [131] The most frequently diagnosed form of bladder cancer (approximately 75 percent) is superficial disease contained in the mucosal and submucosal layers. The remaining patients are diagnosed with muscle-invasive disease, which extends outside the bladder. Bladder cancer patients have shown a 5-year cause-specific survival of more than 95 percent; however, recurrence occurs in more than 50 percent of patients and up to 20 percent develop invasive or metastatic disease. [134]

The most frequent warning sign is painless hematuria, which occurs in 85 percent of patients. Microscopic hematuria may also be present and should be screened for in high-risk patients over the age of 50. Bladder irritability, urinary frequency, urgency and dysuria are common. Patients with advanced disease may experience weight loss and abdominal or bone pain. [131]

In order to plan appropriate patient care, accurate staging must be completed. [131] Bladder cancer is staged using the TNMS (tumor, node, metastasis staging) system, approved by the Union International Contre le Cancer (UICC) in 2002 (Table 5). Tumors identified as “Ta” are considered noninvasive papillary carcinoma, whereas “Tis” refers to carcinoma in situ. [132] Any involvement in the lymph nodes is taken into consideration when staging bladder cancer, as is distant metastasis.

Table 5. Bladder cancer stages

Stage	Description
T1	• subepithelial connective tissue invaded, not muscularis propria[131] • diagnosis difficult, variable prognosis[132]
T2a	• <50% of the depth of muscularis propria invaded[131]
T2b	• >50% of the depth of muscularis propria invaded[131]
T3a	• perivesical tissue invaded (microscopic) [132]
T3b	• perivesical tissue invaded (macroscopic) [132]
T4a	• tumor expansion to prostrate, uterus and vagina[132]
T4b	• tumor expansion to pelvic wall, abdominal wall[132]

Early detection of cancer can help improve survival, which is the main goal of screening. Usually only patients at elevated risk are screened on a regular basis. Ideal screening methods are noninvasive, inexpensive and are highly sensitive and accurate. Tests used to identify bladder cancer include hematuria testing, cystoscopy, bladder imaging, urine cytology and bladder tumor markers. [131] Diagnosis frequently depends on cystoscopic and histologic evaluation of resected tissue. [132] A high rate of incorrectly classified high-grade Ta tumors (grade 3 tumors or higher) has been noted. [135]

The standard method of detection of bladder cancer is cystoscopy. Cystoscopy identifies most superficial disease; however, it is not always successful at detecting small or flat lesions. Additionally, cytoscopy is an invasive procedure. The use of a flexible fiberoscopy is preferred as it is less invasive and provides a clear picture of the bladder interior. Fluorescence endoscopy may also be used for viewing the intravesical area and is reported to have high sensitivity and reasonable specificity, especially for the small or flat lesions frequently missed by conventional cystoscopy. [134] Intravenous urography detects large tumors in the bladder, the upper urinary tract and defects in the kidney. It is unclear if this method of screening is useful as the detection of significant findings is low. [132] Often a combination of methods are required for an accurate diagnosis of small or flat lesions. [135]

Urine cytology is considered a good screening method for high-grade cancers. [131] Sensitivity and specificity are both greater than 90 percent, but it frequently does not detect low-grade papillary tumors. A positive result from urinary cytology indicates a tumor is present in the urinary tract, but does not pinpoint where. [132] Cytology is inexpensive and minimally inconvenient to the patient. [131] There is uncertainty about whether cytology should be performed from a voided urine sample or from a bladder wash sample. Histology from the bladder biopsy is used to make the final diagnosis. [135]

^{18}FDG-PET can be useful for detecting recurrent tumors in the pelvis, distinguishing local

recurrent disease from postsurgical or postirradation fibrosis or necrosis and identifying metastases. However, there is some questions about the value of PET for the management of low-grade or noninvasive tumors due to the excretion of ^{18}FDG by the kidneys and interference with imaging techniques by streak artifacts. Researchers have attempted to limit the amount of ^{18}FDG released in the bladder.

The first treatment for superficial bladder cancer is transurethral resection (TUR) of the tumor. [134] Establishing the correct diagnosis and removing all visible lesions is the aim of TUR. Small tumors may be resected together, while larger tumors need to be resected individually. Frequently bladder tumors are multifocal and therefore there is a risk that tumors may remain after initial TUR. Additionally, tumors may be understaged. [132] A second TUR reduces understaging and the risk of residual disease. [135] Recurrence and progression-free survival may be improved by a second TUR. [132] Followup treatment for TUR is intravesical chemotherapy and it reduces the risk of reoccurrence. Efficacy of chemotherapy agents appears to be similar. In the case of muscle invasive cancer, radical cystectomy is performed. Immunotherapy with Bacillus Clamette-Guérin (BCG) is frequently used after TUR when the disease is confined to the mucosa or submocosa. While the exact mechanism of action is unknown, BCG forms a standard component of treatment for carcinoma in situ. The aim is to eradicate and prevent recurrence of superficial bladder cancer. [134]

1.2. Importance of Key Questions in the Clinical Management of Bladder Cancer

The most important factors in survival from bladder cancer are the stage and the tumor histological grade at diagnosis. Prognosis of bladder cancer is highly dependent on the depth of tumor penetration into the bladder wall. Errors in clinical staging are more likely as the tumor becomes more invasive. Problematic areas for diagnosis and staging of bladder cancer include determination of deep bladder wall invasion and presence of lymph-node metastases. Therefore, accurate staging is pivotal in optimal therapy planning and in avoiding radical surgery in patients with bladder cancer. Some standard imaging methods (e.g., abdominal ultrasonography, CT and MRI) may not provide an accurate basis for therapeutic decisions. For example, tumor involvement is not necessarily detected by changes in the shape or texture of an affected lymph node through CT and MRI. The clinical interpretation of very small lymph nodes on CT and MRI is also problematic as the presence of enlarged regional lymph nodes are not always indicative of metastasis but rather may be reactive to certain procedures such as transurethral biopsy. Procedures such as CT-guided

fine-needle aspiration biopsy can increase the overall staging accuracy but they are subject to sampling errors. ^{18}FDG-PET can be a valuable test for the diagnosis of bladder cancer; however, the evidence about the accuracy and impact of ^{18}FDG-PET on therapeutic decisions and outcomes for bladder cancer patients is scarce. This is partly due to difficulties in interpreting the ^{18}FDG-PET images in the pelvis because ^{18}FDG is excreted by the kidneys and accumulated in ureters and the urinary bladder.

1.3. Results

Three studies[23-25] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for bladder cancer. Each of the three studies evaluated the diagnostic accuracy of ^{18}FDG-PET[23-25] or ^{18}FDG-PET/CT,[24] and one study[24] reported on the diagnostic thinking impact. None of the studies evaluated the effects of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. No economic evaluations of ^{18}FDG-PET or ^{18}FDG-PET/CT for bladder cancer were identified. Characteristics of the populations, conditions of ^{18}FDG-PET administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

1.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in bladder cancer

Characteristics of the studies

Three studies (two prospective, [23,25] one retrospective[24]) evaluated the diagnostic accuracy of ^{18}FDG-PET[23-25] and ^{18}FDG-PET/CT[24] on bladder cancer. Two studies used ^{18}FDG-PET for initial staging[23,25] and one used ^{18}FDG-PET/CT for both staging and restaging purposes. [24]

The studies contained a total of 136 patients with sample sizes ranging from 35 to 55 participants. Participant ages ranged from 33 to 86 years. One study reported the distribution by stage of cancer: clinical stage (CS) I = 16 percent, CS II = 47 percent, CS III = 31 percent and CS IV = 6 percent. [23] ^{18}FDG-PET was compared to a reference standard that varied across the studies. In two studies the reference standard was either histology/biopsy or clinical followup. [23,24] One study established the final diagnosis of all patients using histology/biopsy. [25] One study reported the mean time between last treatment and ^{18}FDG-PET as 37 days. [23] Two studies used a fixed dose of ^{18}FDG (15 MCi[25] and 555 MBq[24]); one study used a weight-based dose (6.5 MBq/kg). [23] The time between injection and PET scan was 60 minutes[23,24] and 20 minutes. [25] Patients fasted for six hours. [23,24] Two

studies[23,24] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose level that was allowed was 120 mg/dL. Methods of interpretation of the images were qualitative in one study[23] and both qualitative and quantitative in a second. [24] Scans were interpreted qualitatively using visual analysis. [23,24] One study[24] reported using SUV but the criterion for abnormality was not reported.

Comparisons

Comparisons for which data were considered for meta-analysis are summarized in Table 6. Pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET for the staging of bladder cancer. Individual study data are summarized in Appendix D.

Table 6. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for bladder cancer

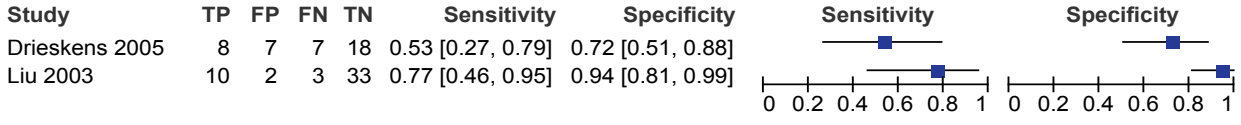
Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Staging	Drieskens 2005[23]	P	FDG-PET	Histology/biopsy or clinical followup	1. FDG-PET v. any reference standard (P studies) [23,25]
	Liu 2003[25]	P	FDG-PET	Histology/biopsy	
Staging and restaging	Jadvar 2008[24]	R	FDG-PET and FDG-PET/CT	Histology/biopsy or clinical followup	No

CT = computed tomography; FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

1. ^{18}FDG-PET for the staging of bladder cancer

Reference standard: any; prospective studies. Two prospective studies[23,25] totaling 88 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to any reference standard for the staging of bladder cancer. Individual 2x2 table results are presented in Figure 2. Sensitivity values in individual studies were 53 percent[23] and 77 percent. [25] Specificity values were 72 percent[23] and 94 percent. [25]

Figure 2. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer



We found that ^{18}FDG-PET had a pooled positive LR of 4.68 (95% CI = 0.65, 33.90) and a pooled negative LR of 0.43 (95% CI = 0.15, 1.19) to accurately detect the stage of bladder cancer (Figures 3

and 4). Neither the positive nor negative LRs were statistically significant, as the 95% CIs included one. Therefore, ^{18}FDG-PET does not seem to be helpful in identifying the stage of the disease. There was considerable heterogeneity in the positive ($p = 0.01$; $I^2 = 84$ percent) and negative ($p = 0.07$, $I^2 = 69$ percent) LRs across the studies. Liu[25] reported statistically significant results for both the positive and negative LRs, whereas results in Drieskens[23] were not statistically significant. It is hard to draw definite conclusions based on the results of two small studies that provide heterogeneous results for the pooled estimates of the accuracy of ^{18}FDG-PET to identify the stage of bladder cancer.

Figure 3. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies)

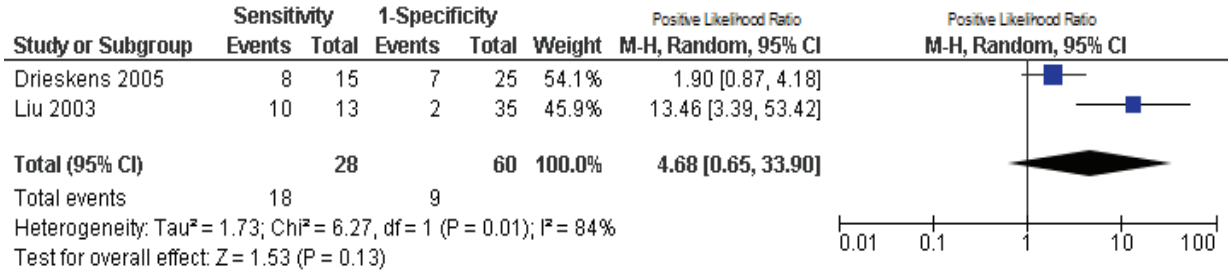


Figure 4. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies)

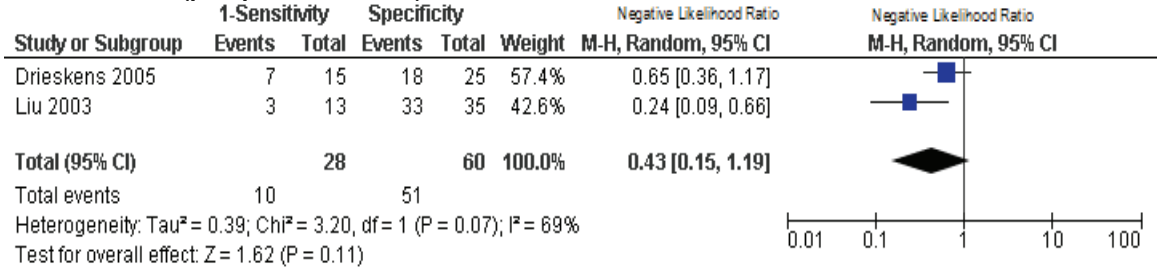
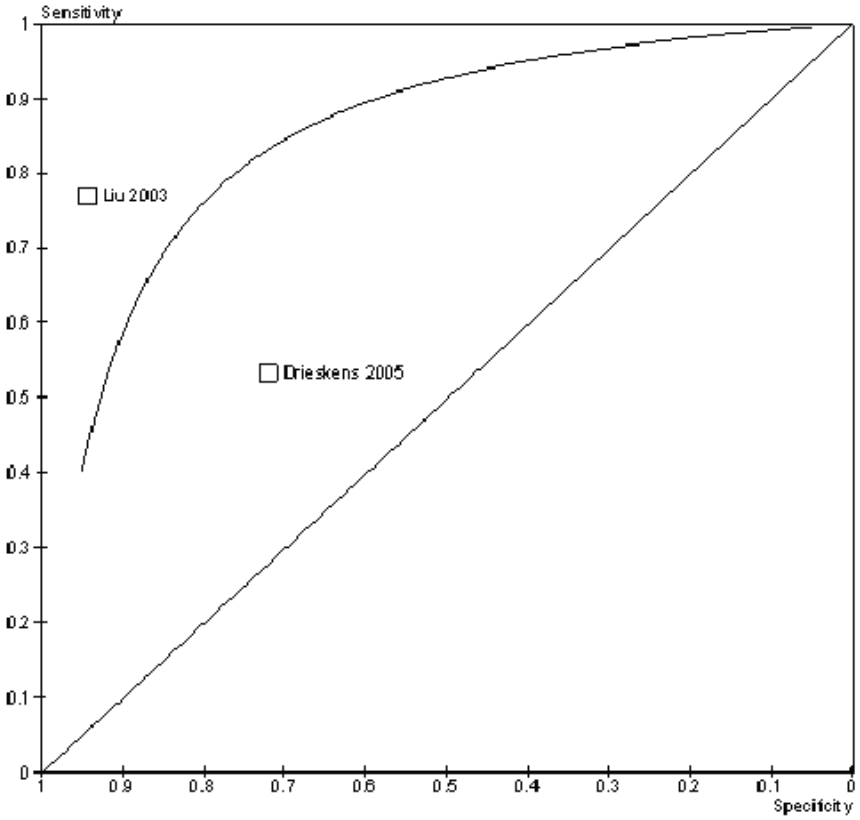


Figure 5 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. any reference standard to identify the stage of bladder cancer.

Figure 5. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the staging of bladder cancer (prospective studies)



Summary of the results

A meta-analysis was calculated to evaluate the accuracy of ^{18}FDG-PET to identify the stage of bladder cancer. The pooled LRs were not statistically significant and therefore, ^{18}FDG-PET does not seem to be helpful for detecting the stage of the disease (Table 7). Heterogeneity across the studies was significant, precluding us from making strong inferences from the pooled overall results.

Table 7. Results of meta-analysis of the accuracy of ^{18}FDG-PET for bladder cancer

PET Purpose	Type	Design	Reference standard	Studies	N	Effect estimate M-H, Random, 95% CI
Staging	FDG-PET	P	Any reference standard	2	88	PLR = 4.68 [0.65, 33.90]
						NLR = 0.43 [0.15, 1.19]

CI = confidence interval; FDG = fluorodeoxyglucose F18; M-H = Mantel Hantzel; NLR = negative likelihood ratio; P = prospective; PET = positron emission tomography; PLR = positive likelihood ratio

1.3.2. Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making with respect to diagnosis and management strategy for patients with bladder cancer

One study evaluated the diagnostic thinking impact of ^{18}FDG-PET ^{18}FDG-PET and ^{18}FDG-PET/CT on the treatment of bladder cancer. A retrospective study by Jadvar et al. [24] evaluated the influence of ^{18}FDG-PET and ^{18}FDG-PET/CT on the management of patients who had been previously treated for transitional cell carcinoma and who were under evaluation for staging and restaging. Both ^{18}FDG-PET and ^{18}FDG-PET/CT were used in this study; however, the results were not separated by the mode of imaging used. The population enrolled was of moderate size (N = 35) and encompassed a wide age range (39-86 years). The subjects were predominately male (71 percent), nondiabetic, with a history of bladder transitional cell carcinoma at initial stages (B2 and C). ^{18}FDG-PET was performed in 17 patients and ^{18}FDG-PET/CT in 18 patients, but mixed results were presented from the two devices.

The diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT was reported as changes in the clinical management of patients. Overall, 17 percent of the patients in the study had their treatment course altered as a result of the ^{18}FDG-PET or and ^{18}FDG-PET/CT imaging analysis. Five patients underwent additional courses of chemotherapy and one patient was under a regime of observation. While the remaining 29 patients did not have their care significantly altered by ^{18}FDG-PET or ^{18}FDG-PET/CT, the authors noted that there was more precise localization of hypermetabolic disease. The authors concluded that combined ^{18}FDG-PET and CT diagnostic information was useful in detecting, localizing and characterizing the extent of metastatic disease.

Overall, the quality of the study was assessed as moderate using the SIGN Methodology Checklist tool. The level of evidence was graded as C (several methodological flaws). Significant issues with the quality of this study included the unblinded interpretation of the ^{18}FDG-PET and ^{18}FDG-PET/CT and the use of multiple modalities to verify the presence of disease (e.g., histology, serial imaging). Additionally, the selection criteria for the patients in this retrospective analysis were not specified, raising the possibility of selection bias. Finally, as there was no clearly defined time period between the ^{18}FDG-PET and ^{18}FDG-PET/CT and the reference standard, disease progression may have occurred between the assessment of the ^{18}FDG-PET/ ^{18}FDG-PET/CT and the final designation of disease status.

Because of the relatively small number of patients included in the Jadvar et al. [24] study, further studies are necessary to assess the role of ^{18}FDG-PET or and ^{18}FDG-PET/CT to make clinical management decisions for bladder cancer patients. Table 8 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on bladder cancer.

Table 8. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for bladder cancer

Study	Results of FDG-PET imaging on patient diagnosis and treatment	Types of bias
Jadvar H, 2008[24] Study type: Retrospective	**Management decision:** Treatment Changes in clinical management after PET/CT: 6/35 (17%): Additional chemotherapy: n = 5 Wait-and-watch regimen: n = 1	Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS) Review bias (PET, unblinded; RS unclear if blinded)

CT = computed tomography; FDG = fluorodeoxyglucose F18; PET = positron emission tomography; RS = reference standard

2. Brain Cancer

2.1. Background

An estimated 21,810 new cases of brain cancer will be diagnosed in the United States in 2008 and 13,070 patients will die from the disease. [133] Primary brain tumors represent a small number of all primary malignant cancers diagnosed, approximately 1.35 percent. Brain cancer as a result of metastases is more common. [136] Brain cancer incidence has increased over time; however, this is largely due to improvements in diagnostic tools, health care, changes in the treatment of elderly patients and changes to the classification of brain tumors. [137] The incidence of brain cancer between 2001 and 2005 was 6.0/100,000[133] in the United States. Caucasians experience certain types of brain tumors (glioma and germ cell tumors) twice as often as African-Americans. In the United States, incidence rates vary from 9.6/100,000 in Virginia to 21.9/100,000 in Colorado. The high rates of brain cancer detected may be linked to greater access to health care and better health care. [137]

Malignant brain tumors encompass a wide range of neoplasms. [136] Patients who are diagnosed with glioblastoma tumors tend to have the shortest survival time of brain cancer patients (less than one third of patients survive one year); similarly, older patients tend to have shorter survival time. [137] The median age for diagnosis of brain cancer is 56 years, while the median age at death is 64 years. [133] According to numbers recorded between 1998 and 2003, 37.7 percent of patients diagnosed with a primary malignant brain tumor survived for 2 years and 30.2 percent survived for 5 years. [137] Primary brain malignancies tend to remain local and rarely spread outside the CNS. [136]

Brain tumors may present differently depending on the location of the lesion, rate of growth and histology. [136] In the initial stages of the disease most symptoms are focalized. As tumor size increases, more generalized symptoms occur. [138] Approximately 50 percent of patients will present with headache, [136] which can last for six months or more. Increased intracranial pressure may cause nausea and vomiting and in patients with low-grade gliomas, seizures are common. [136] Cognitive dysfunction may also occur, which is demonstrated by changes in memory, attention, language use and personality. [138]

Patient prognosis is linked to a number of factors, as are treatment strategies. [137] Increased survival is associated with patient age less than 60, presence of seizures, frontal lobe tumors, low-grade tumors, no tumor necrosis, limited tumor activity, Karnofsky Performance Status scores greater than 70 and total or near-total resection. The World Health Organization (WHO) classifies

brain tumors according to type of cell and histological appearance. [138] The major histological groups are: neuroepithelial tissue or gliomas, tumors of the meninges, germ cell tumors and tumors of sellar regions. [137] More than 80 percent of primary brain tumors are gliomas, tumors of the meninges make up much of the rest. [138]

Diagnosis starts with a complete medical history, physical examination and a careful neurological assessment. [136] In addition, funduscopy and a focused neurologic examination are performed. Appropriate brain imaging is required followed by histopathology to confirm diagnosis. [138]

MRI is used for the initial screening of brain tumors. It produces higher resolution images and can access more areas of the brain than CT scans and is used for neurosurgical planning and risk assessment. [138] To distinguish infiltrative brain tumors from nonneoplastic conditions, high-grade from low-grade tumors, and primary tumors from metastastic tumors, magnetic resonance spectroscopy (MRS) may be used. [136]

Biochemical and metabolic information about tumors and the brain can be determined by MRS. The diagnostic standard is still tissue biopsy. More recently developed stereotactic biopsy techniques are minimally invasive, with decreased morbidity and mortality relative to traditional neurosurgery. Stereotactic biopsy should be obtained to help confirm diagnosis of low-grade gliomas. MRI, MRS and ^{18}FDG-PET assist in tumor localization for biopsy. Testing for biomarkers may also assist in diagnosis, treatment planning and predicting prognosis. [136]

If it is possible to perform a complete resection, surgery is the treatment of choice; [138] there are no clear guidelines on degree or timing of the resection. [136] The decision is based on tumor location, extent, histopathology and comorbid conditions. In the case of high-grade gliomas, a near to total resection aims to decrease tumor burden, by lowering intracranial pressure and improving survival. The patient should be screened for residual tumors within the first three days after surgery. Radiation, chemotherapy or a combination of both frequently follow surgery. [138] As of yet, it is unclear whether it is best to immediately proceed with postoperative radiotherapy, or whether the patient should be observed before proceeding with additional treatment. Early radiation therapy may improve survival times, but it can also lead to radiation-related neurotoxicity. Older patients, whose risk of recurrence is high, may be offered radiotherapy immediately after surgery. [136] Combined chemo- and radiotherapy help improve survival over standard radiation. Patients who are not candidates for surgery or chemotherapy should be considered for palliative care. [138]

2.2. Importance of Key Questions in the Clinical Management of Brain Cancer

Imaging of brain tumors with ^{18}FDG was the first oncologic application of PET for tumor detection, grading of cerebral tumors and assessment of peritumor or remote metabolic alterations. The application of ^{18}FDG-PET for tumor imaging of the brain is based on increased glycolysis in neoplastic cells. ^{18}FDG-PET could be useful to diagnose and grade gliomas and differentiate between tumor recurrence and radiation necrosis. It may be also useful to predict tumor response to chemotherapy compared to radiochemotherapy. [138]There is, however, no consensus regarding the utility of ^{18}FDG-PET in predicting histological grading and survival of brain tumors. Differentiation between inflammatory tissue and malignancies is sometimes difficult due to the high degree of physiologic glucose metabolism in normal brain tissue. This makes the interpretation of increased ^{18}FDG accumulation in both processes difficult. For example, when a hypermetabolic lesion is at the cortical or subcortical gray matter, tumor ^{18}FDG uptake and normal ^{18}FDG uptake are hard to differentiate. Because brain tumors are histologically heterogeneous, CT- or MRI-guided stereotactic brain biopsy does not always yield a valid diagnosis or grading. The correct diagnosis of a relapse is crucial for optimal further treatment. For example, ^{18}FDG may help to distinguish between recurrences and radiation necrosis in cases of glioblastoma multiforme (GBM), and detection of early relapse can help to increase the benefit of interventions such as stereotactic irradiation or gamma knife treatment. Therefore, although the prognosis of GBM tumors remains poor, the use of ^{18}FDG-PET may still have benefits in terms of patient survival time.

2.3. Results

Six studies[26-30,129] provided evidence on the use of ^{18}FDG-PET. We did not find studies that reported on the use of ^{18}FDG-PET/CT for brain cancer. Five studies[26-30] evaluated the diagnostic accuracy of ^{18}FDG-PET for brain cancer. None of the studies reported on the diagnostic thinking impact of ^{18}FDG-PET. One study[129] evaluated the effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes. No economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for brain cancer were identified. Characteristics of the populations, conditions of ^{18}FDG-PET administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

2.3.1. Diagnostic accuracy of ^{18}FDG-PET in brain cancer

Characteristics of the studies

Five studies (three prospective, [26-28] two retrospective[29,30]) evaluated the diagnostic accuracy of ^{18}FDG-PET on brain cancer. ^{18}FDG-PET was used for initial staging in three studies, [27,28,30] for assessment of recurrences in one study, [29] and for establishing both primary diagnosis and recurrences in the remaining study. [26] The studies contained a total of 217 patients with sample sizes ranging from 17 to 81 participants. Participant ages ranged from 20 to 76 years. Four studies reported the distribution by stage of cancer: CS I = 64 percent, CS II = 36 percent; [30] CS II = 22 percent, CS III = 16 percent, CS IV = 42 percent; [26] CS II = 27 percent, CS III = 42 percent, CS IV = 31 percent; [28] and CS I = 7 percent, CS II = 20 percent, CS III = 20 percent, CS IV = 47 percent. [27] ^{18}FDG-PET was compared to a reference standard that varied across the studies. Three studies established the final diagnosis of all patients using histology/biopsy. [27,28,30] In one study the reference standard was either histology/biopsy or clinical followup. [26] One study used MRI and MET-PET (carbon-11 methionine and positron emission tomography) as reference standards in all patients. [29] One study reported the mean time between last treatment and ^{18}FDG-PET as 4 months for chemotherapy, 12 months for radiotherapy and 13 months for surgery. [29] Two studies used a fixed dose of 370 MBq of ^{18}FDG. [28,30] One study used a weight-based dose (2.4 MBq/kg), [26] while another study reported a dose range of 200-300 MBq. [29] The time between injection and PET scan was 30 minutes, [29] 45 minutes, [28] and 60 minutes. [26,30] Patients fasted for four, [28,29] six, [27] or twelve[30] hours. Two studies[29,30] measured glucose levels before administration of ^{18}FDG-PET; the maximum

glucose level that was allowed was "normal levels" [29] and 5.6 mmol/L. [30] Methods of interpretation of the images were qualitative in two studies[27,30] and both qualitative and quantitative in three. [26,28,29] Scans were interpreted qualitatively using visual analysis. [26-29] Two studies[28,29] reported using SUV but the criteria for abnormality were not reported.

Comparisons

Comparisons for which data were considered for a meta-analysis on the accuracy of ^{18}FDG-PET in detecting the stage of brain cancer are summarized in Table 9. Individual study data are summarized in Appendix D.

Table 9. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for brain cancer

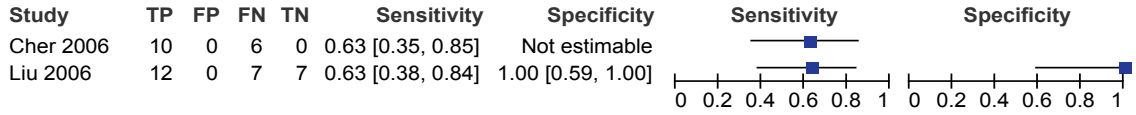
Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Primary diagnosis and recurrences	Chen 2006[26]	P	FDG-PET	Histology/biopsy or clinical followup	No
Recurrences	Potzi 2007[29]	R	FDG-PET	MRI, MET-PET	No
Staging	Cher 2006[27]	P	FDG-PET	Histology/biopsy	1. FDG-PET v. histology/biopsy (P studies) [27,28]
	Liu 2006[28]	P	FDG-PET	Histology/biopsy	
	Stockhammer 2007[30]	R	FDG-PET	Histology/biopsy	

FDG = fluorodeoxyglucose F18; MET = carbon-11 methionine; MRI = magnetic resonance imaging; P = prospective; PET = positron emission tomography; R = retrospective

1. ^{18}FDG-PET for the staging of brain cancer

Reference standard: histology/biopsy; prospective studies. Two prospective studies[27,28] totaling 42 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to histology for the staging of brain cancer. Individual 2x2 table results are presented in Figure 6. The sensitivity value in both of the individual studies was 63 percent. [27,28] Specificity data was provided by one study only[28] (100 percent). We could not calculate a pooled estimate of the positive and negative LRs for the accuracy of the staging of brain cancer because the study by Cher[27] provided sensitivity data only.

Figure 6. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology for the staging of brain cancer



2.3.2. ^{18}FDG-PET as part of a management strategy in brain cancer

One study assessed the impact of ^{18}FDG-PET as part of a management strategy of brain cancer at various stages of treatment. Padma et al. [129] conducted a retrospective study that examined the value of ^{18}FDG-PET results for predicting the survival of patients with brain cancer. The study included 331 patients with a mean age of 47 years (59 percent males), histologically-proven brain tumors according to WHO criteria, and should have been followed up until death or at least one year after ^{18}FDG-PET.

Prognostic value was assessed with respect to the ability of ^{18}FDG-PET to predict the grade of glioma and patient survival. Patients were followed up for an average of 3.6 years after ^{18}FDG-PET. One hundred and thirty-seven (41 percent) of the patients underwent ^{18}FDG-PET prior to histological diagnosis and any therapeutic intervention, while 194 patients underwent ^{18}FDG-PET between 2 months and 10 years following the histological diagnosis.

The influence of ^{18}FDG-PET in predicting survival was found to be significant. Overall, the median survival of patients with high uptake scores on ^{18}FDG-PET was 11 months v. 28 months in patients with low uptake scores. High ^{18}FDG-PET uptake was strongly associated with poor survival; while cases with low uptake had increased likelihood of long term survival (4-5 years). The authors concluded that ^{18}FDG-PET may help in the stratification of patients entered in protocols that evaluate therapeutic strategies in brain tumors. Additionally, the authors discussed the utility of ^{18}FDG-PET v. grading by histology for predicting the survival of patients in whom the ^{18}FDG-PET was done prior to surgery and any mode of therapeutic intervention.

Overall, the study was graded as level D of evidence (multiple flaws in methods). A detailed description of the methodological quality of this study is presented in Appendix H. The issues with the quality in the study included the lack of a comparator group who did not receive the ^{18}FDG-PET as a component of their disease monitoring. The selection criteria were only partially described, raising the possibility of selection bias. While the study had a large population, there is only partial description of the population characteristics. The method of ^{18}FDG-PET testing was not described sufficiently enough to permit reproducibility. Two different types of scans were used over the study period, and it is unknown how this may affect the detection of low or high ^{18}FDG uptake and therefore, affect outcome assessment. Additionally, while this study was relevant to management strategy, a matched design was not employed.

Table 10 provides a summary of the main findings and the types of bias that affected the evidence on ^{18}FDG-PET as part of a management strategy in brain cancer

Table 10. Main findings and types of bias that affected the evidence on ^{18}FDG-PET as part of a management strategy in brain cancer

Study	Patient Centered Outcomes	Types of bias
Padma 2003[129] Study type: Retrospective	FDG-PET used for: Predicting survival High FDG-uptake (n = 166) Low FDG-PET uptake (n = 165)	Selection bias (unclear) Disease progression bias (unclear) Review bias (RS unclear)

Survival	High uptake	Low Uptake
< 1 y	117/165	10/166
> 1 y	48/165	156/166
> 2 y	0/165	104/166
> 3 y	0/165	65/166
4 and 5 y	0/165	49 and 26/166

FDG = Fluorodeoxyglucose F18; PET = positron emission tomography; RS = reference standard; yr = years

3. Cervical Cancer

3.1. Background

In the United States in 2008, 11,070 women are expected to be diagnosed with new cases of cervical cancer and approximately 3,870 will die from the disease. [133] Incidence of cervical cancer varies greatly across subpopulations within the country. [139] Between 2000 and 2004 incidence rates of cervical cancer for Caucasian-American women were 8.5/100,000. For African-American women the numbers increase to 11.4/100,000 and the highest rate occurs in Hispanic-American women at 13.8/100,000. [133] Cervical cancer appears earlier in life than other malignancies. The median age at diagnosis is 48 years and the median age of death is 57 years. On average, cervical cancer accounts for 26.3 years of life lost in women diagnosed with this condition in the United States. [133]

Sexual intercourse at an early age, multiple male sexual partners who also have multiple partners and smoking are considered risks factors associated with the disease. [140] The vast majority of cervical cancer cases (99.7 percent) are associated with human papilloma virus (HPV). [141] There are many different types of HPV. High risk viral subtypes of HPV raise the risk of developing high-grade cervical dysplasia and cancer. Immunosuppression due to renal-allograft transplantation or Hodgkin's disease is also linked to cervical cancer. Precursors to cervical cancer known as cervical intraepithelial neoplasia (CIN) can occur in women less than 40 years of age. [140] HPV vaccines have helped to decrease rates of CIN significantly. Screening for cervical cancer using the Pap smear to assess for abnormal cervical cytology is commonplace in the United States. The Pap smear facilitates detection of precursor lesions, prior to the progression of disease to a more invasive cancer. Abnormal Pap findings require further evaluation with colposcopy and directed biopsies. [141]

If diagnosed in the early stages of disease, a high cure rate can be achieved. However, cervical cancer is often asymptomatic[140] and when left untreated, cervical cancer grows and frequently will metastasize into regional lymph nodes. [142] A patient may report vaginal discharge or postcoital vaginal bleeding. In cases of advanced disease lower extremity edema, deep vein thrombosis or ureteral obstruction may occur. [140] Two thirds of all cervical cancers are composed of squamous cell carcinoma, while much of the remaining 25 percent are adenocarcinoma. Tumors are staged using the International Federation of Gynaecology (FIGO) system, which takes tumor grade, depth, width and extent of invasion into consideration (Table 11). [143]

Table 11. FIGO staging of cervical cancer

Stage	Description
Stage 0	Carcinoma in situ
Stage Ia1	Invasive carcinoma, confined to cervix, lesion ≤3 mm deep, ≤ 7 mm wide
Stage Ia2	Invasive carcinoma, confined to cervix, lesion >3 mm and ≤ 5 mm deep, ≤ 7 mm wide
Stage Ib1	Invasive carcinoma, confined to cervix, lesion ≤ 4 cm
Stage Ib2	Invasive carcinoma, confined to cervix, lesion > 4 cm
Stage IIa	Tumor extended beyond cervix to vagina (but not lower 1/3)
Stage IIb	Tumor extended beyond cervix, parametrial invasion (but not to pelvic side wall or lower 1/3 of vagina)
Stage IIIa	Tumor extended to lower 1/3 of vagina (but not to pelvic side wall)
Stage IIIb	Tumor extended to pelvic side wall, interferes with kidney function
Stage IVa	Tumor extended into bladder or rectum
Stage IVb	Distant metastasis

Taken from Petignat et al. [143]

Limitations to current screening and imaging modalities exist. Pap tests are commonly used to cytologically evaluate the cervix, but are subject to errors occurring during sample collection or evaluation. An alternative to conventional Pap testing is liquid-based cytology. Findings do not consistently demonstrate if liquid-based cytology is more effective than conventional Pap testing. [139] Regardless, screening is a useful tool and has dramatically reduced the incidence and mortality of cervical cancer.

Testing may also be conducted to detect HPV DNA. The United States and some European countries screen for specific biomarkers, which improves efficiency and maximizes sensitivity. It is an adjunctive test with cytology for women 20 years of age or older. Although screening for biomarkers is more sensitive and has high negative predictive values, it suffers from lower specificity than Pap tests as HPV infections are common in sexually active women. [139]

When local disease is diagnosed, screening with CT or MRI is helpful for defining lymph node status and determining the extent of disease. Identifying involved nodes can be difficult as their identification relies on size and morphological criteria. Surgery provides another method for staging. Pelvic lymphadenectomy and para-aortic lymphadenectomy are two techniques frequently used. Many studies demonstrate excellent patient results after surgical staging. Imaging techniques such as CT, MRI and PET, in addition to surgical staging may be more effective in identifying the true extent of disease than clinical testing. However, these techniques have yet to be incorporated into the FIGO staging system. [143]

^{18}FDG-PET could have some advantages over CT in the imaging of cervical cancer. Lesion location can be identified with CT, but ^{18}FDG-PET may be capable of detecting nodal involvement when CT is not. ^{18}FDG-PET may also be useful in diagnosing recurrent and metastatic disease. [144]

Cure can be achieved in 80 to 90 percent of patients with stage I and II disease when treated with surgery or chemoradiotherapy. [143] Surgery is usually performed first followed by chemotherapy or radiotherapy, which helps decrease the risk of reoccurrence. [140] Conisation is performed in woman with stage I disease if fertility is to be preserved; otherwise, simple hysterectomy is performed. Radical hysterectomy is performed for higher grade tumors; radical trachelctomy may provide a surgical option for younger women who wish to preserve fertility. [143] Relapse occurs frequently in patients with stages IIb, III and IV even after treatment with surgery and radiotherapy. [140] Recurrences usually occur within two years after the completion of primary treatment. [143] Approximately 30 percent of women with invasive cancer die from recurrence. [140] The goal for treatment of patients with stage IVb cancer is palliative. How treatment affects quality of life and toxicity influences choice of treatment. [143]

3.2. Importance of Key Questions in the Clinical Management of Cervical Cancer

Cervical cancer spreads directly through the lymphatic system, with pelvic node metastasis preceding aortic node metastasis in the majority of the cases. Sensitive and specific imaging modalities that identify occult lymph node metastasis may help avoid morbid surgical procedures and facilitate treatment planning with novel modalities. Earlier detection of recurrent cervical cancer has the potential to improve survival, since some patients may be salvaged using radiotherapy or radical surgery. Local recurrences may be difficult to detect by anatomical examination because the soft tissue structures are thickened following radiation or surgery. Anatomical imaging techniques such as CT and MRI can be fairly inaccurate in detecting retroperitoneal nodal metastasis and therefore, it is important to explore whether functional imaging methods such as ^{18}FDG-PET can help to improve the accuracy of pretreatment staging and have a positive impact on patient survival. It is important to determine whether the use of ^{18}FDG-PET in patients with cervical cancer can improve patient-centered outcomes by altering the primary management strategies. As the available treatments for cervical cancer recurrence improve, such as radical resection in combination with intraoperative high-dose-rate brachytherapy, the improvement of imaging modalities to identify recurrences early becomes more important.

3.3. Results

Thirty-three studies[31-63] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for cervical cancer. All of the studies evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT for cervical cancer. Six studies reported on the diagnostic thinking impact of ^{18}FDG-PET[33,43,44,61] and ^{18}FDG-PET/CT, [32,38] and two studies[33,43] evaluated the impact of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes. No economic evaluations on the use of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer were identified. Characteristics of the populations, conditions of ^{18}FDG-PET and ^{18}FDG-PET/CT administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

3.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in cervical cancer

Characteristics of the studies

Thirty-three studies (21 prospective, [31-33,35-37,40,42-47,49,51-53,60-63] 12 retrospective[34,38,39,41,48,50,54-59]) evaluated the diagnostic accuracy of ^{18}FDG-PET [33-35,37,39-45,47-51,54-62] and ^{18}FDG-PET/CT [31,32,36,38,46,52,53,59,63] on cervical cancer. Ten studies[37,42,45,47-49,54,55,59,62] used ^{18}FDG-PET for initial staging, one for primary diagnosis and recurrence, [35]11 for recurrence, [33,34,39,41,44,50,51,56,57,60,61] one for restaging, [43] one for staging and recurrence, [40] and one for staging and restaging. [58] Six studies used ^{18}FDG-PET/CT for initial staging, [31,36,46,53,59,63] two for recurrence[38,52] and one for staging and restaging purposes. [32] The studies contained a total of 2,767 patients with sample sizes ranging from 14 to 517 participants. Participant ages ranged from 20 to 87 years. Twenty-seven studies reported the distribution by stage of cancer[32-39,41-44,46,48-53,55-57,59-63] and included variously all stages from IA1 to stage IV. ^{18}FDG-PET was compared to histology/biopsy in all studies, but in 19 studies the reference standard also included clinical followup, [31-35,38-40,43,44,46,50,52,56-58,60-62] and in one[47] imaging followup was also used. Twelve studies[32-34,38,44,47,51-53,56,57,63] reported the mean time between last treatment and ^{18}FDG-PET, which ranged from 7 days[53] to 42 months. [47] Seventeen studies reported using a fixed dose of ^{18}FDG (322MBq, [48] 370MBq, [33-35,37,43-45,47,52,53,57,61,62] 400MBq, [46] 550MBq[55,56]); four studies used a weight-based dose (0.14mCi/kg, [41] 0.22mCi/kg, [38] 5MBq/kg, [48] 5.2MBq/kg[58]). The time between injection and PET scan ranged from 30 minutes[34] to 3 hours. [47] Patients fasted anywhere from 4 hours[31,32,34,38,40,45,51,55,56,63] to overnight. [145] Seven studies[31,32,35,47,52,53,57] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose level permitted was 200 mg/dL. [31,52] Methods of interpretation of the images were

quantitative in one study, [51] qualitative in 16 studies[32-34,37,38,41,44,45,47,50,52,53,56-58,61] and both qualitative and quantitative in 8 studies. [35,36,39,43,48,49,60,62] Scans were interpreted qualitatively using visual analysis in 26 studies. [32-39,41,43-45,47-53,56-62] Four studies[36,39,47,49] reported using both visual analysis and SUV and one[51] used SUV only.

Comparisons

Comparisons for which data were considered for meta-analysis are summarized in Table 12. Pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET in cervical cancer for staging and for detection of recurrences. Pooled data were also obtained to evaluate the accuracy of ^{18}FDG-PET/CT for staging of cervical cancer. Individual study data are summarized in Appendix D.

Table 12. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer

Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Primary diagnosis and recurrences	Chang 2005[35]	P	FDG-PET	Histology/biopsy or clinical followup	No
Staging and recurrences	Grisaru 2004[40]	P	FDG-PET	Histology/biopsy	No
Staging and restaging	Bjurberg 2007[32]	P	FDG-PET/CT	Histology/biopsy or clinical followup	No
	Wong 2004[58]	R	FDG-PET	Histology/biopsy or clinical followup	
Recurrences	Chang 2004[33]	P	FDG-PET	Histology/biopsy or clinical followup (local v. distant)	1. FDG-PET v. histology/biopsy or clinical followup (P studies) [44,60,61]
	Chang 2004[34]	R	FDG-PET	Histology/biopsy or clinical followup (lesion-based)	
	Chung 2007[38]	R	FDG-PET/CT	Histology/biopsy or clinical followup	2. FDG-PET v. histology/biopsy or clinical followup (R studies) [39,50,56]
	Chung 2006[39]	R	FDG-PET	Histology/biopsy or clinical followup	
	Havrilesky 2003[41]	R	FDG-PET	Histology/biopsy (lesion-based)	
	Lin 2006[44]	P	FDG-PET	Histology/biopsy or clinical followup	
	Ryu 2003[50]	R	FDG-PET	Histology/biopsy or clinical followup	
	Sakurai 2006[51]	P	FDG-PET	Histology/biopsy (lesion-based)	
	Sironi 2007[52]	P	FDG-PET/CT	Histology/biopsy or clinical followup	
	Unger 2004[56]	R	FDG-PET	Histology/biopsy or clinical followup	
	Van Der Veldt 2006[57]	R	FDG-PET	Histology/biopsy or clinical followup (lesion-based)	
	Yen 2006[60]	P	FDG-PET	Histology/biopsy or clinical followup	

Table 12. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer (continued)

Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Recurrences (cont')	Yen 2004[61]	P	FDG-PET	Histology/biopsy or clinical followup	
Restaging	Lai 2004[43]	P	FDG-PET	Histology/biopsy or clinical followup	No
Staging	Amit 2006[31]	P	FDG-PET/CT	Histology/biopsy or clinical followup	1. FDG-PET v. any reference standard (P studies) [37,42,45,47,49]
	Choi 2006[36]	P	FDG-PET/CT	Histology/biopsy (lesion-based)	
	Chou 2006[37]	P	FDG-PET	Histology/biopsy	2. FDG-PET v. histology (P studies) [37,42,45,49]
	Hope 2006[42]	P	FDG-PET	Histology/biopsy	
	Lin 2003[45]	P	FDG-PET	Histology/biopsy	
	Loft 2007[46]	P	FDG-PET/CT	Histology/biopsy or clinical followup	3. FDG-PET v. histology (R studies) [48,54,55]
	Ma 2003[47]	P	FDG-PET	Histology/biopsy and imaging followup	4. FDG-PET/CT v. any reference standard (P studies) [31,46,63]
	Park 2005[48]	R	FDG-PET	Histology/biopsy	
	Roh 2005[49]	P	FDG-PET	Histology/biopsy	
	Sironi 2006[53]	P	FDG-PET/CT	Histology/biopsy (node-based)	5. FDG-PET/CT v. histology/biopsy or clinical followup (P studies) [31,46]
	Tran 2003[54]	R	FDG-PET	Histology/biopsy	
	Unger 2005[55]	R	FDG-PET	Histology/biopsy	
	Wright 2005[59]	R	FDG-PET and FDG-PET/CT	Histology/biopsy	
	Yen 2003[62]	P	FDG-PET	Histology/biopsy or clinical followup (lesion-based)	
	Yildirim 2008[63]	P	FDG-PET/CT	Histology/biopsy	

CT = computed tomography; FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

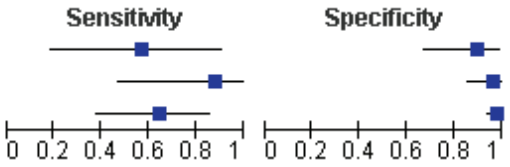
1. ^{18}FDG-PET for recurrences of cervical cancer

Reference standard: histology/biopsy or clinical followup; prospective studies. Three prospective studies[44,60,61] totaling 231 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to histology/biopsy or clinical followup for detecting recurrences of cervical cancer. Recurrences were identified by site including peritoneum, bone, liver/spleen, lung, mediastinal lymph node (MLN), supraclavicular lymph node (SLN), para-aortic lymph node (PALN), pelvic lymph node (PLN), and inguinal lymph node (ILN). Individual 2x2 table results are presented in Figure 7. Sensitivity values in individual studies ranged from 50 percent for bone and PLN sites[44] to 100 percent for liver/spleen, [44,60] MLN[44,60,61] and ILN[44,60] sites. Specificity ranged from 88 percent for MLN[44] to 100 percent for bone, [60] liver/spleen, [44] lung, [44,60] PALN, [60] PLN[44] and ILN. [44,61]

Figure 7. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer

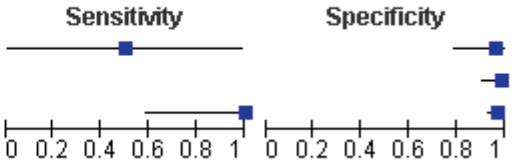
Peritoneum

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	4	2	3	17	0.57 [0.18, 0.90]	0.89 [0.67, 0.99]
Yen 2004	7	2	1	45	0.88 [0.47, 1.00]	0.96 [0.85, 0.99]
Yen 2006	11	3	6	129	0.65 [0.38, 0.86]	0.98 [0.94, 1.00]



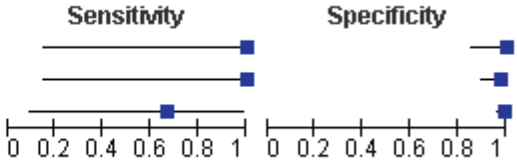
Bone

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	1	1	1	23	0.50 [0.01, 0.99]	0.96 [0.79, 1.00]
Yen 2004	0	1	0	54	Not estimable	0.98 [0.90, 1.00]
Yen 2006	7	4	0	139	1.00 [0.59, 1.00]	0.97 [0.93, 0.99]



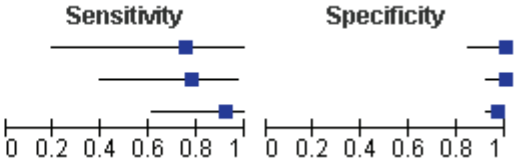
Liver/spleen

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	2	0	0	24	1.00 [0.16, 1.00]	1.00 [0.86, 1.00]
Yen 2004	2	1	0	52	1.00 [0.16, 1.00]	0.98 [0.90, 1.00]
Yen 2006	2	1	1	144	0.67 [0.09, 0.99]	0.99 [0.96, 1.00]



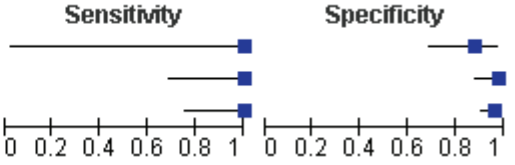
Lung

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	3	0	1	22	0.75 [0.19, 0.99]	1.00 [0.85, 1.00]
Yen 2004	7	0	2	46	0.78 [0.40, 0.97]	1.00 [0.92, 1.00]
Yen 2006	11	4	1	129	0.92 [0.62, 1.00]	0.97 [0.92, 0.99]



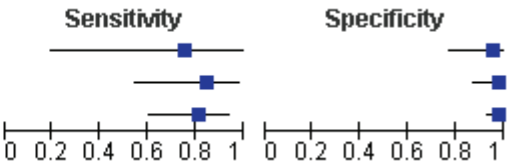
MLN

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	1	3	0	22	1.00 [0.03, 1.00]	0.88 [0.69, 0.97]
Yen 2004	10	1	0	44	1.00 [0.69, 1.00]	0.98 [0.88, 1.00]
Yen 2006	13	5	0	118	1.00 [0.75, 1.00]	0.96 [0.91, 0.99]



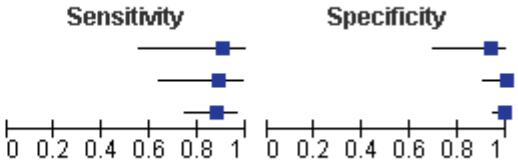
SLN

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	3	1	1	21	0.75 [0.19, 0.99]	0.95 [0.77, 1.00]
Yen 2004	11	1	2	41	0.85 [0.55, 0.98]	0.98 [0.87, 1.00]
Yen 2006	21	3	5	118	0.81 [0.61, 0.93]	0.98 [0.93, 0.99]



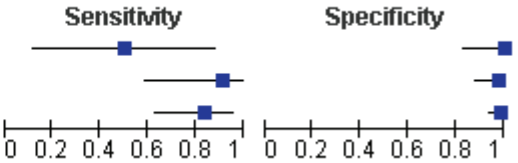
PALN

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	9	1	1	15	0.90 [0.55, 1.00]	0.94 [0.70, 1.00]
Yen 2004	15	0	2	38	0.88 [0.64, 0.99]	1.00 [0.91, 1.00]
Yen 2006	37	1	5	102	0.88 [0.74, 0.96]	0.99 [0.95, 1.00]



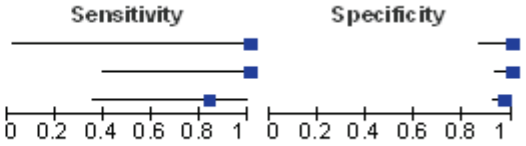
PLN

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	3	0	3	20	0.50 [0.12, 0.88]	1.00 [0.83, 1.00]
Yen 2004	10	1	1	43	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]
Yen 2006	20	2	4	117	0.83 [0.63, 0.95]	0.98 [0.94, 1.00]



ILN

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lin 2006	1	0	0	25	1.00 [0.03, 1.00]	1.00 [0.86, 1.00]
Yen 2004	4	0	0	51	1.00 [0.40, 1.00]	1.00 [0.93, 1.00]
Yen 2006	5	5	1	136	0.83 [0.36, 1.00]	0.96 [0.92, 0.99]



Figures 8 and 9 present the positive and negative LRs of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer. We found that all the positive LRs by site of recurrence were statistically significant ranging from 15.24 (95% CI = 5.63, 41.27) for MLN to 45.89 (95% CI = 14.09, 149.49) for liver/spleen. Overall, the positive LRs across the studies were homogeneous except for MLN, where moderate heterogeneity was found across the studies ($p = 0.12$; $I^2 = 53$ percent). All the negative LRs by site of recurrence were statistically significant except for the identification of bone recurrences. Negative LRs ranged from 0.09 (95% CI = 0.02, 0.40) for MLN to 0.37 (95% CI = 0.22, 0.60) for peritoneum. The negative LRs across the studies were homogeneous except for the identification of recurrences in bone ($p = 0.07$; $I^2 = 70$ percent) and PLN ($p = 0.03$; $I^2 = 71$ percent).

Figure 8. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site)

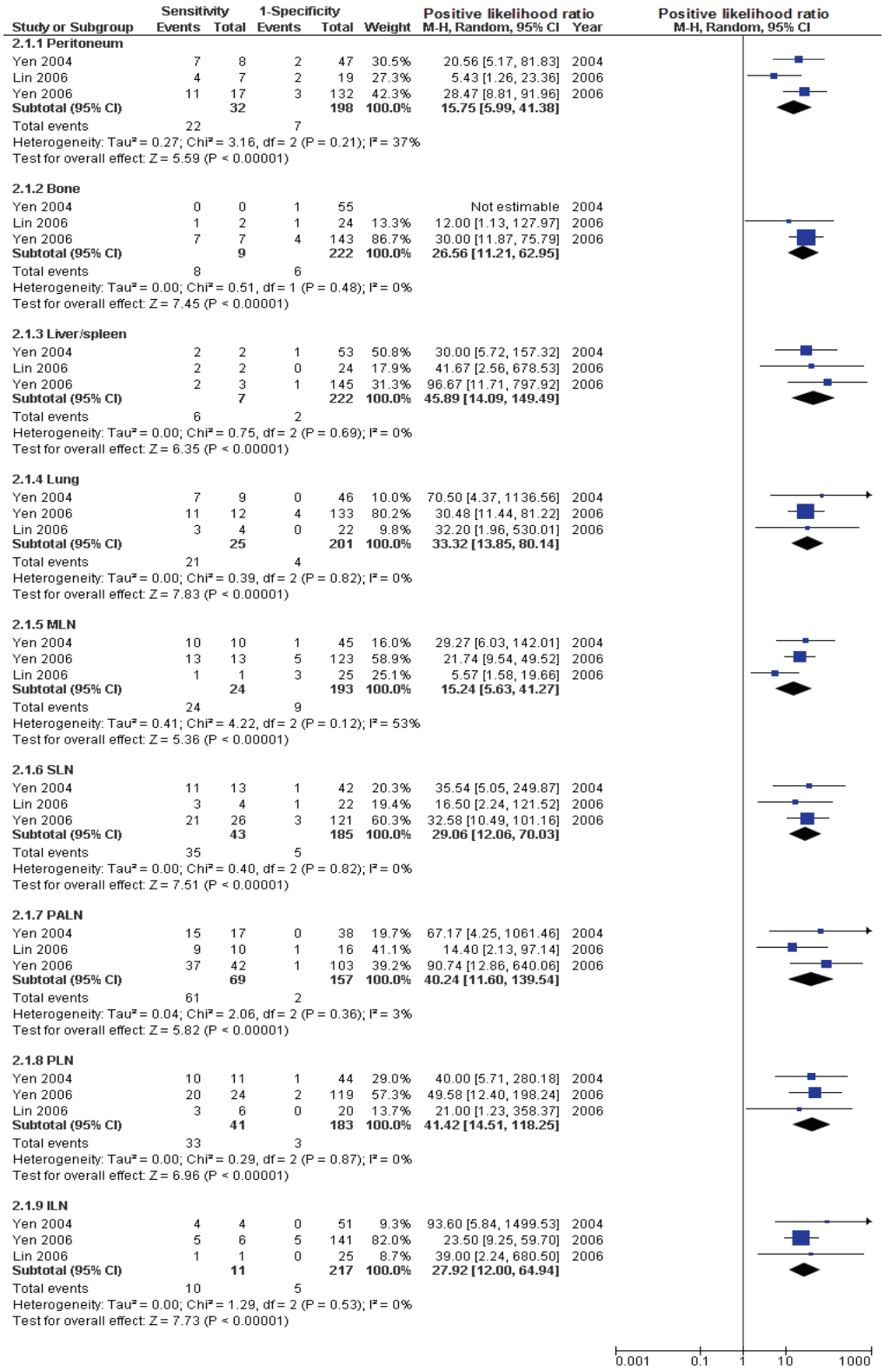


Figure 9. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site)

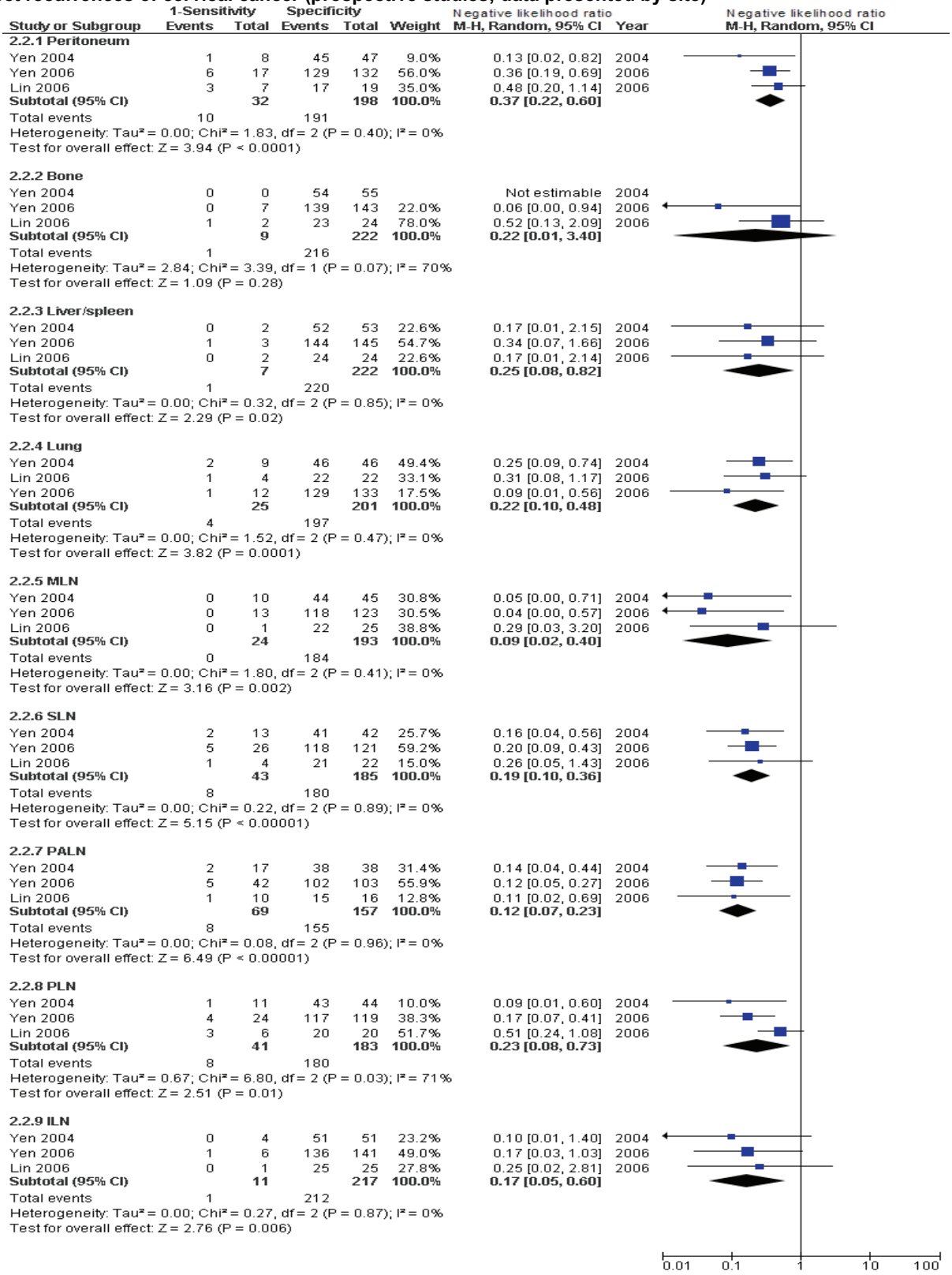
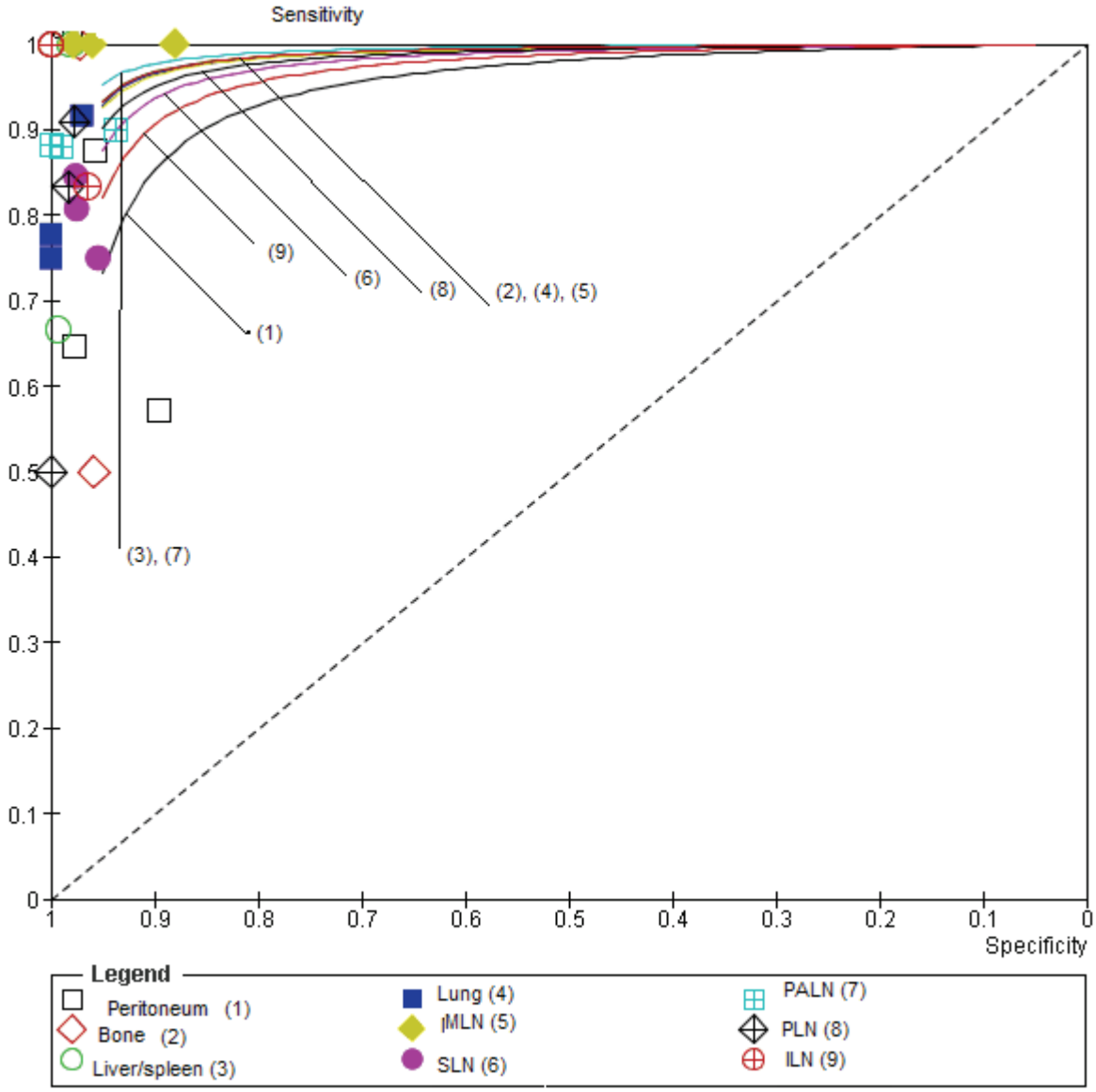


Figure 10 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical per site of lesion cancer based on prospective studies.

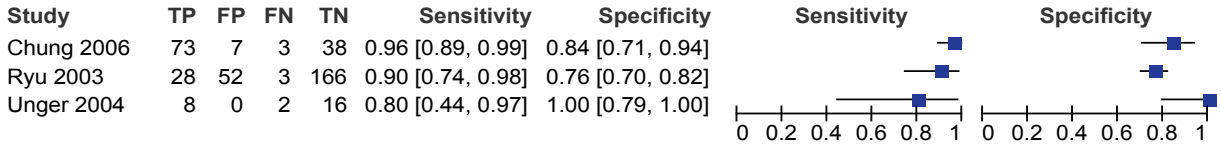
Figure 10. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of cervical cancer (prospective studies; data presented by site)



Reference standard: histology/biopsy or clinical followup; retrospective studies. Separate meta-analyses were conducted for retrospective studies. Three retrospective studies[39,50,56] totaling 396 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to histology/biopsy or clinical followup for detecting recurrences of cervical cancer. Individual 2x2

table results are presented in Figure 11. Sensitivity values in individual studies ranged from 80 percent[56] to 96 percent. [39] Specificity ranged from 76 percent[50] to 100 percent. [56]

Figure 11. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer



Based on the analysis of retrospective studies, we found that ^{18}FDG-PET had a pooled positive LR of 5.33 (95% CI = 2.36, 12.05) and a pooled negative LR of 0.11 (95% CI = 0.04, 0.28) to accurately detect recurrences of cervical cancer (Figures 12 and 13). The positive and negative LRs were consistently positive and statistically significant and therefore, ^{18}FDG-PET seems to be helpful to identify recurrences of the disease. However, the magnitude of both the positive ($p = 0.03$; $I^2 = 70$ percent) and the negative ($p = 0.12$; $I^2 = 53$ percent) LRs were heterogeneous across the studies, therefore the magnitude of the effect is less certain.

Figure 12. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer (retrospective studies)

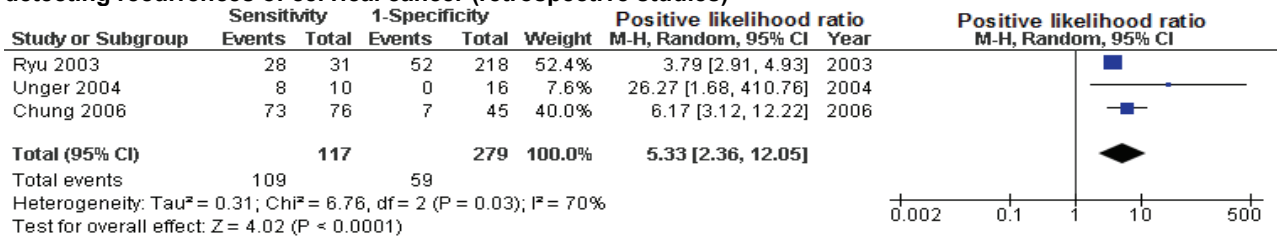


Figure 13. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer (retrospective studies)

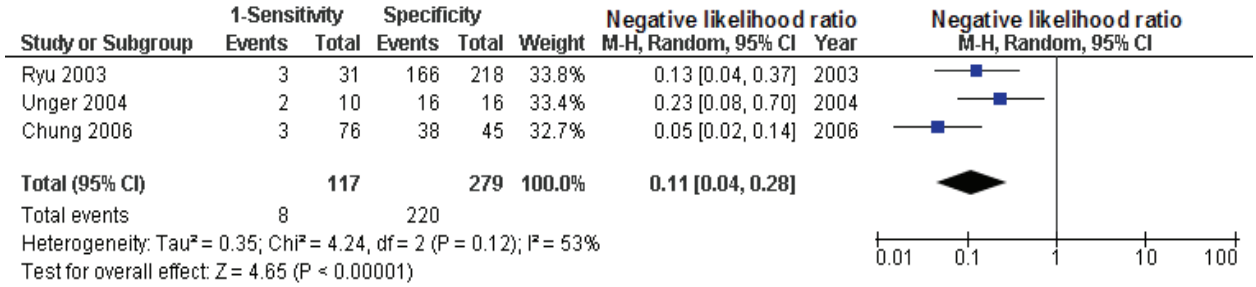
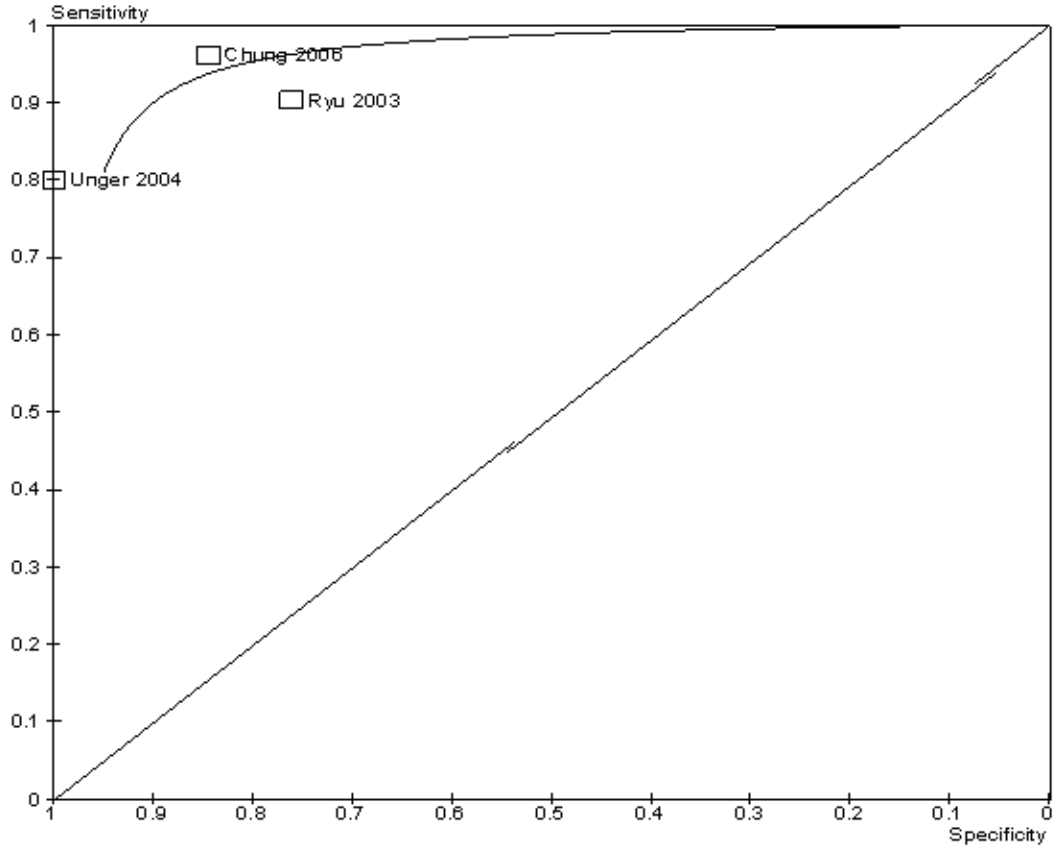


Figure 14 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for detecting recurrences of cervical cancer.

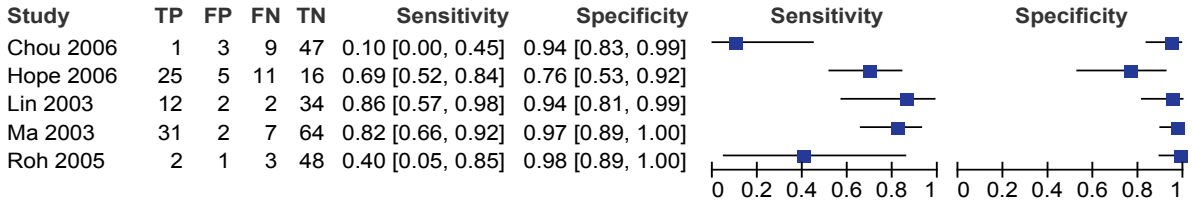
Figure 14. Summary ROC plot of ^{18}FDG-PET v. any reference standard for detecting recurrences of cervical cancer (retrospective studies)



2. ^{18}FDG-PET for the staging of cervical cancer

Reference standard: any; prospective studies. Five prospective studies[37,42,45,47,49] totaling 325 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to a variety of reference standards for the staging of cervical cancer. Individual 2x2 table results are presented in Figure 15. Sensitivity ranged from 10 percent[37] to 86 percent. [45] Specificity ranged from 76 percent[42] to 98 percent. [49]

Figure 15. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer



We found that ^{18}FDG-PET had a pooled positive LR of 8.22 (95% CI = 2.59, 26.08) and a pooled negative LR of 0.38 (95% CI = 0.12, 1.20) to accurately identify the stage of cervical cancer (Figures 16 and 17). The pooled positive LR was statistically significant and therefore, ^{18}FDG-PET seems to be helpful to detect the stage of the disease. The negative LR was not statistically significant and therefore, ^{18}FDG-PET does not seem to be helpful in ruling out the presence of particular stages of the disease. There was high heterogeneity across the studies in the positive LR ($p = 0.01$; $I^2 = 69$ percent) and negative LR ($p < 0.000001$; $I^2 = 95$ percent).

Figure 16. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies)

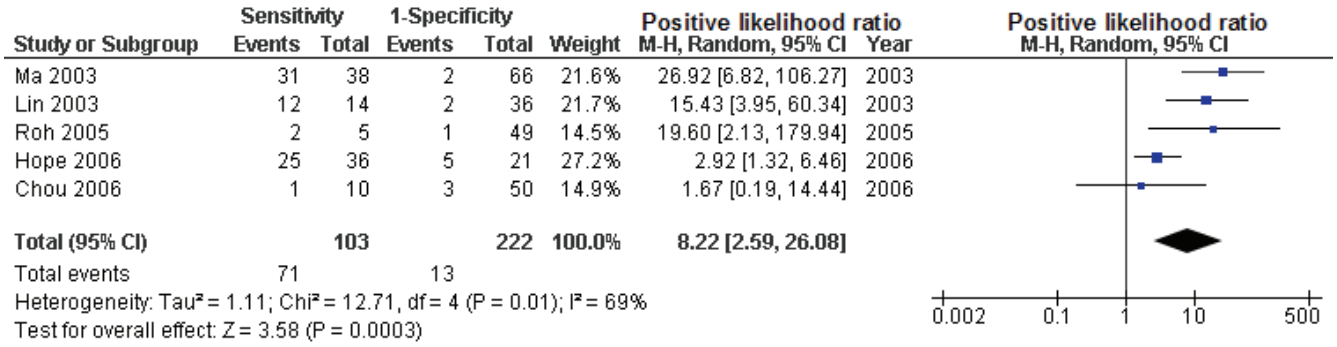


Figure 17. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies)

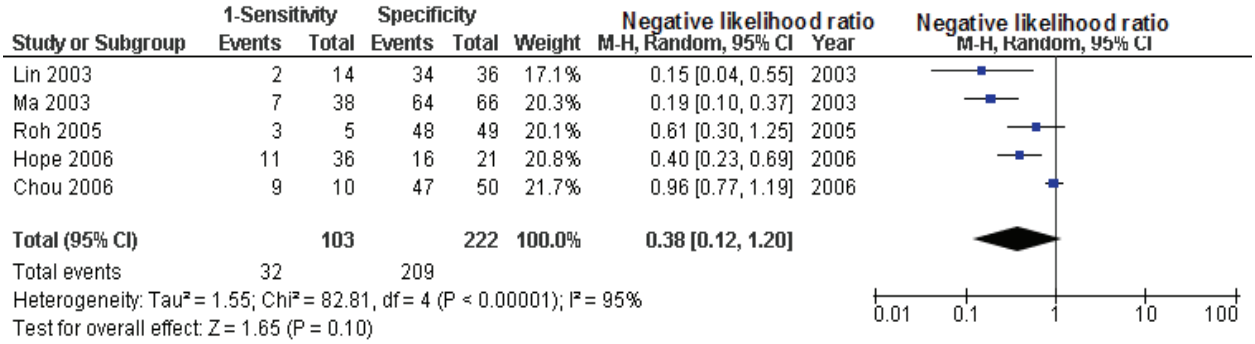
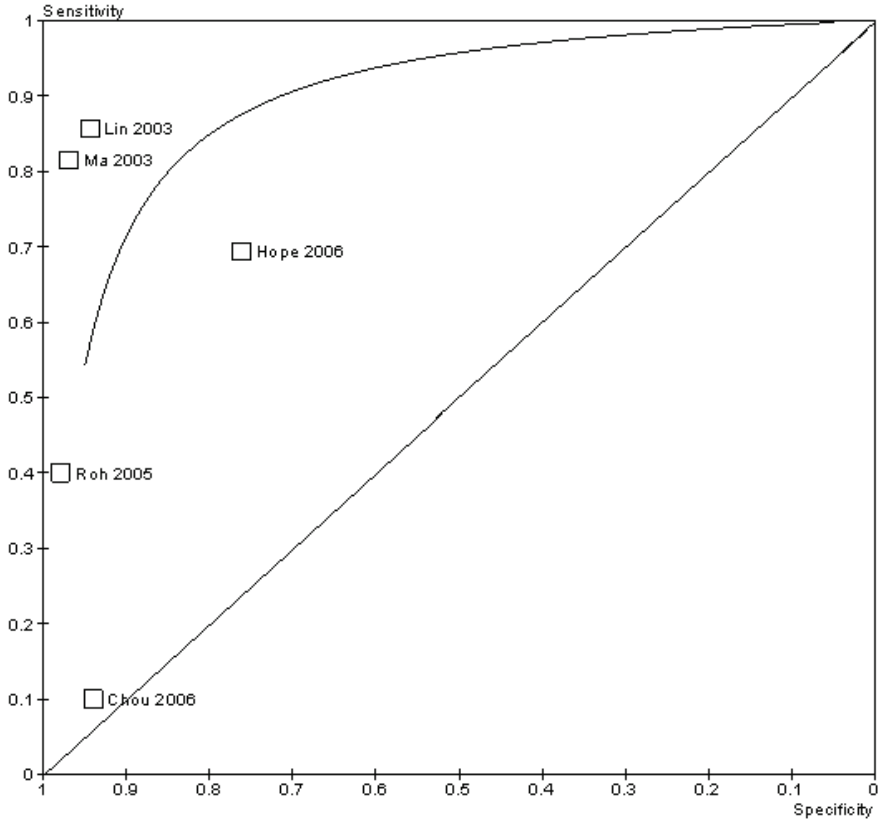


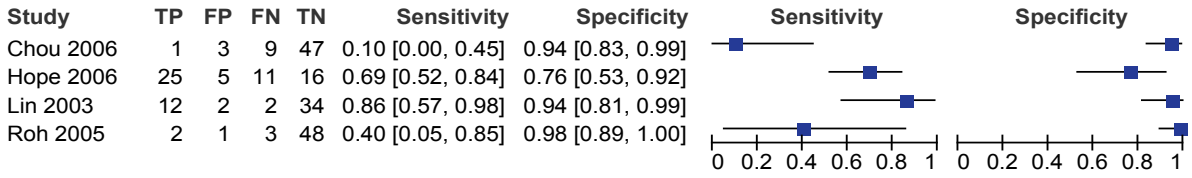
Figure 18 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. any reference standard for the staging of cervical cancer based on prospective studies.

Figure 18. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the staging of cervical cancer (prospective studies)



Reference standard: histology/biopsy; prospective studies. Four prospective studies[37,42,45,49] totaling 221 participants provided data for a subgroup analysis of the accuracy of ^{18}FDG-PET when histology/biopsy or clinical followup were used as the reference standard for the staging of cervical cancer. Individual 2x2 table results are presented in Figure 19. Sensitivity ranged from 10 percent[37] to 86 percent. [45] Specificity ranged from 76 percent[42] to 98 percent. [49]

Figure 19. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer



We found that when only histology/biopsy were considered as the reference standard, ^{18}FDG-PET had a pooled positive LR of 5.79 (95% CI = 1.88, 17.88) and a pooled negative LR of 0.47 (95% CI = 0.17, 1.32) to accurately identify the staging of cervical cancer (Figures 20 and 21). Both the positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET seems to help to classify the stage of the disease. However, both the positive ($p = 0.08$; $I^2 = 56$ percent) and the negative ($p < 0.000001$; $I^2 = 92$ percent) LRs were heterogeneous across the studies precluding firm conclusions based on these results.

Figure 20. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)

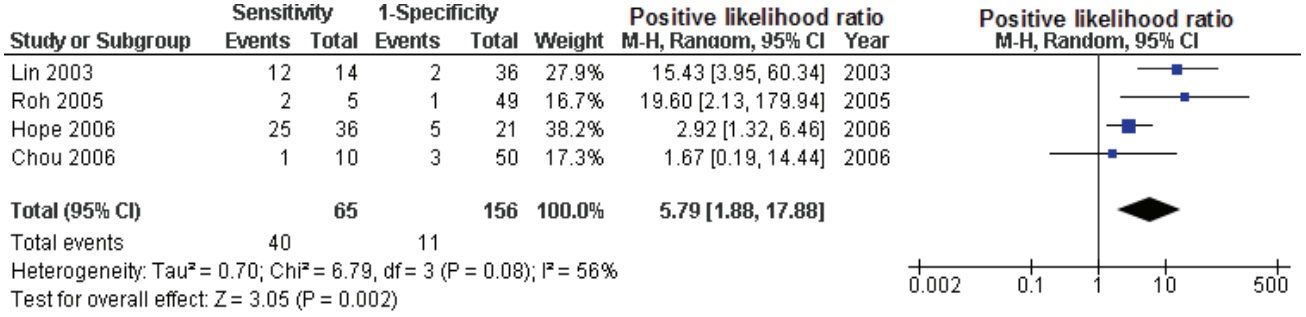


Figure 21. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)

Study or Subgroup	1-Sensitivity Events	Total	Specificity Events	Total	Weight	Negative likelihood ratio M-H, Random, 95% CI	Year
Lin 2003	2	14	34	36	19.9%	0.15 [0.04, 0.55]	2003
Roh 2005	3	5	48	49	25.2%	0.61 [0.30, 1.25]	2005
Hope 2006	11	36	16	21	26.6%	0.40 [0.23, 0.69]	2006
Chou 2006	9	10	47	50	28.3%	0.96 [0.77, 1.19]	2006
Total (95% CI)		**65**		**156**	**100.0%**	**0.47 [0.17, 1.32]**	
Total events	25		145				

Heterogeneity: Tau² = 0.97; Chi² = 39.79, df = 3 (P < 0.00001); I² = 92%

Test for overall effect: Z = 1.43 (P = 0.15)

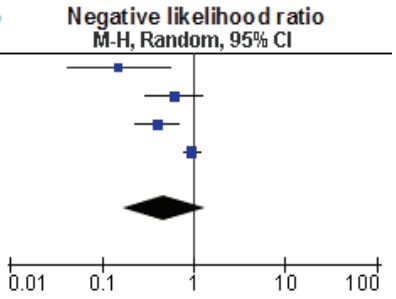
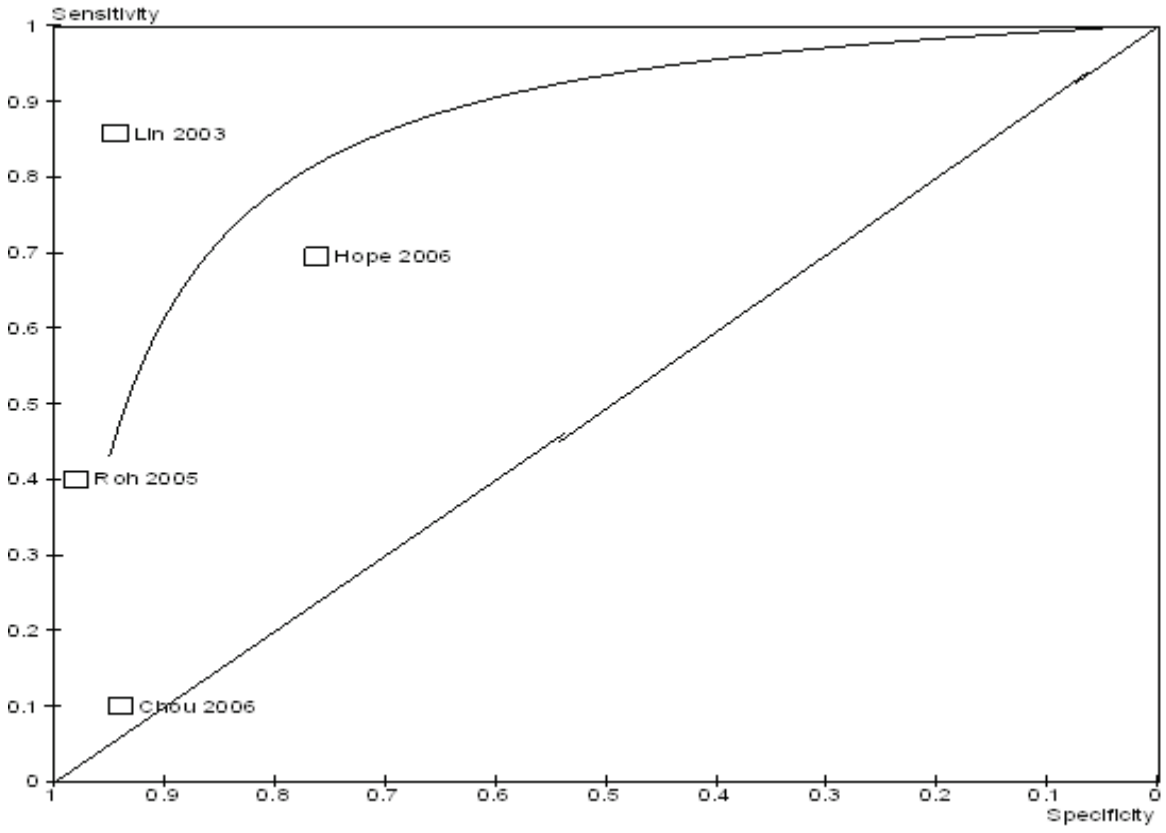


Figure 22 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer.

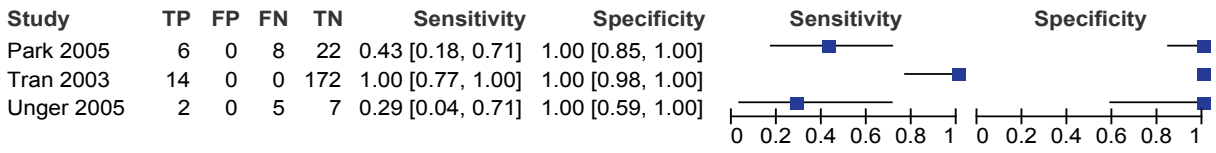
Figure 22. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)



Reference standard: histology/biopsy; retrospective studies. A separate meta-analysis was conducted for retrospective studies of ^{18}FDG-PET compared to histology/biopsy for the staging of

cervical cancer. Three retrospective studies[48,54,55] totaling 236 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET compared to histology/biopsy for the staging of cervical cancer. Individual 2x2 table results are presented in Figure 23. Sensitivity values ranged from 29 percent[55] to 100 percent. [54] Specificity was 100 percent across the three studies.

Figure 23. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer



Based on the analysis of the retrospective studies, we found that ^{18}FDG-PET had a pooled positive LR of 32.90 (95% CI = 2.89, 375.25) and a pooled negative LR of 0.41 (95% CI = 0.11, 1.55) to accurately identify the stage of cervical cancer (Figures 24 and 25). The positive LR was statistically significant; however, the results were moderately heterogeneous across the studies (p = 0.11; I^2 = 55 percent). The negative LR was not statistically significant and therefore, the test does not seem to be helpful in ruling out the presence of particular stages of the disease.

Figure 24. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies)

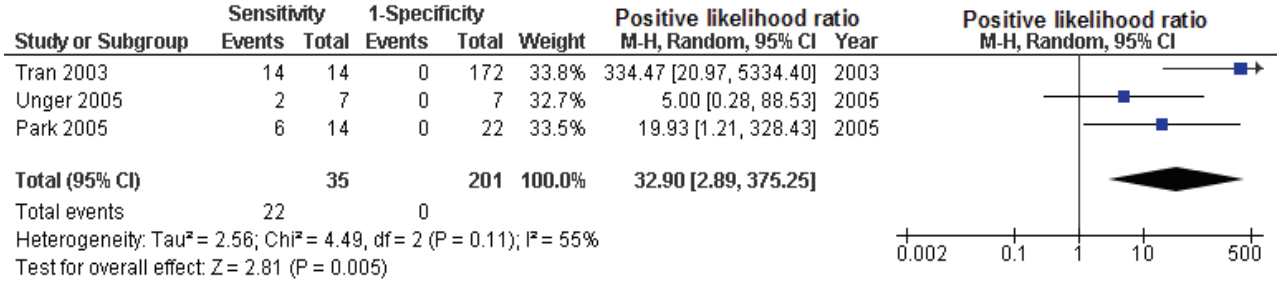


Figure 25. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies)

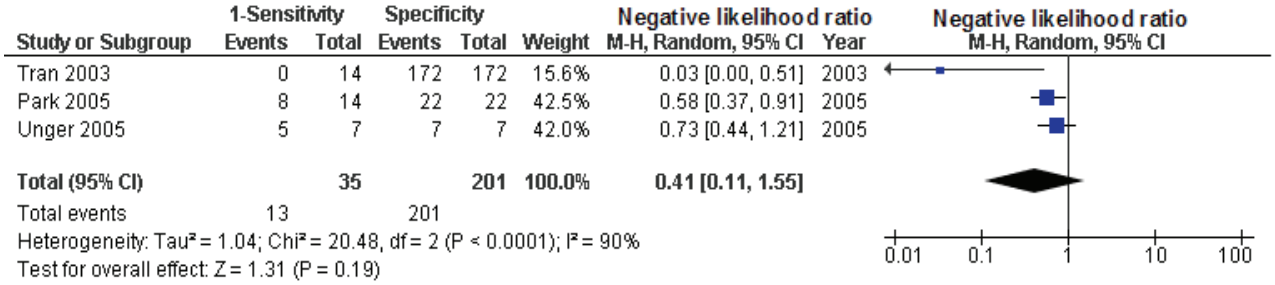
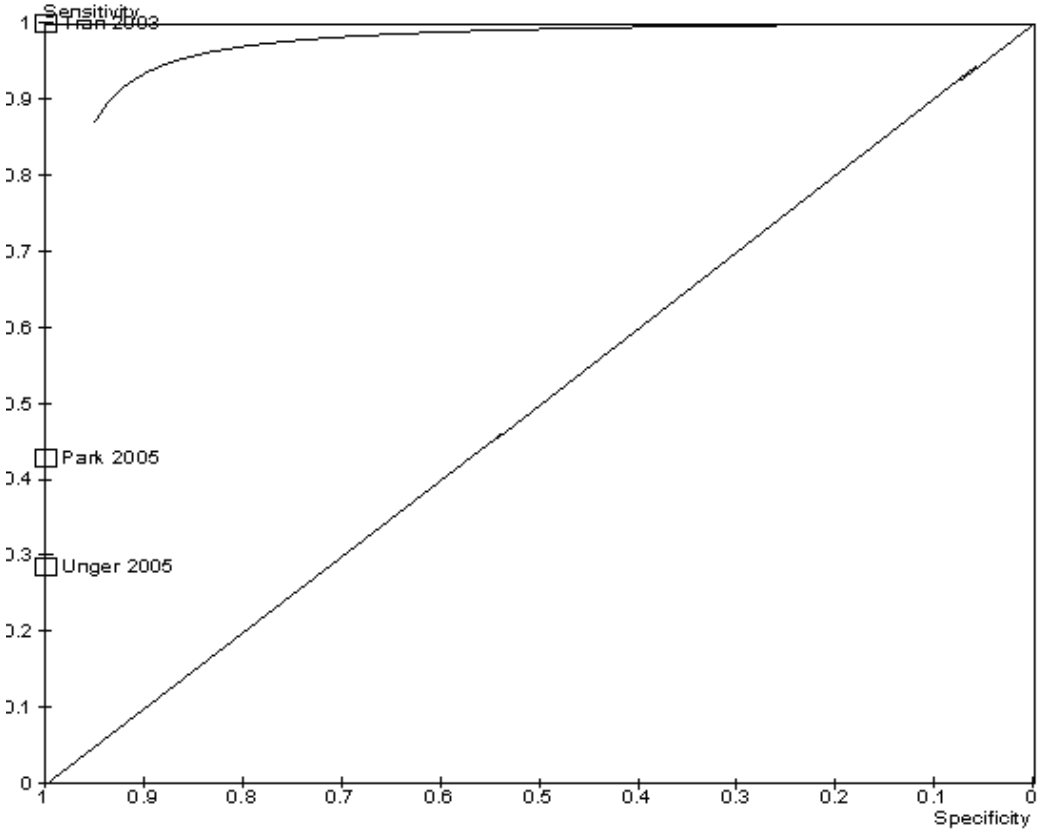


Figure 26 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer based on retrospective studies.

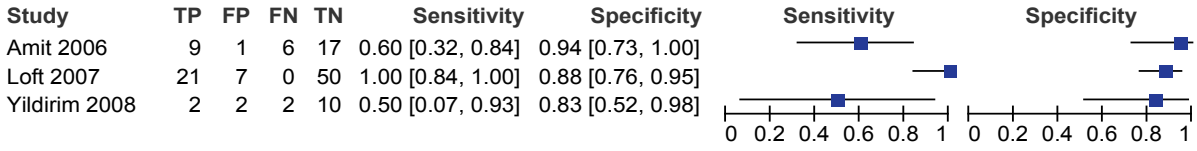
Figure 26. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the staging of cervical cancer (retrospective studies)



3. ^{18}FDG-PET/CT for the staging of cervical cancer

Reference standard: any; prospective studies. Three prospective studies[31,46,63] totaling 127 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET/CT compared to a variety of reference standards for the staging of cervical cancer. Individual 2x2 table results are presented in Figure 27. Sensitivity values ranged from 50 percent[63] to 100 percent. [46] Specificity ranged from 83 percent[63] to 94 percent. [31]

Figure 27. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer



We found that ^{18}FDG-PET/CT had a pooled positive LR of 6.89 (95% CI = 3.82, 12.42) and a pooled negative LR of 0.28 (95% CI = 0.06, 1.38) to accurately identify the stage of cervical cancer (Figures 28 and 29). The positive LR was statistically significant and the results were homogeneous across the studies and therefore, ^{18}FDG-PET/CT seems to be helpful to identify the stage of the disease. The negative LR was not statistically significant and the results were quite heterogeneous across the studies ($p = 0.004$; $I^2 = 82$ percent) and therefore, ^{18}FDG-PET/CT does not seem to be helpful in ruling out particular stages of the disease.

Figure 28. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies)

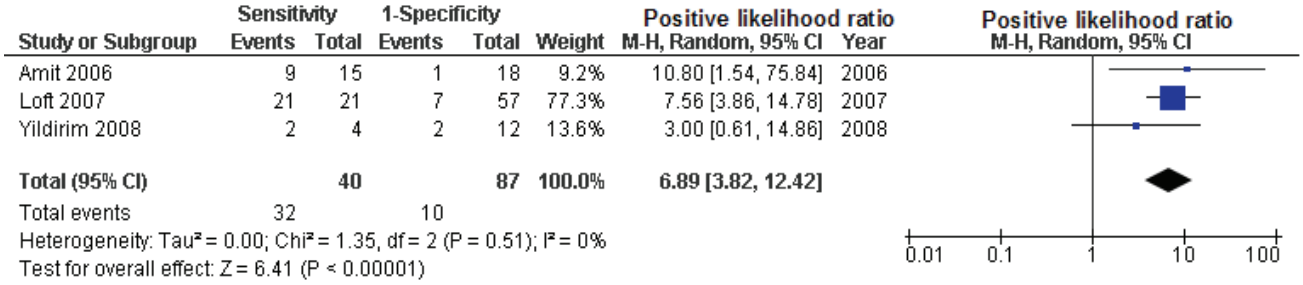


Figure 29. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies)

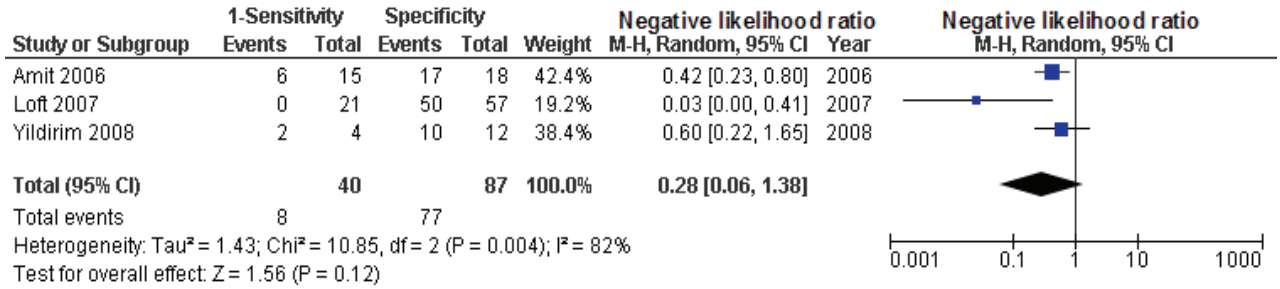
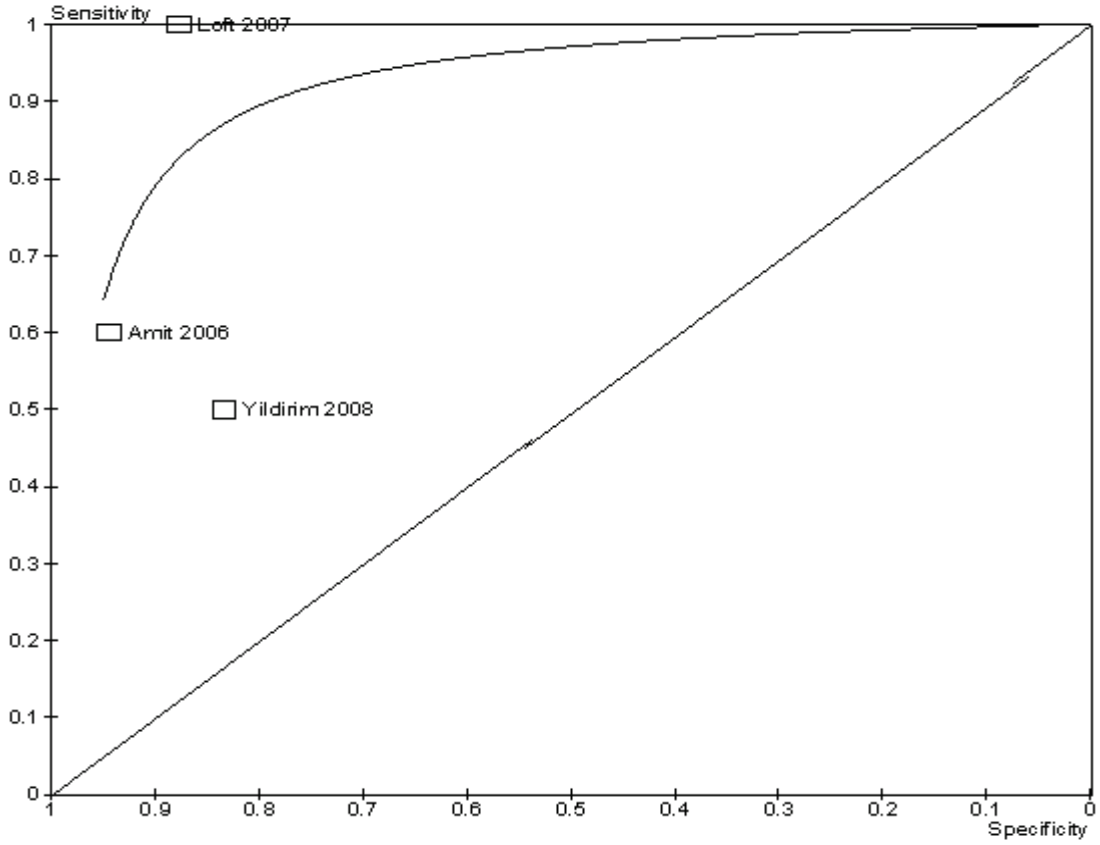


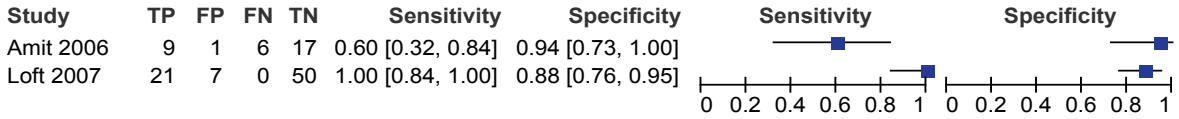
Figure 30 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer.

Figure 30. Summary ROC plot of ^{18}FDG-PET/CT v. any reference standard for the staging of cervical cancer (prospective studies)



Reference standard: histology/biopsy or clinical followup; prospective studies. Two prospective studies[31,46] totaling 111 participants provided data for a subgroup analysis of the accuracy of ^{18}FDG-PET/CT compared to histology/biopsy or clinical followup for the staging of cervical cancer. Individual 2x2 table results are presented in Figure 31. Sensitivity values were 60 percent[31] and 100 percent. [46] Specificity values were 88 percent[46] and 94 percent. [31]

Figure 31. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer



We found that when only histology/biopsy were considered as the reference standard, ^{18}FDG-PET/CT had a pooled positive LR of 7.85 (95% CI = 4.16, 14.80) and a pooled negative LR of 0.12 (95% CI = 0.00, 10.08) to accurately identify the stage of cervical cancer (Figures 32 and 33). The positive LR was statistically significant and the results were homogeneous across the studies and therefore, ^{18}FDG-PET/CT seems to be helpful to identify the stage of the disease. The negative LR was not statistically significant and the results were quite heterogeneous across the studies ($p = 0.002$; $I^2 = 90$ percent) and therefore, ^{18}FDG-PET/CT does not seem to be helpful to rule out particular stages of the disease.

Figure 32. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)

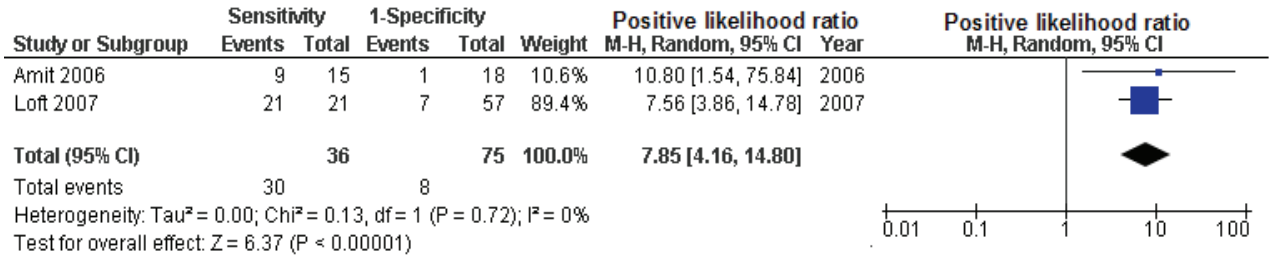


Figure 33. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)

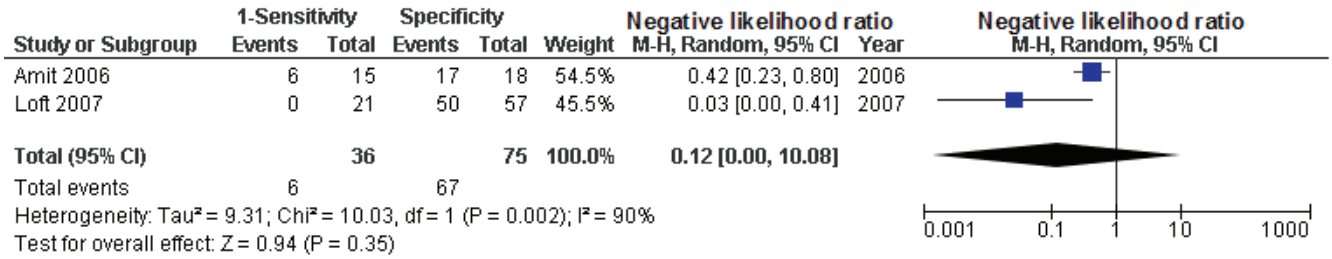
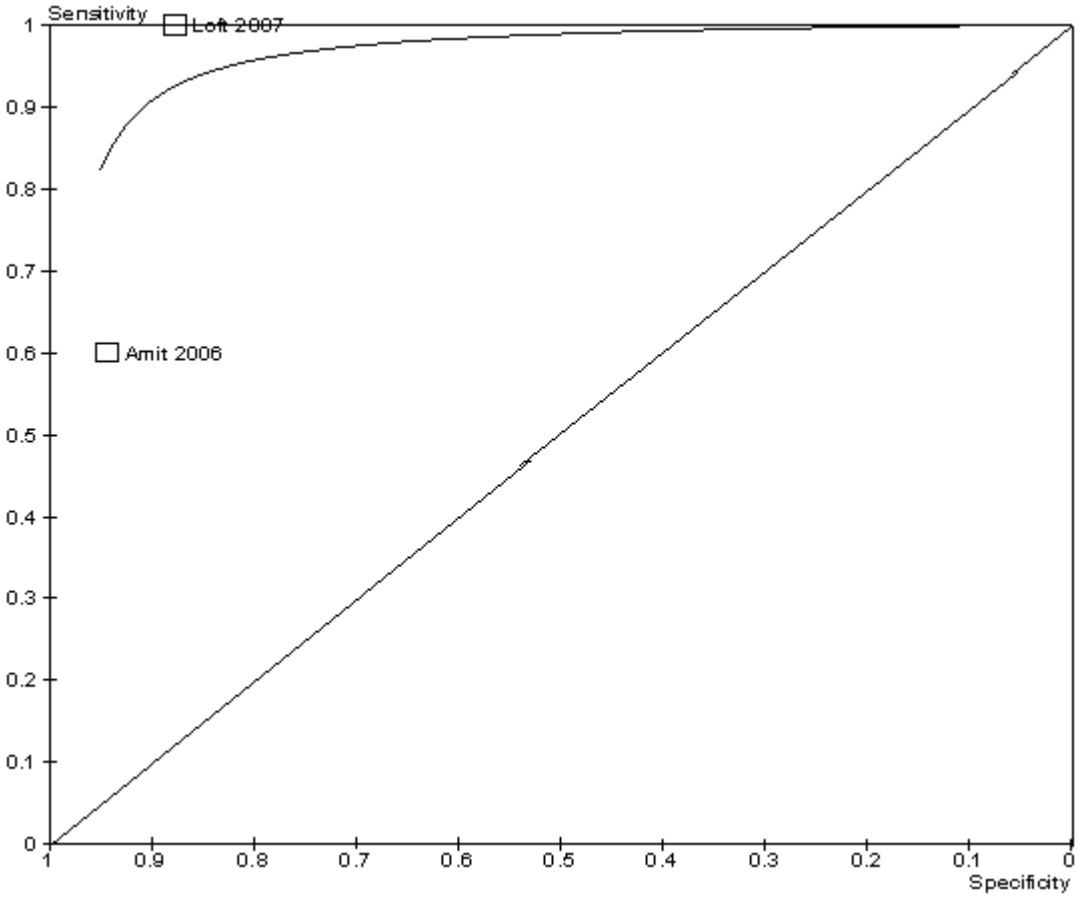


Figure 34 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer.

Figure 34. Summary ROC plot of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the staging of cervical cancer (prospective studies)



Summary of the results

Meta-analyses were calculated to evaluate the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT to detect recurrences and identify the staging of cervical cancer (Table 13). It appears that overall ^{18}FDG-PET is useful to detect or rule out recurrences, although there is some variation in the magnitude of the likelihood ratios across sites. The findings are consistent across each of the sites of recurrence in terms of being statistically significant, as well as for both prospective and retrospective designs. The consistency may provide some robustness for the overall findings. When ^{18}FDG-PET and ^{18}FDG-PET/CT were evaluated for staging purposes, we found that the values in the positive and negative LRs were similar for both techniques. Significant results were reported for the positive LR, indicating that both techniques seem to be useful to detect the stage of the disease. The results for the negative LR were not statistically significant and therefore, it appears that a negative result both in ^{18}FDG-PET and ^{18}FDG-PET/CT is not useful to identify the stage of cervical cancer. The

findings were consistent across the different reference standards and study designs (i.e., retrospective v. prospective).

Table 13. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer

PET Purpose	Type of PET	Reference standard	Design	Studies	N	Effect estimate M-H, Random, 95% CI
Recurrences	FDG-PET	Histology/biopsy or clinical followup	P	3	230	Peritoneum: PLR = 15.75 [5.99, 41.38]; NLR = 0.37 [0.22, 0.60]
						Bone: PLR = 26.56 [11.21, 62.95]; NLR = 0.22 [0.01, 3.40]
						Liver/spleen: PLR = 45.89 [14.09, 149.49]; NLR = 0.25 [0.08, 0.82]
						Lung: PLR = 33.32 [13.85, 80.14]; NLR = 0.22 [0.10, 0.48]
						MLN: PLR = 15.24 [5.63, 41.27]; NLR = 0.09 [0.02, 0.40]
						SLN: PLR = 29.06 [12.06, 70.03]; NLR = 0.19 [0.10, 0.36]
						PALN: PLR = 40.24 [11.60, 139.54]; NLR = 0.12 [0.07, 0.23]
						PLN: PLR = 41.42 [14.51, 118.25]; NLR = 0.23 [0.08, 0.73]
						ILN: PLR = 27.92 [12.00, 64.94]; NLR = 0.17 [0.05, 0.60]
			R	3	396	PLN = 5.33 [2.36, 12.05]
						NLR = 0.11 [0.04, 0.28]
Staging	FDG-PET	Any reference standard	P	5	325	PLR = 8.22 [2.59, 26.08]
						NLR = 0.38 [0.12, 1.20]
		Histology/biopsy	P	4	221	PLR = 5.79 [1.88, 17.88]
						NLR = 0.47 [0.17, 1.32]
			R	3	236	PLR = 32.90 [2.89, 375.25]
						NLR = 0.41 [0.11, 1.55]
	FDG-PET/CT	Any reference standard	P	3	127	PLR = 6.89 [3.82, 12.42]
						NLR = 0.28 [0.06, 1.38]
		Histology/biopsy or clinical followup	P	2	111	PLR = 7.85 [4.16, 14.80]
						NLR = 0.12 [0.00, 10.08]

CI = confidence interval; FDG = fluorodeoxyglucose F18; ILN = inguinal lymph node; M-H = Mantel Hantzel; MLN = mediastinal lymph node; NLR = negative likelihood ratio; P = prospective; PALN = para-aortic lymph node; PET = positron emission tomography; PLN = pelvic lymph node; PLR = positive likelihood ratio; R = retrospective; SLN = supraclavicular lymph node

3.3.2. Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making with respect to diagnosis and management strategy for patients with cervical cancer

Six studies reported on the diagnostic thinking impact of ^{18}FDG-PET[33,43,44,61] and ^{18}FDG-PET/CT. [32,38] Chang et al. [33] evaluated the treatment decision impact of ^{18}FDG-PET on disease recurrence of cervical cancer. Consecutive outpatients were prospectively enrolled between February 2001 and January 2003. A historical control group that did not undergo ^{18}FDG-PET was used for

comparison. Eligible patients had a history of histologically confirmed squamous cell carcinoma, adenocarcinoma, or adenosquamous carcinoma of the uterine cervix and had complete remission after primary treatment or salvage therapy. In addition, patients had elevated serum SCC-Ag levels greater than 2.0 ng/mL over the last 2 weeks of the study, but no evidence of recurrent disease on physical examination, Pap smear, chest X-ray, CT or MRI of the pelvis and abdomen, or histological evaluation. Patients who received cytotoxic therapy within the previous 3 months, were previously diagnosed with malignant disease other than nonmelanoma skin malignancy, or had skin or pulmonary lesions, were ineligible. The study population consisted of 27 females, with a mean age of 53.9 years (range 34.8 to 75.8). The disease stage at initial diagnosis was primarily stage I (44 percent) and II (42 percent). It should be noted that 15 of the patients in this study, who had documented relapse after PET, were also included as a subset of the population in another study by Yen et al., [61] reported below. Although these studies were conducted by the same authors and institutions and have overlap in the patient populations, different outcomes were reported and therefore both studies were included in our review. Final diagnosis was established through histological or cytological confirmation via CT or sonar-guided biopsy before treatment. Laparoscopic or exploratory surgery was performed if it was judged as potentially useful for patient management purposes. Patients with inconsistent findings underwent clinical followup for 6 months. ^{18}FDG-PET images were interpreted through visual analysis. Diagnosis and treatment decisions were made by consensus among a multidisciplinary panel.

Of the 17 patients with recurrent disease identified by ^{18}FDG-PET, seven received therapy with curative intent, four received palliative chemotherapy, and six received supportive care. One patient with negative PET findings had recurrent disease diagnosed at 2 months of clinical followup; she received palliative therapy. Forty-one percent (7/17) of patients with recurrence based on PET findings received treatment with curative intent, compared to 53 percent (16/30) of patients in the historical control group.

The authors concluded that ^{18}FDG-PET is a valuable tool for detecting recurrent disease. ^{18}FDG-PET findings enabled the selection of patients for treatment with curative intent and also avoided administering unnecessary treatment to patients with incurable disease. Finally, they concluded that ^{18}FDG-PET has the possibility of improving the survival as well as the quality of life in patients with recurrent cervical malignancies.

Overall, the quality of this study was assessed as high. Components that were well reported include selection criteria, choice and execution of reference standard, execution of index test, intermediate test results and withdrawals from the study. The period between the index test and reference standard was sufficiently short to prevent disease progression. Furthermore, a reference standard was applied to all patients, albeit not the same standard across all patients. The spectrum of patients enrolled in the study was only partially described, therefore the possibility of spectrum bias cannot be ruled out. This study is also vulnerable to review bias, as the reference standard was not blinded to the findings of the index test and also, it is unclear whether the interpretation of the index test was blinded to the reference.

A study by Lai et al. [43] examined the treatment decision impact of ^{18}FDG-PET for assessing restaging and recurrence of cervical cancer. The study population consisted of 45 females with a median age of 51 years (range 25 to 87) whose initial diagnosis was primarily stage I (33 percent) and II (50 percent). Along with a subset of patients from the study by Chang et al., [33] all patients were also included in the study by Yen et al., [61] reported below. Although these studies were conducted by the same authors and institutions, and the patient populations overlap, different outcomes were reported and therefore both are included in our review. Of the 40 patients included in the analysis, 22 (55 percent) had a change in treatment planning as a result of ^{18}FDG-PET findings. Fifteen patients had their management shifted from curative to palliative treatment, while seven continued to be treated with curative intent but had a change in their treatment field or modality. Prior to PET scanning, 23 patients planned to undergo concurrent chemotherapy and radiotherapy (CCRT), 17 planned to undergo surgery, and none were scheduled to receive treatment with palliative intent. After PET, 12 patients received CCRT, 13 received surgery and 15 received treatment with palliative intent. In addition, 11 patients underwent a guided biopsy and three patients underwent exploratory surgery due to the findings of ^{18}FDG-PET.

The authors concluded the ^{18}FDG-PET is better than CT/MRI in restaging cancer recurrence and may significantly reduce the number of unnecessary salvage attempts compared to conventional assessment. They also concluded that use of ^{18}FDG-PET in restaging allows clinicians to offer optimal management of recurrent cervical carcinoma.

The quality of this prospective study was assessed as moderate. The selection criteria were clearly described, as was the choice of reference standard and the execution of both index and reference tests. All test results and study participants were accounted for. The period between the

index and reference test was sufficiently short that it is unlikely that disease progressed between tests. The whole sample received disease verification using a reference standard; however, the same reference standard was not used for all patients, raising the possibility of verification bias. There was only a partial description of the spectrum of patients included in the study. Finally, it was unclear whether the index test or the reference standard was interpreted while blind to other test results. Due to the lack of clarity in reporting, it is uncertain to what degree review bias may have affected the study findings.

Lin et al. [44] prospectively investigated the benefit of adding ^{18}FDG-PET to the diagnostic workup in patients with histologically documented re-recurrent cervical cancer after curative salvage or unexplained elevations in tumor markers. The study sample consisted of 26 females with a median age of 56 years. The disease stage at initial diagnosis was mainly stage I (42 percent) and II (38 percent). Of the 26 patients, 24 had a second recurrence, and two had a third recurrence. The median time between salvage therapy and documented re-recurrence was 12.8 months. ^{18}FDG-PET had a positive clinical impact on 12 patients (46 percent). Among these, nine were changed from curative to palliative treatment and three had an isolated in field failure successfully resected due to ^{18}FDG-PET. In contrast, ^{18}FDG-PET led to unnecessary and invasive additional procedures, such as biopsies, in four patients. As a result of these additional procedures, ^{18}FDG-PET was stated to have had an overall negative impact in the management of two patients. The authors concluded that ^{18}FDG-PET may facilitate the selection of suitable management strategies for individual patients with re-recurrent cervical cancer.

This prospective study was assessed as being of high quality. Both the selection criteria and choice of reference standard were clearly described, and all test results and study participants were accounted for. The period between the index and reference test was short enough that it is unlikely that disease progressed between tests. The whole sample received disease verification using a reference standard; however, the same reference standard was not used for all patients, raising the possibility of verification bias. The execution of the reference test was well described and the execution of the index test was partially described. There was insufficient detail in reporting the recruitment of patients into the study; therefore it is uncertain whether spectrum bias may have occurred. Finally, the index test results were interpreted without knowledge of the results of the reference standard; however, it is unclear whether the interpretation of the reference standard was

also blinded. Therefore, it remains unclear whether review bias may have affected the results of this study.

Yen et al. [61] investigated the treatment decision impact of ^{18}FDG-PET for assessing the recurrence of cervical cancer. A total of 55 patients were enrolled from two separate prospective studies examining the role of ^{18}FDG-PET in cervical cancer patients. There were 40 patients who had documented treatment failure (Lai et al. [43]) and 15 patients with unexplained elevated tumor marker squamous cell carcinoma antigen or carcinoembryonic antigen serum level (Chang et al. [33]). The median age of patients was 51 years (range 25 to 86). Forty-five percent of patients had initial stages of Ib or IIa, while 55 percent had stages between IIb and IVa.

Of the 55 enrolled patients, 36 (65 percent) had their treatment plans modified based on the findings of ^{18}FDG-PET, while 19 (35 percent) were treated according to their pre-PET plan. Among these 36 patients, nine (25 percent) had treatment that remained curative in intent although the field or modality of radiation changed, while 27 (75 percent) received palliative therapy. Three of the nine patients whose treatment was changed were downstaged.

A prognostic scoring system was used to categorize patients as having low, moderate or high risk of mortality. Based on the findings of ^{18}FDG-PET, 10 patients in the low-risk group were changed to palliative treatment, while 17 stayed at curative treatment (seven with changes in treatment plan, 10 with no changes). In the intermediate-risk group, 12 patients were changed to palliative care, and seven stayed on curative therapy (two with changes in treatment plan, five with no changes). Five patients were changed to palliative treatment in the high risk group, whereas one patient stayed on curative treatment.

The authors concluded that ^{18}FDG-PET is useful in the management of recurrent cervical cancer because it allows for a more precise restaging than CT or MRI. ^{18}FDG-PET may offer maximal benefit by identifying patients suitable for therapy with precise restaging information.

This prospective study was assessed to be of moderate quality. The selection criteria and choice of reference standard were clearly described, and all test results and study participants were accounted for. The time interval between the index and reference test was sufficiently brief that it is unlikely that disease progressed between tests. The whole sample received disease verification using a reference standard; however, the same reference standard was not used for all patients, raising the possibility of verification bias. The execution of both the index and reference standard was only partially described. In addition, it was unclear whether there was blind interpretation of the tests,

which may have introduced review bias. Finally, there was lack of clarity in the description of how patients were recruited into the study, creating the potential for spectrum bias.

The study by Bjurberg et al. [32] evaluated the treatment decision and diagnostic testing impact of ^{18}FDG-PET for assessment of staging and restaging of cervical cancer. This prospective study enrolled 42 patients with biopsy-proven cervical cancer that were included in three sub-groups for analysis: 1) early disease, followup after surgical treatment (n = 10), FIGO stage Ia:2-IIa; 2) locally advanced disease scheduled for radical radiotherapy (n = 17), FIGO stage Ib:2-IVb; and 3) relapsing disease (n = 15). The mean age of the patients in the three groups was 38.8 years, 55.5 years and 50.3 years, respectively. Changes in treatment management were reported in groups 2) and 3). Of the 17 patients in group 2), four had their treatment strategy changed following ^{18}FDG-PET detection of new metastases (24 percent). In the 15 patients from group 3), three cases were assessed as benign by ^{18}FDG-PET imaging. Subsequent followup testing verified that ^{18}FDG-PET had correctly identified the patients to be free from recurrent disease. ^{18}FDG-PET led to a change in the treatment plan for three of the 12 patients assessed as positive for recurrent disease (25 percent). Additional diagnostic testing was performed in six of the 12 recurrent cases.

The authors concluded that ^{18}FDG-PET provided detail about the extent of disease, which contributed to the restaging and appropriate management of the patients with locally advanced or recurrent cervical cancer. However, they felt that there was no added value to patient management when ^{18}FDG-PET was used in followup for patients with early stage disease. They based this statement on their current results, and pointed out that the followup period was short and the number of patients was very small.

This prospective study was assessed to be of moderate quality. The selection criteria and choice of reference standard were clearly described, and all test results and study participants were accounted for. The whole sample received disease verification using a reference standard; however, the same reference standard was not used for all patients, raising the possibility of verification bias. The time interval between the index and reference test was not clearly reported, and the execution of both the index and reference standard was only partially described. It was unclear whether there was blind interpretation of the tests, which may have introduced review bias.

In a retrospective review of the medical records of 52 women, Chung et al. [38] examined the treatment decision impact of ^{18}FDG-PET/CT on assessing the recurrence of cervical cancer. The mean age of patients was 53 years (range 32 to 77), with primarily stage I (50 percent) and II (40

percent) cancer. Treatment management was altered on the basis of ^{18}FDG-PET/CT findings for 12 patients (23 percent). In three patients, previously unplanned treatment was initiated, while in five changes were made to the previously planned therapeutic approach. For the remaining three patients, previously planned diagnostic procedures were not required. In addition, ^{18}FDG-PET/CT provided valuable information for 12 patients by identifying the exact location of lymph nodes (five patients), showing precise location of pelvic wall or bone infiltration (five patients), and the exact location of distant metastases. In nine patients, ^{18}FDG-PET/CT guided additional invasive diagnostic procedures.

The authors also reported the prognostic outcomes of patients undergoing ^{18}FDG-PET/CT. The 2-year disease-free survival rate of patients who had negative ^{18}FDG-PET/CT results was significantly better than for patients who tested positive on ^{18}FDG-PET/CT (85 percent v. 10.9 percent, $p = 0.002$).

The authors concluded that ^{18}FDG-PET/CT provided good anatomical and functional localization of suspicious lesions. The superior diagnostic interpretation of ^{18}FDG-PET/CT had a positive impact on clinical management, treatment planning and on disease-free survival rate. The quality of this retrospective study was rated moderate. The greatest methodological weaknesses included only partial reporting of the spectrum of patients enrolled and lack of clarity on the period between the index and reference standard. Although the whole sample received disease verification using a reference standard, the same reference standard was not used for all patients, raising the possibility of verification bias. The selection criteria and choice of reference standard were clearly described, and all test results and study participants were accounted for.

3.3.3. ^{18}FDG-PET and ^{18}FDG-PET/CT as part of a management strategy in cervical cancer

Two studies[33,43] evaluated the impact of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes. Chang et al. [33] compared the mean overall survival between a group that underwent ^{18}FDG-PET and a historical control group. The followup times were similar for both groups (11.9 months and 14 months, respectively). Characteristics of the population and ^{18}FDG-PET have been described above. Compared to the historical control group, overall survival improved for patients who had ^{18}FDG-PET as part of their diagnostic work-up (^{18}FDG-PET group = 22.0 months [95% CI: 17.3 to 26.7 months]; historical cohort: 12.7 months [95% CI: 7.9 to 17.5; $p = 0.02$]. This

was an observational study; therefore, reliable conclusions cannot be made regarding the effectiveness of ^{18}FDG-PET as part of a management strategy to improve patients' overall survival. It is unknown whether factors other than the exposure to the intervention (i.e., ^{18}FDG-PET) are equally distributed among the groups.

The study by Lai et al. [43] compared the 2-year overall survival rate for patients who underwent ^{18}FDG-PET as part of their diagnostic work-up compared to a group of patients who did not undergo disease restaging with PET. Characteristics of the population and ^{18}FDG-PET have been described above. All seven patients who continued with curative treatment, but had their treatment field altered, remained alive. In the primary surgery group, the 2-year overall survival rate was significantly greater among the ^{18}FDG-PET group compared to the historical cohort whose disease was restaged without ^{18}FDG-PET (hazard ratio: [HR]: 0.21; 95% CI: 0.05-0.83; p = 0.02). Among patients receiving primary RT or CCRT, there was no difference between the two groups in 2-year overall survival (HR: 0.99; 95% CI: 0.53, 1.85; p = 0.99). The authors reported that clinical characteristics were similar for study participants and historical control patients in the primary surgery group and they suggested that it was unlikely that the observed benefit in overall survival probably were due to other prognostic factors. This was an observational study and therefore, conclusions that can be made about the effectiveness of ^{18}FDG-PET on patient-survival are limited.

Summary of the results

There were six studies of moderate to high methodological quality that evaluated the use of ^{18}FDG-PET[33,43,44,61] or ^{18}FDG-PET/CT[32,38] to assess the staging/restaging[32,43] or recurrence[33,38,44,61] of cervical cancer. The findings were generally positive with authors concluding that both ^{18}FDG-PET and ^{18}FDG-PET/CT assisted in guiding the management strategy by allowing for more precise restaging. Notably, the use of ^{18}FDG-PET and ^{18}FDG-PET/CT most often altered the management course from curative to palliative care, thus avoiding unnecessary treatment. Further, two prospective studies[33,43] found improved survival among patients who had ^{18}FDG-PET as part of their management strategy.

Table 14 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer.

Table 14. Main findings and types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for cervical cancer

Study	Results of FDG-PET imaging on patient diagnosis, treatment and outcomes	Types of bias
Bjurberg 2007[32] Study type: Prospective	**Management decision:** Treatment and diagnostic testing impact Treatment strategy changed due to identification of new metastasis for 4/17 cases (24%) PET did not confirm clinical suspicion of recurrence. PET deemed to be true negative on followup 3/15 cases; Treatment strategy changed for 3/12 positive recurrence cases (25%) Additional diagnostic testing occurred in 6/12 positive recurrence cases	Spectrum bias (unclear) Selection bias (unclear) Verification bias (>1 RS) Review bias (PET; unblinded; RS, unblinded)
Chang 2004[33] Study type: Prospective	**Management decision:** Treatment Treatment plan changed in 17/27 cases (63%) Curative therapy (n = 7) Palliative chemotherapy (n = 4) Supportive care (n = 6) 7/18 (39%) patients with recurrence received curative therapy based on PET, compared to 53% (16/30) in historical control group **Patient-centered outcomes:** Mean overall survival PET group: 22 mo. (95% CI: 17.3, 26.7) v. historical control: 12.7 mo. (95% CI: 7.9,17.5)	Spectrum bias (unclear) Verification bias (>1 RS) Review bias (PET; unclear if blinded; RS, unblinded)
Chung 2007[38] Study type: Retrospective	**Management decision:** Treatment and diagnostic testing impact Treatment management change in 12 patients (23%): Initiated previously unplanned treatment (n = 4), Changed previously planned therapeutic approach (n = 5) Eliminate previously planned diagnostic procedure (n = 3) PET/CT guided additional invasive diagnostic procedures (n = 9).	Spectrum bias (unclear) Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS) Review bias (RS, unclear)
Lai 2004[43] Study type: Prospective	**Management decision:** Treatment and diagnostic testing impact Treatment plan change in 22/40 patients (55%): Shifted from curative to palliative treatment (n = 15), Curative treatment continued; altered treatment field or modality (n = 7) Diagnostic testing impact due to PET findings in 14 patients: Additional guided biopsy (n = 11); exploratory surgery (n = 3) **Patient-centered outcomes** Patients treated with altered treatment field remained alive (n = 7). Primary surgery group, had a significant 2-yea overall survival rate in PET group compared to those restaged without PET. Patients receiving primary RT or CCRT had no significant differences among the two groups.	Spectrum bias (unclear) Verification bias (>1 RS) Review bias (PET and RS; unclear if blinded)
Lin 2006[44] Study type: Prospective	**Management decision:** Treatment PET had positive clinical impact on 12/26 patients (46%); Changed from curative to palliative treatment (n = 9), Isolated in field failure successfully resected due to PET (n = 3) PET led to unnecessary and invasive additional procedures (e.g., biopsies) (n = 4) PET stated to have had overall negative impact in management (n = 2)	Spectrum bias (unclear) Verification bias (>1 RS) Review bias (RS; unclear if blinded)
Yen 2004[61] Study type: Prospective	**Management decision:** Treatment Treatment plans modified based on PET in 36/55 patients (65%): Field or modality of radiation changed (n = 9) Changed from curative to palliative therapy (n = 27)	Spectrum bias (unclear,) Verification bias (>1 RS) Review bias (PET and RS; unclear if blinded)

CCRT = concurrent chemotherapy and radiotherapy; CI = confidence interval; FDG = fluorodeoxyglucose F18; mo = months; PET = positron emission tomography; RS = reference standard; RT = radiotherapy

4. Kidney Cancer

4.1. Background

Approximately 54,390 new cases of kidney and renal pelvis cancer will be diagnosed in the United States during 2008; 60 percent of these cases will occur in men. [133] The National Cancer Institute indicates that kidney cancer has been increasing at a rate of 2 percent per year for the last 65 years. Mortality has also increased, but to a lesser degree than incidence, during this same time period. [133] An estimated 13,101 deaths will be caused by kidney cancer in 2008. [133] Native-Americans suffer the highest rates of kidney cancer (20.9 and 10.0 cases per 100,000 for men and women respectively); however, African-Americans also show higher rates of cancer incidence than what is observed among Caucasian-Americans. [146] The median age of onset for renal cancer is 65 years. [133] In Europe, 5-year survival is 71 percent in individuals under 45 years of age where kidney cancer is rare but decreases to 45 percent in patients over 74 years in Europe. [147]

The early diagnosis of renal carcinoma is hampered by the observation that tumors can grow quite large before the patient exhibits any symptoms. Pain, hematuria and flank masses have been traditional symptoms but these appear in only 9 percent of patients and are often indicative of advanced disease. [147] Hyprochromic anaemia, fever, cachexia, fatigue, and weight loss may also be symptomatic of renal carcinoma. [147] Approximately 30% of patients have with metastatic disease. [147]

Specific early screening programs for kidney cancer are not realistic as the populations that would be targeted are too large and currently there is no evidence to support a population-based screening program. [147] Small renal masses are detected through routine imaging with increasing frequency, but imaging is not specific enough to accurately discriminate between benign and malignant tumors. Local symptoms are considered the best predictive tool in determining malignancy. [146]

CT seems to be an effective tool in the staging of renal carcinomas with a sensitivity of 90 percent for small tumors and 95 percent for tumors larger then 3 centimeters. [147] Ultrasonography is often used as well, with sensitivity of 60 percent for detecting small tumors and 85 percent for larger tumors. [147] MRI has not been shown to be effective in characterising tumors in patients with renal carcinoma but may still be employed to provide information about the tumor involvement with the vena cava or when surgical removal of tumors is being planned. [147] Staging of renal carcinoma is commonly done using the Robson classification scheme within the United States. While this staging

system is uncomplicated, the main weakness is that it combines stages that may have widely varied survival prognoses. The Robson stages of renal carcinoma are outlined below in Table 15.

Table 15. Robson stages of renal cancer

Stage	Description
I	Renal carcinoma is localized to the kidney only
II	Cancer extends to renal capsule but is confined to the Gerota's fascia
III	Tumor is associated with the inferior vena cava or renal vein (stage IIIa) or local hiliar lymph notes (stage IIIb)
IV	Cancer has spread to other local organs or distant sites

Taken from Corgna et al. [147]

The TNM classification offers more complete stratification of patients and a more accurate assessment of their prognosis. In TNM classification, the T refers to the tumor size and whether or not it has spread to adjacent tissues, the N represents whether or not there has been spread to the lymph nodes, and the M indicates whether the cancer has metastasised.

Surgical resection is the primary method of curative therapy for kidney cancer. Radical nephrectomy is the main operation performed, but new organ-sparing approaches such as laparoscopic nephrectomy are used increasingly. [146,147] Palliative surgery is also frequent. [147] There are no standard chemotherapy or immunotherapy treatments for renal cell carcinoma. [147] Radiation therapy in patients with metastatic disease may resolve symptoms in some patients. [147]

4.2. Importance of Key Questions in the Clinical Management of Kidney Cancer

Survival is very much related to the stage of cancer when it is diagnosed. The ability to detect and characterize renal masses more accurately and to stage malignant renal tumors is crucial for the management of patients. The identification of a single or several metastatic lesions can lead to differing therapeutic approaches (e.g., surgery or systemic treatment). If detected early, renal tumors can be treated with alternatives other than standard radical nephrectomy, such as minimally invasive surgery and partial nephrectomy. Although solid renal masses are considered malignant tumors, benign solid renal masses are not uncommon. Morphological imaging methods present several diagnostic problems in differentiating between benign and malignant solid renal tumors. They also show some limitations in evaluating kidney cancer with regard to local spread and distant disease. Improving the diagnostic yield of these investigations while precluding the need for obtaining a tissue diagnosis would have obvious implications in management. The role of ^{18}FDG-PET in the diagnosis, staging, and management of kidney cancer has not been clearly defined.

4.3. Results

Eight studies[64-71] provided evidence on the use of ^{18}FDG-PET. We did not find studies that reported on the use of ^{18}FDG-PET/CT for kidney cancer. All eight studies evaluated the diagnostic accuracy of ^{18}FDG-PET for kidney cancer. Three studies[67,69,70] reported on the diagnostic thinking impact of ^{18}FDG-PET. None of the studies evaluated the impact of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. No economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for kidney cancer were identified. Characteristics of the populations, conditions of ^{18}FDG-PET administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

4.3.1. Diagnostic accuracy of ^{18}FDG-PET in kidney cancer

Characteristics of the studies

Eight studies (three prospective, [64,65,67] five retrospective[66,68-71]) evaluated the diagnostic accuracy of ^{18}FDG-PET. One study used ^{18}FDG-PET for initial staging, [67] one for primary diagnosis, [65] one for restaging, [68] one for initial staging and recurrences, [71] and four used ^{18}FDG-PET for both primary diagnosis and initial staging. [64,66,69,70] The studies contained a total of 250 patients with sample sizes ranging from 15 to 66 participants. Participant ages ranged from 23 to 87 years. In five studies ^{18}FDG-PET was compared to histology/biopsy or clinical followup as the reference standard; [64,67-70] in the three remaining studies histology/biopsy was used exclusively as the reference standard. [65,66,71] One study reported the mean time between last treatment and ^{18}FDG-PET as 3 to 24 months. [68] Three studies used a fixed dose of ^{18}FDG (1.5 mCi, [67] 395.9 MBq, [71] 370 MBq[66]); two studies used a weight-based dose (2.516 to 5.2 MBq/kg[70] and 2 MBq/kg[64]); two studies reported a dose range (370 to 444 MBq[65] and 370 to 555 MBq[68]); and one study did not report on dosing. [69] The time between ^{18}FDG injection and PET scan was 45 minutes, [69] 50 minutes, [66] 60 minutes, [64,65,67,70] or ranged between 45 to 60 minutes. [68,71] Patients fasted for four hours, [67,70] six hours, [64-66] and overnight. [71] Three studies measured maximum glucose levels of 135 mg/Dl, [65] 140 mg/dL[70] and 150 mg percent[66] before administration of ^{18}FDG-PET. Methods of interpretation were qualitative in five studies[64,65,68,69,71] and both qualitative and quantitative in two studies. [66,70] Scans were interpreted qualitatively using visual analysis in all studies. [64-66,68-71] Both studies[66,70] used SUV

values for the quantitative interpretation of the PET images with one study reporting the criterion for abnormality as an SUV > 2.5 g/mL. [66]

Comparisons

Comparisons for which data were considered for meta-analysis are summarized in Table 16. Pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET for the primary diagnosis and staging of kidney cancer. Individual study data are summarized in Appendix D.

Table 16. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for kidney cancer

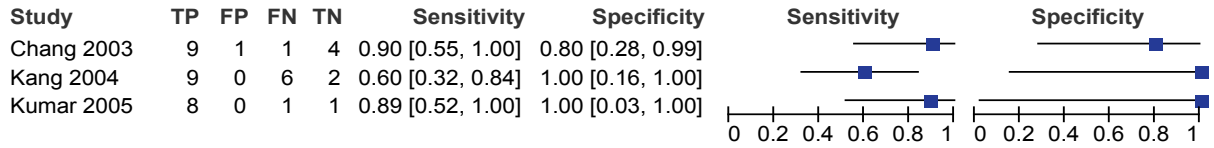
Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Primary diagnosis and staging	Aide 2003[64]	P	FDG-PET	Histology/biopsy or clinical followup	1. FDG-PET v. any reference standard (R studies) [66,69,70] 2. FDG-PET v. histology/biopsy or clinical followup (R studies) [69,70]
	Chang 2003[66]	R	FDG-PET	Histology/biopsy	
	Kang 2004[69]	R	FDG-PET	Histology/biopsy or clinical followup	
	Kumar 2005[70]	R	FDG-PET	Histology/biopsy or clinical followup	
Staging and recurrences	Majhail 2003[71]	R	FDG-PET	Histology/biopsy	No
Primary diagnosis	Ak 2005[65]	P	FDG-PET	Histology/biopsy	No
Restaging	Jadvar 2003[68]	R	FDG-PET	Histology/biopsy or clinical followup	No
Staging	Dilhuydy 2006[67]	P	FDG-PET	Histology/biopsy or clinical followup	No

FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

1. ^{18}FDG-PET for the primary diagnosis and staging of kidney cancer

Reference standard: any; retrospective studies. A meta-analysis of retrospective studies was conducted to evaluate the accuracy of ^{18}FDG-PET for the primary diagnosis and staging of kidney cancer. Three retrospective studies[66,69,70] totaling 42 participants compared ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer. Individual 2x2 table results are presented in Figure 35. Sensitivity ranged from 60 percent[69] to 90 percent. [66] Specificity ranged from 80 percent[66] to 100 percent. [69,70]

Figure 35. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer



We found that ^{18}FDG-PET had a pooled positive LR of 3.95 (95% CI = 1.14, 13.73) and a pooled negative LR of 0.30 (95% CI = 0.12, 0.79) to help in the diagnosis and staging of kidney cancer (Figures 36 and 37). Both the positive and negative LRs were statistically significant and the results were fairly homogeneous across the studies. Therefore, ^{18}FDG-PET seems to be helpful in the primary diagnosis and detection of staging of kidney cancer.

Figure 36. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies)

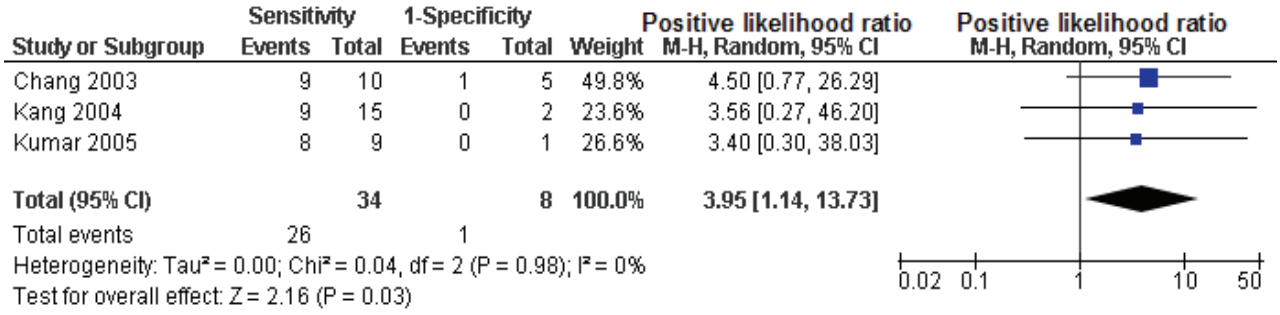


Figure 37. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies)

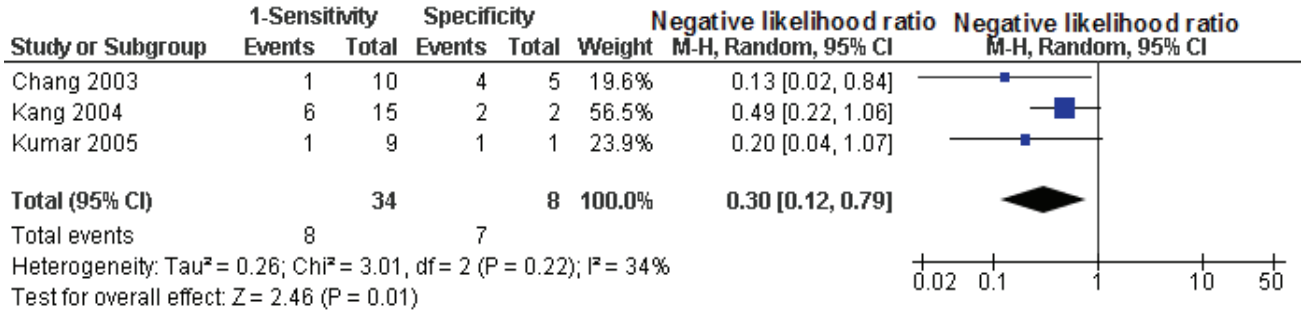
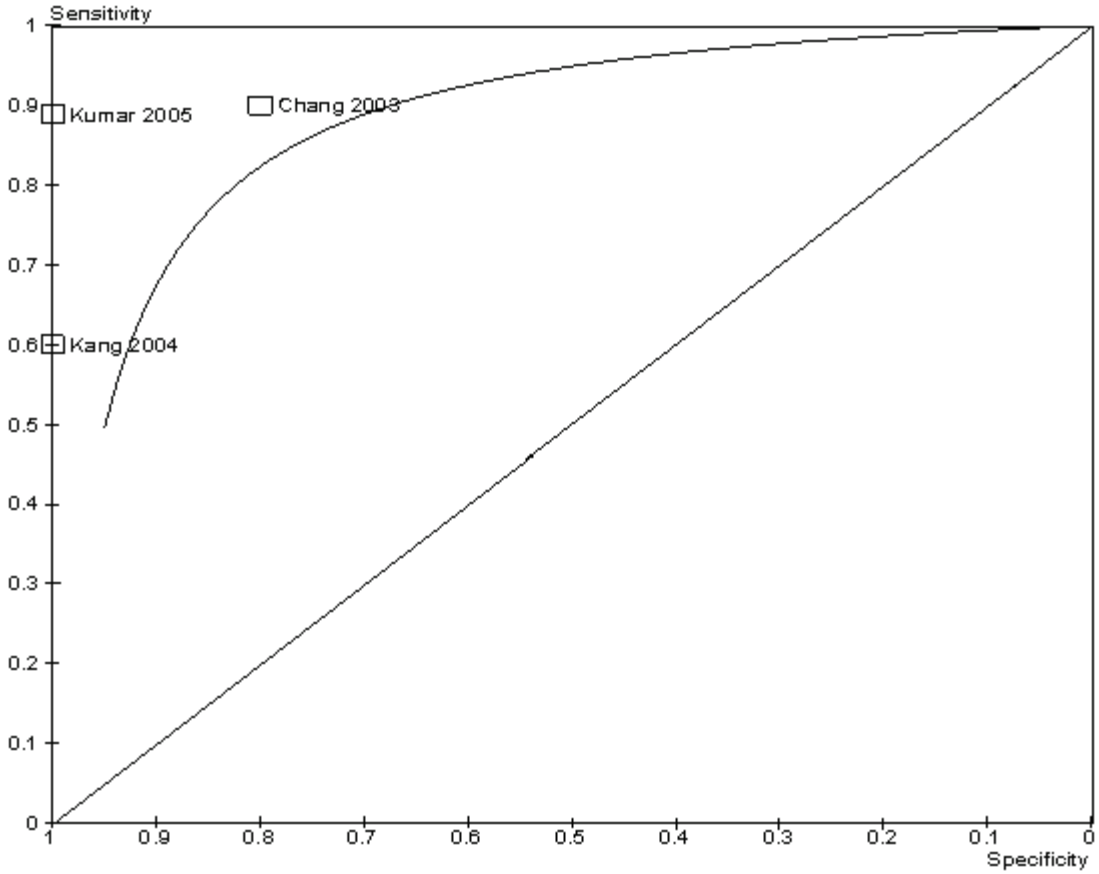


Figure 38 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer based on retrospective studies.

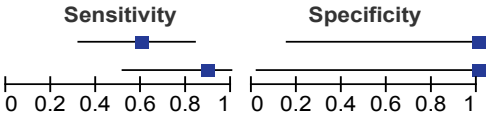
Figure 38. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the primary diagnosis and staging of kidney cancer (retrospective studies)



Reference standard: histology/biopsy or clinical followup; retrospective studies. Two retrospective studies[69,70] totaling 27 participants provided data for a subgroup analysis of the accuracy of ^{18}FDG-PET when histology/biopsy or clinical followup were used as the reference standard for the primary diagnosis and staging of kidney cancer. Individual 2x2 table results are presented in Figure 39. Sensitivities were 60 percent[69] and 89 percent; [70] specificity was 100 percent in both studies.

Figure 39. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Kang 2004	9	0	6	2	0.60 [0.32, 0.84]	1.00 [0.16, 1.00]
Kumar 2005	8	0	1	1	0.89 [0.52, 1.00]	1.00 [0.03, 1.00]



When histology or biopsy was considered as the reference standard, ^{18}FDG-PET had a pooled positive LR of 3.48 (95% CI = 0.60, 20.15) and a pooled negative LR of 0.42 (95% CI = 0.21, 0.84) to help in the diagnosis and staging of kidney cancer (Figures 40 and 41). The positive LR was not statistically significant and therefore, is inconsistent with a previous meta-analysis of three studies (see Figure 38), suggesting that no firm conclusions can be drawn. The negative LR was statistically significant and homogeneous across the studies and therefore, it can be said that ^{18}FDG-PET may be useful to rule out a diagnosis of kidney cancer.

Figure 40. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies)

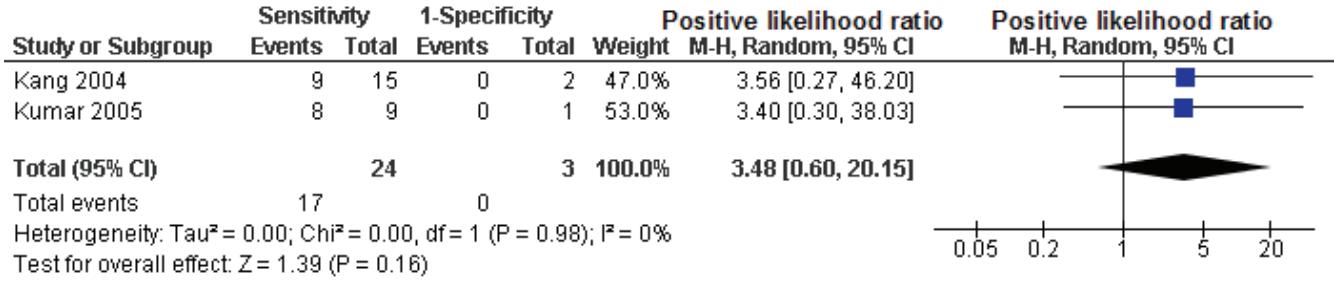


Figure 41. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies)

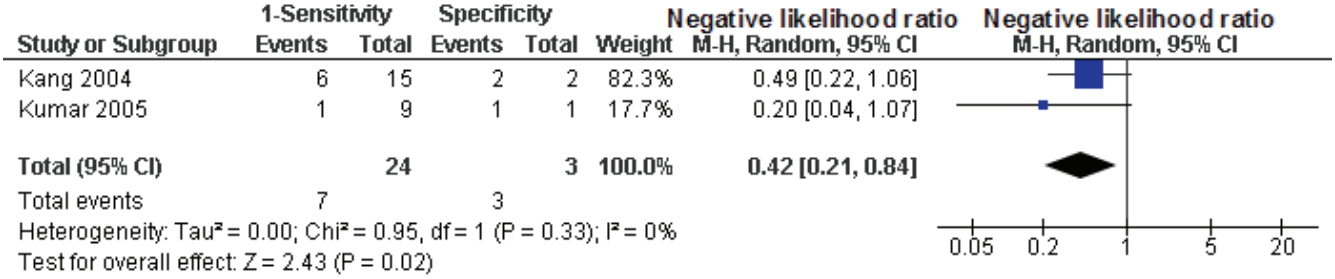
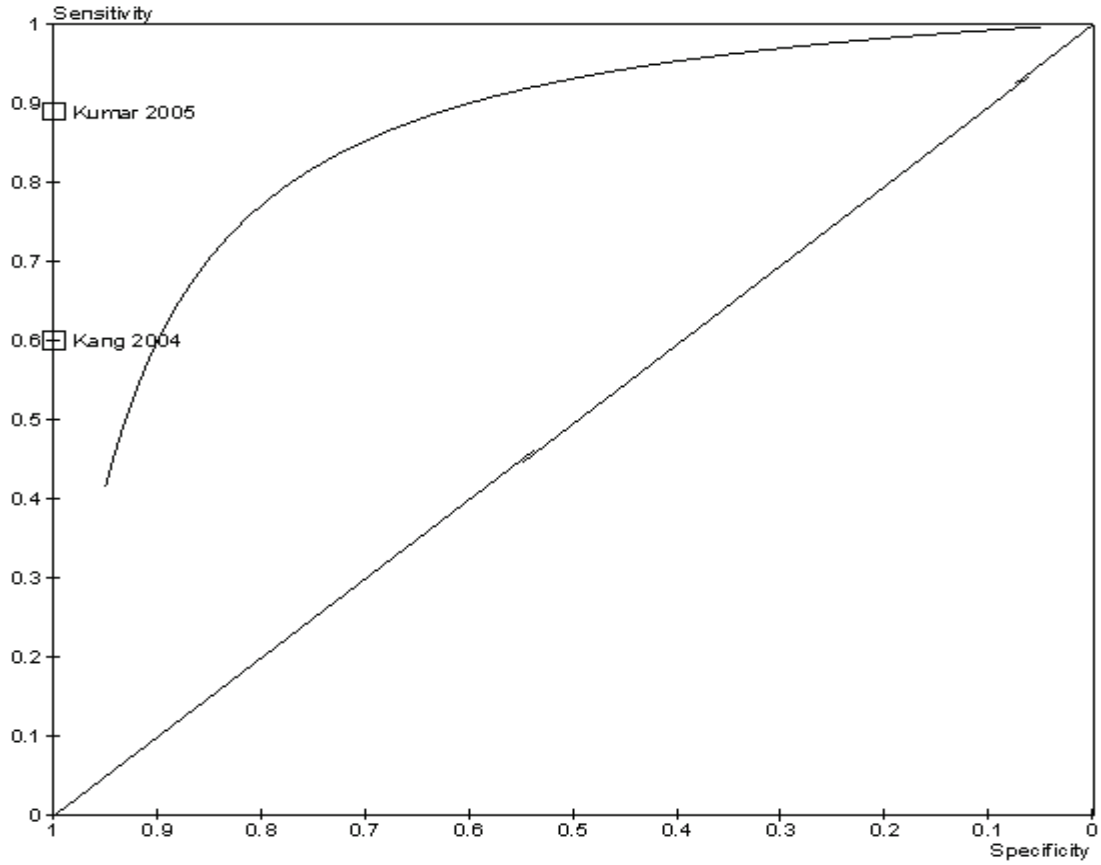


Figure 42 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer based on retrospective studies.

Figure 42. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of kidney cancer (retrospective studies)



Summary of the results

Meta-analyses were calculated to evaluate the accuracy of ^{18}FDG-PET for the diagnosis and staging of kidney cancer (Table 17). When ^{18}FDG-PET was compared against any reference standard, retrospective studies reported a statistically significant result for the positive LR. However, when a subset of these studies, which used histology/biopsy or clinical followup as the reference standard, were considered, the pooled results were not statistically significant. Since the non-significant positive LR in the latter meta-analysis may be due to insufficient sample size rather than the utility of the test, no conclusions can be drawn regarding the ability of ^{18}FDG-PET to rule in a diagnosis or stage of kidney cancer. The pooled negative LRs were similar and statistically significant in both meta-analyses, suggesting that ^{18}FDG-PET may be useful to rule out a diagnosis of kidney cancer or particular stages of the disease.

Table 17. Results of meta-analyses of the accuracy of ^{18}FDG-PET for kidney cancer

PET Purpose	Type of PET	Reference standard	Design	Studies	N	Effect estimate M-H, Random, 95% CI
Primary diagnosis and staging	FDG-PET	Any reference standard	P	3	42	PLR = 3.95 [1.14, 13.73]
						NLR = 0.30 [0.12, 0.79]
			R	2	27	PLR = 3.48 [0.60, 20.15]
						NLR = 0.42 [0.21, 0.84]

CI = confidence interval; FDG = fluorodeoxyglucose F18; M-H = Mantel Hantzel; NLR = negative likelihood ratio; P = prospective; PET = positron emission tomography; PLR = positive likelihood ratio; R = retrospective

4.3.2. Diagnostic thinking impact of ^{18}FDG-PET on physician decision making with respect to diagnosis and management strategy for patients with kidney cancer

Three studies[67,69,70] evaluated the use of ^{18}FDG-PET on physician decision-making with regard to patient management and diagnostic work-up of renal cell carcinomas (RCC). The studies considered patients with suspected but undiagnosed primary RCCs and patients with recurrent or metastatic disease. The imaging by ^{18}FDG-PET was used for both initial diagnostic and staging purposes.

Dilhuydy et al. [67] investigated the treatment decision impact of ^{18}FDG-PET imaging on the restaging and management of patients suffering from RCC with metastatic disease. Participants included 24 patients who underwent a total of 26 PET scans. Overall, there were five changes (21 percent) to the management strategy. Evaluation of the changes to clinical management was subdivided by the type of assessment the patients were undergoing. There were 20 ^{18}FDG-PET scans in patients to assess limited or solitary tumor sites. Of these, the treatment plan was modified after ^{18}FDG-PET in only three patients (15 percent). Additionally, five ^{18}FDG-PET scans were performed in patients who appeared to have had a complete response to treatment. Of these, two scans were positive for ^{18}FDG uptake, prompting a change in therapeutic management. Thus, the impact of the ^{18}FDG-PET imaging appeared to be greater in the assessment of patients thought to have complete response following treatment. However, the number of patients in this group was small, so this result should be interpreted with caution. The five changes resulting from the ^{18}FDG-PET imaging were: from observation to surgery (n = 2) or immunotherapy (n = 2) and from surgery to immunotherapy (n = 1).

The authors concluded that positive ^{18}FDG-PET images may lead to modification of the treatment decisions made; however, negative ^{18}FDG-PET results should not alter treatment planning.

Particular value of ^{18}FDG-PET imaging was found in the identification of distant metastatic sites, justifying the addition of complementary treatment in addition to surgery.

This retrospective study was determined to be of moderate quality. The choice and administration of the independent reference standard were well reported. Additionally, equivocal results were reported, and there was satisfactory explanation for withdrawals. However, given the retrospective nature of the study, interpretation of the reference standard was not blinded, which may have introduced review bias. Furthermore, there was an incomplete description of the spectrum of included patients and the inclusion criteria, making it difficult to rule out selection bias. There was more than one reference test, which may have introduced verification bias.

Kang et al. [69] evaluated the accuracy of ^{18}FDG-PET imaging on a mixed population of patients undergoing initial diagnosis and staging or restaging of RCC. The subsequent impact on treatment decisions and diagnostic workup was assessed. This was a retrospective review of 66 consecutive patients who underwent 90 ^{18}FDG-PET scans. The sample included two types of patients: those with suspicion of primary RCC who had not undergone nephrectomy (n = 17, 17 scans); and for restaging of patients with RCC who had undergone nephrectomy (n = 54, 73 scans). A total of 17 patients underwent multiple ^{18}FDG-PET scans. The treatment plan was revised in 12 cases (13 percent) of the total 90 scans in this study. There was minor impact on the additional diagnostic work-up in one case in which the ^{18}FDG-PET scan led to the order for an abdominal MRI to confirm the presence of a primary RCC. Changes made in treatment plans included two cases in which surgery was indicated as a result of ^{18}FDG-PET imaging. Additionally, in nine cases the ^{18}FDG-PET analyses lead to reinterpretation of conventional imaging. Within the subgroup of 17 patients with no history of nephrectomy, two were accurately identified as having benign cysts by ^{18}FDG-PET, however, 6/15 (40 percent) disease positive individuals were not captured by ^{18}FDG-PET imaging, yielding to a lower sensitivity than conventional CT imaging. Of the patients with a history of disease who had undergone nephrectomy, ^{18}FDG-PET detected 64 percent of all soft tissue metastasis and 79 percent of bone metastasis. For 87 of the 90 ^{18}FDG-PET studies, there was at least one associated conventional image available (e.g., CT scan). When compared to the associated conventional images, ^{18}FDG-PET studies showed a lack of sensitivity for detection of metastatic lesions. ^{18}FDG-PET imaging failed to identify all lesions detected by conventional imaging in 39 scans (45 percent). However, ^{18}FDG-PET images did identify previously unknown lesions in 11 scans (13 percent).

The prognostic value of ^{18}FDG-PET imaging was assessed by following the progression of metastatic lesions present on ^{18}FDG-PET imaging prior to immunotherapy. Of 31 lesions which progressed, 25 (81 percent) had been positively identified on the initial ^{18}FDG-PET scan. There were 42 lesions that remained stable, of which only 28 (67 percent) had been positive on the initial ^{18}FDG-PET scan.

Overall, the authors concluded that although ^{18}FDG-PET imaging was more specific than conventional imaging, its use was limited by its low sensitivity for detecting RCC. It was thought that ^{18}FDG-PET imaging holds value as a complementary tool, particularly in suspicious or equivocal cases.

This retrospective study was assessed as being of moderate quality. The spectrum of patients and selection criteria, the choice and administration of the independent reference standard, and intermediate test results were all well reported. Additionally, all cases were verified by a reference standard and there were no withdrawals. There was inadequate reporting on some aspects, including detail about the execution of the ^{18}FDG-PET scan and reference tests. Due to the retrospective design of the study, the reference standard did not undergo blind interpretation, thus leading to the possibility of review bias. There was no one reference standard; rather a combination of methods was used (histological or clinical followup), which may have introduced verification bias in the validation of true disease status.

Kumar et al. [70] retrospectively evaluated the impact of ^{18}FDG-PET imaging on a mixed population of patients with suspected or known RCC who were undergoing assessment for diagnosis and staging of their disease. The impact on subsequent treatment management was assessed. Twenty-four patients who underwent ^{18}FDG-PET imaging were included in this analysis. In the 24 patients, a total of 28 solid renal masses were assessed. There were 10 patients with primary renal tumors and 14 metastatic renal tumors. ^{18}FDG-PET results led to changes in 3 of the 10 patients with primary tumors (30 percent). These changes included avoidance of surgery in the case of a mass determined to be benign, proceeding with surgery in a case where lung metastasis was ruled out, and cancellation of surgery due to detection of unsuspected bone metastases. There were no changes in treatment management reported for the 14 metastatic renal tumors imaged by ^{18}FDG-PET.

The authors concluded that ^{18}FDG-PET was useful as a complementary modality to CT scans for staging and treatment management of primary malignant renal tumors, as well as for characterization of renal masses resulting from metastases of other primary cancers. They identified that the study

was limited by the population, as the included cases had known renal masses previously detected by conventional imaging by CT scan or MRI.

This retrospective study was assessed as being of moderate quality. Methodological strengths included: a clear description of the selection criteria and spectrum of included patients, appropriate choice of reference standard. Additionally, intermediate results were reported and there were no withdrawals. There was inadequate reporting of some aspects, including detail about the execution of the ^{18}FDG-PET scan and reference tests. As only patients with known renal masses were included, the generalizability of the study may be limited. The interpretation of the index and reference tests was not reported to be blinded, thus introducing the possibility of review bias. Finally, it is crucial to note that not only was there no single reference standard, but PET imaging also formed a part of the reference standard in some instances. This methodological flaw may have introduced verification bias in the validation of the true disease status.

Summary of the results

Three retrospective studies of moderate methodological quality evaluated the use of ^{18}FDG-PET[67,69,70] on physician decision-making for the initial diagnosis[69,70] and staging[67,69,70] of kidney cancer. The usefulness of ^{18}FDG-PET seems less apparent for kidney cancer; notably, the proportion of cases where management changed as a result of the ^{18}FDG-PET findings was generally small. This could be a due to the fact that kidney cancer is not often diagnosed until tumors are large or more advanced. Alternatively, the minimal impact of ^{18}FDG-PET results on physician decision-making may result from the small number of studies with small overall number of participants. Moreover, most of the authors recommend ^{18}FDG-PET as a complementary tool but one author qualified that it may be most useful in selecting cases (e.g., suspicious or equivocal cases). More work is needed to delineate in what situations ^{18}FDG-PET would be most useful. General conclusions cannot be made, since the evidence base is limited, consisting of a small number of retrospective studies with moderate methodological quality.

Table 18 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET on kidney cancer.

Table 18. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET for kidney cancer

Study	Results of FDG-PET imaging on patient diagnosis and treatment	Types of bias
Dilhuydy 2006[67] Study type: Retrospective	**Management decision:** Treatment and diagnostic testing impact Management strategy changed 5/24 (21%) Treatment instead of monitoring strategy changed (n = 4): Received surgery (n = 2) or immunotherapy (n = 2) Type treatment altered (n = 1) (surgery instead of immunotherapy) Management changed in 2/5 patients assessed as "complete response" to prior treatment by conventional CT + bone scans	Spectrum bias (unclear) Selection bias (unclear) Verification bias (>1 RS) Review bias (PET, unblinded; RS, unclear)
Kang 2004[69]	**Management decision:** Treatment and diagnostic testing impact 66 patients received 90 PET scans Management strategy changed in 12/90 (13%) Recurrences identified lead to surgery (n = 2) (Treatment); Additional diagnostic by MRI ordered (n = 1) (Diagnostic Imaging); Reinterpretation of previous imaging (n = 9) (Diagnostic Imaging) **Prognostic value** for immunotherapy: Accuracy of metastatic lesion detection by PET assessed: 81% of PET positive lesions progressed v. 67% of PET negative lesions	Verification bias (>1 RS) Review bias (PET, unblinded; RS, unclear if blinded)
Kumar 2005[70]	**Management decision:** Treatment Treatment strategy changed for 3/10 (30%) primary renal tumor cases. No changes were mentioned in the 14 cases of renal cancer metastasis. Thus, overall 3/24 cases changed (13%): -Identified to have a benign mass, and surgery avoided (n = 1) -Unsuspected bone metastasis, radical surgery cancelled (n = 1) -Ruled out lung metastasis, surgery proceeded (n = 1)	Verification bias (>1 RS) (PET, unclear if blinded; RS, unblinded)

CT = computed tomography; FDG = fluorodeoxyglucose F18; MRI = magnetic resonance imaging; PET = positron emission tomography; RS = reference standard

5. Ovarian Cancer

5.1. Background

Ovarian cancer is the fifth most common malignancy and the fifth leading cause of cancer mortality in American women. It also leads to more deaths than any other gynecological malignancy. [148] In the United States, an estimated 21,650 new cases will be diagnosed and 15,520 women with die from ovarian cancer in 2008. [133] Caucasian-American women experience higher incidence rates than African-American or Asian-American women. [148]

Between 1996 and 2004, the 5-year survival rate for women with ovarian cancer in the United States was 45.5 percent. [133] Poor outcomes are associated with the lack of effective methods for prevention and early detection. If diagnosed early, survival rates improve dramatically, to approximately 95 percent. [149] For 80 to 90 percent of women, diagnosis occurs after 40 years of age; less than 1 percent are diagnosed before 20 years of age. [148] The median age at first diagnosis is 62 years. [133]

Approximately 90 percent of ovarian cancers are derived from the epithelial cells of the ovaries. [148] Of epithelial ovarian neoplasms, 10 to 20 percent tend to develop into borderline or low malignant potential tumors. Epithelial tumors are divided into five categories: serous, mucinous, endometrioid, clear-cell and Brenner. Mixed forms of tumor cells where there is a second or third cell type in addition to the main tumor cell are possible. [150] Nonepithelial tumors include: sex cord-stromal, germ-cell and indeterminate tumors. [148]

Before proceeding with treatment, the extent of disease must be determined. Ovarian cancer is classified into four stages. Tumors are staged using the International Federation of Gynaecology (FIGO) system, which takes tumor grade, depth, width and extent of invasion into consideration (Table 19). [151] At the time of diagnosis approximately 75 percent of patients present with advanced disease. [149]

Table 19. FIGO staging of ovarian cancer

Stage	Description
Stage I	Limited to ovaries
Ia	One ovary
Ib	Both ovaries
Ic	Ruptured capsule, surface tumor or positive washings
Stage II	Pelvic extension
IIa	Uterus, tube(s)
IIb	Other pelvic tissue
IIc	Positive washings, ascites
Stage III	Abdominal extension and/or regional lymph nodes
IIIa	Microscopic peritoneal metastases
IIIb	Macroscopic peritoneal metastases ≤ 2cm
IIIc	Macroscopic peritoneal metastases > 2 cm and/or regional lymph nodes
Stage IV	Distant metastases outside peritoneal cavity

Taken from Aebi et al. [151]

Two techniques are used in screening for ovarian cancer: serum tumor marker cancer antigen 125 (CA-125) and transvaginal ultrasonography (TVUS), neither of which seem to identify ovarian cancer at an early, potentially curable stage. Unlike cervical cancer, there has been no success in identifying precancerous lesions through screening techniques. The links between current epidemiological, biological and pathological data are not fully understood and there is a lack of animal models. Moreover, the disease is virulent and frequently diagnosed only in the advanced stages of disease. [149]

Treatment for ovarian cancer typically involves surgery, the extent of which depends on the stage of disease. Total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy and peritoneal biopsies are required. Preserving fertility may be considered for younger patients with less advanced disease. For women with stage Ia or b, surgery alone should be adequate to treat the disease; however, patients with stage Ic or IIb may also require adjuvant chemotherapy. For more developed stages of disease, additional surgical goals include cytoreduction aiming at leaving no residual disease. Chemotherapy is also required and typically incorporates a platinum-based regime. [151] A second surgery to determine if further therapy in required may be performed. [144]

5.2. Importance of Key Questions in the Clinical Management of Ovarian Cancer

The diagnostic work-up currently used for the characterization of ovarian lesions includes gynecological examination and TVUS. It has been reported that TVUS is not accurate enough to guarantee a precise differential diagnosis because benign and malignant ovarian lesions may present similar morphological characteristics. Measurement of specific serum tumor markers such as CA-125 is often used to detect recurrences; however, this does not allow localization of the recurrence or

differentiation between local and diffused disease. Furthermore, nonrecurrent conditions like infections will often produce elevated CA-125 titers. Conventional imaging modalities such as MRI and helical CT with contrast enhancement are often used in conjunction with CA-125 to detect recurrences. However, detection of recurrences in small peritoneal lesions or differentiation of peritoneal abnormalities can be challenging. Early detection of recurrence in ovarian cancer may allow different therapeutic interventions that could improve outcomes and increase the chances of prolonged remission and survival. Additionally, ^{18}FDG-PET may be capable of detecting early, small regions of relapse when other tests do not detect disease. [144] There is a need to evaluate the evidence on the use of ^{18}FDG-PET in differentiating malignant from benign disease, staging and grading malignant disease, differentiating recurrent disease from therapy-induced changes and monitoring response to therapy in ovarian cancer.

5.3. Results

Twenty-four studies[40,72-90,126-128,130] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for ovarian cancer. Twenty studies[40,72-90] evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT, five studies[75,89,126-128] reported on the diagnostic thinking impact of ^{18}FDG-PET/CT, and one study[130] evaluated the effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for ovarian cancer. Characteristics of the populations, conditions of ^{18}FDG-PET and ^{18}FDG-PET/CT administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

5.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in ovarian cancer

Characteristics of the studies

Twenty studies (fourteen prospective, [40,73-76,78,79,81,82,84,85,87,88,90] six retrospective[72,77,80,83,86,89]) evaluated the diagnostic accuracy of ^{18}FDG-PET[40,77,79,81,84,88,90] or ^{18}FDG-PET/CT. [74] Twelve studies used ^{18}FDG-PET to assess recurrences, [72,73,75,77,78,80-83,86,88,89] two for primary diagnosis, [79,85] two for initial staging, [76,90] two for restaging purposes, [84,87] one for primary diagnosis and initial staging, [74] and one for initial staging and assessing recurrences. The studies contained a total of 871 patients with sample sizes ranging from 13 to 101 participants. The participants ages ranged in age from 17 to 89 years. ^{18}FDG-PET was compared to a reference standard that varied across the studies. In ten studies the reference standard was exclusively histology/biopsy, [72-74,76,79,83-85,87,90] in nine studies it was either histology/biopsy or clinical followup, [75,77,78,80-82,86,88,89] and in one study the reference standard was histology/biopsy or conventional imaging. [40] Seven studies reported the mean time between last treatment and ^{18}FDG-PET as 6 months or more, [72,73] less than 6 months, [77] 3 months or more, [83] 3.6 months, [80] 30 days[84] and 29 days. [87] Six studies used a fixed dose of ^{18}FDG-PET of 350 MBq[78] or 370 MBq, [79,81,82,87,90] five studies used a weight-based dose of 6.5 MBq/kg, [76] 5.5 MBq/kg, [74] 5.2 MBq/kg, [84] or 0.22 mCi/kg; [75,83] six studies reported doses ranging from 260 to 666 MBq. [40,77,80,85,86,89] When reported, the time between injection and PET scan ranged from 45 to 90 minutes. Patients fasted for four hours, [40,72,73,75,83] six hour, [74,76,77,80-82,84-89] or twelve hours; [79,90] one study[78] did not indicate fasting. Thirteen studies[72,73,75,77,78,80-84,86,87,89] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose levels allowed were normal levels, [78,80,82] 200 mg/dL, [72,73,83,86,89] 140 mg/dL, [81,87] and 7.5 mmol/L. [77] Methods of interpretation of the images were qualitative in nine studies[72,73,77,80,81,85-87,89] and both qualitative and quantitative in seven studies. [74,75,78,79,82,84,90] Scans were interpreted qualitatively using visual analysis in all studies. SUV values were reported in five studies for the interpretation of the PET images. The criterion for abnormality was SUV greater than 3 g/mL[74,75,87,88] or greater than 2.5 g/mL. [78]

Comparisons

Comparisons for which data were considered for meta-analysis are summarized in Table 20. Pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for assessing recurrences of ovarian cancer. Individual study data are summarized in Appendix D.

Table 20. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer

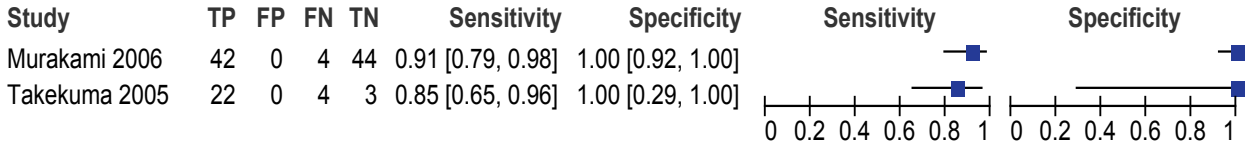
Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Primary diagnosis and staging	Castellucci 2007[74]	P	FDG-PET/CT	Histology/biopsy	No
Staging and recurrences	Grisaru 2004[40]	P	FDG-PET	Histology/biopsy or clinical followup	No
Primary diagnosis	Kawahara 2004[79]	P	FDG-PET	Histology/biopsy	No
	Risum 2007[85]	P	FDG-PET/CT	Histology/biopsy	
Recurrences	Bristow 2003[73]	P	FDG-PET/CT	Histology/biopsy	1. FDG-PET v. histology/biopsy or clinical followup (P studies) [81,88] 2. FDG-PET/CT v. any reference standard (P studies) [73,75,78,82] 3. FDG-PET/CT v. histology/biopsy or clinical followup (P studies) [75,78,82] 4. FDG-PET/CT v. any reference standard (R studies) [72,80,83,86,89] 5. FDG-PET/CT v. histology/biopsy or clinical followup (R studies) [80,86,89] 6. FDG-PET/CT v. histology/biopsy (R studies) [72,83]
	Bristow 2005[72]	R	FDG-PET/CT	Histology/biopsy	
	Chung 2007[75]	P	FDG-PET/CT	Histology/biopsy or clinical followup	
	Garcia-Velloso 2007[77]	R	FDG-PET	Histology/biopsy or clinical followup	
	Hauth 2005[78]	P	FDG-PET/CT	Histology/biopsy or clinical followup	
	Kim 2007[80]	R	FDG-PET/CT	Histology/biopsy or clinical followup	
	Murakami 2006[81]	P	FDG-PET	Histology/biopsy or clinical followup	
	Nanni 2005[82]	P	FDG-PET/CT	Histology/biopsy or clinical followup	
	Pannu 2004[83]	R	FDG-PET/CT	Histology/biopsy	
	Sebastian 2008[86]	R	FDG-PET/CT	Histology/biopsy or clinical followup	
	Takekuma 2005[88]	P	FDG-PET	Histology/biopsy or clinical followup	
	Thrall 2007[89]	R	FDG-PET/CT	Histology/biopsy or clinical followup	
Restaging	Picchio 2003[84]	P	FDG-PET	Histology/biopsy	No
	Sironi 2004[87]	P	FDG-PET/CT	Histology/biopsy	
Staging	Drieskens 2003[76]	P	FDG-PET/CT	Histology/biopsy	No
	Yoshida 2004[90]	P	FDG-PET	Histology/biopsy	

CT = computed tomography; FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

1. ^{18}FDG-PET for recurrences of ovarian cancer

Reference standard: histology/biopsy or clinical followup; prospective studies. Two prospective studies[81,88] totaling 119 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure 43. Sensitivities were 91 percent[81] and 85 percent; [88] specificity was 100 percent.

Figure 43. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer



We found that ^{18}FDG-PET had a pooled positive LR of 22.4 (95% CI = 1.64, 305.38) and a pooled negative LR of 0.13 (95% CI = 0.06, 0.29) to accurately detect recurrences of ovarian cancer (Figures 44 and 45). Both the positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET seems to be helpful for identifying recurrences of the disease. There was moderate heterogeneity in the positive LR ($p = 0.17$; $I^2 = 47$ percent); negative LR was homogeneous across studies.

Figure 44. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)

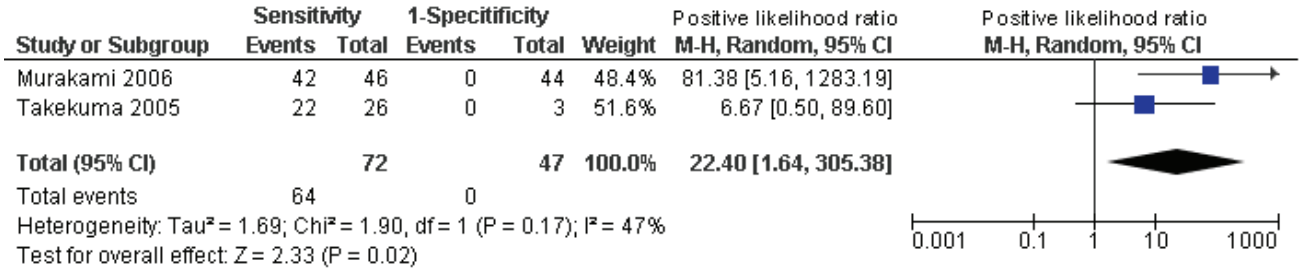


Figure 45. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)

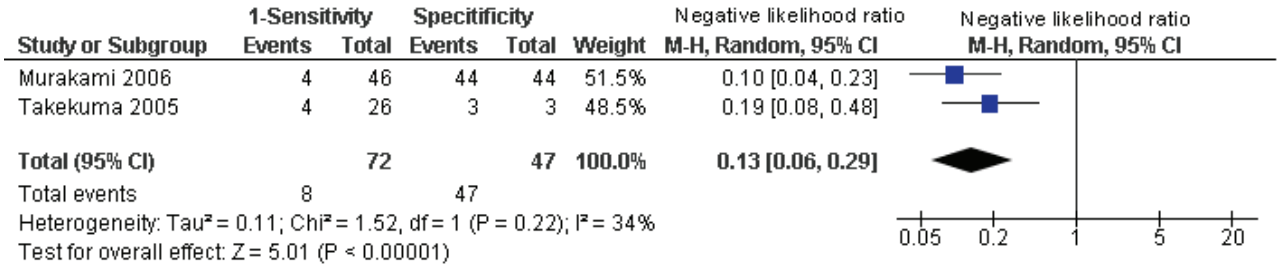
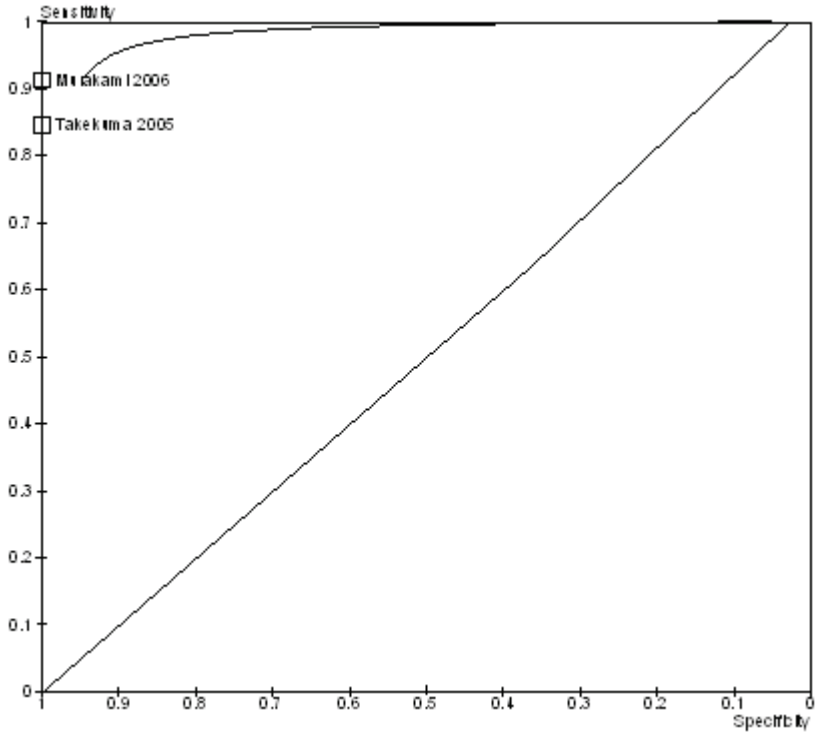


Figure 46 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer based on prospective studies.

Figure 46. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)

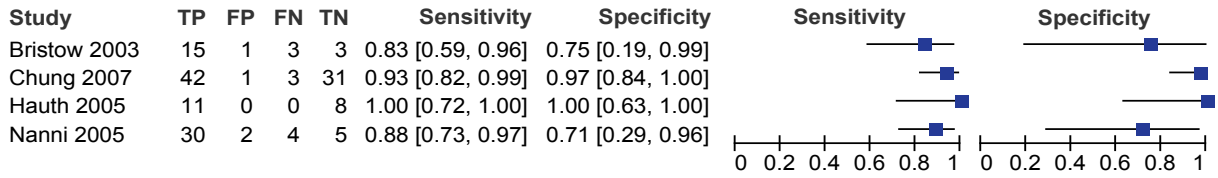


2. ^{18}FDG-PET/CT for recurrences of ovarian cancer

Reference standard: any; prospective studies. Four prospective studies[73,75,78,82] totaling 159 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure

47. Sensitivity ranged from 83 percent[73] to 100 percent. [78] Specificity ranged from 71 percent[82] to 100 percent. [78]

Figure 47. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer



We found that ^{18}FDG-PET/CT had a pooled positive LR of 6.97 (95% CI = 1.64, 25) and a pooled negative LR of 0.12 (95% CI = 0.06, 0.26) to accurately detect recurrences of ovarian cancer (Figures 48 and 49). Both the positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET/CT seems to be helpful for identifying recurrences of the disease. There was moderate heterogeneity in the positive LR ($p = 0.10$; $I^2 = 52$ percent); negative LR was homogeneous across studies.

Figure 48. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies)

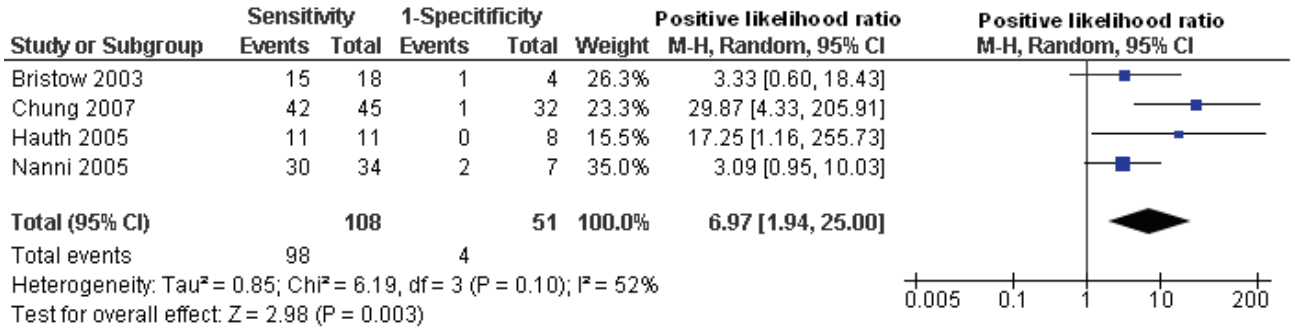


Figure 49. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies)

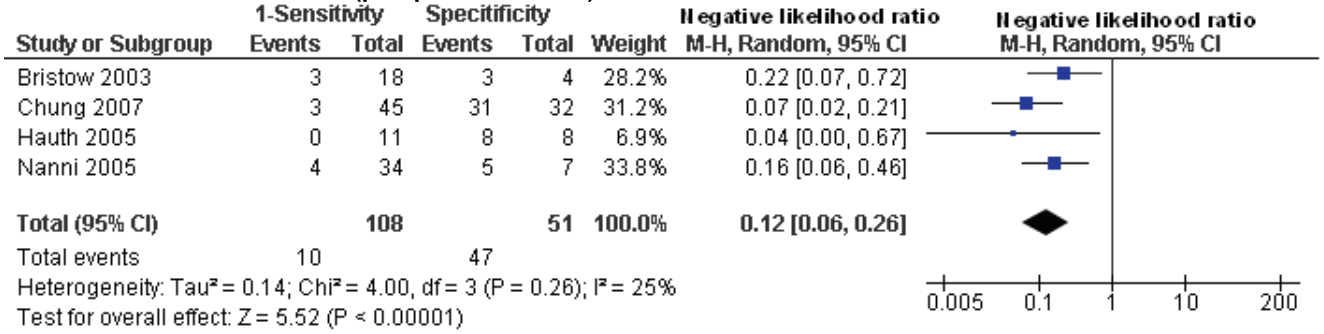
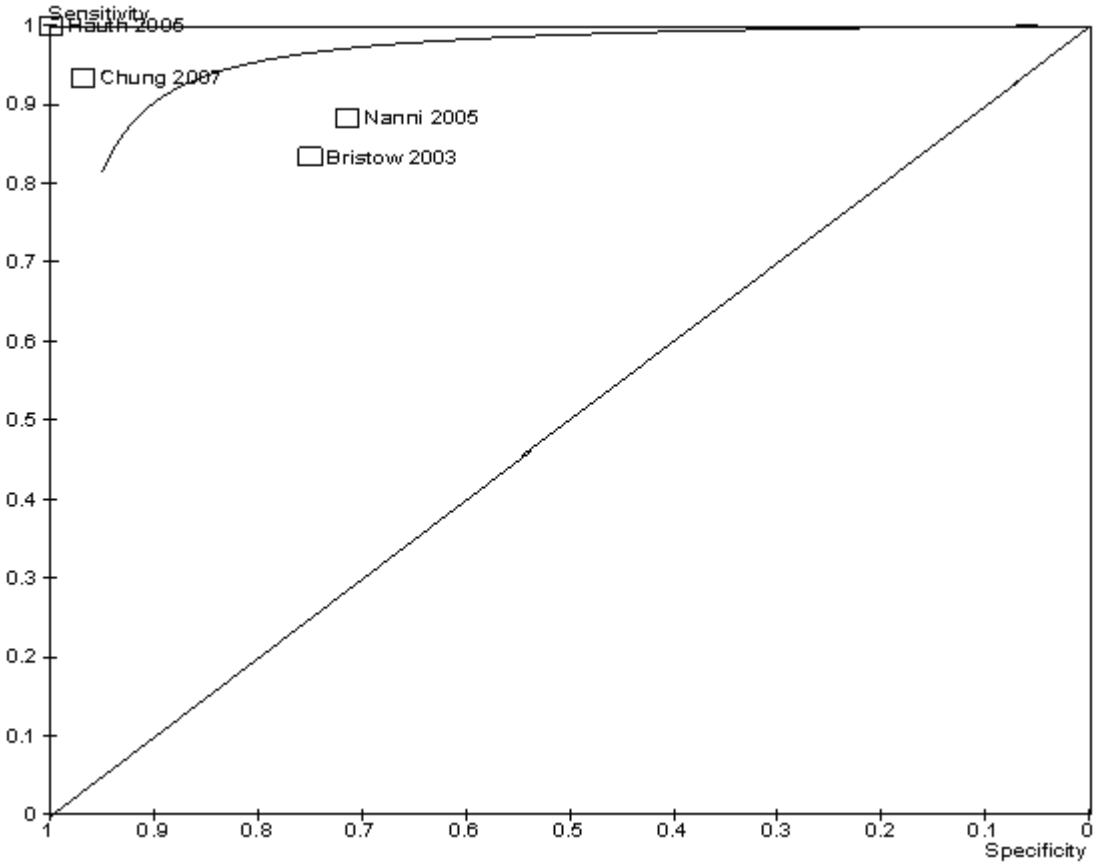


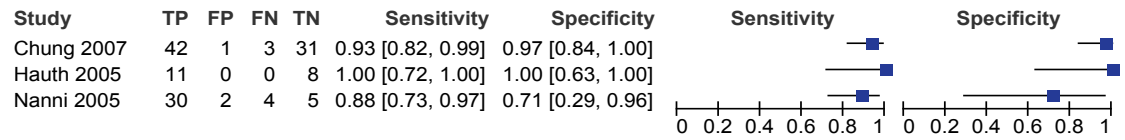
Figure 50 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer based on prospective studies.

Figure 50. Summary ROC plot of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (prospective studies)



Reference standard: histology/biopsy or clinical followup, prospective studies (subgroup analysis). Three prospective studies[75,78,82] totaling 137 participants provided data for a subgroup analysis of the diagnostic accuracy of ^{18}FDG-PET/CT when histology/biopsy or clinical followup were used as the reference standard to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure 51. Sensitivity ranged from 88 percent[82] to 100 percent. [78] Specificity ranged from 71 percent[82] to 100 percent. [78]

Figure 51. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (subgroup analysis)



We found that when only histology/biopsy were considered as reference standard, ^{18}FDG-PET/CT had a pooled positive LR of 9.84 (95% CI = 1.64, 59.15) and a pooled negative LR of 0.10 (95% CI = 0.05, 0.22) to accurately detect recurrences of ovarian cancer (Figures 52 and 53). Both the positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET/CT seems to be helpful for identifying recurrences of the disease. The positive LR was moderately heterogeneous ($p = 0.06$, $I^2 = 64$ percent); the negative LR was homogeneous across the studies.

Figure 52. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (prospective studies)

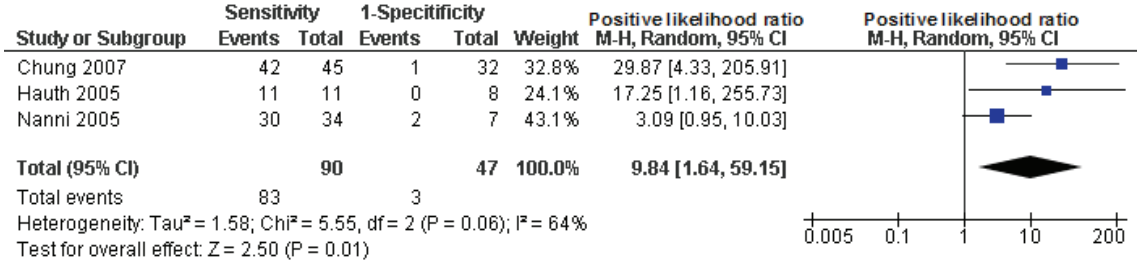


Figure 53. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for detecting recurrences of ovarian cancer (prospective studies)

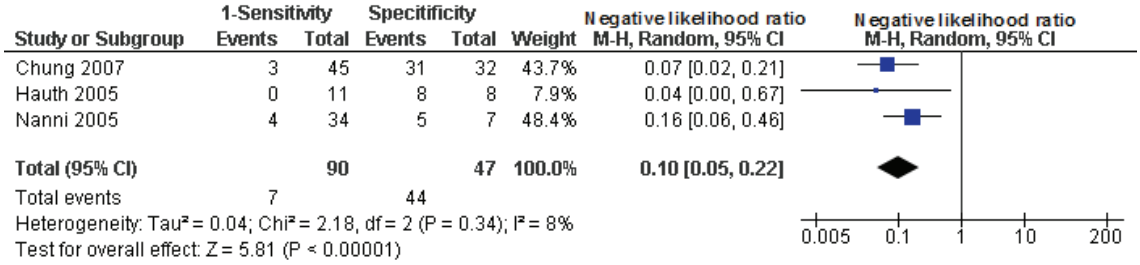
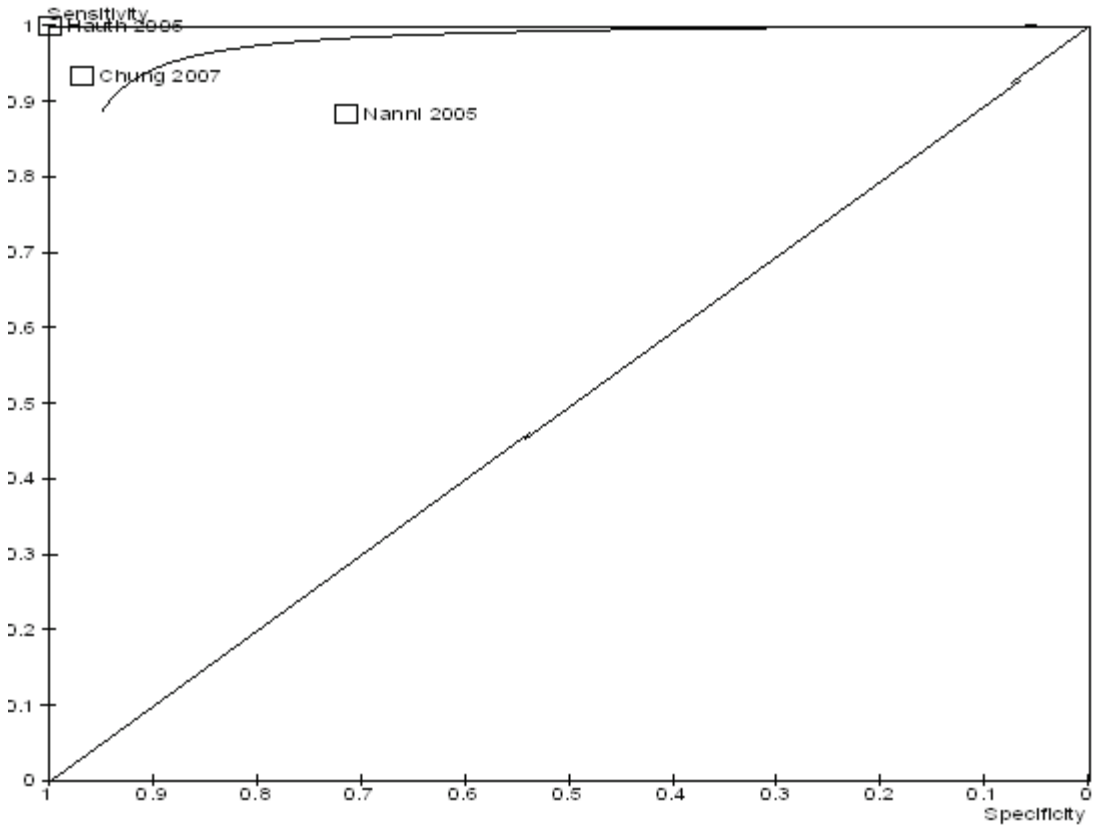


Figure 54 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer based on prospective studies.

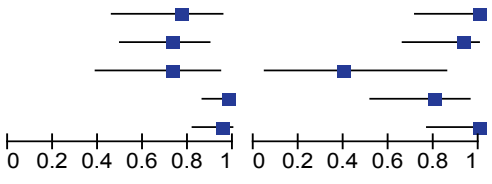
Figure 54. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (prospective studies)



Reference standard: any; retrospective studies. Separate meta-analyses were conducted for retrospective studies. Five retrospective studies[72,80,83,86,89] totaling 180 participants provided data for a meta-analysis of the accuracy of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure 55. Sensitivity ranged from 73 percent[80,83] to 97 percent. [86] Specificity ranged from 40 percent[83] to 100 percent. [72,89]

Figure 55. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Bristow 2005	10	0	3	11	0.77 [0.46, 0.95]	1.00 [0.72, 1.00]
Kim 2007	16	1	6	13	0.73 [0.50, 0.89]	0.93 [0.66, 1.00]
Pannu 2004	8	3	3	2	0.73 [0.39, 0.94]	0.40 [0.05, 0.85]
Sebastian 2008	37	3	1	12	0.97 [0.86, 1.00]	0.80 [0.52, 0.96]
Thrall 2007	35	0	2	14	0.95 [0.82, 0.99]	1.00 [0.77, 1.00]



Based on the analysis of retrospective studies, we found that ^{18}FDG-PET/CT had a pooled positive LR of 6.02 (95% CI = 1.40, 25.91) and a pooled negative LR of 0.19 (95% CI = 0.08, 0.45)

to accurately detect recurrences of ovarian cancer (Figures 56 and 57). The positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET/CT seems to be helpful for identifying recurrences of the disease. However, both the positive ($p = 0.001$; $I^2 = 78$ percent) and the negative ($p = 0.02$; $I^2 = 66$ percent) LRs were heterogeneous across the studies precluding firm conclusions based on these results.

Figure 56. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (retrospective studies)

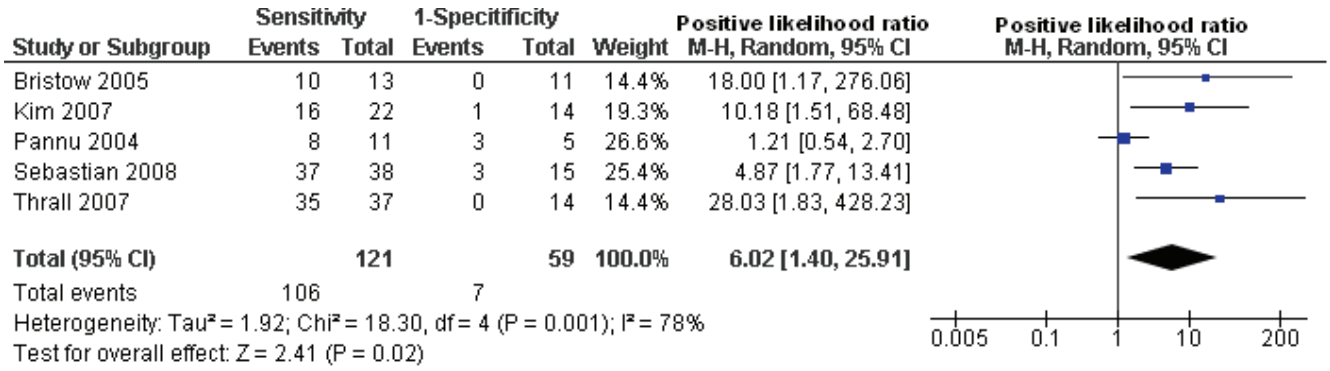


Figure 57. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer (retrospective studies)

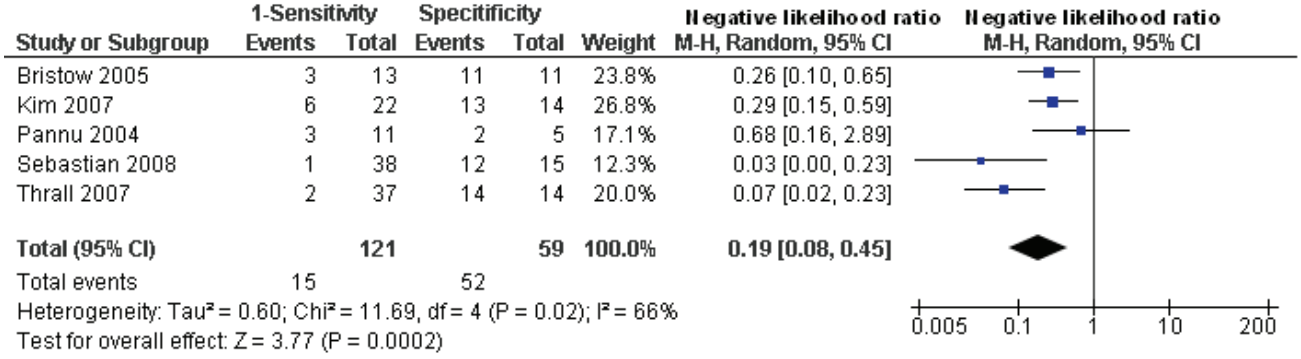
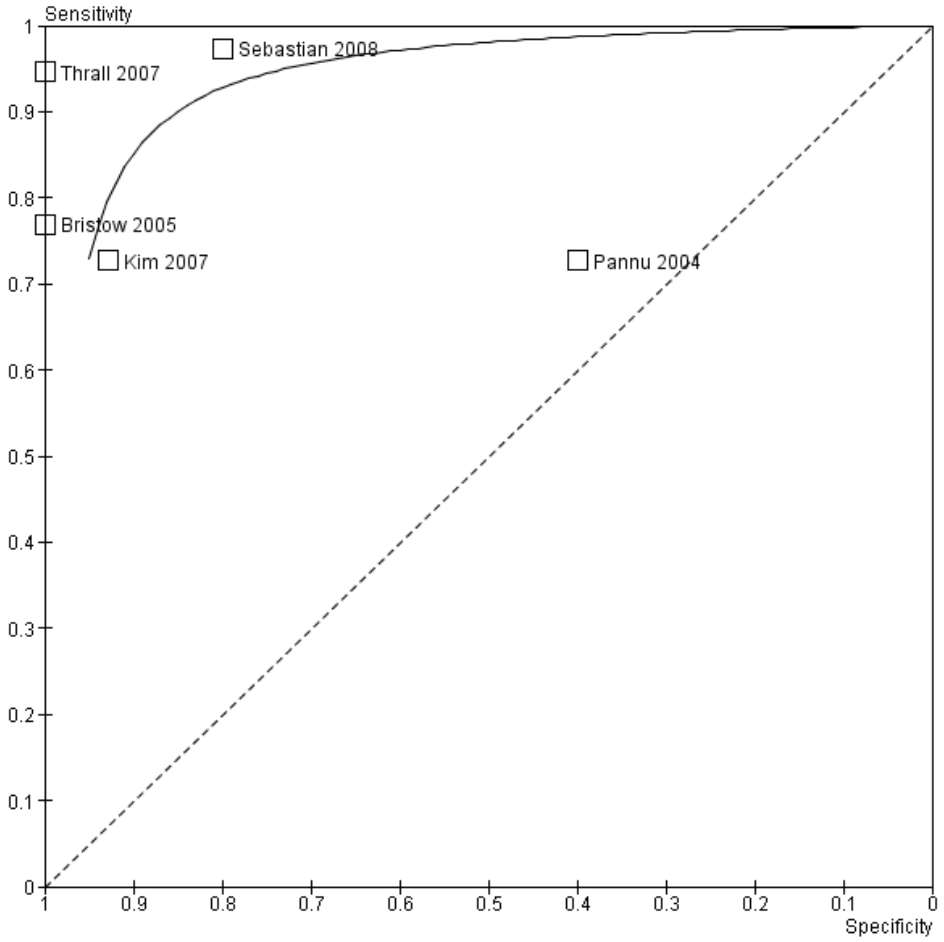


Figure 58 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. any reference standard to detect recurrences of ovarian cancer based on retrospective studies.

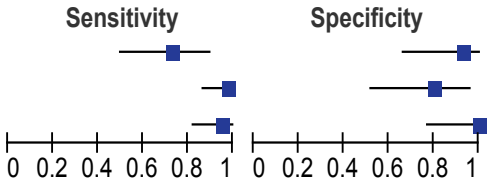
Figure 58. Summary ROC plot of ^{18}FDG-PET v. any reference standard to detect recurrences of ovarian cancer (retrospective studies)



Reference standard: histology/biopsy or clinical followup, retrospective studies (subgroup analysis). Three retrospective studies[80,86,89] totaling 140 participants provided data for a subgroup analysis of the diagnostic accuracy of ^{18}FDG-PET/CT when histology/biopsy or clinical followup were used as the reference standard to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure 59. Sensitivity ranged from 73 percent[80] to 97 percent. [86] Specificity ranged from 80 percent[86] to 100 percent. [89]

Figure 59. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (subgroup analysis)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Kim 2007	16	1	6	13	0.73 [0.50, 0.89]	0.93 [0.66, 1.00]
Sebastian 2008	37	3	1	12	0.97 [0.86, 1.00]	0.80 [0.52, 0.96]
Thrall 2007	35	0	2	14	0.95 [0.82, 0.99]	1.00 [0.77, 1.00]



We found that when only histology/biopsy or clinical followup were considered as reference standard in retrospective studies, ^{18}FDG-PET/CT had a pooled positive LR of 6.68 (95% CI = 2.86, 15.64) and a pooled negative LR of 0.10 (95% CI = 0.02, 0.44) to accurately detect recurrences of ovarian cancer (Figures 60 and 61). The positive and negative LRs were statistically significant and therefore, ^{18}FDG-PET/CT seems to be helpful for identifying recurrences of the disease. The positive LR was homogeneous across the studies; however, the negative LR was heterogeneous across the studies (p = 0.01; I^2 = 77 percent).

Figure 60. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies)

Study or Subgroup	Sensitivity Events	Sensitivity Total	1-Specitificity Events	1-Specitificity Total	Weight	Positive likelihood ratio M-H, Random, 95% CI
Kim 2007	16	22	1	14	19.9%	10.18 [1.51, 68.48]
Sebastian 2008	37	38	3	15	70.4%	4.87 [1.77, 13.41]
Thrall 2007	35	37	0	14	9.7%	28.03 [1.83, 428.23]
Total (95% CI)		**97**		**43**	**100.0%**	**6.68 [2.86, 15.64]**
Total events	88		4			

Heterogeneity: Tau² = 0.00; Chi² = 1.95, df = 2 (P = 0.38); I² = 0%
Test for overall effect: Z = 4.38 (P < 0.0001)

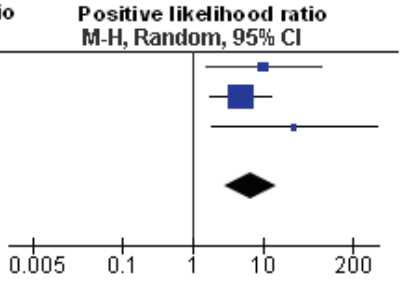


Figure 61. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies)

Study or Subgroup	1-Sensitivity Events	1-Sensitivity Total	Specitificity Events	Specitificity Total	Weight	Negative likelihood ratio M-H, Random, 95% CI
Kim 2007	6	22	13	14	40.8%	0.29 [0.15, 0.59]
Sebastian 2008	1	38	12	15	24.8%	0.03 [0.00, 0.23]
Thrall 2007	2	37	14	14	34.4%	0.07 [0.02, 0.23]
Total (95% CI)		**97**		**43**	**100.0%**	**0.10 [0.02, 0.44]**
Total events	9		39			

Heterogeneity: Tau² = 1.21; Chi² = 8.64, df = 2 (P = 0.01); I² = 77%
Test for overall effect: Z = 3.07 (P = 0.002)

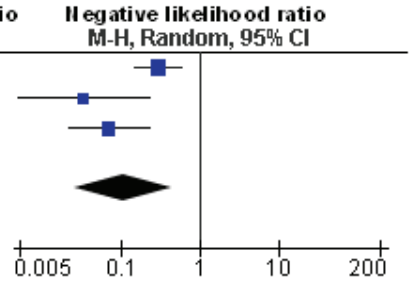
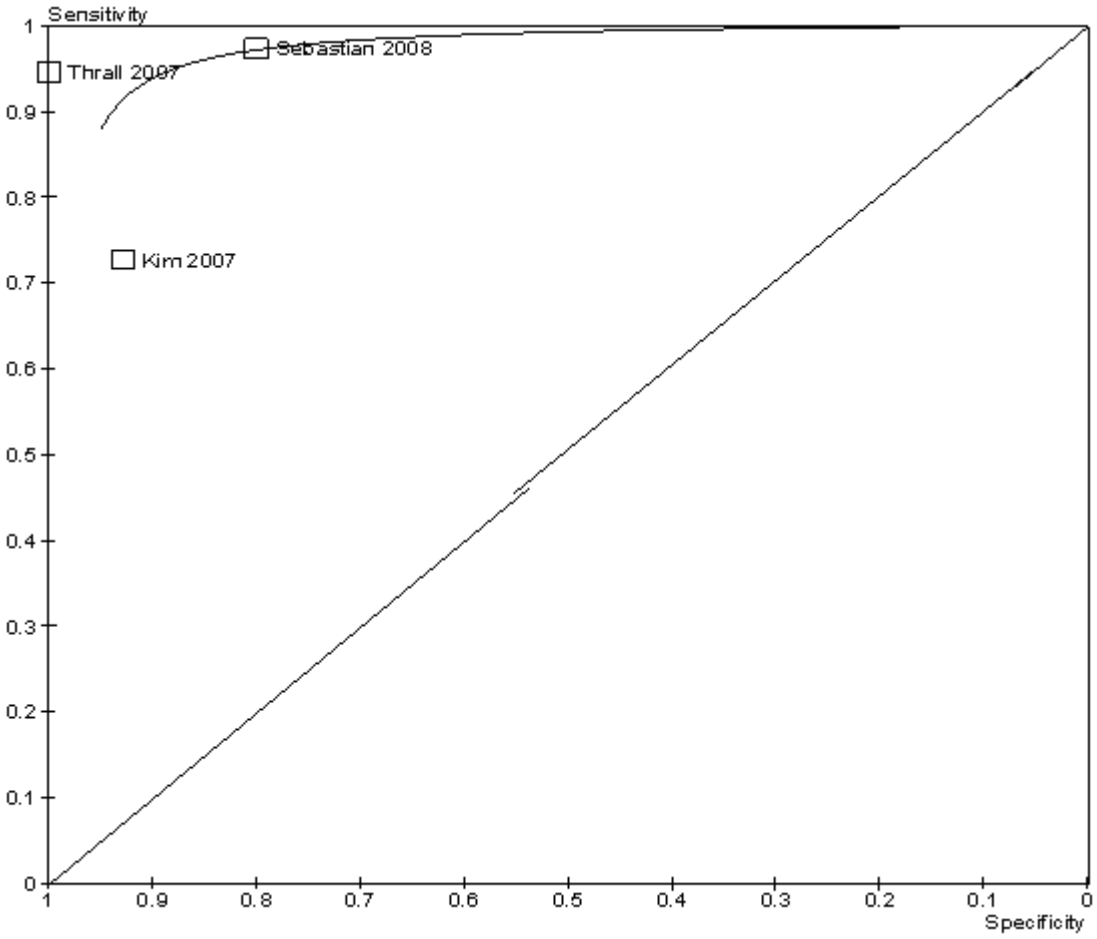


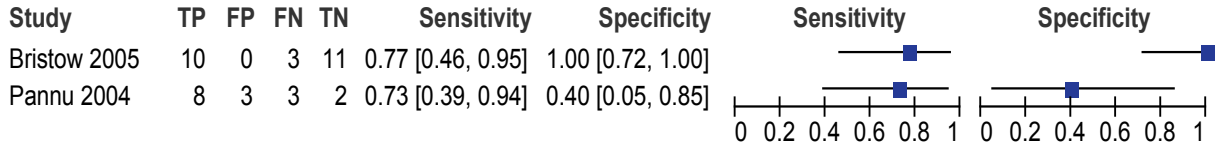
Figure 62 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer based on retrospective studies.

Figure 62. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (retrospective studies)



Reference standard: histology/biopsy, retrospective studies (subgroup analysis). Two retrospective studies[72,83] totaling 40 participants provided data for a subgroup analysis of the diagnostic accuracy of ^{18}FDG-PET/CT when only histology/biopsy was used as the reference standard to detect recurrences of ovarian cancer. Individual 2x2 table results are presented in Figure 63. Sensitivities were 73 percent[83] and 77 percent; [72] specificities were 40 percent[83] and 100 percent.[72]

Figure 63. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup to detect recurrences of ovarian cancer (subgroup analysis)



We found that when only histology/biopsy was considered as the reference standard in retrospective studies, ^{18}FDG-PET/CT had a pooled positive LR of 3.97 (95% CI = 0.09, 167.47) and a pooled negative LR of 0.36 (95% CI = 0.15, 0.86) to accurately detect recurrences of the disease (Figures 64 and 65). The positive LR was not statistically significant and was heterogeneous across the studies (p = 0.008; I^2 = 86 percent), precluding any reliable interpretation from the results.

Figure 64. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies)

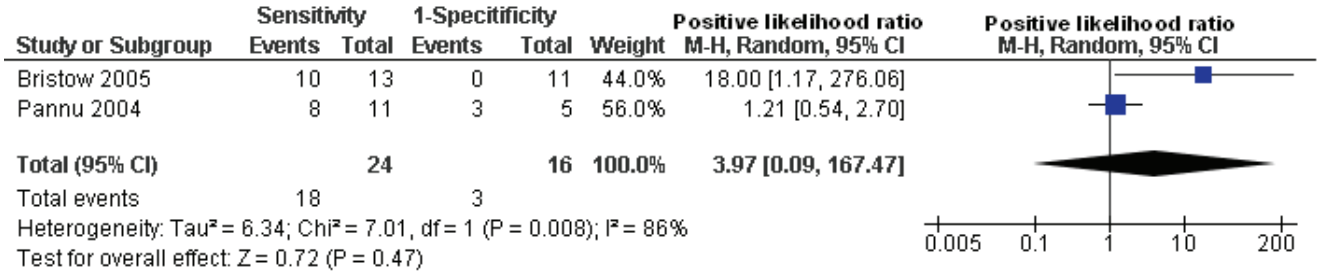


Figure 65. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies)

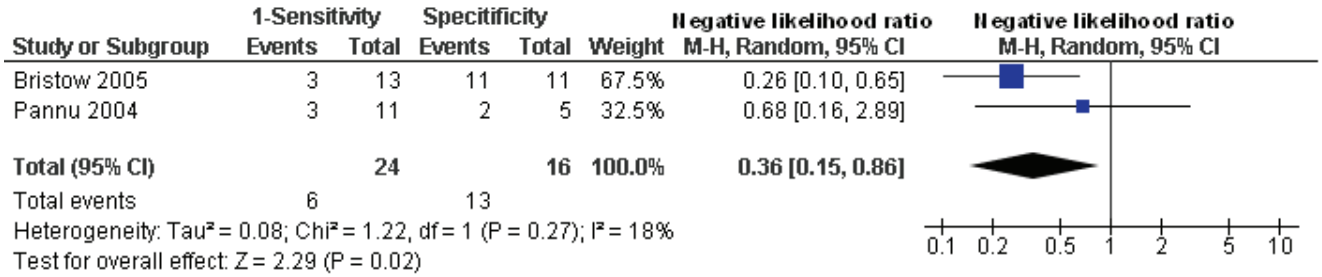
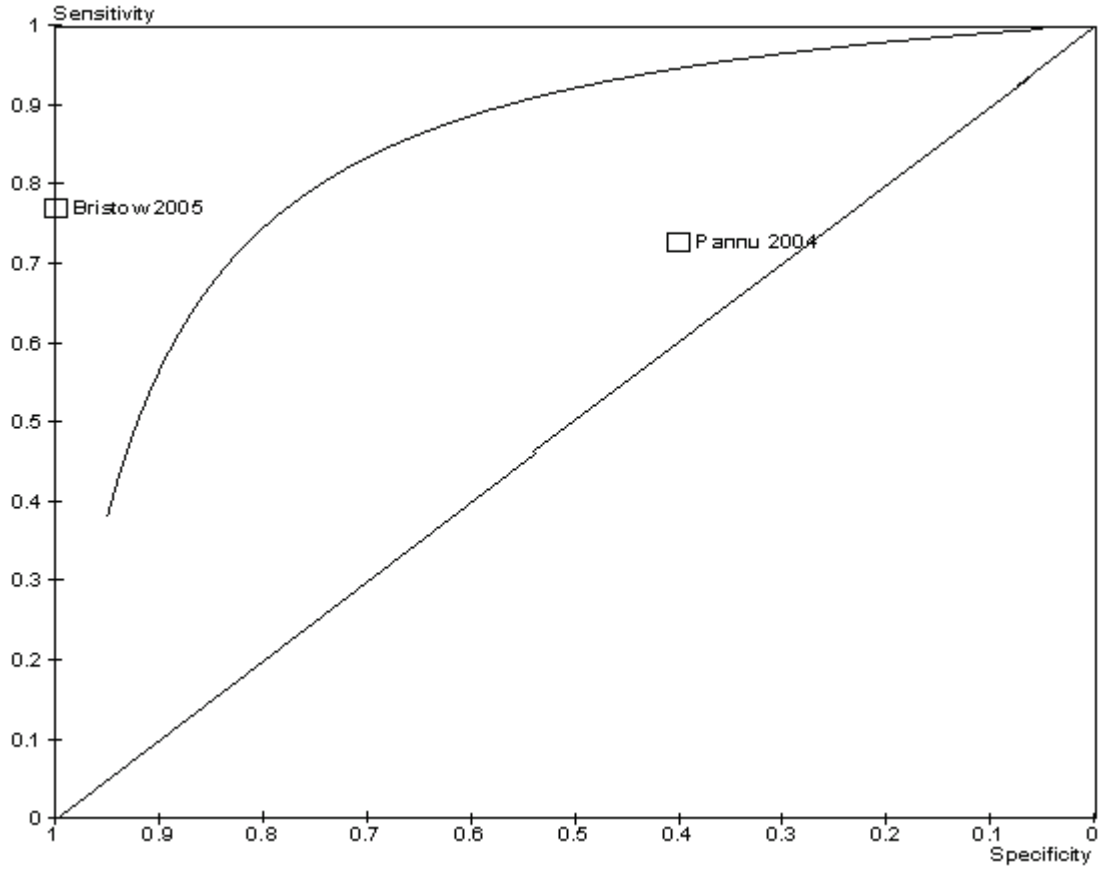


Figure 66 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. histology/biopsy to detect recurrences of ovarian cancer based on retrospective studies.

Figure 66. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy to detect recurrences of ovarian cancer (retrospective studies)



Summary of the results

Meta-analyses were calculated to evaluate the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in detecting recurrences of ovarian cancer (Table 21). Pooled positive and negative LRs showed a consistent, statistically significant effect across a range of reference standards and study designs, providing evidence in support of the usefulness of ^{18}FDG-PET and ^{18}FDG-PET/CT in detecting recurrences. Only one comparison did not yield a statistically significant positive LR, however this result could be attributed to smaller numbers of studies in the analysis and smaller samples (i.e. n=40 v. n>100 in other comparisons).

Table 21. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer

PET Purpose	Type of PET	Reference standard	Design	Studies	N	Effect estimate M-H, Random, 95% CI
Recurrences	FDG-PET	Histology/biopsy or clinical followup	P	2	119	PLR = 22.40 [1.64, 305.38]
						NLR = 0.13 [0.06, 0.29]
	FDG-PET/CT	Any reference standard	P	4	159	PLR = 6.97 [1.94, 25.00]
						NLR = 0.12 [0.06, 0.26]
		Histology/biopsy or clinical followup		3	137	PLR = 9.84 [1.64, 59.15]
						NLR = 0.10 [0.05, 0.22]
		Any reference standard	R	5	180	PLR = 6.02 [1.40, 25.91]
						NLR = 0.19 [0.08, 0.45]
		Histology/biopsy or clinical followup		3	140	PLR = 6.68 [2.86, 15.64]
						NLR = 0.10 [0.02, 0.44]
		Histology/biopsy		2	40	PLR = 3.97 [0.09, 167.47]
						NLR = 0.36 [0.15, 0.86]

CI = confidence interval; FDG= fluorodeoxyglucose F18; M-H = Mantel Hantzel; NLR=negative likelihood ratio; P=prospective; PET=positron emission tomography; PLR=positive likelihood ratio; R=retrospective

5.3.2. Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making with respect to diagnosis and management strategy for patients with ovarian cancer

All five studies[75,89,126-128] considered integrated ^{18}FDG-PET/CT imaging to assess for ovarian cancer recurrence. None of the studies evaluated the diagnostic thinking impact of ^{18}FDG-PET alone. The included studies assessed the impact of ^{18}FDG-PET/CT imaging on the management of patients suspected of suffering a recurrence of cancer as well as for surveillance of recurrences in patients following treatment for their primary cancer.

Chung et al. [75] evaluated the accuracy of ^{18}FDG-PET/CT imaging on diagnosis of suspected recurrences of ovarian cancer and whether the results altered either treatment decisions or additional diagnostic testing regimes for the participants. This prospective study enrolled 77 women with suspected cancer recurrence. ^{18}FDG-PET/CT impacted either the diagnostic followup or the treatment plan in 19 cases (25 percent). The majority of changes (11/19) occurred in cases with a positive ^{18}FDG-PET/CT scan and no other clinical indicators of tumor recurrence. These 11 patients had normal CA-125 and no clinical symptoms. The treatment plan was changed from observation to a regime of chemotherapy. Eight patients whose CA-125 levels were elevated were shown to have only physiological or inflammatory ^{18}FDG uptake. The need for additional diagnostic procedures was eliminated and the patients were followed by normal observation. Overall, ^{18}FDG-PET/CT was determined to have high impact on patient care.

The authors concluded that ^{18}FDG-PET/CT was a sensitive surveillance method to identify cervical cancer recurrences and that it allowed for the optimization and customization of an appropriate treatment plan.

This prospective study was assessed as being of moderate quality. The spectrum of patients that were included and the details of the selection criteria, as well as the choice and administration of the reference standard were well described. There was also a good description of the ^{18}FDG-PET/CT procedure and intermediate test results. Reporting was inadequate on some aspects, notably the time delay between the ^{18}FDG-PET/CT scan and the reference verification of disease status, the lack of detail of the reference test procedure, and the unblinded interpretation of the final reference standard. Additionally, because the disease status of patients was not verified by one reference standard, there is the possibility of verification bias.

Mangili et al. [126] evaluated the impact of ^{18}FDG-PET/CT imaging on the management decisions for patients suspected to have ovarian cancer recurrence. Data from a chart review of 32 patients who underwent ^{18}FDG-PET/CT scans were used in this study. All patients had previously undergone surgery and chemotherapy for ovarian cancer. The suspicion of recurrence was based on a number of followup measures, including elevated serum CA-125 or abnormalities on an annual chest X-ray or abdominal ultrasound. The impact on patient management was assessed retrospectively in a blinded fashion. All pertinent information (e.g., diagnosis, staging, previous treatment of primary tumor, followup including CA-125 values and imaging studies, and the detailed CT report) from the patient charts was collected and distributed to two teams of oncologists. Each team completed a pre-test questionnaire regarding clinical management. Where there was a discrepancy between the care plans, the teams were asked to reach consensus regarding patient management. The ^{18}FDG-PET/CT scans were then distributed with the coded patient information and in a different sequence than the pre-test charts. Both teams came to a post-test consensus on the most appropriate course of clinical management. The addition of ^{18}FDG-PET/CT images to the patient information resulted in changes to the treatment plans for 14 patients (44 percent). The treatment modality was altered for eight patients, including receiving chemotherapy rather than surgery or vice versa (n = 4), undergoing further examination (e.g., CT) rather than diagnostic surgery (n = 1), or chemotherapy rather than diagnostic surgery (n = 3). Prior to ^{18}FDG-PET/CT, seven patients were designated to the observation approach; however, with the addition of ^{18}FDG-PET/CT, this decision was changed for six patients. Of these, two underwent further diagnostic procedures, while four underwent changes to

their treatment plans to undergo chemotherapy (n = 1) or surgery (n = 3). Thus, for seven patients (22 percent), the ^{18}FDG-PET/CT imaging facilitated a change in management from either further invasive diagnostics or a "watch and see" approach to a definitive treatment plan. There was an increase in the overall number of patients undergoing chemotherapy (10/32 to 16/32) as a result of the discovery of more disseminated disease based on the ^{18}FDG-PET/CT images.

The authors concluded that ^{18}FDG-PET/CT imaging for detection of ovarian cancer recurrence demonstrated a higher level of accuracy compared to conventional contrast enhanced CT. The authors suggested that use of integrated ^{18}FDG-PET/CT with a fully diagnostic CT could replace the current approach of using multiple imaging modalities from a number of sessions to restage ovarian cancer recurrences.

This retrospective chart review was determined to be of moderate quality. The spectrum of included patients, choice and administration of the reference standard and the inclusion of intermediate test results were well reported. Additionally, all cases were verified by a reference standard that was interpreted without knowledge of the ^{18}FDG-PET/CT results. Descriptions provided for withdrawals were satisfactory. There was inadequate reporting on some aspects, including the inclusion criteria, the blinded interpretation of the ^{18}FDG-PET/CT imaging, and the period of time between the administration of the ^{18}FDG-PET/CT scan and the reference standard. There was more than one reference standard test used to verify disease status, which may have introduced verification bias.

Simcock et al. [127] investigated the treatment decision impact of integrated ^{18}FDG-PET/CT imaging on the restaging and management of recurrent ovarian cancer in a population of 61 patients. Of these, 56 had sufficient followup data to be included in the analysis (median total followup: 21.6 months). Collectively, these patients underwent 66 ^{18}FDG-PET/CT scans; the majority of women (86 percent) had one scan.

The disease state immediately prior to each of the 66 ^{18}FDG-PET/CT scans were described as follows: "uncertain" (n = 30), "suspected local recurrence" (n = 15), "suspected systemic disease" (n = 14), and "surveillance with no evidence of disease" (n = 7). The impact of the ^{18}FDG-PET/CT imaging on the patient management plans was "high" for 32 patients (57 percent) who received a total of 33 ^{18}FDG-PET/CT scans. Twenty-nine scans had a low impact on patient management. Of the 32 high-impact management changes, 20 occurred in patients who were not assigned a disease state using conventional assessment. As there were 30 patients in this "uncertain" category following

their conventional staging, two-thirds of the patients were diagnosed following the ^{18}FDG-PET/CT. The ^{18}FDG-PET/CT facilitated changes ranging from altering an active treatment approach to observation (n = 6), or from observation to treatment by radiation, chemotherapy or surgery (n = 7). Thus, in 13 patients (23 percent) the results of the ^{18}FDG-PET/CT determined whether or not they received treatment. Other high-impact modality changes included switching from surgery to chemotherapy (n = 6), chemotherapy to radiation or a combination of therapies (n = 4), biopsy to chemotherapy or surgery (n = 4), radiation to chemotherapy or enlargement of radiation target fields (n = 1), and switching from a combination of therapies to radiation or chemotherapy alone (n = 2).

Additionally, prognostic outcomes were reported. The survival of patients was analyzed according to their ^{18}FDG-PET/CT determined disease status (systemic, localized, or no disease/equivocal). While there was no significant difference in overall survival among the three groups, there was significantly lower survival in patients with ^{18}FDG-PET/CT designated systemic disease v. the combined ^{18}FDG-PET/CT designated localized and no disease/equivocal patient survival.

The authors concluded that the use of ^{18}FDG-PET/CT in monitoring patients with recurrent or suspected recurrent ovarian cancer significantly modified the assessment of the cancer state. They also concluded that ^{18}FDG-PET/CT altered management in a substantial proportion of patients.

This prospective study was determined to be of moderate quality. The selection criteria, choice and execution of the reference standard, and intermediate test results were well reported. Additionally, all cases were verified by a reference standard, and there was satisfactory explanation for withdrawals. However, reporting was inadequate on some aspects, notably of whether or not the spectrum of patients was representative of typical clinical practice, which raises the possibility of selection bias. In addition, the lack of blinded interpretation of either the ^{18}FDG-PET/CT scans or the reference standards may have introduced review bias to the interpretation of results. Furthermore, the period of time between the administration of the ^{18}FDG-PET/CT scan and the reference standard was unclear, and there was no standard reference test; rather, a combination of methods was used (histological or clinical followup). This variation in validation of the diagnosis may have lead to verification bias

Soussan et al. [128] investigated the impact of ^{18}FDG-PET/CT on treatment decision for the management of possible recurrent ovarian cancer. This prospective study enrolled 29 outpatients who underwent ^{18}FDG-PET/CT scans. All patients had previously undergone surgery and

chemotherapy. Two questionnaires were completed by the treating oncologists to determine the impact of ^{18}FDG-PET/CT on management decisions; the first was completed following the independent CT scan but prior to the ^{18}FDG-PET/CT scan, and the second was completed upon receipt of the ^{18}FDG-PET/CT data. Followup data were also collected from the referring oncologists.

The final therapeutic decision was changed based on the results of the ^{18}FDG-PET/CT for 10 patients (34 percent). The modality of therapy for the patient changed in three cases (e.g., chemotherapy to chemotherapy plus surgery); for six patients there was a change in plan from an approach of observation to treatment. One patient was switched from chemotherapy to observation. There was major modification in the assessment of disease distribution in 15 patients (52 percent); of these, 11 had more advanced disease and four had more limited disease. A minor change in distribution was found in one patient. Of these patients, nine had their treatment plan altered as a result.

The authors concluded that the use of ^{18}FDG-PET/CT in evaluating patients with suspicion of recurrent ovarian cancer significantly modified treatment decisions. The impact was particularly important for the management of cases that were determined to be positive by ^{18}FDG-PET/CT despite being assessed as negative by CT alone.

This prospective study was determined to be of high quality. The choice and execution of the reference standard, description of the ^{18}FDG-PET/CT interpretation, and inclusion of intermediate test results were well reported. Additionally, all cases were verified by a reference standard, and all enrolled participants were accounted for satisfactorily. However, there was inadequate reporting on some aspects, including whether or not the spectrum of patients was representative of typical clinical practice. The reference standard used to verify disease included a number of measures, which varied among patients, thus leading to potential verification bias. It was unclear whether the interpretation of either the ^{18}FDG-PET/CT or the reference standard was blinded.

Thrall et al. [89] retrospectively investigated the treatment decision impact for 39 patients who underwent integrated ^{18}FDG-PET/CT scans. All patients had confirmed ovarian cancer; the majority (69 percent) assessed to have stage III. All patients had undergone cytoreductive surgery and platinum-based chemotherapy. The analysis of possible recurrences included a total of 59 ^{18}FDG-PET/CT scans. Indications for undergoing ^{18}FDG-PET/CT imaging ranged from a routine component with no clinical or imaging abnormalities (n = 4 scans) to abnormalities such as elevated serum CA-125 (n = 24), clinical symptoms of recurrence (n = 9), abnormal CT scan (n = 14), and

assessment of treatment response (n = 8). Twenty-five patients had one ^{18}FDG-PET/CT scan, 10 patients had two ^{18}FDG-PET/CT scans, two patients had either three or four scans.

During the followup period, 33 patients (85 percent) had a cancer recurrence. Of 24 ^{18}FDG-PET/CT scans performed in 22 patients, 18 scans correctly identified recurrences (75 percent). The correct determination of recurrence was associated with subsequent clinical management decisions. Overall, ^{18}FDG-PET/CT imaging resulted in changes to the treatment plans of 14 patients (36 percent) with known disease recurrence. In four (29 percent) ^{18}FDG-PET/CT imaging identified distant metastases, prompting a change from treatment with curative intent to palliative care. Additionally, of the eight ^{18}FDG-PET/CT scans completed in five patients to assess treatment response, there was one case of nonresponsive progressive disease correctly identified by ^{18}FDG-PET/CT, one patient had stable disease, and three were responsive to treatment. None of these patients had clearly identifiable disease by conventional CT imaging at baseline.

The authors concluded that the use of ^{18}FDG-PET/CT was most valuable in assessment of patients with rising CA-125 levels despite negative or equivocal CT scans. They found ^{18}FDG-PET/CT to be useful in optimal selection of patients for planning of appropriate surgery and radiation therapy.

This retrospective study was determined to be of high quality. The spectrum of participants and their selection criteria were clearly defined. In addition, the choice and execution of the reference standard, the description of the ^{18}FDG-PET/CT imaging procedure and reporting intermediate test results were well detailed. However, due to the retrospective nature of the study, there was no blinded interpretation of either the ^{18}FDG-PET/CT scans or the reference standards. The reference standard used to verify disease status consisted of a combination of methods (histological, clinical followup, surgical findings), which raises concerns about the possibility of verification bias.

5.3.3. ^{18}FDG-PET and as part of a management strategy in ovarian cancer

None of the studies evaluated ^{18}FDG-PET/CT as part of a management strategy. Kim et al. [130] retrospectively assessed the value of ^{18}FDG-PET compared to second-look laparotomy (SLL) in the prognosis and detection of recurrences in patients with advanced ovarian cancer following primary chemotherapy. The study population consisted of 55 patients aged 25 to 78 years of age (mean age: 49.2, SD = 12.1), primarily with stage III (49 percent) and IV (44 percent). All patients were treated with a regime of chemotherapy, following which they were divided into two groups for followup.

One group (n = 30) underwent SLL while the second (n = 25) underwent ^{18}FDG-PET imaging during the followup period. The ^{18}FDG-PET was performed a median of 6.8 months after the initial laparotomy and was visually interpreted by two nuclear physicians. Malignancy status was based on quantitative assessment by calculation of SUV values. The median length of followup for both groups was 35 months, and disease recurrence was verified by a variety of methods, including histology, physical exam, additional imaging, and CA-125 levels. The prognostic indicators investigated were the progression free interval, the disease free interval and the incidence of disease recurrence.

Overall, there was evidence of recurrence of ovarian cancer in 37 patients (67 percent). In the ^{18}FDG-PET group there were 17 cases with recurrent cancer; of these, 13 (76 percent) were detected by ^{18}FDG-PET. There were no significant differences in prognostic indicators between the two groups. When the progression-free interval was compared between the two groups overall, there was no significant difference in the duration of the disease free period. The same was true for the disease-free interval for the subset of patients found to be negative or positive for recurrent disease in either the ^{18}FDG-PET or SLL group.

The authors concluded that neither SLL nor ^{18}FDG-PET was clearly advantageous for indicating disease prognosis and that ^{18}FDG-PET can be used as an alternate followup modality for patients with ovarian cancer.

Overall, this study was determined to be of low quality. The retrospective nature of the prognostic indicator portion of the research, and the lack of a definitive method for determining of an ^{18}FDG-uptake positive lesion were some areas of concern. As there was no clear definition of a positive lesion, the study was subject to threshold bias due to variation in interpretation among assessors. Additionally, there was no uniform reference standard administered to the participants, introducing the possibility of verification bias. Outcomes were adequately described and there was an appropriate description of the general characteristics of the participants. As well, all the patients were accounting for. Detailed inclusion criteria are lacking, which makes it difficult to rule out the potential for participant selection bias.

Summary of the results

There were five studies of moderate to high methodological quality that evaluated the use of ^{18}FDG-PET/CT [75,89,126-128] to assess the restaging[126-128] or recurrence[75,89] of ovarian cancer. The five

studies consistently demonstrated that ^{18}FDG-PET/CT altered management plans in an important proportion of patients. The change in management varied, but slightly more often resulted in more aggressive treatment. Only study[130] of poor quality examined the effect of ^{18}FDG-PET on patient-centered outcomes and prognosis and prohibited any definitive conclusions.

Table 22 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer.

Table 22. Main findings and types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for ovarian cancer

Study	Results of FDG-PET imaging on patient diagnosis, treatment and outcomes	Types of bias
Chung 2007[75] Study type: Prospective	**Management decision:** Treatment and diagnostic imaging impact Management strategy changed 19/77 cases (25%) 11 cases without clinical symptoms or abnormal CA-125 were changed from observation to chemotherapy (Treatment) 8 cases with elevated CA-125 had negative PET/CT, so additional diagnostic tests were cancelled (Diagnostics)	Disease progression bias (unclear) Verification bias (>1 RS) Review bias (RS, unclear if blinded)
Kim 2004[130] Study type: Retrospective	**Patient centered outcomes and prognosis:** Progression-free interval: PET: 28.8 mo (SD 12.7 m) for 25 cases; SLL: 30.6 mo (SD 13.7 mo) for 30 cases Disease free interval in pts. with negative test results: PET: 40.5 mo (SD 11.6 mo); SLL: 48.6 mo (SD 12.1 mo) Disease free interval in pts. with positive test results: PET: 23.7 mo (SD 5.3 mo); SLL: 26.2 mo (SD 6.7 mo)	Selection bias (unclear) Review bias (RS, unclear if blinded)
Mangili 2007[126] Study type: Retrospective	**Management decision:** Treatment and diagnostic imaging impact Management strategy changed 14/32 cases (44%) Changed from observation to treatment or further diagnostics (n = 6) Changed to surgery (n = 3) Underwent further diagnostic examination (n = 2) Changed to chemotherapy (n = 1) Treatment modality changed (n = 8) Surgery to chemotherapy (n = 3) Diagnostic surgery to chemotherapy (n = 3) Chemotherapy to surgery (n = 1) Chemotherapy to additional diagnostic examination (n = 1)	Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS) Review bias (RS, unclear if blinded)
Simcock 2006[127] Study type: Prospective	**Management decision:** Treatment 32 cases (33 PET/CT scans) had a high impact of on management (57%) Observation changed to treatment (n = 7) Active treatment changed to observation (n = 6) Surgery changed to chemotherapy (n = 6) Biopsy changed to treatment (e.g., chemotherapy) (n = 4) Changed between various other treatment modalities (n = 8) (e.g., radiation, chemotherapy, surgery) Changed from treatment to biopsy (n = 1) 29 PET/CT scans had a low impact of on management	Spectrum bias (possible) Selection bias (unclear) Disease progression bias (lengthy interval) Verification bias (>1 RS) Review bias (RS unblinded)
Soussan 2008[128] Study type: Prospective	**Management decision:** Treatment 16 cases had their diagnosis altered by PET (52%) Upstaged (n = 11); downstaged (n = 4); different disease distribution (n = 1) Management strategy changed 10/29 (34%) Changed from observation to chemotherapy (n = 6) Additional treatment modality added to care plan (n = 2) Changed from chemotherapy to observation (n = 1)	Spectrum bias (possible) Selection bias (unclear) Review bias (RS, not blinded)
Thrall M 2007[89] Study type: Retrospective	**Management decision:** Treatment and diagnostic imaging impact Assisted treatment planning of known recurrences 14/39 (36%) Changed from treatment to palliative (n = 4) Assisted with treatment modality plan (n = 10) In cases with no clinical symptoms and normal CA-125, 3 recurrences identified by PET (8% of population) Negative PET allowed cancellation of SSL in 4 surveillance cases	Disease progression bias (lengthy interval) Verification bias (>1 RS) Review bias (RS, not blinded)

CA-125 = cancer antigen 125; CT = computed tomography; FDG = fluorodeoxyglucose F18; mo = months; PET = positron emission tomography; pts. = patients; RS = reference standard; SLL = second-look laparotomy

6. Pancreatic Cancer

6.1. Background

In 2008 there will be an estimated 37,680 new cases of pancreatic cancer and an estimated 34,290 deaths from the same disease. [133] In 2006 pancreatic cancer was the fourth leading cause of cancer deaths in the United States with a survival rate of four percent. [152] The majority of cases occur in people over the age of 50 and the median age at diagnosis is 72 years. Screening for pancreatic cancer is difficult primarily because there are no tumor markers that can be screened at an early stage of disease. [153]

Pancreatic cancers are staged using both the TNM classification and by clinical assessment (i.e., resectable, locally advanced or metastatic). [153] Assignment of the disease stage is typically determined by a combination of history and physical examination, coupled with CT imaging. In patients deemed to be at high risk for metastasis or for whom the staging is indeterminate, the diagnosis is confirmed by fine needle aspiration or laparoscopy.

The majority of pancreatic cancers are diagnosed at a late stage of disease hampering efforts to provide curative therapy. Early symptoms may include weight loss, jaundice, pain, anorexia, dark urine, nausea, vomiting, and weakness. [152] The majority of patients are treated with palliative care at diagnosis although up to 20 percent of patients present with surgically resectable disease. [154] Painless obstructive jaundice has traditionally been associated with resectable disease. [152] For those patients optimally staged and who have surgery, only 20 percent are expected to survive to 5 years. [154] Chemotherapy has shown to be ineffective in treating metastatic disease and has only limited palliative benefit. [154] Surgical resection remains the only treatment that is potentially curative. [152]

Diagnosing pancreatic cancer at an early stage is critical to providing curative therapy. Research on the molecular aspects of the disease has observed common genetic changes in pancreatic cancer cells which may be useful in developing future screening and treatment technologies. [153] Ultrasonography is often the first diagnostic test performed when patients present with suspected disease with a sensitivity of 95 percent in tumors greater than 3 centimeters. [155] Sensitivity decreases with smaller tumors. [155] Conventional CT is also used for initial imaging; dual-phase helical CT scans have the highest sensitivity (98 percent) in detecting pancreatic malignancies and metastases. [152]

6.2. Importance of Key Questions in the Clinical Management of Pancreatic Cancer

The high mortality rates of pancreatic cancer are associated with the lack of specificity of symptoms that lead to late presentations at the time of diagnosis, the aggressive nature of the disease, and the limitations of current diagnostic procedures. Accurate staging, particularly identification of distant metastases, is of paramount importance in order to properly select patients who are the most likely to benefit from surgery. Currently, dynamic CT or endoscopic pancreatocholangiography are used in the diagnosis of patients with suspected pancreatic cancer. Mass-forming pancreatitis occurs when the pancreatitis-associated inflammation affects a portion of the pancreas, creating the appearance of a mass on imaging tests. Chronic pancreatitis is a risk factor for pancreatic cancer, so mass-forming pancreatitis is frequently found in those patients with suspected pancreatic cancer. Differential diagnosis between pancreatic cancer and pancreatitis is a common problem with imaging modalities. Another problem of CT scans relates to monitoring the treatment response. Pancreatic cancer usually presents cancer cells sparsely scattered in active desmoplastic background and frequently invades major organs including celiac trunk or superior mesenteric vessels with small primary mass. These characteristics of pancreatic cancer make it difficult to determine tumor response following chemotherapy or chemoradiotherapy. Other recent technological advances for the diagnosis of pancreatic cancer include magnetic resonance cholangiopancreatography (MRCP) which may be useful for the noninvasive demonstration of the morphologic contours of the pancreatic duct. However, this tool cannot always detect tumor progression, especially when the tumor is not large enough to be identified accurately, and it is difficult to know the biological activity of the tumor. Early studies have suggested that combined ^{18}FDG-PET/CT without contrast enhancement does not provide additional benefit compared to other diagnostic imaging techniques in the diagnosis of pancreatic cancer; [155] but further evidence needs to be evaluated before firm recommendations are made.

6.3. Results

Seventeen studies[91-107] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for pancreatic cancer. All of the studies evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT, and five studies reported on the diagnostic thinking impact of ^{18}FDG-PET[91,101,103,105] and ^{18}FDG-PET/CT. [95] One study[91] evaluated the effects of ^{18}FDG-PET as part of a management strategy on patient centered outcomes. Finally, one study[95] conducted an economic evaluation on the use of ^{18}FDG-PET/CT for pancreatic cancer. Characteristics of the populations, conditions of

^{18}FDG-PET and ^{18}FDG-PET/CT administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

6.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in pancreatic cancer

Characteristics of the studies

Seventeen studies (fourteen prospective, [91,93-98,100-106] three retrospective[92,99,107]) evaluated the diagnostic accuracy of ^{18}FDG-PET[91,92,94,97-107] and ^{18}FDG-PET/CT[95,96] or both[93] on pancreatic cancer. ^{18}FDG-PET was used for primary diagnosis in five studies, [94,99,100,102,106] for initial staging in two studies, [101,107] for assessing recurrence in one study, [104] and for both primary diagnosis and staging in six studies. [91,92,97,98,103,105] Two studies used ^{18}FDG-PET/CT for both diagnosis and staging, [95,96] while one study used ^{18}FDG-PET and ^{18}FDG-PET/CT for diagnosis, staging and restaging. [93] The studies contained a total of 1,051 patients with sample sizes ranging from 15 to 112 participants. Participant ages ranged from 21 to 93 years. One study reported the distribution by stage of cancer: CS I = 6 percent, CS II = 23 percent, CS III = 65 percent and CS IV = 6 percent. [104] ^{18}FDG-PET was compared to a reference standard that varied across the studies. In fifteen studies the reference standard was either histology/biopsy or clinical followup. [91-101,103-106] Two studies established the final diagnosis of all patients using histology/biopsy. [102,107] One study reported the mean time between last treatment and ^{18}FDG-PET as 12 months. [104] Seven studies used a fixed dose of ^{18}FDG (5 MCi, [107] 120 MBq, [94] 200 MBq, [98] 370 MBq, [91] 400 MBq, [97,102] 444 MBq[105]). Seven studies used a weight-based dose (3 MBq/kg, [100,101] 3.7 MBq/kg, [98] 4 MBq/kg, [93] 5 MBq/kg[96,103,104]). Three studies reported a dose range for ^{18}FDG: 200 to 220 MBq, [106] 260 to 370 MBq, [92] 350 to 450 MBq. [95] The time between injection and PET scan was 20 minutes, [97] 35 to 50 minutes, [102] 60 minutes, [91,93-95,98,100,101,105-107] 60 to 90 minutes, [96] 60 to 120 minutes, [92] and 90 minutes. [103,104] Patients fasted for the following durations: 4 hours, [91,95,107] 5 hours, [98,100] 6 hours, [93,97,101,106] 8 hours, [103,104] 12 hours, [94] and overnight. [92,105] Eight studies[92,93,96,97,100,103-105] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose level allowed was 10 mmol/L, [97] 110 mg/dL, [96,103,104] 120 mg/dL, [105] and 200 mg/dL. [93,100] Methods of interpretation of the images were qualitative in three studies[93,95,101] and both qualitative and quantitative in fourteen. [91,92,94,96-100,102-107] Scans were interpreted qualitatively using visual analysis. [91-98,100-107] Seven studies reported using

SUV, where the criterion for abnormality was SUV > 2.5 g/mL[92,105] SUV > 3 g/mL, [98,102] and SUV > 3.5 g/mL. [96,100,103]

Comparisons

Comparisons for which data were considered for meta-analysis are summarized in Table 23. Pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET for both primary diagnosis and staging purposes, and for primary diagnosis purposes separately. Pooled data were also obtained to evaluate the accuracy of ^{18}FDG-PET/CT for both the primary diagnosis and staging of pancreatic cancer. Individual data are summarized in Appendix D.

Table 23. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer

Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Primary diagnosis and staging	Bang 2006[91]	P	FDG-PET	Histology/biopsy or clinical followup	1. FDG-PET v. histology/biopsy or clinical followup (P studies) [91,93,96-98,103,105] 2. FDG-PET/CT v. histology/biopsy or clinical followup (P studies) [93,95,96]
	Borbath 2005[92]	R	FDG-PET	Histology/biopsy or clinical followup	
	Heinrich 2005[95]	P	FDG-PET/CT	Histology/biopsy or clinical followup	
	Lemke 2004[96]	P	FDG-PET and FDG-PET/CT	Histology/biopsy or clinical followup	
	Lytras 2005[97]	P	FDG-PET	Histology/biopsy or clinical followup	
	Maemura 2006[98]	P	FDG-PET	Histology/biopsy or clinical followup	
	Ruf 2006[103]	P	FDG-PET	Histology/biopsy or clinical followup	
	Sperti 2007[105]	P	FDG-PET	Histology/biopsy or clinical followup	
	Casneuf 2007[93]	P	FDG-PET and FDG-PET/CT	Histology/biopsy or clinical followup	
Primary diagnosis	Giorgi 2004[94]	P	FDG-PET	Histology/biopsy or clinical followup	1. FDG-PET v. all comparators (P studies) [94,100,102,106] 2. FDG-PET v. histology/biopsy or clinical followup (P studies) [94,100,106]
	Mansour 2006[99]	R	FDG-PET	Histology/biopsy or clinical followup	
	Nishiyama 2005[100]	P	FDG-PET	Histology/biopsy or clinical followup	
	Rasmussen 2004[102]	P	FDG-PET	Histology/biopsy	
	van Kouwen 2005[106]	P	FDG-PET	Histology/biopsy or clinical followup	
Recurrences	Ruf 2005[104]	P	FDG-PET	Histology/biopsy or clinical followup	No
Staging	Nishiyama 2005[101]	P	FDG-PET	Histology/biopsy or clinical followup	No
	Wakabayashi 2008[107]	R	FDG-PET	Histology/biopsy	

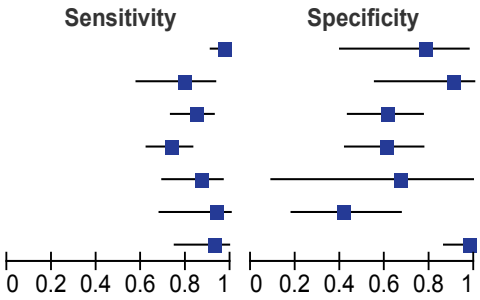
FDG = fluorodeoxyglucose F18; P = prospective; PET=positron emission tomography; R = retrospective

1. ^{18}FDG-PET for the primary diagnosis and staging of pancreatic cancer

Reference standard: histology/biopsy or clinical followup; prospective studies. Seven prospective studies[91,93,96-98,103,105] totaling 479 participants provided data for a meta-analysis of the diagnostic accuracy of ^{18}FDG-PET v. histology/biopsy or clinical followup for the diagnosis and staging of pancreatic cancer. Individual 2x2 table results are presented in Figure 67. Sensitivity ranged from 73 percent[97] to 97 percent. [91] Specificity ranged from 41 percent[103] to 97 percent. [105]

Figure 67. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Bang 2006	90	2	3	7	0.97 [0.91, 0.99]	0.78 [0.40, 0.97]
Casneuf 2007	19	1	5	9	0.79 [0.58, 0.93]	0.90 [0.55, 1.00]
Lemke 2004	54	14	10	22	0.84 [0.73, 0.92]	0.61 [0.43, 0.77]
Lytras 2005	58	13	21	20	0.73 [0.62, 0.83]	0.61 [0.42, 0.77]
Maemura 2006	26	1	4	2	0.87 [0.69, 0.96]	0.67 [0.09, 0.99]
Ruf 2006	14	10	1	7	0.93 [0.68, 1.00]	0.41 [0.18, 0.67]
Sperti 2007	24	1	2	37	0.92 [0.75, 0.99]	0.97 [0.86, 1.00]



We found that, ^{18}FDG-PET had a pooled positive LR of 2.77 (95% CI = 1.62, 4.73) and a pooled negative LR of 0.19 (95% CI = 0.10, 0.34) to accurately diagnose and identify the stage of pancreatic cancer (Figures 68 and 69). There was considerable heterogeneity in the positive (p = 0.003; I^2 = 70 percent) and negative (p = 0.004, I^2 = 68 percent) LRs across the studies, which limits our ability to draw conclusions about the overall accuracy of ^{18}FDG-PET.

Figure 68. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

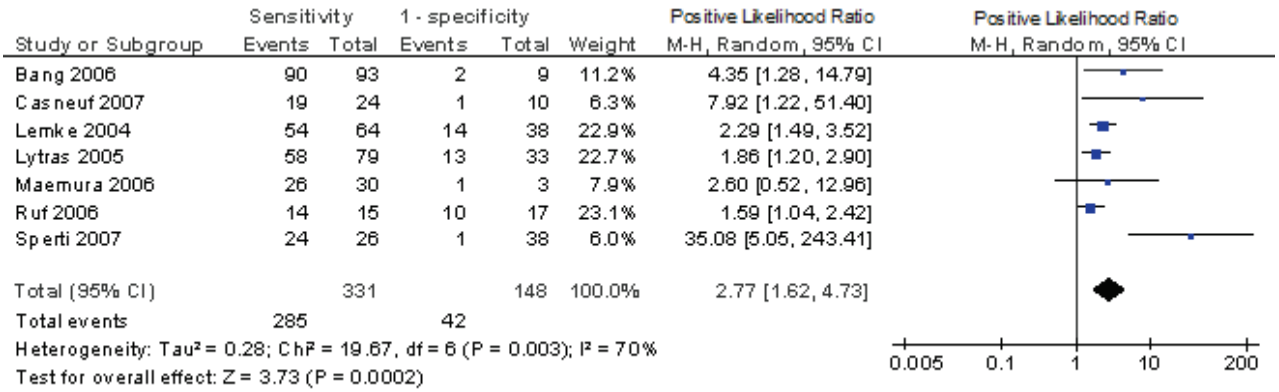


Figure 69. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

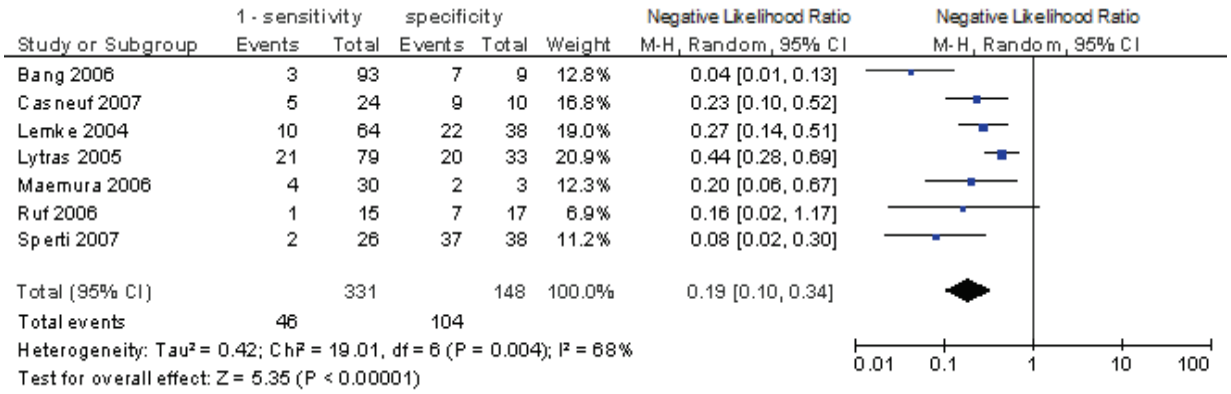
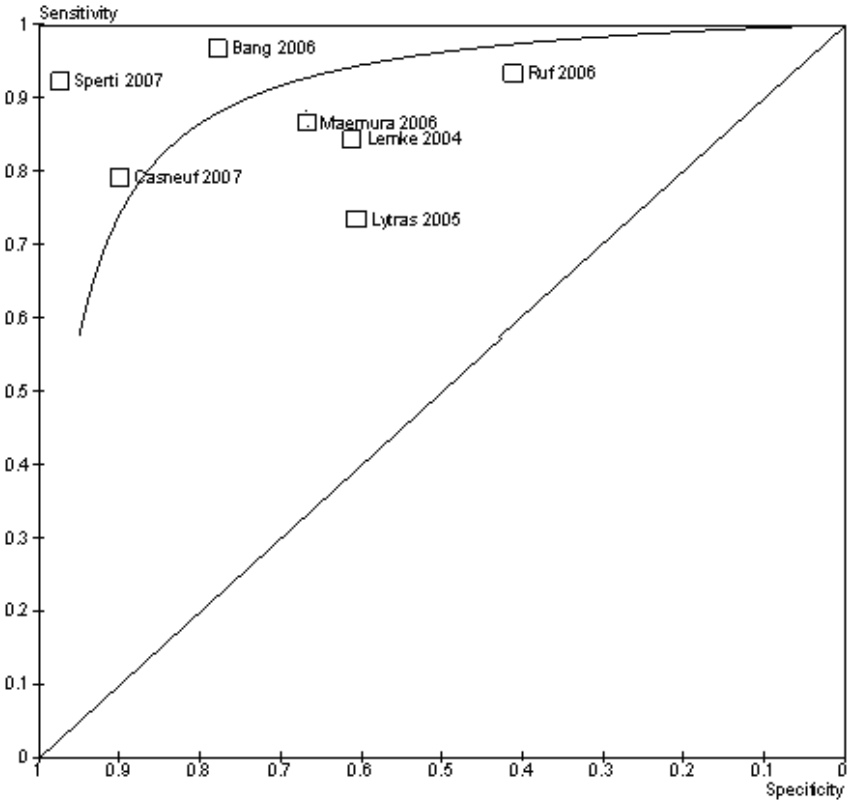


Figure 70 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer based on prospective studies.

Figure 70. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

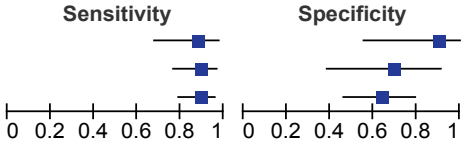


2. ^{18}FDG-PET/CT for the primary diagnosis and staging of pancreatic cancer

Reference standard: histology/biopsy or clinical followup; prospective studies. Three prospective studies[93,95,96] totaling 193 participants provided data for a meta-analysis of the diagnostic accuracy of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the diagnosis and staging of pancreatic cancer. Individual 2x2 table results are presented in Figure 71. Sensitivity was 89 percent in all the individual studies. Specificity ranged from 64 percent[96] to 90 percent. [93]

Figure 71. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Casneuf 2007	21	1	3	9	0.88 [0.68, 0.97]	0.90 [0.55, 1.00]
Heinrich 2005	41	4	5	9	0.89 [0.76, 0.96]	0.69 [0.39, 0.91]
Lemke 2004	57	13	7	23	0.89 [0.79, 0.95]	0.64 [0.46, 0.79]



We found that ^{18}FDG-PET/CT had a pooled positive LR of 2.69 (95% CI = 1.84, 3.94) and a pooled negative LR of 0.16 (95% CI = 0.10, 0.26) to accurately diagnose and identify the stage of the disease (Figures 72 and 73). Both positive and negative LRs were homogeneous across the studies.

Figure 72. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

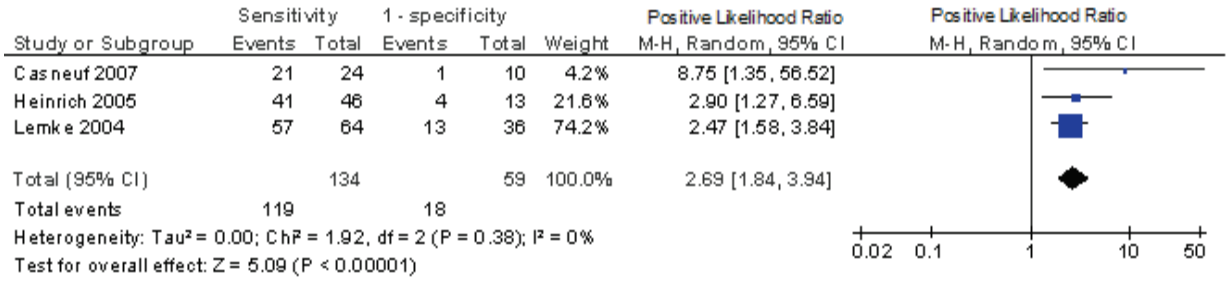


Figure 73. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

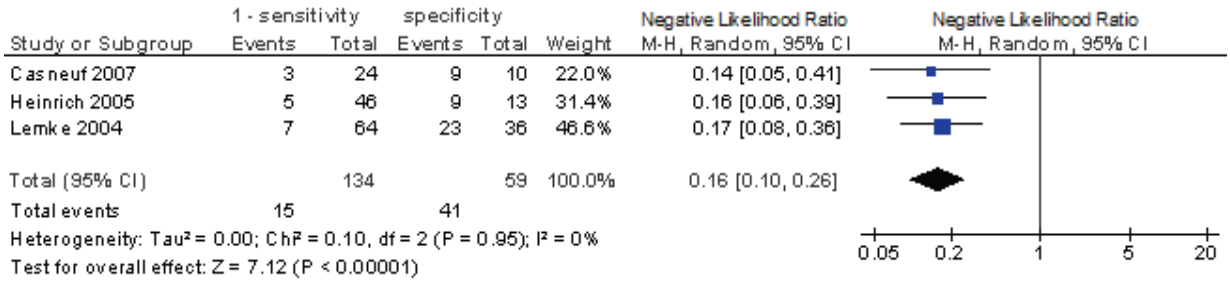
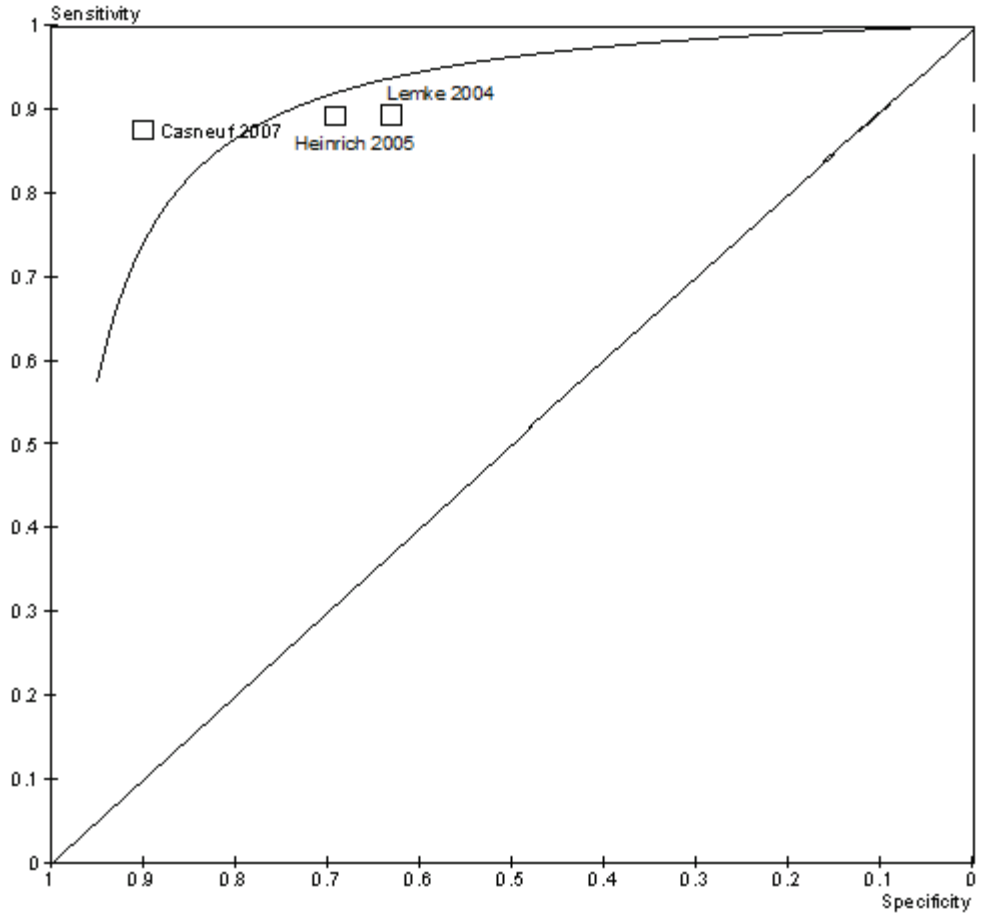


Figure 74 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer, based on prospective studies

Figure 74. Summary ROC plot of ^{18}FDG-PET/CT v. histology/biopsy or clinical followup for the primary diagnosis and staging of pancreatic cancer (prospective studies)

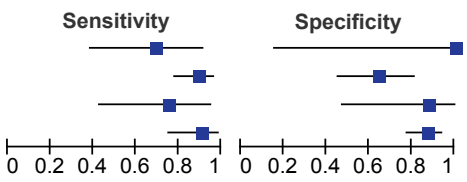


3. ^{18}FDG-PET for the primary diagnosis of pancreatic cancer

Reference standard: any; prospective studies. Four prospective studies[94,100,102,106] totaling 230 participants provided data for a meta-analysis of the diagnostic accuracy of ^{18}FDG-PET compared to a variety of reference standards for the primary diagnosis of pancreatic cancer. Individual 2x2 table results are presented in Figure 75. Sensitivity ranged from 69 percent[94] to 91 percent. [106] Specificity ranged from 65 percent[100] to 100 percent. [94]

Figure 75. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Giorgi 2004	9	0	4	2	0.69 [0.39, 0.91]	1.00 [0.16, 1.00]
Nishiyama 2005	49	11	6	20	0.89 [0.78, 0.96]	0.65 [0.45, 0.81]
Rasmussen 2004	9	1	3	7	0.75 [0.43, 0.95]	0.88 [0.47, 1.00]
van Kouwen 2005	29	10	3	67	0.91 [0.75, 0.98]	0.87 [0.77, 0.94]



We found that, when all the reference standards were considered, ^{18}FDG-PET had a pooled positive LR of 4.28 (95% CI = 2.07, 8.86) and a pooled negative LR of 0.21 (95% CI = 0.12, 0.40) to accurately diagnose the disease (Figures 76 and 77). There was moderate heterogeneity in the positive ($p = 0.07$; $I^2 = 58$ percent) and negative ($p = 0.16$, $I^2 = 42$ percent) LRs across the studies, making it difficult to draw conclusions about the overall accuracy of ^{18}FDG-PET.

Figure 76. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies)

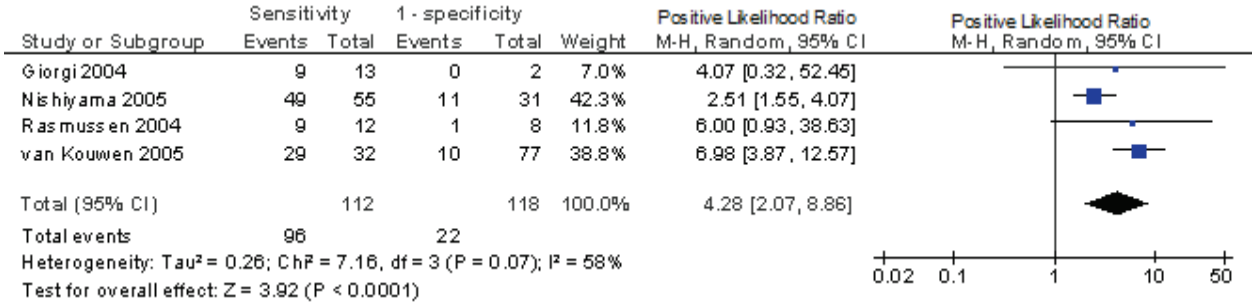


Figure 77. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies)

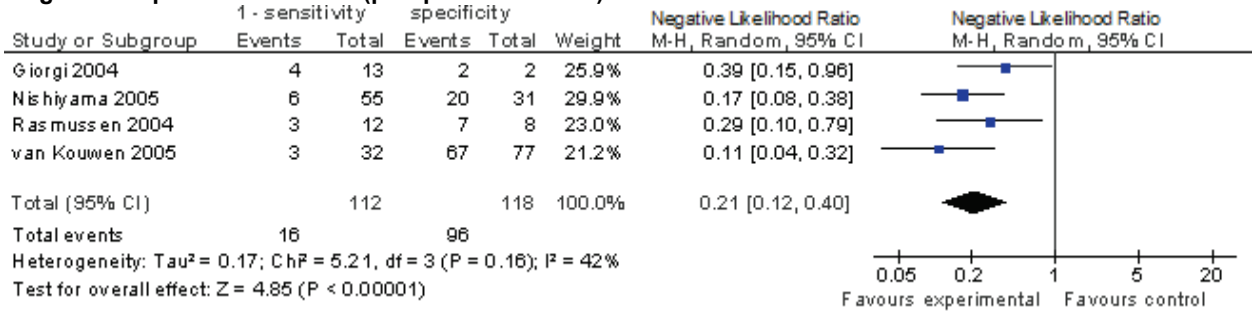
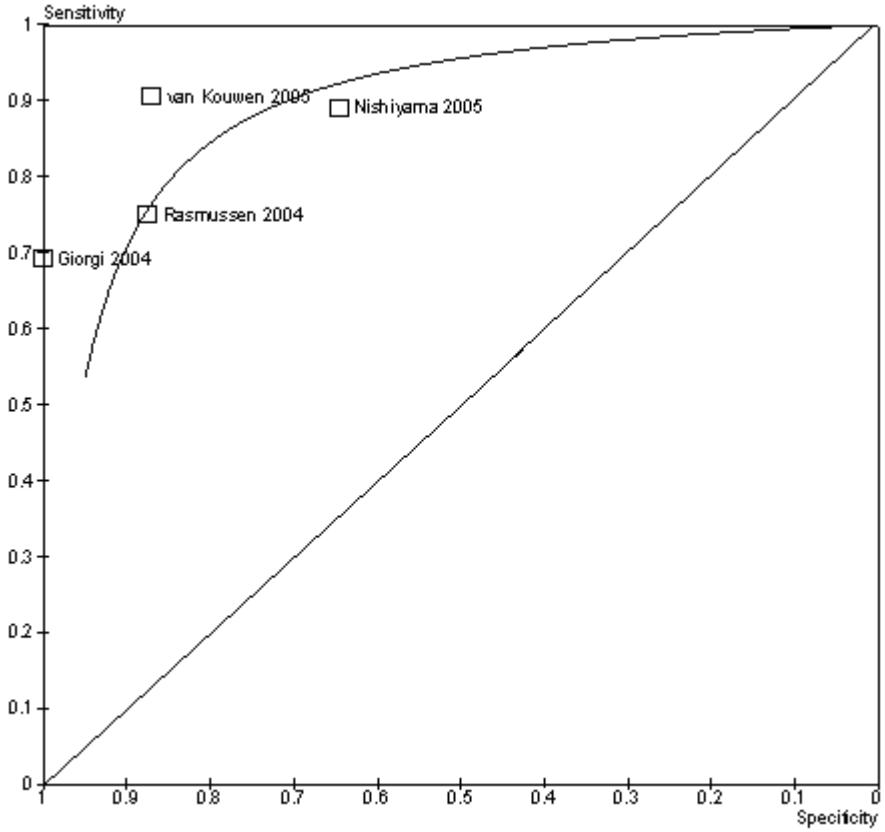


Figure 78 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. all the reference standards for the primary diagnosis of pancreatic cancer, based on prospective studies.

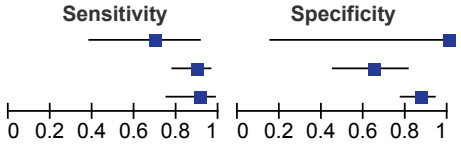
Figure 78. Summary ROC plot of ^{18}FDG-PET v. any reference standard for the primary diagnosis of pancreatic cancer (prospective studies)



Reference standard: histology/biopsy or clinical followup, prospective studies (subgroup analysis). Three prospective studies[94,100,106] totaling 210 participants provided data for a subgroup analysis of the diagnostic accuracy of ^{18}FDG-PET when only histology/biopsy or clinical followup was used as the reference standard for the primary diagnosis of pancreatic cancer. Individual 2x2 table results are presented in Figure 79. Sensitivity ranged from 69 percent[94] to 91 percent. [106] Specificity ranged from 41 percent[103] to 97 percent. [105]

Figure 79. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer

Study	TP	FP	FN	TN	Sensitivity	Specificity
Giorgi 2004	9	0	4	2	0.69 [0.39, 0.91]	1.00 [0.16, 1.00]
Nishiyama 2005	49	11	6	20	0.89 [0.78, 0.96]	0.65 [0.45, 0.81]
van Kouwen 2005	29	10	3	67	0.91 [0.75, 0.98]	0.87 [0.77, 0.94]



When ^{18}FDG-PET was compared to histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer, the pooled positive LR was 4.11 (95% CI = 1.74, 9.70) and the

pooled negative LR was 0.9 (95% CI = 0.09, 0.44) (Figures 80 and 81). There was considerable heterogeneity in the positive (p = 0.03; I^2 = 71 percent) and negative (p = 0.09, I^2 = 58 percent) LRs across the studies, making it difficult to draw conclusions about the overall accuracy of ^{18}FDG-PET.

Figure 80. Meta-analysis of the positive likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies)

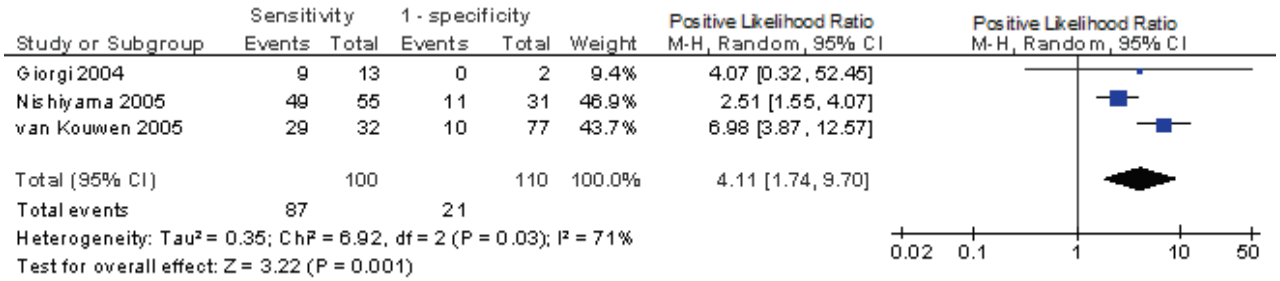


Figure 81. Meta-analysis of the negative likelihood ratio of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies)

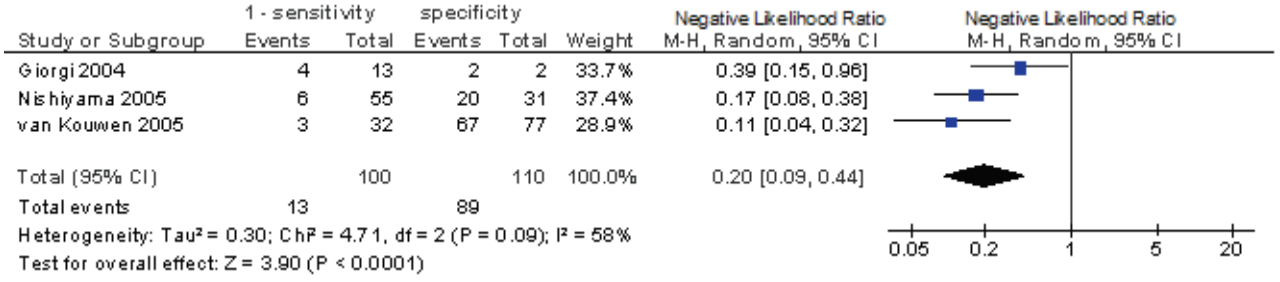
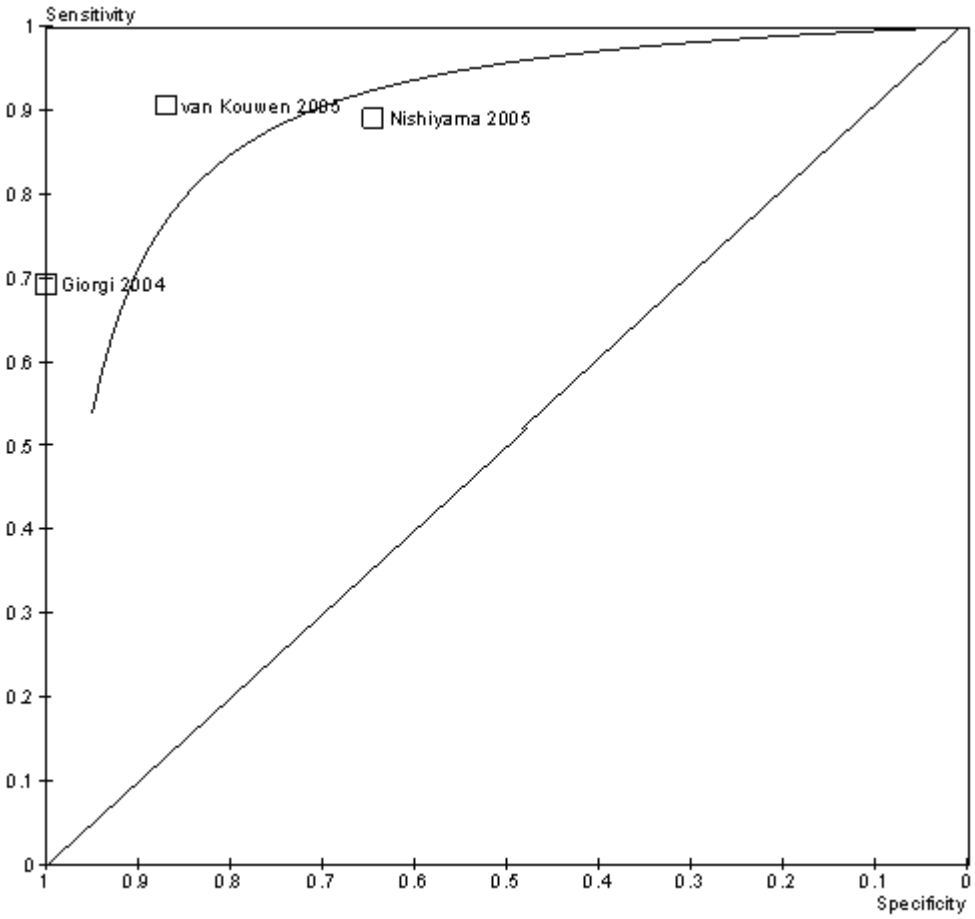


Figure 82 shows the estimates of sensitivity and 1-specificity plotted in ROC space for ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer

Figure 82. Summary ROC plot of ^{18}FDG-PET v. histology/biopsy or clinical followup for the primary diagnosis of pancreatic cancer (prospective studies)



Summary of the results

Meta-analyses were calculated to evaluate the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for both the primary diagnosis and staging of pancreatic cancer (Table 24). The findings were consistently significant, suggesting that both ^{18}FDG-PET and ^{18}FDG-PET/CT are useful for primary diagnosis and staging. However, the pooled results are, for the most part, heterogeneous; therefore the magnitude of the effect is uncertain. Further, there is no clear evidence for the choice of ^{18}FDG-PET v. ^{18}FDG-PET/CT, since the observed heterogeneity indicates considerable uncertainty in these estimates. When ^{18}FDG-PET was evaluated for primary diagnosis purposes only, the positive likelihood radio was slightly better for ruling in the disease, but the negative LR remained almost the same. Evidence on the use of ^{18}FDG-PET for recurrences and staging is derived from individual

study data and therefore, firm conclusions about the utility of ^{18}FDG-PET for these indications cannot be made.

Table 24. Results of meta-analyses of the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer

PET Purpose	Type of PET	Reference standard	Design	Studies	N	Effect estimate M-H, Random, 95% CI
Primary diagnosis/ staging	FDG-PET	Histology/biopsy or clinical followup	P	7	479	PLR = 2.77 [1.62, 4.73]
						NLR = 0.19 [0.10, 0.34]`
	FDG-PET/CT			3	193	PLR = 2.69 [1.84, 3.94]
						NLR = 0.16 [0.10, 0.26]
Primary diagnosis	FDG-PET	Any reference standard	P	4	230	PLR = 4.28 [2.07, 8.86]
						NLR = 0.21 [0.12, 0.40]
		Histology/biopsy or clinical followup		3	210	PLR = 4.11 [1.74, 9.70]
						NLR = 0.20 [0.09, 0.44]

CI = confidence interval; FDG = fluorodeoxyglucose F18; M-H = Mantel Hantzel; NLR = negative likelihood ratio; P = prospective; PET = positron emission tomography; PLR = positive likelihood ratio

6.3.2. Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making with respect to diagnosis and management strategy for patients with pancreatic cancer

Five studies evaluated the diagnostic thinking impact of ^{18}FDG-PET[91,101,103,105] and ^{18}FDG-PET/CT. [95] with regards to patient management and diagnostic work-up of pancreatic cancer. The imaging by ^{18}FDG-PET was used for both initial diagnostic (^{18}FDG-PET, 3 studies;[91,103,10518]FDG-PET/CT, 1 study[95] and staging purposes (^{18}FDG-PET, 4 studies;[91,101,103,105] ^{18}FDG-PET/CT, 1 study[95]).

Bang et al. [91] evaluated the clinical impact of using ^{18}FDG-PET on the diagnosis and staging of pancreatic cancer and on monitoring tumor response to chemoradiation. One hundred and two patients undergoing evaluation for suspected primary pancreatic cancer were prospectively enrolled in this study. Of these, 93 patients with confirmed pancreatic cancer were assessed for the diagnostic thinking impact of ^{18}FDG-PET. ^{18}FDG-PET findings led to a change in the pre-treatment staging in 25 patients (27 percent). For 20 patients, the treatment management was altered by changing the resectability status. The majority of changes (17/20) were due to the identification of previously unsuspected metastases, and resulted in the cancellation of previously planned surgical resection. Of particular note, 8/17 of the newly identified metastases were sites in the liver and had not been detected by the initial dynamic CT scan. Three cases previously considered un-resectable were downstaged and considered to be treatable following ^{18}FDG-PET. These findings were subsequently confirmed by biopsy.

The authors concluded that ^{18}FDG-PET was a sensitive and specific imaging modality which would be a good adjunct to conventional imaging techniques. They noted that it was particularly sensitive in the detection of small unsuspected hepatic lesions relative to conventional imaging by CT scan or ultrasonography. ^{18}FDG-PET was useful in the reassessment of conventionally staged tumors and treatment decisions regarding resectability were amended, which allowed for the cancellation of unnecessary surgeries.

This prospective study was determined to be of moderate quality. Methodological strengths included: a clear description of the selection criteria, as well as sufficient description of the choice of the standard, blinded interpretation of the index test, accounting for all participants and intermediate test results. Weaknesses included a partial description of the spectrum of included patients, as well as an incomplete description of the execution of the index test or reference standards and lack of clarity about the period between the execution of ^{18}FDG-PET and the reference standard. These weaknesses could have led to spectrum and disease progression bias. Additionally, there was more than one standard used for confirmation of the true disease status, which may have introduced verification bias. Of particular concern is that the physicians interpreting the results of the reference standard were not described as blinded to the results of the ^{18}FDG-PET, possibly introducing review bias.

Heinrich et al. [95] investigated the treatment decision impact of integrated ^{18}FDG-PET/CT on the diagnosis and staging of pancreatic cancer. Fifty-nine consecutive patients with focal lesions in the pancreas were prospectively enrolled in this study. Of the 37 patients who were judged to have resectable pancreatic cancer, treatment management changed in six patients (16 percent) as a result of ^{18}FDG-PET/CT findings. In addition, of the 46 patients who were found by the reference standard to have malignant pancreatic lesions, ^{18}FDG-PET/CT findings changed management in 15 patients (33 percent). ^{18}FDG-PET/CT resulted in changes to significantly more treatment decisions compared to standard staging (9/46 cases; 20 percent [$p = 0.03$]). ^{18}FDG-PET/CT also identified 17 benign lesions. Although the detection of these lesions did not impact treatment, scan results occasionally prompted further diagnostic evaluation, including biopsies. In addition, ^{18}FDG-PET/CT improved detection of distant metastases; these were diagnosed in 13 patients, of which five were solely identified by ^{18}FDG-PET/CT. In two patients, cancer was found by ^{18}FDG-PET/CT only and had not been previously identified on physical examination. As a result, the surgical treatment was changed for both patients.

The authors concluded that ^{18}FDG-PET/CT significantly changed the overall management of patients with pancreatic cancer when compared to standard staging. Based on their clinical and economic evaluation, the authors stated that preoperative staging use of ^{18}FDG-PET/CT is beneficial and may advance standard staging.

This prospective study of consecutively enrolled patients was determined to be of moderate quality. The methodological strengths of this study include a short time period between the ^{18}FDG-PET/CT and reference standard, as well as sufficient description of the choice of the standard, the execution of the index test, accounting for all participants and intermediate test results. However, the execution of the reference standard, spectrum of patients included in the study, and selection criteria were not described in adequate detail. Thus, both spectrum bias and selection bias may have affected the results of this study. Furthermore, all patients did not receive the same reference standard, which may have introduced verification bias. Of particular concern is that the physicians interpreting the results of the reference standard were not blinded to the results of the ^{18}FDG-PET/CT, and blinding of the ^{18}FDG-PET/CT interpretation was not clearly reported.

Nishiyama et al. [101] examined the impact of ^{18}FDG-PET used in the diagnosis and staging of pancreatic cancer on treatment decisions. Forty-two consecutive patients with histopathologically confirmed pancreatic cancer and no previous treatment were prospectively enrolled. ^{18}FDG-PET had an impact on treatment management in 5 of 42 patients (12 percent). Three patients were altered from curative to palliative treatment, while two other patients were changed from palliative to curative treatment.

The authors recommended routine ^{18}FDG-PET for preoperative staging of patients with potentially resectable pancreatic cancer. This could result in a marked increased detection of metastases. However, the authors also stated the CT and ^{18}FDG-PET have a complementary role in the identification of distant metastases and ^{18}FDG-PET alone does not provide sufficient information for staging. This article was written prior to readily available access to integrated ^{18}FDG-PET/CT scanners and the authors speculated that such hybrid scanners would increase detection and localization.

This prospective study of consecutively enrolled patients was assessed to be of high quality. Components that were well addressed included description of the spectrum of patients, choice of reference standard, details of the execution of index test and reference standard, and clear reporting of all participants and test results. In addition, the time interval between the index and reference

standard was judged to be sufficiently brief to avoid significant change in patients' conditions. Although all patients received verification of their disease status by a reference test, they did not all receive the same standard, introducing the possibility of verification bias. The index test was interpreted without knowledge of the result of the reference standard, but the interpretation of the reference standard was unblinded. An additional weakness was the limited description of the selection criteria, raising uncertainty as to what these criteria were and how they were applied.

Ruf et al. [103] evaluated the treatment decision influence of ^{18}FDG-PET/MRI fusion on the diagnosis and staging of pancreatic cancer. The study prospectively enrolled 32 adult patients suspected of having a pancreatic mass. In 8 of 32 patients (25 percent), topographical assignment and interpretation of ^{18}FDG-PET foci was improved through fusion of ^{18}FDG-PET and MRI images. However, image fusion only resulted in a change of treatment in one patient, for whom surgery was performed with a curative intent. The remaining seven patients did not have a change in treatment as a result of ^{18}FDG-PET/MRI image fusion due to inoperability (n = 2), other medical reasons (n = 1), other metastases being present in other regions thus preventing curative surgery (n = 2), and image fusion having no influence on the planned palliative surgery (n = 2).

The authors concluded the ^{18}FDG-PET/MRI improved assignment of foci but had minimal therapeutic consequences. This was mainly attributed to the small number of patients most of whom had multiple lesions that prevented curative treatment. It is possible that a larger treatment impact would be observed in patients with small, resectable primaries and only peripancreatic lymph node manifestations.

This prospective study was assessed to be of moderate quality. Many elements of the study were well described, including the selection criteria, choice of reference standard, as well as the period between execution of the index test and reference standard. In addition, intermediate test results and study withdrawals were reported. The spectrum of patients included in the study was not well documented and it was unclear from where they were recruited or referred. Therefore, it is possible that spectrum bias may have occurred. Also, both the index and reference standard tests were interpreted in an unblinded manner, thereby increasing the risk of review bias in this study.

A final study by Sperti et al. [105] investigated the treatment decision impact of integrated ^{18}FDG-PET on the diagnosis of pancreatic cancer. A prospective sample of 71 patients with suspected intraductal papillary mucinous neoplasms (IPMN) of the pancreas underwent ^{18}FDG-PET scans. Of these, 64 had ^{18}FDG-PET scans available and were included in the analysis. The treatment plan was

substantially altered in 44 (69 percent). Positive ^{18}FDG-PET results impacted treatment decisions in 10 patients; seven (11 percent) underwent surgical resection, two patients with hepatic metastases not evident in CT avoided laparotomy, and one patient underwent resection of a borderline IPMN associated with unsuspected colon cancer. Negative ^{18}FDG-PET results prompted changes in treatment management in 34 patients; 19 (30 percent) were scheduled for followup and 15 (23 percent) underwent more limited resection (six had a more conservative resection and nine avoided splenectomy).

The authors concluded that ^{18}FDG-PET was superior to conventional imaging techniques in its ability to identify patients with pancreatic IPMN for surgical intervention or followup. This is particularly true for asymptomatic patients.

This prospective study was assessed to be of high quality. The spectrum of patients was representative of those who would receive this test in practice and the choice of reference standard was appropriate and independent of the index test. There was good reporting of the execution of the index and reference tests, any intermediate test results and study withdrawals. Both the index test and reference standard were interpreted in a blinded manner. Although all patients received a reference standard, this reference was not the same for all patients; some received histological verification of disease and others clinical followup. Therefore, verification bias may have affected the results of this study. The main weaknesses of the study were lack of clarity on the duration between index and reference tests and only partial description of the selection criteria.

6.3.3. ^{18}FDG-PET as part of a management strategy in pancreatic cancer

Bang et al. [91] additionally examined using ^{18}FDG-PET to monitor patient response to concurrent chemoradiation. The characteristics of the study population (n=93) and quality have been discussed in detail in the section above. The outcomes of a subset of 15 (16 percent) patients diagnosed with pancreatic cancer who were followed with pre- and post-treatment imaging were included in this analysis. The remaining 78 patients did not receive concurrent chemoradiation for reasons not specified. ^{18}FDG-PET was compared to dynamic CT scans, serial serum CA19-9 measurements and a clinical benefit score determined by a series of quantitative and qualitative measurements (intensity of pain, analgesic use, Karnofsky performance scale and body weight). The authors evaluated whether treatment response could be determined by followup CT and ^{18}FDG-PET assessments in all patients, and correlated patient response status with the time to disease progression. Response was

judged as complete if disease sites disappeared and partial if lesions were reduced in size (CT scan) or ^{18}FDG uptake (^{18}FDG-PET).

No patients were considered to be "responders" based on the CT scan results; however, five cases were judged to be "responders" by ^{18}FDG-PET. The time to disease progression was significantly longer in the ^{18}FDG-PET "responders" group as compared to the ^{18}FDG-PET "non-responders." The mean time to progression in the "responders" was 399 days (95% CI, 282 to 526) v. 233 days (95% CI, 181 to 235) in the "non-responders." The clinical benefit score and serial changes in CA19-9 did not correlate significantly with the results of either imaging modality.

The authors concluded that ^{18}FDG-PET was more accurate than dynamic CT scan for determining treatment response to concurrent chemoradiotherapy. Neither clinical benefit score nor serum CA19-9 measurements were found to predict treatment response. However, it should be noted that while the overall study sample was large, there was only a small number of patients (n = 15) included in this analysis of ^{18}FDG-PET impact on management strategy relating patient-centered outcomes.

6.3.4. Cost-effectiveness of ^{18}FDG-PET/CT for pancreatic cancer

Heinrich et al. [95] examined cost savings with use of ^{18}FDG-PET/CT in addition to routine diagnostic procedures to determine staging and eligibility for surgery among patients with presumed resectable pancreatic cancer. The authors conducted a secondary analysis of patient data that had been collected as part of a phase II clinical trial evaluating neoadjuvant chemotherapy for resectable pancreatic cancer. The sample included 59 patients who had a focal pancreatic lesion or suspected pancreatic cancer and had undergone ^{18}FDG-PET/CT. Accuracy data for the diagnostic tests were derived from the trial data; diagnosis was confirmed through intraoperative findings and results of histology or biopsy. Cost data were obtained from the hospital accounting department. A cost-benefit analysis considering direct costs during the staging and peri-operative period was conducted from a hospital perspective. Among the 59 patients, ^{18}FDG-PET/CT detected metastasis in five patients who were then deemed ineligible for surgery. This resulted in cost savings of US$1,066 per patient. Cost savings were higher (US$2,844 per patient) with selective use of ^{18}FDG-PET/CT among patients identified as surgical candidates through standard, routine staging procedures. Results of sensitivity analyses for shorter length of stay, type of fine-needle aspiration, and surgical confirmation of metastasis were consistent in demonstrating cost savings. The study was based on

Swiss data and practice patterns; however, the authors suggested that results may be generalizable to other centers in Europe and the United States. The authors presented results and conclusions within the stated objectives and given data. The primary and sensitivity analyses were restricted to a limited number of costs and outcomes.

Summary of the results

Five prospective studies of moderate to high quality (four on ^{18}FDG-PET[91,101,103,105] and one[95] on ^{18}FDG-PET/CT) provided evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT). One study[91] evaluated the use of ^{18}FDG-PET as part of a management strategy for pancreatic cancer. ^{18}FDG-PET and ^{18}FDG-PET/CT were used for both the primary diagnosis and staging of pancreatic cancer in these studies, with the exception of one study, [101] which evaluated ^{18}FDG-PET for staging purposes only. The management plan was altered in an important number of patients with both ^{18}FDG-PET and ^{18}FDG-PET/CT (up to 69 percent), more often resulting in a conservative course of management thus avoiding unnecessary surgery. ^{18}FDG-PET and ^{18}FDG-PET/CT may be used for the appropriate selection of surgical candidates through more accurate identification of metastases or identification of resectable disease in otherwise asymptomatic patients. One study demonstrated that ^{18}FDG-PET/CT for the identification of surgical candidates would result in cost savings. [95] One study[103] showed no advantage in terms of therapeutic consequences with the addition of MRI images to FDG-PET for diagnosis and staging. Only one study involving 15 patients evaluated patient-centered outcomes, [91] hence the value of FDG-PET in terms of patient-centered outcomes in pancreatic cancer remains unclear.

Table 25 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact, effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT and economic outcomes for pancreatic cancer.

Table 25. Main findings and types of bias that affected the evidence on the diagnostic thinking impact, effect on patient-centered outcomes and economic outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for pancreatic cancer

Study	Results of FDG-PET imaging on patient diagnosis, treatment and outcomes	Types of bias
Bang 2006[91] Study type: Prospective	Treatment strategy and staging was impacted for 25/93 cases (27%): Upstaged: 20/25 changes Downstaged: 5/25 changes Treatment modality changed in 20/25 cases (80%): Upstaged and deemed to be unresectable: 17/20 Downstaged and deemed to be resectable: 3/20 Previously unidentified distant metastases were found in the 17 cases determined to be unresectable	Spectrum bias (unclear) Verification bias (>1 RS)
Heinrich 2005[95] Study type: Prospective	**Management decision**: Treatment and diagnostic testing impact Treatment strategy changed for 6/37 patients (16%) judged to have resectable cancer. Distant metastasis detected by PET/CT only (n = 5) Simultaneous cancer found and led to change in surgery (n = 2, one with curative intent, one palliative) PET/CT enabled minimally invasive histological assessment by exact anatomic delineation of lesions Detected benign lesions in 17 patients, 10 of which were not identified by conventional CT. Some lesions required further diagnostic evaluation and no change in treatment made **Economic evaluation** Alternatives compared: a) Standard, routine staging; b) FDG-PET/CT + standard staging PET/CT identified metastasis and avoided surgery in 5/59 patients Total net savings from PET/CT: $62,912 ($1,066 per patient) Total net savings for patients eligible for surgery after routine staging: $105,262 ($2,844 per patient)	Spectrum bias (unclear) Selection bias (unclear) Verification bias (>1 RS) Review bias (PET, unclear; RS, unblinded)
Nishiyama 2005[101] Study type: Prospective	**Management decision:** Treatment Treatment management impacted in 5/42 patients (12%) Changed from curative to palliative treatment (n = 3) Changed from palliative to curative treatment (n = 2)	Selection bias (unclear) Verification bias (>1 RS) Review bias (RS, unclear if blinded)
Ruf 2006[103] Study type: Prospective	**Management decision:** Treatment and diagnostic testing impact Interpretation of PET foci improved through fusion of PET/MRI images 8/32 patients (25%) Image fusion resulted in a change of treatment in only 1 patient (surgery was expanded to curative)	Spectrum bias (unclear) Verification bias (>1 RS) Review bias (PET and RS unblinded)
Sperti 2007[105] Study type: Prospective	**Management decision:** Treatment Treatment plans were altered in 44/64 patients (69%) Positive PET results impacted treatment in 10 patients Negative PET results impacted management in 34 patients	Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS)

CT = computed tomography; FDG = fluorodeoxyglucose F18; MRI = magnetic resonance imaging; PET = positron emission tomography; RS = reference standard

7. Prostate Cancer

7.1. Background

Cancer of the prostate is the most common cancer in men. [156] In developed countries, it is the second most frequently diagnosed cancer and the third most common cause of death from cancer in men. In 2004, the incidence of invasive prostate cancer was 145.3 per 100,000 with a death rate of 25.4 per 100,000. [157] It is estimated that 186,320 new cases will be diagnosed in the United States in 2008[133] and there will be 28,660 attributable deaths. [133] Mortality rates in African-American men are more than twice the rates observed in other racial and ethnic groups in the United States. [158]

Diagnosis of prostate cancer begins with the assessment of general health and comorbidities. [156] Prostate cancer screening is controversial because of the lack of definitive evidence of benefit. [159] The digital rectal examination (DRE) was the test traditionally used for screening; however, two other procedures, transrectal ultrasound (TRUS) imaging and serum prostate-specific antigen (PSA) concentrations are now commonly used. Imaging procedures, such as ultrasound, CT and MRI, have been suggested as possible screening modalities. Each modality has relative merits and disadvantages for distinguishing different features of prostate cancer.

The PSA test has been widely adopted in the United States in the management of prostate cancer. This test, which measures the amount of PSA protein in the blood (often elevated in patients with prostate cancer), [158] is used as a disease marker although no specific cut-off point for normal PSA has been defined. [160] This screening method can lead to overdiagnose the disease, [156,158] and the effect of early detection and treatment on mortality is not fully understood because of the long natural history of prostate cancer and the inherent delay in measurable treatment effects. [160] Bone scintigraphy may be performed if bone metastases are suspected. [156]

The most common grading system used in the United States is the Gleason grading system. [161] Biopsy material is needed to assess the Gleason score. The system uses a summary score between 2 and 10 (10 being the most aggressive) of the two most common patterns of tumor growth in a biopsy specimen (one for more than 50 percent of the growth and one for the majority of the remaining tumor growth), which are each given a score of 1 to 5 (5 being the most aggressive). [160] The summary score reflects the addition of these two scores, with a higher score being indicative of more disordered growth and aggressive cancer.

Two other systems that are used for the staging of prostate cancer are the Jewett system (stages A through D) [162] and a revised TNM system that employs the same broad T stage categories as the Jewett system but includes subcategories of T stage, such as a stage to describe patients diagnosed through PSA screening. This revised TNM system is clinically useful and more precisely stratifies newly diagnosed patients. [163] Local staging (T stage) is evaluated by DRE. [156,160] Pelvic imaging using MRI or CT is performed before radical treatment when Partin tables (probabilities of disease extension and progression) indicate more than 15 percent risk of nodal involvement. [156]

Rigorous evaluation of any prostate cancer screening modality is desirable because the natural history of the disease is variable and appropriate treatment is not clearly defined. [159] Clinical practice guidelines on the management of clinically localized prostate cancer demonstrate major differences in their specific recommendations, [164] and there is no general consensus as to what constitutes best treatment for localized disease. [156] Little high-quality evidence is available to guide decisions regarding the comparative effectiveness and harms of treatments for clinically localized prostate cancer, especially in men with PSA-detected disease. [165]

Radical prostatectomy (removal of the entire prostate and potential removal of nearby lymph nodes), [166] radiotherapy[160] and hormone therapy (androgen suppression or bicalutamide monotherapy) [156] are the main treatments for locally advanced prostate cancer. Cryosurgery is a surgical technique under development that involves destruction of prostate cancer cells by intermittent freezing of the prostate tissue with cryoprobes, followed by thawing. It is less well established than standard prostatectomy and long-term outcomes are not as well established as with prostatectomy or radiation therapy. [167]

Improvements in brachytherapy have made it an effective radiotherapy for early-stage prostate cancer. Advances in hormonal therapy for prostate cancer have included the development of gonadotropin-releasing hormone (GnRH) agonists, which inhibit the ability of the pituitary gland to stimulate the testes to make testosterone. Additional approaches include bilateral orchiectomy, estrogen therapy, antiandrogens, ketoconazole, and aminoglutethimide. [159] Advances have also been made in chemotherapy for advanced prostate cancer. [158] There is high-quality evidence from one RCT in favor of surgery over watchful waiting with palliative intent for non-high grade localized prostate cancer. [168] Data from RCTs indicate that men with Gleason scores of 8 to 10 were likely to have evidence of biochemical recurrence, regardless of whether treatment was radical prostatectomy alone or was combined with androgen deprivation. High-dose electron beam radiation therapy

(EBRT) was more effective than conventional-dose EBRT in controlling biochemical failure in both low-risk disease and higher-risk disease. [165]

7.2. Importance of Key Questions in the Clinical Management of Prostate Cancer

Tumor grading of prostate cancer is a fundamental determinant of disease biology and prognosis. Implementation of an accurate noninvasive imaging technique to detect recurrent and metastatic prostate cancer is critical for the effective management of these patients. Current imaging tests in prostate cancer include ultrasound, CT, MRI, and In-111 capromab pendetide scan. There are still controversies regarding the value of ^{18}FDG-PET to identify local recurrences, metastases or nodal and soft tissue lesions. ^{18}FDG-PET imaging in prostate cancer can be problematic because ^{18}FDG tracer undergoes renal excretion with subsequent accumulation in the urinary bladder, causing image artifacts in the lower pelvis. The close proximity of excreted ^{18}FDG to sites of potential local recurrence (i.e., the prostate bed and adjacent lymph nodes) complicates the interpretation of ^{18}FDG-PET images of the pelvis. Furthermore, ^{18}FDG accumulation in the primary prostate cancer is generally low, and may overlap with the uptake in benign prostatic hyperplasia (BPH) and uptake in the normal gland. It is important to evaluate the utility of ^{18}FDG-PET in patients with suspected or known prostate cancer, and the impact on management and patient outcomes.

7.3. Results

Four studies[108-111] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for prostate cancer. All of them evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT. None of the studies reported on the diagnostic thinking impact of ^{18}FDG-PET or ^{18}FDG-PET/CT or evaluated the impact of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET and ^{18}FDG-PET/CT for prostate cancer. Characteristics of the populations, conditions of ^{18}FDG-PET and ^{18}FDG-PET/CT administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

7.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in prostate cancer

Characteristics of the studies

Four studies (two prospective, [109,110] two retrospective[108,111]) evaluated the diagnostic accuracy of ^{18}FDG-PET[108-111] and ^{18}FDG-PET/CT[111] on prostate cancer. ^{18}FDG-PET was used for restaging in one study, [108] for the assessment of recurrences in one study, [110] for both restaging and recurrence purposes in two studies. [109,111] One study also used ^{18}FDG-PET/CT for both staging and assessment of recurrences. [111] The studies contained a total of 173 patients with sample sizes ranging from 12 to 91 participants. Participant ages ranged from 49 to 81 years. One study reported the distribution by stage of cancer: T1N0M0 = 54 percent, T2N0M0 = 46 percent. [108] ^{18}FDG-PET was compared to a reference standard that varied across the studies. In two studies the reference standard was either histology/biopsy or clinical followup, [109,111] while the reference standard in one study was either histology/biopsy or CT/bone scintigraphy. [110] One study established the final diagnosis of all patients using histology/biopsy. [108] Three studies reported the mean time between last treatment and ^{18}FDG-PET as 6 months, [109] 43.2 months, [111] and 3.2 years. [108] Three studies used a fixed dose of ^{18}FDG (10 MCi[108] and 555 MBq[110,111]); one study reported a dose range of 370 to 555 MBq. [109] The time between injection and PET scan was 30 to 45 minutes, [108] 45 to 60 minutes, [109,111] and 40 to 90 minutes. [110] Patients fasted for four hours. [108-110] One study[110] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose level that was allowed was not reported. Methods of interpretation of the images were qualitative in three studies[108-110] and both qualitative and quantitative in the remaining study. [111] Scans were interpreted qualitatively using visual analysis. [108-111] One study[111] reported using SUV but the criterion for abnormality was not reported.

Comparisons

No pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET for prostate cancer for any of the clinical indications considered: staging, recurrences, considered together or separately (Table 26). Individual data are summarized in Appendix D.

Table 26. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for prostate cancer

Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Restaging and recurrences	Jadvar 2003[109]	P	FDG-PET	Histology/biopsy or clinical followup	No
	Schoder 2005[111]	R	FDG-PET and FDG-PET/CT	Histology/biopsy or clinical followup	
Recurrences	Oyama 2003[110]	P	FDG-PET	Histology/biopsy, CT, bone scintigraphy	No
Restaging	Chang 2003[108]	R	FDG-PET	Histology/biopsy	No

FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

Summary of the results

Four studies of poor to moderate methodological quality evaluated the use of ^{18}FDG-PET[109-111] or ^{18}FDG-PET/CT[111] to assess both restaging and recurrences, [109,111] recurrences[109-111] or restaging[108] alone of prostate cancer. Due to heterogeneity across studies in terms of their design, the type of PET and indications for its use, no pooled estimate of the accuracy of ^{18}FDG-PET could be obtained. The impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on diagnostic thinking or patient-centered outcomes was not assessed in any studies. There is currently insufficient data to recommend the introduction of ^{18}FDG-PET or ^{18}FDG-PET/CT imaging for this indication.

8. Small Cell Lung Cancer

8.1. Background

The leading cause of cancer death world wide is lung cancer. [169] Every year 15 percent of new lung cancer diagnoses are classified as small cell lung cancer (SCLC), which accounts for up to 25 percent of lung cancer deaths. In the United States there were an estimated 213,380 new cases of lung cancer diagnosed (all types) and 160,390 deaths in 2007. [170] Rates peaked in 1986 with 17.4 percent of new cancers diagnosed as SCLC. In the early 1970s, 72.4 percent of those diagnosed with SCLC occurred in male patients. By 2002, the male to female ratio of patients diagnosed with SCLC was 1:1 in 2002. [171]

Approximately 95 percent of all cases are due to cigarette smoking, although environmental factors may also play a role. Decreasing incidence and mortality rates may be related to the declining number of smokers[170] (in the United States, 36 percent of population smoked in 1950 compared to 20 percent of population in 1990) [171] and the development of low-tar filters. [170] Additionally, smoking during treatment tends to shorten patient survival times. The risk of all types of lung cancer can be decreased by smoking cessation. [172]

SCLC is difficult to treat due to its rapid growth and quick development of widespread metastasis. There is an initial dramatic response to treatment; however, the majority of patients die from recurrent disease. [171] Less than 5 percent of patients survive 5 years following diagnosis. [172] Disease usually reoccurs at the primary site in the lung or lymph nodes. Factors that improve prognosis are: small tumor size, no lymph node involvement and possibility of lobectomy. [173] Dyspnea, persistent cough and hemoptysis are the most common presenting symptoms and postobstructive pneumonia may also occur. Common sites of metastases include bone, liver, lymph nodes, central nervous system, adrenal glands, subcutaneous tissue and pleura. Disease that has metastasized can produce pain, headache, malaise, seizures, fatigue, anorexia and weight loss. [170] Extent of disease, performance status, gender and age are the strongest clinical prognostic factors. [172]

Traditionally staging of SCLC uses a system developed by the Veterans Administration Lung Cancer Study Group (VALCSG). There are two stages, limited-stage disease (LD) and extensive-stage disease (ED). LD is defined as disease confined to one hemithorax with the tumor encompassed in one radiation port. Approximately 30 percent of patients are staged with LD at diagnosis. [170] Half of these patients achieve remission compared to the 20 to 40 percent of patients

with ED. [172] ED is any cancer that does not fit into the LD category and represents patients with more disseminated disease. [170] Generally, ED patients have a poorer prognosis and palliative chemotherapy and radiation treatment aims to provide relief of symptoms with minimal toxicity. [172] The International Association for the Study of Lung Cancer has proposed TNM groupings for the clinical staging of SCLC since such systems have shown to be a good tool for predicting outcome. [174] However, TNM classifications are not typically used in SCLC as they require surgery to confirm accuracy and SCLC patients are often poor candidates for surgery.

Establishing a diagnosis of SCLC is a multi-step process that includes a detailed history, physical examination and testing involving a complete blood count, electrolyte panel and histology or cytology. [170] Contrast-enhanced CT scan of the chest and abdomen, bone scan, and CT scan or MRI of the brain are also performed. [118]

PET has been used for the assessment of single pulmonary nodules and for the evaluation of the mediastinum in patients with non-SCLC. Initial studies suggest that the ^{18}FDG tracer is avidly absorbed by SCLC tumors and that staging evaluation with ^{18}FDG-PET may be an effective complement to conventional staging methods. [118]

When treating SCLC the control of symptoms and improvements to patient quality of life should be considered. [172] In patients with LD, combination regimens of chemo and radiotherapy achieve better responses and longer survival than single agents. [170] Dosing schedules of chemo and radiotherapy may be concurrent, sequential or alternating. [169] Patients diagnosed with ED receive chemotherapy as their mainstay treatment. Patient response rate is high, at 60 to 80 percent, but the median survival time is only 8 to 10 months. [170] Palliative radiotherapy may be provided to patients with relapsed ED to help control symptoms. [172] If surgery is planned, it must be a part of a multidisciplinary approach and chemotherapy should still be considered the primary treatment. [173]

8.2. Importance of Key Questions in the Clinical Management of Small Cell Lung Cancer

SCLC has a very aggressive biological behavior. Exact staging of SCLC has an important impact on survival and treatment decisions. The primary role of diagnostic imaging in SCLC is to accurately distinguish between LD and ED. Patients with LD are often offered concomitant chemotherapy and radiotherapy, whereas chemotherapy alone is the standard treatment of patients with ED. Thus, accurate staging is pivotal to reserve the combined modality treatment to those patients who actually might benefit from it.

Chest radiography, thorax and upper abdomen CT scan, MRI, thoracoscopy, bone scans, and bone marrow biopsy are routinely used for staging. However, the use of these diagnostic procedures may result in difficulties identifying tumor tissue in some settings (e.g., in normal-sized lymph nodes). Furthermore, anatomic imaging modalities are mostly used to evaluate a given region of the body rather than the entire body and therefore, it is likely that metastases outside the imaging field are not diagnosed. In contrast to the dependence primarily on anatomic imaging features, ^{18}FDG-PET depends on the metabolic characteristics of a tissue for the detection of disease. As ^{18}FDG preferentially accumulates in viable tumor cells and not in fibrotic or necrotic tissue, a change in ^{18}FDG-uptake on PET scan might be a better parameter for monitoring the response and it might be able to assess response before structural changes occur.

8.3. Results

Ten studies[112-121] provided evidence on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for SCLC. All of them evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT. Three studies[112,113,117] reported on the diagnostic thinking impact of ^{18}FDG-PET and one[117] also evaluated the impact of ^{18}FDG-PET/CT. None of the studies evaluated the impact of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for SCLC. Characteristics of the populations, conditions of ^{18}FDG-PET administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

8.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in small cell lung cancer

Characteristics of the studies

Ten studies (six prospective, [113-118] four retrospective[112,119-121]) evaluated the diagnostic accuracy of ^{18}FDG-PET on SCLC. Seven studies used ^{18}FDG-PET for initial staging[113-115,118-121] and three used ^{18}FDG-PET for both initial staging and restaging. [112,116,117] The studies contained a total of 471 patients with sample sizes ranging from 21 to 120 participants. Participant ages ranged from 33 to 90 years. ^{18}FDG-PET was compared to a reference standard that varied across the studies. In four studies the reference standard was either histology/biopsy or clinical followup, [112,113,117,120] three studies used clinical followup and conventional imaging. [116,119,121] One study established the final

diagnosis by histology/biopsy and conventional imaging, [114] one study used histology/biopsy, [115] and one study used conventional imaging. [118] Two studies reported the median time between last treatment and ^{18}FDG-PET as 207 days[120] and 4 days; [119] one reported the time as greater than two weeks. [118] Five studies used a fixed dose of ^{18}FDG (400MBq, [115,116] 370MBq, [120] 300MBq, [119] and 15MCi[118]); three used a dose range (300 to 400MBq, [117] 10 to 15MCi, [113] and 15 to 20MCi[121]); one used a weight-based dose (5MBq/kg[114]); and one did not report on dosing. [112] The time between ^{18}FDG injection and PET scan was 50 minutes, [113] 60 minutes, [115,118-120] a median of 84 minutes, [116] 90 minutes, [114] and two studies reported ranges (45 to 60 minutes[121] and 50 to 60 minutes[117]). In nine studies patient fasting was reported between four and twelve hours, [112-120] with five of these studies measuring maximum glucose levels before administration of ^{18}FDG-PET (4.6 mmol/L, [116] 4.7 mmol/L, [115] 6 mmol/L, [114] and 150 mg/dL[113,118]). Methods of interpretation were qualitative in four studies[114,115,118,119] and both qualitative and quantitative in four studies. [112,113,116,120] Scans were interpreted qualitatively using visual analysis in all eight studies. One study[112] used a marker of lesions greater than 10 mm in transverse diameter for quantitative interpretation of the PET images.

Comparisons

Comparisons for which data were considered for direct meta-analysis are summarized in Table 27. Statistical pooling was considered to evaluate the accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT for the staging of SCLC. Individual study data are summarized in Appendix D.

Table 27. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for SCLC

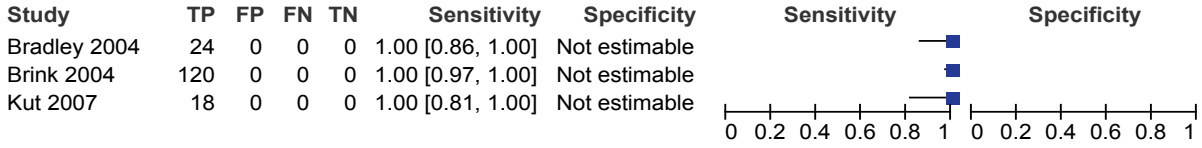
Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Staging and restaging	Blum 2004[112]	R	FDG-PET	Histology/biopsy or clinical followup	No
	Fischer 2006[116]	P	FDG-PET/CT	Clinical followup	
	Kamel 2003[117]	P	FDG-PET and FDG-PET/CT	Histology/biopsy or clinical followup	
Staging	Bradley 2004[113]	P	FDG-PET	Histology/biopsy or clinical followup	1. FDG-PET v. all comparators (P studies) [113,114,118]
	Brink 2004[114]	P	FDG-PET	Histology/biopsy or conventional staging	2. FDG-PET or FDG-PET/CT v. all comparators (R studies) [119,121]
	Fischer 2007[115]	P	FDG-PET and FDG-PET/CT	Histology/biopsy	
	Kut 2007[118]	P	FDG-PET	Conventional staging	
	Niho 2007[119]	R	FDG-PET and FDG-PET/CT	Clinical followup or conventional staging	
	Pandit 2003[120]	R	FDG-PET	Histology/biopsy or clinical followup	
	Vinjamuri 2008[121]	R	FDG-PET and FDG-PET/CT	Clinical followup	

FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

1. ^{18}FDG-PET for the staging of SCLC

Reference standard: any; prospective studies. Three prospective studies[113,114,118] totaling 162 participants provided data to analyze the accuracy of ^{18}FDG-PET v. any reference standard for identifying the stage of SCLC. Individual 2x2 table results are presented in Figure 83. Sensitivity in the three studies was 100 percent. The studies did not provide data to calculate specificity and therefore, pooled estimates of the positive and negative LRs were not obtained.

Figure 83. Results from 2x2 tables of individual prospective studies of ^{18}FDG-PET v. any reference standard for the staging of SCLC

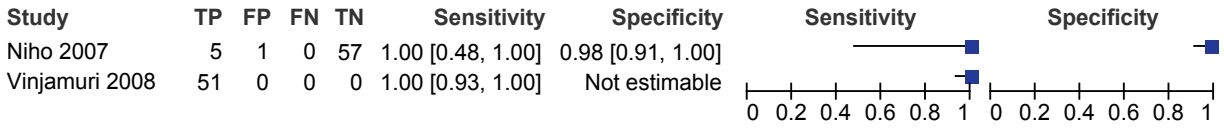


2. ^{18}FDG-PET/CT for the staging of SCLC

Reference standard: any; retrospective studies. Two retrospective studies[119,121] totaling 114 participants provided data to analyze the accuracy of ^{18}FDG-PET/CT v. any reference standard for identifying the stage of SCLC. Individual 2x2 table results are presented in Figure 84. Sensitivity was 100% in the two studies.[119,121] The specificity was 0.98 in one study, [119] however the second

study did not provide data to calculate specificity and therefore, pooled estimates of the positive and negative LRs were not obtained.

Figure 84. Results from 2x2 tables of individual retrospective studies of ^{18}FDG-PET/CT v. any reference standard for the staging of SCLC (retrospective studies)



Summary of the results

No conclusions can be drawn regarding the utility of ^{18}FDG-PET or ^{18}FDG-PET/CT for the staging of SCLC. No pooled estimates of the positive and negative LRs were obtained. Information about the specificity of the test is not available from the studies included in this analysis.

8.3.2. Diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT on physician decision making with respect to diagnosis and management strategy for patients with small cell lung cancer

Three studies reported on the diagnostic thinking impact of ^{18}FDG-PET. Blum et al. [112] evaluated the treatment decision impact of ^{18}FDG-PET on staging and restaging patients with SCLC. Thirty-six consecutive outpatients who had undergone 47 ^{18}FDG-PET scans were retrospectively enrolled in this study. Of the 36 patients, 15 underwent ^{18}FDG-PET for staging and 25 for restaging, of which four patients had previously undergone staging. The treatment plan was considerably altered for 17 (43 percent) of all cases. Seven of the 15 (47 percent) patients who underwent ^{18}FDG-PET for initial staging had changes to their treatment plans due to upstaging in their disease identified by ^{18}FDG-PET. Five of these patients had their management altered from radical concurrent chemoradiotherapy to palliative chemotherapy alone or the later addition of palliative radiotherapy. The remaining two patients had their radiotherapy target volume changed to include additional disease shown by ^{18}FDG-PET. In addition, 10 of the 25 patients (40 percent) who underwent ^{18}FDG-PET for restaging had their treatment plans changed. Five of these patients were upstaged based on ^{18}FDG-PET and therefore avoided prophylactic cranial irradiation (PCI) (n = 3) or changed from chemotherapy to observation alone (n = 2). Three patients were downstaged and went on to have

PCI. An additional two patients were reported to have changes in their treatment plan but the nature of this change was not described.

Prognostic outcomes were also reported in the study. Patients who achieved a complete metabolic response on ^{18}FDG-PET had a median time to progression of 13.7 months, compared to 9.7 months for patients who did not achieve a complete response. In addition, of the 16 patients with an incomplete response, five were still alive with a median followup of 19 months. It is possible that salvage treatments had a favorable impact on patients found to have residual disease by ^{18}FDG-PET.

The authors concluded that ^{18}FDG-PET could be used in conjunction with conventional imaging to improve staging and better ensure that patients receive the most appropriate management.

This retrospective study reviewed consecutive patients and was assessed to be of moderate quality. The spectrum of included patients, the choice of reference standard, any intermediate results and withdrawals were well reported. Although all patients received a reference standard, the reference standards were not the same across all patients; some received histological confirmation of disease and others were followed up clinically. The multiple methods used to validate disease status may have lead to verification bias. In addition, the selection criteria and execution of the index and reference tests were only partially described and the duration between the index and reference tests was unclear. The results of the index test were interpreted in an unblinded manner and the blinding of the reference test interpretation was unclear.

Bradley et al. [113] evaluated the treatment decision impact of ^{18}FDG-PET on the staging of SCLC. A prospective sample of 25 outpatients with newly diagnosed, untreated, histologically or cytologically confirmed SCLC underwent ^{18}FDG-PET scans. Of these, 24 were included in the analysis and one patient withdrew from the study. ^{18}FDG-PET scans contributed to a change in the diagnosis of seven patients (29 percent), all of whom were upstaged. An unsuspected primary tumor or regional nodal metastasis was identified by ^{18}FDG-PET in seven patients (29 percent), six of whom had nodes that were not considered enlarged by CT criteria but showed ^{18}FDG uptake on PET. This resulted in a significant alteration to the radiation therapy portal and the nodes were included in the high-dose region for each of these patients. Further, the addition of ^{18}FDG-PET identified two patients (8 percent) with ED SCLC who were thought to have LD based on conventional staging.

The authors concluded that ^{18}FDG-PET had high sensitivity and appeared to be of value for staging and treatment planning in patients presumed to have LD SCLC.

This study was determined to be of high quality. The selection criteria, choice of the reference standard, intermediate test results and study withdrawals were well reported. In addition, the reference standard was independent of the index test and all the tests were conducted within a sufficiently brief time period. However, reporting of the recruitment of patients was inadequate and the execution of the index and reference tests. Since cases were not all verified using the same reference standard, there is risk of verification bias. Although the reference standard was interpreted while blinded to the results of the ^{18}FDG-PET, the results of the reference tests were used in the interpretation of the ^{18}FDG-PET scans, which raises the possibility of review bias.

Kamel et al. [117] investigated the treatment decision impact of ^{18}FDG-PET and integrated ^{18}FDG-PET/CT on staging and restaging in patients with SCLC. Forty-five consecutive outpatients who underwent ^{18}FDG-PET imaging were retrospectively enrolled in the study; however, three patients were excluded due to incomplete data, therefore 42 patients were included in the analysis. Of these, 24 patients were referred for ^{18}FDG-PET for initial staging and 20 for restaging, where two patients were included for both staging and restaging. No description regarding the interpretation of the scans was provided. The treatment management was altered in 12 of 42 patients (29 percent). Nine patients with ^{18}FDG-PET for initial staging had a change in treatment; three were given palliative chemotherapy and one patient was given curative surgery since ^{18}FDG-PET findings excluded mediastinal involvement and distant metastases. Three patients had a change in radiation field and two patients had a change in radiation volume. Therefore, of the patients who received ^{18}FDG-PET for initial staging, three were upstaged, one was downstaged and five had minor changes to their diagnosis which influenced their treatment plan. Three of the 20 patients (15 percent) with ^{18}FDG-PET for restaging after therapy had a change in treatment management; one patient had chemotherapy reinstituted, while chemotherapy was discontinued in two patients. Four patients had a change in diagnosis, including the three patients with changes in regards to chemotherapy and one patient with a false negative from ^{18}FDG-PET. Three additional patients were identified as having progressive disease by ^{18}FDG-PET and therefore had a minor change to their diagnosis that did not impact their treatment.

The authors concluded that ^{18}FDG-PET imaging has the potential to improve the outcomes of combined chemoradiotherapy by preventing futile treatment of patients with distant metastases or advanced locoregional disease not identified by conventional imaging. In addition, ^{18}FDG-PET may

optimize radiation treatment for patients with LD through its accurate definition of radiation field and volume.

This retrospective study of consecutive patients was assessed to be of high quality. The spectrum of included patients, the execution of the index test, intermediate results and study withdrawals were well reported. In addition, patients received an appropriate reference standard, which was consistent across all patients. Although there was blinded interpretation of the reference standard, the index test was interpreted with knowledge of the results of the standard, introducing risk of review bias. Other weaknesses of this study included only partial description of the selection criteria, raising the possibility that the results are not generalizable to other patient populations. In addition, disease progression bias cannot be ruled out as there was lack of clarity with regard to the time between the diagnostic tests. Finally, results were not presented separately for ^{18}FDG-PET and ^{18}FDG-PET/CT.

Summary of the results

There is some evidence from three studies of moderate to high quality that ^{18}FDG-PET[112,113,117] or ^{18}FDG-PET/CT[117] is useful in the staging of SCLC resulting in more appropriate management and in cases of upstaging avoidance of 'futile' therapy. The studies consistently demonstrated that ^{18}FDG-PET altered management plans in an important proportion of patients. The change in management varied, but patients were more frequently upstaged, resulting in more aggressive treatment or switching to palliative care. One study[112] examined prognostic outcomes and found that complete metabolic responders on ^{18}FDG-PET had a median time to progression of 13.7 months, compared to 9.7 months for patients who did not achieve a complete response.

Table 28 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact and effect on patient-centered outcomes of ^{18}FDG-PET and ^{18}FDG-PET/CT for SCLC.

Table 28. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT for SCLC

Study	Results of FDG-PET imaging on patient diagnosis and treatment	Types of bias
Blum 2004[112] Study type: Retrospective	**Management decision:** Treatment Treatment plans altered for 17/36 patients (47%) overall Initial staging: 7/15 plans changed (all upstage) Radical concurrent chemotherapy to palliative therapy (n = 5) Radiotherapy target volume increased (n = 2) Restaging: 10/25 plans changed (3 upstage, 5 downstage, 2 ND) PCI in patients with positive CT but negative FDG uptake (n = 3) PCI omitted in cases that did not have complete response (n = 3) Observation in cases with no FDG uptake but positive CT (n = 2) **Prognostic outcomes:** Complete metabolic responders on PET had a longer median time to progression (13.7 mo v. 9.7 mo)	Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS) Review bias (PET unblinded; RS, unclear)
Bradley 2004[113] Study type: Prospective	**Management decision:** Treatment Major change in diagnosis of 7/25 patients (28%); all upstaged Among these, unsuspected primary tumor identified in 6 patients (not detected by CT), leading to significant change to radiation therapy portal Identification of 2 patients with extensive-stage disease, who were diagnosed as limited-stage SCLC by conventional staging	Spectrum bias (unclear) Selection bias (unclear) Verification bias (>1 RS) Review bias (RS, unclear if blinded)
Kamel 2003[117] Study type: Retrospective	**Management decision:** Treatment Treatment altered in 12/42 patients (29%) overall. Initial staging: 9/24 changes in management. Upstaged and palliative chemotherapy (n = 3); downstaged and curative resection (n = 1); minor change to diagnosis and altered radiation field (n = 5) Restaging after therapy, 3/20 changes in management: chemotherapy reinstituted (n = 1); discontinued (n = 2)	Selection bias (unclear) Disease progression bias Review bias (PET and RS unblinded)

CT = computed tomography; FDG = fluorodeoxyglucose F18; ND = no data; PET = positron emission tomography; RS = reference standard

9. Testicular Cancer

9.1. Background

Testicular cancer is characterized by malignant cells in one or both testicles. The majority (95 percent) of testicular neoplasms are germ cell tumors (GCTs) with other neoplasms, such as sex-chord stromal tumors and lymphomas occurring only rarely. [175] GCTs are broadly separated into two groups: seminomas and nonseminomas, each comprising approximately 50 percent of cases. [175] Seminomas originate from the sperm-producing germ cells of the testes and may be one of three types: classic, anaplastic, or spermatocytic. Nonseminomas are also germ cell tumors but appear very different histologically. Types of nonseminomas include choriocarcinoma, embryonal carcinoma, teratoma, and yolk sac tumors. Testicular tumors may contain both seminoma and nonseminoma cells. [176]

Testicular cancer is the most common cancer in men between the ages of 15 and 35 years, accounting for one to two percent of all neoplasms in men. [177] In 2004, the incidence of invasive testicular cancer was 5.2 per 100,000 and the attributable death rate 0.2 per 100,000 (measures adjusted by age to the 2000 United States standard population). [157] National cancer statistics estimate that there will be 8,090 new cases and 380 deaths in 2008. [133] Testicular cancer occurs most often in men between the ages of 20 and 39. [176]

Testicular changes symptomatic of GCT are usually found during self-examination, after testicular trauma or by a sex partner. [177] Signs and symptoms of testicular cancer include acute pain in the testicle or scrotum, dull ache in the scrotum or abdomen, scrotal heaviness, and firmness of the testicle. A physical examination includes palpation of the testes and is accompanied by blood tests that measure the levels of tumor markers[178] such as alpha-fetoprotein (AFP), beta-human chorionic gonadotropin (ß-HCG), and lactate dehydrogenase (LDH). Higher than normal levels of these markers may suggest the presence of a testicular tumor, even if it is too small to be detected by physical exams or imaging tests. [176] Scrotal ultrasonography determines whether a suspected mass is intra- or extratesticular. Intratesticular masses are presumed to be cancerous until proven otherwise. [177] Final diagnosis is often made by radical orchiectomy (surgical removal of the testicle through an incision in the groin). PET scans are not yet recommended outside clinical trials as part of routine staging procedures because the procedure has not conclusively demonstrated improved sensitivity of staging compared with CT scanning alone. [178]

After testicular cancer is diagnosed, a patient may receive CT of the abdomen and pelvis to detect metastasis to the retroperitoneal lymph nodes and chest radiography. [177] Patients with neurologic symptoms may receive CT or MRI of the brain. [177]

Testicular cancer is categorized using the TNMS system. Staging is determined based on how much the primary tumor has spread to tissues surrounding the testicles, on extent of spread to regional lymph nodes, on metastasis to other organs, and on serum levels of proteins produced by certain types of testicular cancer. [177]

With overall cure rates of more than 95 percent (80 percent for metastatic disease), testicular GCT are considered the model for curable cancer. [175] Nonseminomas tend to grow and spread more quickly; seminomas are more sensitive to radiation. As a result, treatment options differ slightly depending on the characterization of the cancer. If the tumor contains both seminoma and nonseminoma cells, it is treated as a nonseminoma. [176]

The primary treatment for all testicular tumors is radical inguinal orchiectomy. [177,178] Almost all seminomas are curable with orchiectomy with or without radiation, and only occasionally do these cancers require chemotherapy. [175] Nonseminomatous GCTs are less sensitive to radiation and, when metastatic, frequently require both chemotherapy and surgery. [175] A surveillance strategy is an option for patients with stage I seminomas, [178] radiation therapy for seminomas stage I and IIa, and lymph node dissection for stage I and II nonseminoma. [177] Though limited data exist to guide the choice of high-dose (2-3 cycles of etoposide and carboplatin with or without cyclophosphamide or ifosfamide) or conventional-dose chemotherapy for initial salvage treatment, [175] chemotherapy is a treatment option for seminoma stage II and III and all stage II nonseminoma. [177]

Multiple studies have demonstrated the importance of resecting residual masses following first-line of salvage chemotherapy for nonseminoma GCTs. [175] Postchemotherapy surgical resection of seminoma is technically more difficult and carries a higher morbidity due to the desmoplastic reaction frequently induced by treatment. [175] PET scan can be used to guide surgical decisions in this setting. [175]

9.2. Importance of Key Questions in the Clinical Management of Testicular Cancer

The role of ^{18}FDG-PET in the diagnosis, staging and followup of germ cell tumors is still a matter of debate and there is a need to define optimal indications for ^{18}FDG-PET in testicular cancer. ^{18}FDG-PET may have a role in distinguishing between benign and malignant tissue by

characterizing the metabolic activity of the tissue rather than the anatomical size only. ^{18}FDG-PET may also offer the potential to detect residual malignancy after primary curative therapy for testicular cancer. For many solid tumors the early detection of recurrent or residual disease may not confer a clinical benefit to patients, because further curative treatment options may not be available. However, residual or recurrent germ cell malignancy can be cured by further treatment and hence ^{18}FDG-PET may have an important clinical role for patients with such tumors. Existing imaging methods, such as CT scan, chest X-ray and tumor marker evaluation with AFP and β-HCG may be insufficient to identify absent, residual or recurrent disease. Identification of these characteristics may influence subsequent patient management policy

9.3. Results

Four studies[122-125] provided evidence on the use of ^{18}FDG-PET for testicular cancer. We did not find studies that reported on the use of ^{18}FDG-PET/CT for testicular cancer. All the four studies[122-125] evaluated the diagnostic accuracy of ^{18}FDG-PET, one study[124] reported on the diagnostic thinking impact of ^{18}FDG-PET. None of the studies evaluated the effects of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET or ^{18}FDG-PET/CT for testicular cancer. Characteristics of the populations, conditions of ^{18}FDG-PET administration, interpretation of results and methodological quality of the studies are summarized in Appendices D to J.

9.3.1. Diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT in testicular cancer

Characteristics of the studies

Four studies (three prospective, [122,123,125] one retrospective[124]) evaluated the diagnostic accuracy of ^{18}FDG-PET on testicular cancer. One study used ^{18}FDG-PET for initial staging, [125] one for restaging purposes, [122] and two to assess recurrences. [123,124] The studies contained a total of 135 patients with sample sizes ranging from 15 to 54 participants. Participant ages ranged from 20 to 62 years. Two studies reported the distribution by clinical stage of cancer. One included patients at CS I (20 percent), CS II (47 percent) and CS III (33 percent); [124] the other included patients at CS IIb (10 percent), CS IIc (70 percent) and CS III (20 percent). [123] ^{18}FDG-PET was compared to a reference standard that varied across the studies. In two studies the reference standard was either histology/biopsy or clinical followup. [122,125] One study established the final diagnosis of all patients using histology/biopsy[123] and one study used clinical followup for final diagnosis. [124] Two studies

reported the median time between last treatment and ^{18}FDG-PET as 29 days[123] and 45 days; [124] one reported the time since last treatment as 4 to 12 weeks. [122] All studies used a fixed dose of ^{18}FDG, which ranged from 320 MBq to 400 MBq. When reported, the time between injection and PET scan was 45 minutes. [122,125] Patients fasted for four hours[122,123] to six hours. [124,125] Two studies[122,123] measured glucose levels before administration of ^{18}FDG-PET; the maximum glucose level that was allowed was normal levels. Methods of interpretation of the images were qualitative in three studies[122,124,125] and both qualitative and quantitative in one. [123] Scans were interpreted qualitatively using visual analysis in all studies. SUV values were reported in one study[123] for interpretation of the PET images. The criterion for abnormality was SUV greater than 2 g/mL.

Comparisons

No pooled data were obtained to evaluate the accuracy of ^{18}FDG-PET for testicular cancer for any of the clinical indications considered (i.e., staging, recurrences, and restaging) (Table 29). Individual data are summarized in Appendix D.

Table 29. Summary of comparisons considered for meta-analyses of the accuracy of ^{18}FDG-PET for testicular cancer

Indication	Studies	Design	Type of PET	Reference standard	Meta-analysis
Staging	Lassen 2003[125]	P	FDG-PET	Histology/biopsy or clinical followup	No
Recurrences	Hinz 2008[123]	P	FDG-PET	Histology/biopsy	No
	Karapetis 2003[124]	R	FDG-PET	Clinical followup	No
Restaging	Becherer 2005[122]	P	FDG-PET	Histology/biopsy or clinical followup	No

FDG = fluorodeoxyglucose F18; P = prospective; PET = positron emission tomography; R = retrospective

Summary of the results

Four studies of moderate methodological quality evaluated the diagnostic accuracy of ^{18}FDG-PET[122-125] in the staging, [125] restaging, [122] or recurrence[123,124] of testicular cancer. Due to heterogeneity in the study designs and indications for which ^{18}FDG-PET was used, no pooled estimate of the diagnostic test performance could be obtained.

9.3.2. Diagnostic thinking impact of ^{18}FDG-PET on physician decision making with respect to diagnosis and management strategy for patients with testicular cancer

One study[124] reported on the diagnostic thinking impact of ^{18}FDG-PET. The retrospective study by Karapetis et al. [124] examined the treatment decision impact of ^{18}FDG-PET imaging on assessing the recurrence of testicular cancer. A series of 15 patients with metastatic or extragonadal germ cell tumors, who had undergone ^{18}FDG-PET scanning were enrolled. The treatment plan was altered in only one patient on the basis of the ^{18}FDG-PET scan; management was changed from observation to surgical excision of residual mass. Normal ^{18}FDG-PET scans provided confirmation in four patients with small residual masses; however, did not alter their subsequent treatment. Seven patients had more than one ^{18}FDG-PET scan. The subsequent ^{18}FDG-PET scans supported, but did not change, treatment management plans.

The authors concluded that ^{18}FDG-PET scanning did not have a discernable impact on treatment decisions for the majority of patients. However, ^{18}FDG-PET often provided support for management decisions made on the basis of the results of other clinical assessments. The authors recommended that ^{18}FDG-PET scans should be arranged with a clear aim in patient management and should not be interpreted in isolation of other assessments.

This study was assessed to be of moderate quality. The spectrum of patients included was representative of patients who would receive the test in practice, a reference standard was applied to the whole sample and was independent of the index test, and all patients and test results were accounted for. However, both the choice of reference standard and time period between tests was unclear. The selection criteria were only partially described, raising the possibility of selection bias. The authors acknowledge the risk of selection bias that is inherent to this study, particularly relevant given the retrospective nature of data collection. In addition, patients did not all receive the same reference standard test and the execution of the index and reference tests was not described sufficiently. Although the index test was blindly interpreted, the reference standard was interpreted using the results of the index, which may have introduced review bias.

Summary of the results

The evidence for ^{18}FDG-PET in testicular cancer is very limited and inconclusive. One small retrospective study[124] of moderate methodological quality evaluated the physician decision-making impact when ^{18}FDG-PET imaging is used in the assessment of the recurrence of testicular cancer.

The management plan was changed in only a small proportion of the patients. Small residual masses were confirmed in 27% (4/15) of patients; however had no effect on subsequent treatment decisions. Evidence from larger, prospective studies of higher quality is needed before conclusions can be made regarding the impact of ^{18}FDG-PET imaging on patient management for testicular caner.

Table 30 provides a summary of the main findings and the types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET for testicular cancer.

Table 30. Main findings and types of bias that affected the evidence on the diagnostic thinking impact of ^{18}FDG-PET for testicular cancer

Study	Results of FDG-PET imaging on patient diagnosis and treatment	Types of bias
Karapetis 2003[124] Study type: Retrospective	**Management decision:** Treatment Management plan altered in only 1/15 patients (7%) Changed from observation to surgical excisions of residual Confirmation of small residual masses in 4/15, subsequent treatment not altered	Selection bias (unclear) Disease progression bias (unclear) Verification bias (>1 RS) Review bias (RS unblinded)

FDG = fluorodeoxyglucose F18; PET = positron emission tomography; RS = reference standard

Chapter 4. Discussion

1. Bladder Cancer

Evidence from three studies is available on the diagnostic performance of ^{18}FDG-PET and ^{18}FDG-PET/CT for bladder cancer. Data from a meta-analysis of two prospective studies totalling 88 participants showed that ^{18}FDG-PET does not seem to be helpful in identifying the stage of the disease. The diagnostic accuracy of ^{18}FDG-PET and ^{18}FDG-PET/CT has not been evaluated for other clinical situations. These results are in agreement with findings reported by other researchers in this field. A recent review by Machtens et al[179] evaluated the utility of ^{18}FDG-PET as a diagnostic tool in malignant urological tumors of the small pelvis (e.g., prostate, testicular and bladder tumors). The authors found that, compared to other urologic malignancies, the value of ^{18}FDG-PET in the imaging of bladder cancer has been investigated the least. No additional role for PET in comparison with conventional imaging in tumor detection and local staging was found for any of the indications. This review also included three studies published before 2003 that evaluated the diagnostic accuracy of ^{18}FDG-PET in the imaging of bladder cancer and yielded similar conclusions. The methods used to identify the primary studies in the Machtens review were not stated and the results were likely affected by selection bias. The Duke report[14] did not evaluate the evidence on the use of ^{18}FDG-PET for bladder cancer and therefore, it is unknown whether further studies on this topic had been published prior to the search period that was covered in our report.

Only one study of moderate sample size and moderate quality evaluated the diagnostic thinking impact of ^{18}FDG-PET on the treatment of bladder cancer and reported that 17 percent of the treatment decisions were changed after knowing the results of ^{18}FDG-PET. Since the amount and quality of the evidence is limited, firm conclusions about the diagnostic thinking impact of ^{18}FDG-PET in bladder cancer cannot be drawn. The American College of Radiology (ACR) has recently developed guidelines to evaluate the diagnostic and therapeutic value of a variety of imaging examinations for the pretreatment staging of invasive bladder cancer. [180] Using a 9-point rating scale (1 = least appropriate, 9 = most appropriate), ^{18}FDG-PET whole body scanning was rated below other diagnostic options, such as chest X-ray, CT urography, and pelvis MRI (rated 2, 9, 8, and 8 points, respectively).

2. Brain Cancer

Six studies provided evidence on the use of ^{18}FDG-PET for brain cancer. We did not find studies that reported on the use of ^{18}FDG-PET/CT for brain cancer. The majority of the studies evaluated the diagnostic accuracy of ^{18}FDG-PET. ^{18}FDG-PET does not seem to be highly discriminative in identifying the stage of the disease, and in distinguishing between necrosis and recurrences. The sensitivity and specificity values of the studies were modest and had wide confidence intervals, precluding firm conclusions about the diagnostic utility of ^{18}FDG-PET for brain cancer. These results are in accordance with the findings reported for brain cancer in the Duke report. [14] Evidence on the diagnostic thinking impact and the economic value of ^{18}FDG-PET and ^{18}FDG-PET/CT for brain cancer is not available. The effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes continue to be scarcely evaluated. There is limited evidence from low quality studies identified here and in the Duke report that suggest that the best indication of ^{18}FDG-PET seems to be differentiating between high and low grade gliomas. There is, however, no consensus regarding the utility of ^{18}FDG-PET in predicting histological grading and survival of brain tumors.

3. Cervical Cancer

^{18}FDG-PET and ^{18}FDG-PET/CT have been evaluated for a variety of clinical indications in the assessment of patients with cervical cancer. These include 1) initial staging, 2) detection of recurrence, and 3) restaging, including planning for salvage therapy. There is evidence around the diagnostic accuracy for each of these indications, and also limited evidence for the diagnostic thinking impact. Only two studies provided evidence for the effects of ^{18}FDG-PET on patient-centered outcomes and these addressed problems of detection of recurrence and restaging. There is no data on the cost effectiveness of either ^{18}FDG-PET or ^{18}FDG-PET/CT.

Initial Staging

^{18}FDG-PET or ^{18}FDG-PET/CT has not been studied in women with early stage (FIGO 0 to Ia) cervical cancer who are treated with local therapy and do not usually undergo staging examinations.

Women with locally advanced carcinoma of the cervix (FIGO Stages Ib to IV) are treated with combined modality therapy consisting of radiation and chemotherapy. The usual practice prior to therapy is to stage with imaging tests (CT or MRI) of the abdomen and pelvis to guide treatment. A

systematic review of the earlier literature[7] suggested that the sensitivity of ^{18}FDG-PET is superior to CT and/or MR imaging. PET imaging has the potential to identify additional sites of disease resulting in either more accurate delineation of radiation therapy fields and possibly improved patient-centered outcomes, or to identify patients with wide-spread disease who could then be spared futile combined modality therapy with its associated toxicity.

This current technology assessment identified that, when either ^{18}FDG-PET or ^{18}FDG-PET/CT were evaluated for staging purposes, the values for the positive and negative LRs were similar for both techniques. Results were also consistent when different reference standards were used and for both retrospective and prospective study designs. Significant results were reported for the positive LR, indicating that both techniques were useful to detect the stage of the disease. The results of the negative LR were not statistically significant. Only one study[32] provided information indicating that improved diagnostic accuracy impacts positively on diagnostic thinking for initial staging. Although of moderate quality, the sample size of this prospective study was small at 42, with only 17 patients studied for staging of locally advanced disease. There are no studies evaluating the effect of ^{18}FDG-PET or ^{18}FDG-PET/CT on patient-centered outcomes. Additional high quality studies of the impact on diagnostic thinking and patient-centered outcomes, within the context of staging of locally advanced cervical cancer with ^{18}FDG-PET/CT, are required.

Detection of Recurrence and Restaging

Recurrent or persistent cervical cancer following radiation therapy may be treated with pelvic exenteration which is radical surgery removing the bladder, cervix, uterus, tubes and ovaries, para-cervical tissues and upper vagina. The surgery requires a team of surgeons from different specialties 6 to 12 hours of operative time. The surgery is only considered when there is no evidence of systemic or lymphatic spread of cancer. Restaging prior to surgery is currently performed with CT and pelvic MRI; however, the overall accuracy of detecting pelvic lymphadenopathy for CT scan is 72 percent (pooled sensitivity 0.47; 95% CI 0.21-0.73) and MRI 78 percent (pooled sensitivity 0.72; 95% CI 0.53-0.87). [181] Suboptimal restaging results in a relatively high rate of aborted exenterative procedures. Prospective studies of ^{18}FDG-PET showed sensitivities of 50 percent to 100 percent in the detection of pelvic and abdominal lymphadenopathy and specificities of 94 percent to 100 percent; retrospective studies showed sensitivities of 80 percent to 96 percent and specificities of 76 percent to 100 percent. In both cases, the meta-analysis demonstrated statistically significant positive

and negative LRs. The high sensitivity of ^{18}FDG-PET and ^{18}FDG-PET/CT has the potential to increase completion rates of radical surgery in women with recurrent or persistent cervical cancer through more appropriate selection of surgical candidates; this needs to be confirmed in future studies.

In summary, the use of ^{18}FDG-PET or ^{18}FDG-PET/CT shows promise to improve the care of patients with cervical cancer for both the staging of locally advanced disease and the detection and restaging of recurrent disease. The majority of studies in this area have provided evidence to support the value of FDG-PET or PET/CT for enhanced diagnostic accuracy. Future research should focus on generating more evidence for the impact on diagnostic thinking and patient-centered outcomes in order to confirm or refute the preliminary findings observed here. Cost studies are needed to clarify the relative costs and benefits of implementing FDG-PET in practice.

4. Kidney Cancer

Due to the potential problem of physiological excretion of ^{18}FDG through the kidneys interfering with abnormal uptake of ^{18}FDG within a primary renal tumor, the usefulness of ^{18}FDG-PET has been explored mainly in the detection of local and distant metastases at initial staging and at restaging. There is some evidence to suggest that the accuracy of ^{18}FDG-PET may be sufficient to support its use in the initial staging of renal cancer. When ^{18}FDG-PET was compared against any reference standard in prospective studies, statistically significant results were obtained for both the positive and negative likelihood ratios. In retrospective studies, statistically significant results were only obtained for the negative likelihood ratio.

However the diagnostic thinking impact of this technology has been explore in only one study for patients with recurrent disease and in two studies for initial staging and detection of recurrence. All studies were retrospective and of moderate quality. The impact of the tests varied from 0 percent to 30 percent, the lowest impact being in restaging.

Because of the high sensitivities and specificities reported in prospective studies of diagnostic accuracy, the application of ^{18}FDG-PET to at least initial staging of renal cancer seems to be worthy of additional study by well-designed prospective trials. There is insufficient evidence to support its widespread adoption at this time.

5. Ovarian Cancer

^{18}FDG-PET and ^{18}FDG-PET/CT have been evaluated for a variety of clinical indications in the assessment of patients with ovarian cancer. These include 1) primary diagnosis, 2) initial staging, 3) detection of recurrence and 3) restaging. There is evidence around the diagnostic accuracy for each of these indications, and also limited evidence for the diagnostic thinking impact. Only one study provided evidence for the effects of ^{18}FDG-PET as part of a management strategy on patient-centered outcomes, and there is no data on the cost effectiveness of either ^{18}FDG-PET or ^{18}FDG-PET/CT for ovarian cancer. Most of the studies have evaluated the use of ^{18}FDG-PET/CT.

^{18}FDG-PET has not been studied in women with suspected ovarian cancer for primary diagnosic purposes. The evidence on the accuracy of ^{18}FDG-PET/CT for the primary diagnosis of ovarian cancer is limited to two high-quality studies that reported individual sensitivity values ranging from 87% to 100%. Although the results are promising, no firm conclusions should be made based on these estimates and further studies should evaluate the value of ^{18}FDG-PET/CT as part of the initial workup to diagnose ovarian cancer.

The evidence on the efficacy of ^{18}FDG-PET and ^{18}FDG-PET/CT for the initial staging of ovarian cancer is very limited and no firm conclusions can be made regarding their utility for this indication.

The clinical indication for which ^{18}FDG-PET has been evaluated the most is for the detection of recurrences following treatment. Meta-analyses of the diagnostic accuracy of both ^{18}FDG-PET and ^{18}FDG-PET/CT showed a consistent, statistically significant effect in both the positive and negative LRs across a range of reference standards and study designs, providing evidence to support the usefulness of ^{18}FDG-PET and ^{18}FDG-PET/CT in detecting ovarian cancer recurrences. Only one comparison did not yield a statistically significant positive LR, however this result could be attributed to smaller numbers of studies in the analysis and smaller samples.

Previous evaluations of the earlier literature on the diagnostic performance of ^{18}FDG-PET for ovarian cancer[7,14] suggested that ^{18}FDG-PET is less useful for detecting microscopic residual ovarian cancer, but has fair sensitivity for detecting disease recurrence. This current technology assessment identified that, when either ^{18}FDG-PET or ^{18}FDG-PET/CT were evaluated for the detection of recurrences, the values for the pooled positive and negative LRs were similar for both techniques.

The results of this technology report indicate that ^{18}FDG-PET, especially when combined with CT, is a potentially useful tool for detecting recurrent ovarian tumors. These results agree with the

findings reported by other researchers that have synthesized the evidence on the clinical efficacy of ^{18}FDG-PET or ^{18}FDG-PET/CT. [182,183] Particularly, the American College of Radiology (ACR) has recently developed evidence-based guidelines to evaluate the diagnostic and therapeutic value of a variety of imaging examinations for ovarian cancer. [183] After an analysis of the current literature and expert panel consensus, ACR recommendations for the staging of ovarian cancer do not include ^{18}FDG-PET/CT, and priority is given to the use of abdominal and pelvic CT and CA–125. Alternatively, ^{18}FDG-PET/CT is considered to rule out recurrent ovarian cancer along with these two options and in some instances, it can substitute for CT.

The evidence on the diagnostic thinking impact of PET is limited to integrated ^{18}FDG-PET/CT imaging for assessing recurrence of ovarian cancer. The five studies consistently demonstrated that ^{18}FDG-PET/CT altered management plans in an important proportion of patients and confirmed the results of previous evaluations. [14] Only one study of poor quality examined the effect of ^{18}FDG-PET/CT on patient-centered outcomes and prognosis, and prohibited any definitive conclusions about the value of ^{18}FDG-PET/CT as part of a management strategy.

6. Pancreatic Cancer

Seventeen studies evaluated the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT. The majority of studies are on ^{18}FDG-PET but some of them have evaluated ^{18}FDG-PET/CT. The findings were consistently significant, suggesting that both ^{18}FDG-PET and ^{18}FDG-PET/CT are useful in diagnosing and establishing initial stage of the disease. However, the pooled results are, for the most part, heterogeneous; therefore, the magnitude of the effect is uncertain. Further, there is no clear evidence for the choice of ^{18}FDG-PET v. ^{18}FDG-PET/CT, since the observed heterogeneity indicates considerable uncertainty in these estimates. We found that when ^{18}FDG-PET was evaluated for primary diagnostic purposes, the positive LR was slightly better for ruling in the disease, but the negative LR remained almost the same.

A review by the Blue Cross Blue Shield Association[184] compared the use of ^{18}FDG-PET with the use of conventional diagnostic workup (i.e., CT, MRI, and ultrasonography) and ^{201}Tl SPECT. The review concluded that ^{18}FDG-PET helps to differentiate benign from malignant lesions in patients with suspected pancreatic malignancies.

A review by Higashi et al[185] of the literature published from 1999 to 2003 on the utility of ^{18}FDG-PET for the diagnosis of pancreatic cancer concluded that ^{18}FDG-PET can effectively differentiate pancreatic cancer from benign lesions with a high degree of accuracy. A systematic review by Orlando et al[8] evaluated the evidence published between 1966 and 2003 about the diagnostic accuracy of ^{18}FDG-PET/CT compared with CT alone for the detection of pancreatic cancer. The authors found that adding ^{18}FDG-PET/CT to the diagnostic workup can be of benefit in detecting pancreatic malignancies depending of the pretest probability of the patient. These conclusions are supported by the findings of the Duke report, [14] in which ^{18}FDG-PET sensitivity and specificity were found to be slightly better than CT alone. Evidence on the use of ^{18}FDG-PET for recurrences and staging is derived from individual study data and therefore, firm conclusions about the utility of ^{18}FDG-PET for these indications cannot be made.

There is some evidence on the diagnostic thinking impact of ^{18}FDG-PET and ^{18}FDG-PET/CT that indicates that the management plan is altered in an important number of patients (up to 69 percent), more often resulting in a conservative course of management thus avoiding unnecessary surgery. The review by the Blue Cross Blue Shield Association[184] also addressed the question about how ^{18}FDG-PET can alter diagnosis and management decisions; however, no firm conclusions were made because results of primary studies differed considerably. The review by Higashi et al[185] concluded that ^{18}FDG-PET is useful and cost-beneficial in the preoperative staging of pancreatic cancer. According to the authors, ^{18}FDG-PET may be able to detect unexpected distant metastases in about 40% of the cases. Only one study reported on the effects of ^{18}FDG-PET on patient-centered outcomes and there is limited evidence that ^{18}FDG-PET can improve time to disease progression.

Finally, one study conducted a cost-minimization analysis on the use of ^{18}FDG-PET/CT for pancreatic cancer. The analysis, however, is insufficient to demonstrate changes in costs relative to changes in clinical effects due to the implementation of ^{18}FDG-PET/CT in the diagnostic workup of pancreatic cancer. The study does not provide a common diagnostic efficacy denominator for ^{18}FDG-PET/CT, therefore precluding a cost-effectiveness analysis.

7. Prostate Cancer

The studies identified for analysis in this technology assessment addressed only the detection of recurrence in patients with an increased PSA following primary therapy for prostate cancer. No

study addressed the diagnosis or staging of disease. Due to heterogeneity across studies in terms of study design, type of PET and indications for its use, no pooled estimate of the accuracy of ^{18}FDG-PET could be obtained. The impact of ^{18}FDG-PET on diagnostic thinking or patient-centered outcomes was not assessed in any studies. Furthermore, there is no data available describing the use of ^{18}FDG-PET/CT technology, which might be expected to improve diagnostic accuracy in pelvic imaging compared to ^{18}FDG-PET, since false positive results might occur due to urinary excretion of ^{18}FDG.

PSA relapse after primary therapy for prostate cancer is a common clinical problem. Frequently the PSA begins to increase before recurrent disease can be localized clinically or by imaging studies. The average time between PSA relapse and the clinical presentation of metastases is eight years. [186] There is controversy regarding when treatment should be initiated, and which treatment modality is most appropriate. [187] At present, the need for a highly sensitive test to restage disease and aid in the prescription of therapy has not been confirmed. Furthermore, the role of ^{18}FDG-PET or ^{18}FDG-PET/CT as compared to existing anatomical imaging has not been fully characterized. There is currently insufficient data to recommend the introduction of ^{18}FDG-PET or ^{18}FDG-PET/CT imaging for this indication.

8. Small Cell Lung Cancer

There is no evidence to support the role of ^{18}FDG-PET or ^{18}FDG-PET/CT in the diagnosis of SCLC and no cost-effectiveness data is available for this tumour for any indication. However, there is preliminary evidence from three cohort studies of moderate quality supporting the technique's use in staging and restaging.

The current staging system in SCLC reflects treatment algorithms that recommend thoracic radiation only in patients with limited stage disease, where all known disease can be encompassed by radiation. Meta-analyses have shown that radiation under these circumstances results in a modest improvement in survival.[188] The correct identification of this subgroup of patients, with accurate delineation of radiation fields, might be expected to improve outcomes.

Information on the specificity of ^{18}FDG-PET or ^{18}FDG-PET/CT was not available; therefore pooled estimates of the positive and negative likelihood ratios could not be obtained. However, the sensitivities of the ^{18}FDG-PET and ^{18}FDG-PET/CT were consistently high across studies. Studies

suggest that staging or restaging with FDG-PET can improve on the accuracy of conventional staging investigations such as CT, MRI and radionuclide bone scans.[112,113,117] The data suggests that between 5 percent and 10 percent of patients with SCLC may be upstaged to extensive stage disease[112-114,116,117,119] and a small number of patients with extensive stage disease may be downstaged to limited stage[112,114,117] based on ^{18}FDG-PET or ^{18}FDG-PET/CT findings. These findings have potential to alter the decision about whether or not to use radiation treatment. In addition, the radiation treatment plan may be influenced in 15 percent to 30 percent of patients.[112,113,117,119] These results would suggest that FDG-PET should be a part of the staging investigations of SCLC.

These data are limited in that the studies are generally small and subject to bias. Nevertheless there is consistency in the findings that FDG-PET appears more accurate than conventional staging investigations for SCLC. There is a lack of information about patient-centered outcomes from any of the studies to date. None of the studies report on whether treatment changes directed by staging with FDG-PET changes the patient-centered outcomes for patients with SCLC. Future research should attempt to address this issue.

9. Testicular Cancer

The recent evidence for ^{18}FDG-PET in testicular cancer is very limited and inconclusive. Four studies of moderate quality evaluated the diagnostic accuracy of ^{18}FDG-PET for testicular cancer. ^{18}FDG-PET/CT has not been evaluated for this condition. The clinical indications that were evaluated included initial staging, restaging and recurrences; however, a pooled analyzes of the data was precluded due to the limited number of studies.

Previous evaluation have found direct and consistent evidence about a higher sensitivity and specificity of ^{18}FDG-PET compared to CT for the initial staging of patients with germ cell tumors.[14] However, the clinical relevance of these results have been questioned due to heterogeneity in the clinical populations evaluated. A recent narrative review by Machtens et al[179] has discussed the utility of ^{18}FDG-PET as a diagnostic tool in malignant urological tumors of the small pelvis (e.g., prostate, testicular and bladder tumors). The authors found no additional role for PET in comparison with conventional imaging in tumor detection and local staging for testicular cancer; however, they suggested that the use of FDG-PET can be considered in the restaging of seminomatous germ cell

tumors after chemotherapy. This review did not use a systematic approach in the selection and evaluation of the primary studies and therefore, a variety of bias may affect the validity of the results and conclusions.

One small retrospective study of moderate methodological quality evaluated the physician decision-making impact when ^{18}FDG-PET imaging is used in the assessment of the recurrence of testicular cancer. The management plan was changed in only a small proportion of the patients, suggesting that ^{18}FDG-PET imaging had minimal effect on subsequent treatment decisions. Evidence from larger, prospective studies of higher quality is needed before conclusions can be made regarding the impact of ^{18}FDG-PET imaging on patient management for testicular cancer.

Our findings agree with those reported by a synthesis of the evidence reported by the American College of Radiology (ACR) on the diagnostic and therapeutic value of a variety of imaging examinations for the staging of testicular malignancies. [189] Using a 9-point rating scale (1 = least appropriate, 9 = most appropriate), ^{18}FDG-PET whole body scanning was rated below other diagnostic options, such as chest X-ray, abdomen and pelvis CT, chest CT, and abdomen and pelvis MRI (rated 4, 9, 8, 7, and 5 points, respectively). ^{18}FDG-PET is likely best suited for following up recurrent disease and differentiating residual nonseminomatous tumors from mature teratoma.

Finally, none of the studies evaluated the effects of ^{18}FDG-PET or ^{18}FDG-PET/CT as part of a management strategy on patient-centered outcomes. There were no economic evaluations on the use of ^{18}FDG-PET and therefore, no firm conclusions on the clinical and economic impact of ^{18}FDG-PET or ^{18}FDG-PET/CT can be made. There is no evidence that the preliminary results provided by a limited number of small primary studies can be translated into meaningful advances in the management of patients with testicular malignancies. Further well-designed prospective studies are required to assess the value of ^{18}FDG-PET and ^{18}FDG-PET/CT in areas that remain unevaluated, such as clinical management and treatment response.

Chapter 5. Conclusions

Strengths and Limitations

The strengths of this technology report pertain to its rigour in terms of searching the literature, the criterion-based selection of relevant evidence, the rigorous appraisal of validity, the quantitative summary of data, and the evidence-based inferences.

Our search strategy is likely to have identified the majority of the available literature on ^{18}FDG-PET or ^{18}FDG-PET/CT for the nine types of cancer addressed in this report.

We particularly targeted the indexed literature, however we acknowledge the possibility that the review may not be fully comprehensive, as we did not include evidence presented in abstracts of scientific meetings.

We provided detailed information on the review procedures (e.g., duplicate study selection, quality assessment by two assessors, and verification of data extraction to increase review reliability.)

We adopted a comprehensive strategy to appraise the methodological quality of the included studies. Our approach to quality focused mainly on an assessment of the internal validity of the studies as recommended by several researchers. [190]

Our decisions to include or exclude studies into meta-analysis of the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT were transparent and documented. For each study, summary statistics of test performance (e.g., sensitivity and specificity) were used to report the results of individual studies.

One of the limitations of this review is the restriction of included studies to English-language publications. We did not include foreign language literature because of the difficulties in translation. Twenty-one studies were excluded for this reason, with the majority of them being published in Japanese and Chinese languages. The magnitude of bias that the exclusion of foreign literature may have produced in the results of the pooled estimates of the positive and negative LRs is unknown.

The review included only studies with sample sizes of 12 or more adult participants. There are advantages and disadvantages to including small studies in systematic reviews of diagnostic tests. By including small studies, the power and precision of the pooled estimates would increase (in fact, some may argue that this is precisely the strength behind any meta-analysis). Likewise, avoiding

sample size considerations may be one of the main strengths of a systematic review. However, there are some disadvantages of including studies with very small numbers: they can introduce numeric instability in the accuracy estimates, and since small studies may be more prone to selection bias (due to non-consecutive patient sampling) and publication bias, their inclusion may incorporate biased data into the summary estimates. Limiting inclusion of studies based on sample size considerations seems to be a common practice when conducting systematic reviews of diagnostic tests.[190] To note, only 32 studies were excluded from the review based on sample-size considerations. Further research is needed to identify the best approach to solve the problems associated with the inclusion of small studies in systematic reviews and meta-analyses of diagnostic tests.

Another potential limitation is that only studies published from 2003 and forward were included in the review. Most of the scientific literature published before 2003 was covered in the Duke report [14] for six of the cancers considered in our review (i.e., brain, cervical, ovarian, pancreatic, SCLC, and testicular); however, a number of studies published before 2003 for three types of cancers that were not considered in the Duke report (i.e., prostate, bladder and kidney) have been left out. However, the technology around FDG PET has evolved dramatically over the last few years and it is expected that there is substantial variation between "old" and "new" studies in the manufacture and procedures for this technology.

Finally, a variety of diagnostic tests can be used in isolation or in combination with ^{18}FDG-PET and ^{18}FDG-PET/CT for the nine cancers that were considered in this review. The value of ^{18}FDG-PET and ^{18}FDG-PET/CT may differ by patient depending on the results of previous imaging, and the presence of various clinical signs and symptoms. The studies included in the review reported on individual values of ^{18}FDG-PET and ^{18}FDG-PET/CT and therefore, the review does not provide information about the aggregated diagnostic value from adding other tests to inform important therapeutic decisions. When interpreting the impact and value that ^{18}FDG-PET and ^{18}FDG-PET/CT may have on patient management decisions for a given patient, the clinical context must be taken into consideration.

Conclusions and Future Areas of Research

For some type of cancers (e.g., cervical, ovarian, and pancreatic cancer), there is some evidence of the utility of ^{18}FDG-PET or ^{18}FDG-PET/CT for diagnosing, staging, or detecting recurrences, but additional studies are required to augment the evidence base. Particularly, further studies are needed to reach firm conclusions about the clinical effectiveness of ^{18}FDG-PET and ^{18}FDG-PET/CT in terms of the impact on diagnosis and treatment options, patient-centered outcomes, and economic costs. It is still unclear how ^{18}FDG-PET and ^{18}FDG-PET/CT affect patient treatment and ultimately their outcome. For other types of cancer examined in the review (e.g., bladder, kidney, prostate, SCLC, and testicular) the utility of these tests is inconclusive and requires more careful study.

The strongest evidence for the diagnostic accuracy of ^{18}FDG-PET or ^{18}FDG-PET/CT has been produced for staging of locally advanced cervical cancer and the detection and restaging of recurrent disease, the detection of ovarian cancer recurrences following treatment, and the diagnosing and initial staging of pancreatic cancer. These and other indications require further research to show the impact of^{18}FDG-PET or ^{18}FDG-PET/CT on patient management or added value in the diagnostic and therapeutic process.

Although PET technology development appears to have reached maturity with the fusion of ^{18}FDG-PET and CT in an integrated system, imaging protocols will continue to be refined over the next few years. Further evaluations of the utility of this technology should be done with developments concentrating on enhancing patient throughput and establishing new and more focused clinical applications in various subpopulations of patients. Because it may be quite challenging to enrol patients into a study for low-incidence cancers, multicenter studies will be required to adequately address these important issues.

Finally, some of the most important roles of ^{18}FDG-PET and ^{18}FDG-PET/CT have not been sufficiently explored (e.g., estimating prognosis, selecting and changing treatment modalities, estimating their role in the evaluation of tumor burden regardless of histology). Evaluations of the procedures or therapies forestalled or cancelled based on ^{8}FDG-PET and ^{18}FDG-PET/CT must be explored further. If the total clinical contributions of ^{18}FDG-PET and ^{18}FDG-PET/CT have to be evaluated to inform policy decisions, these information gaps need to be filled with new methodological approaches.

Abbreviations

Abbreviation	Description
^{18}FDG	18F-fluorodeoxyglucose
AC	carbon-11 acetate
AD	adenocarcinoma
AFP	alpha-fetoprotein
AHRQ	Agency for Healthcare Research and Quality
ASC	adenosquamous carcinoma
ß-HCG	beta-human chorionic gonadotropin
BCG	Bacillus Clamette-Geurin
BEP	bleomycin, etoposide, and platinum
BMI	body mass index
BPH	benign prostatic hyperplasia
CA-125	cancer antigen 125
CCRT	concurrent chemotherapy and radiotherapy
CEA	carcinoembryonic antigen
CHEC	Consensus on Health Economic Criteria
CI	confidence interval
CIN	cervical intraepithelial neoplasia
CMS	Centers for Medicaid and Medicare Services
CNS	central nervous system
COI	conflict of interest
CP	chronic pancreatitis
CS	clinical stage
CT	computed tomography
D	days
DM	diabetes mellitus
DRE	digital rectal examination
EBRT	electron beam radiation therapy
ECOG	Eastern Cooperative Oncology Group
ED	extensive-disease
ERCP	endoscopic retrograde cholangiopancreatography
EPC	Evidence-based Practice Centers
EUS	endoscopic ultrasound
F-FMISO	18F-fluoromisonidazole
FIGO	Federation Internationale de Gynecologie et d'Obstetrique
FNA	fine needle aspiration
FOV	field of view
GBM	glioblastoma multiforme
GCT	germ cell tumors
GLUT	glucose transport proteins
GnRH	gonadotropin-releasing hormone
H	hours
HPV	human papilloma virus
HR	hazard ratio
ILN	Inguinal lymph node
IPMN	intraductal papillary mucinous neoplasms
IQR	interquartile range
IV	intravenous
LD	limited-disease
LDH	lactate dehydrogenase
LD	lymph nodes
LR	likelihood ratio
MAX	maximum
MET	carbon-11 methionine
M-H	Mantel-Hantzel

MIN	minutes
MLN	mediastinal lymph node
MRCP	magnetic resonance cholangiopancreatography
MO	months
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NA	not applicable
ND	not described
NLR	negative likelihood ratio
OSEM	ordered subset expectation maximization
PALN	para-aortic lymph node
PCI	prophylactic cranial irradiation
PET	positron emission tomography
PLN	pelvic lymph node
PLR	positive likelihood ratio
PO	oral
PSA	prostate-specific antigen
PTC	percutaneous transhepatic cholangiography
QUADAS	Quality Assessment of Studies of Diagnostic Accuracy
RCC	renal cell carcinoma
RCT	randomized controlled trial
RH-PLND	radical hysterectomy + pelvic lymphadenectomy
RI	retention index
RMI	risk of malignancy index
ROC	receiver operating characteristic
ROI	region of interest
RT	radiotherapy
SCC-Ag	squamous cell carcinoma antigen
SCLC	small cell lung cancer
SD	standard deviation
SEC	seconds
SIGN	Scottish Intercollegiate Guidelines Network
SLL	second-look laparotomy
SLN	supraclavicular lymph node
SROC	summary receiver operating characteristic
SUV	standardized uptake value
TA	technology assessment
TNMS	tumor, node, metastasis staging
TRUS	transrectal ultrasound
TUR	transurethral resection
TVUS	transvaginal ultrasonography
UICC	Union International Contre le Cancer
US	Ultrasound
VALCSG	Veterans Administration Lung Cancer Study Group
VATAP	Veterans Affairs Technology Assessment Program
WK	weeks
WHO	World Health Organization
YR	years

References and Included Studies

1. Robert G, Milne R. Positron emission tomography: establishing priorities for health technology assessment. Health Technol Assess 1999;3(16).

2. Al-Nahhas AM, Win Z, Singh A, et al. The role of 18F-FDG PET in oncology: Clinical and resource implications. Nucl Med Rev 2006;9:1-5.

3. Laupacis AD, Alter M, Mamdami M, et al. Health technology assessment of positron emission tomography (PET) a systematic review. ICES Investigative Report. Toronto: Institute for Clinical and Evaluative Sciences, 2001.

4. Jaeschke R, Guyatt G, Sackett D . Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1994;271(5):389-91.

5. EMEA (European Agency for the Evaluation of Medicinal Products). Points to consider on the evaluation of diagnostic agents. London (UK): EMEA, 2001.

6. Ioannidis JA, Lau J. 18F-FDG PET for the diagnosis and grading of soft-tissue sarcome: a meta-analysis. J Nucl Med 2003;44:717-24.

7. Havrilesky LJ , Kuasingam SL, Matchar DB, et al. FDG-PET for management of cervical and ovarian cancer. Gynecol Oncol 2005; 97:183-91.

8. Orlando LA, Kulasingam SL, Matchar DB. Meta-analysis: the detection of pancreatic malignancy with positron emission tomography. Aliment Pharmacol Ther 2004;20(10):1063-70.

9. Delgado-Bolton RC, Carreras JL, Perez-Castejon MAJ. A systematic review of the efficacy of 18F-FDG PET in unknown primary tumors. Curr Med Imaging Rev 2006;2(2):215-25.

10. Medicare National Coverage Determinations Manual. Chapter 1, Part 4 (Sections 200 310.1) Coverage Determinations. Revised 2007. [web page] Available at: http://www.cms.hhs.gov/manuals/downloads/ncd103c1_Part4.pdf. Accessed: April 9, 2008.

11. Evers S, Goosens M, de Vet H, et al. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care 2005;21(2):240-5.

12. Berger M, Gould MK, Barnett PG. The cost of positron emission tomography in six United States Veterans Affairs Hospitals and two academic medical centres. AJR 2003;181:359-65.

13. Virnig B, Durham S. Post coverage analysis: use of positron emission tomography (PET) scans. Minneapolis (MN): Research Data Assistance Center, University of Minnesota, 2004.

14. Natchar DB, Kulasingam SL, Havrilesky L, et al. Positron emission testing for six cancers (brain, cervicall, small cel lung, ovarian, pacreatic and testicular). Rockville (MD): Agency for Health Care Research and Quality, 2004.

15. Hillner BE, Siegel BA, Liu D, et al. Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the national oncologic PET registry. J Clin Oncol 2008;26(13):2155-61.

16. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88-94.

17. Scottish Intercollegiate Guidelines Network. Methodology checklist 5: studies of diagnostic accuracy [web page]. Available at: http://www.sign.ac.uk/guidelines/fulltext/50/checklist5.html. Accessed Mar 20, 2008.

18. Whiting J, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. Health Tech Assess 2004;8(25).

19. DerSimonian R. Meta-analysis in the design and monitoring of clinical trials. Stat Med 1996;15:1237-52.

20. Halkin A, Reichman J, Schwaber M, et al. Likelihood ratios: getting diagnostic testing into perspective. QJM 1998;91:247-58.

21. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med 1993;12:1293-316.

22. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327(557-60).

23. Drieskens O, Oyen R, Van Poppel H, et al. FDG-PET for preoperative staging of bladder cancer. Eur J Nucl Med Mol Imaging 2005;32(12):1412-7.

24. Jadvar H, Quan V, Henderson RW, et al. [F-18]-Fluorodeoxyglucose PET and PET-CT in diagnostic imaging evaluation of locally recurrent and metastatic bladder transitional cell carcinoma. Int J Clin Oncol 2008;13(1):42-7.

25. Liu IJ, Segall GM, Nino-Murcia M, et al. Fluorodeoxyglucose positron emission tomography studies in the diagnosis and staging of transitional cell carcinoma. Adv Exp Med Biol 2003;539(Pt A):129-42.

26. Chen W, Silverman DH, Delaloye S, et al. 18F-FDOPA PET imaging of brain tumors: comparison study with 18F-FDG PET and evaluation of diagnostic accuracy. J Nucl Med 2006;47(6):904-11.

27. Cher LM, Murone C, Lawrentschuk N, et al. Correlation of hypoxic cell fraction and angiogenesis with glucose metabolic rate in gliomas using 18F-fluoromisonidazole, 18F-FDG PET, and immunohistochemical studies. J Nucl Med 2006;47(3):410-8.

28. Liu RS, Chang CP, Chu LS, et al. PET imaging of brain astrocytoma with 1-11C-acetate. Eur J Nucl Med Mol Imaging 2006;33(4):420-7.

29. Potzi C, Becherer A, Marosi C, et al. [11C] methionine and [18F] fluorodeoxyglucose PET in the follow-up of glioblastoma multiforme. J Neuro-Oncol 2007;84(3):305-14.

30. Stockhammer F, Thomale UW, Plotkin M, et al. Association between fluorine-18-labeled fluorodeoxyglucose uptake and 1p and 19q loss of heterozygosity in World Health Organization Grade II gliomas. J Neurosurg 2007;106(4):633-7.

31. Amit A, Beck D, Lowenstein L, et al. The role of hybrid PET/CT in the evaluation of patients with cervical cancer. Gynecol Oncol 2006;100(1):65-9.

32. Bjurberg M, Kjellen E, Ohlsson T, et al. FDG-PET in cervical cancer: staging, re-staging and follow-up. Acta Obstet Gynecol Sc 2007;86(11):1385-91.

33. Chang TC, Law KS, Hong JH, et al. Positron emission tomography for unexplained elevation of serum squamous cell carcinoma antigen levels during follow-up for patients with cervical malignancies: a phase II study. Cancer 2004;101(1):164-71.

34. Chang WC, Hung YC, Lin CC, et al. Usefulness of FDG-PET to detect recurrent cervical cancer based on asymptomatically elevated tumor marker serum levels: a preliminary report. Cancer Invest 2004;22(2):180-4.

35. Chang YC, Yen TC, Ng KK, et al . Does diabetes mellitus influence the efficacy of FDG-PET in the diagnosis of cervical cancer? Eur J Nucl Med Mol Imaging 2005;32(6):647-52.

36. Choi HJ, Roh JW, Seo SS, et al. Comparison of the accuracy of magnetic resonance imaging and positron emission tomography/computed tomography in the presurgical detection of lymph node metastases in patients with uterine cervical carcinoma: a prospective study. Cancer 2006;106(4):914-22.

37. Chou HH, Chang TC, Yen TC, et al. Low value of [18F]-fluoro-2-deoxy-D-glucose positron emission tomography in primary staging of early-stage cervical cancer before radical hysterectomy. J Clin Oncol 2006;24(1):123-8.

38. Chung HH, Jo H, Kang WJ, et al . Clinical impact of integrated PET/CT on the management of suspected cervical cancer recurrence. Gynecol Oncol 2007;104(3):529-34.

39. Chung HH, Kim SK, Kim TH, et al. Clinical impact of FDG-PET imaging in post-therapy surveillance of uterine cervical cancer: from diagnosis to prognosis. Gynecol Oncol 2006;103(1):165-70.

40. Grisaru D, Almog B, Levine C, et al. The diagnostic accuracy of 18F-Fluorodeoxyglucose PET/CT in patients with gynecological malignancies. Gynecol Oncol 2004;94(3):680-4.

41. Havrilesky LJ, Wong TZ, Secord AA, et al. The role of PET scanning in the detection of recurrent cervical cancer. Gynecol Oncol 2003;90(1):186-90.

42. Hope AJ, Saha P, Grigsby PW. FDG-PET in carcinoma of the uterine cervix with endometrial extension. Cancer 2006;106(1):196-200.

43. Lai CH, Huang KG, See LC, et al. Restaging of recurrent cervical carcinoma with dual-phase [18F]fluoro-2-deoxy-D-glucose positron emission tomography. Cancer 2004;100(3):544-52.

44. Lin CT, Yen TC, Chang TC, et al. Role of [18F]fluoro-2-deoxy-D-glucose positron emission tomography in re-recurrent cervical cancer. Int J Gynecological Cancer 2006;16(6):1994-2003.

45. Lin WC, Hung YC, Yeh LS, et al . Usefulness of (18)F-fluorodeoxyglucose positron emission tomography to detect para-aortic lymph nodal metastasis in advanced cervical cancer with negative computed tomography findings. Gynecol Oncol 2003;89(1):73-6.

46. Loft A, Berthelsen AK, Roed H, et al. The diagnostic value of PET/CT scanning in patients with cervical cancer: a prospective study. Gynecol Oncol 2007;106(1):29-34.

47. Ma SY, See LC, Lai CH, et al. Delayed (18)F-FDG PET for detection of paraaortic lymph node metastases in cervical cancer patients. J Nucl Med 2003;44(11):1775-83.

48. Park W, Park YJ, Huh SJ, et al . The usefulness of MRI and PET imaging for the detection of parametrial involvement and lymph node metastasis in patients with cervical cancer. Jpn J Clin Oncol 2005;35(5):260-4.

49. Roh JW, Seo SS, Lee S, et al. Role of positron emission tomography in pretreatment lymph node staging of uterine cervical cancer: a prospective surgicopathologic correlation study. Eur J Cancer 2005;41(14):2086-92.

50. Ryu SY, Kim MH, Choi SC, et al . Detection of early recurrence with 18F-FDG PET in patients with cervical cancer. J Nucl Med 2003;44(3):347-52.

51. Sakurai H, Suzuki Y, Nonaka T, et al. FDG-PET in the detection of recurrence of uterine cervical carcinoma following radiation therapy--tumor volume and FDG uptake value. Gynecol Oncol 2006;100(3):601-7.

52. Sironi S, Picchio M, Landoni C, et al. Post-therapy surveillance of patients with uterine cancers: value of integrated FDG PET/CT in the detection of recurrence. Eur J Nucl Med Mol Imaging 2007;34(4):472-9.

53. Sironi S, Buda A, Picchio M, et al. Lymph node metastasis in patients with clinical early-stage cervical cancer: detection with integrated FDG PET/CT. Radiology 2006;238(1):272-9.

54. Tran BN, Grigsby PW, Dehdashti F, et al. Occult supraclavicular lymph node metastasis identified by FDG-PET in patients with carcinoma of the uterine cervix. Gynecol Oncol 2003;90(3):572-6.

55. Unger JB, Ivy JJ, Ramaswamy MR, et al. Whole-body [18F]fluoro-2-deoxyglucose positron emission tomography scan staging prior to planned radical hysterectomy and pelvic lymphadenectomy. Int J Gynecological Cancer 2005;15(6):1060-4.

56. Unger JB, Ivy JJ, Connor P, et al. Detection of recurrent cervical cancer by whole-body FDG PET scan in asymptomatic and symptomatic women. Gynecol Oncol 2004;94(1):212-6.

57. Van Der Veldt AAM, Hooft L, Van Diest PJ, et al. Microvessel density and p53 in detecting cervical cancer by FDG PET in cases of suspected recurrence. Eur J Nucl Med Mol Imaging 2006;33(12):1408-16.

58. Wong TZ, Jones EL, Coleman RE. Positron emission tomography with 2-deoxy-2-[(18)F]fluoro-D-glucose for evaluating local and distant disease in patients with cervical cancer. Mol Imaging Biol 2004;6(1):55-62.

59. Wright JD, Dehdashti F, Herzog TJ, et al. Preoperative lymph node staging of early-stage cervical carcinoma by [18F]-fluoro-2-deoxy-D-glucose-positron emission tomography. Cancer 2005;104(11):2484-91.

60. Yen TC, Lai CH, Ma SY, et al. Comparative benefits and limitations of [18]F-FDG PET and CT-MRI in documented or suspected recurrent cervical cancer. Eur J Nucl Med Mol Imaging 2006;33(12):1399-407.

61. Yen TC, See LC, Chang TC, et al. Defining the priority of using 18F-FDG PET for recurrent cervical cancer. J Nucl Med 2004;45(10):1632-9.

62. Yen TC, Ng KK, Ma SY, et al. Value of dual-phase 2-fluoro-2-deoxy-d-glucose positron emission tomography in cervical cancer.[erratum appears in J Clin Oncol. 2004 Jan 1;22(1):209]. J Clin Oncol 2003;21(19):3651-8.

63. Yildirim Y, Sehirali S, Avci ME, et al. Integrated PET/CT for the evaluation of para-aortic nodal metastasis in locally advanced cervical cancer patients with negative conventional CT findings. Gynecol Oncol 2008;108(1):154-9.

64. Aide N, Cappele O, Bottet P, et al. Efficiency of [(18)F]FDG PET in characterising renal cancer and detecting distant metastases: a comparison with CT. Eur J Nucl Med Mol Imaging 2003;30(9):1236-45.

65. Ak I, Can C. F-18 FDG PET in detecting renal cell carcinoma. Acta Radiol 2005;46(8):895-9.

66. Chang CH, Shiau YC, Shen YY, et al. Differentiating solitary pulmonary metastases in patients with renal cell carcinomas by 18F-fluoro-2-deoxyglucose positron emission tomography: a preliminary report. Urol Int 2003;71(3):306-9.

67. Dilhuydy MS, Durieux A, Pariente A, et al. PET scans for decision-making in metastatic renal cell carcinoma: a single-institution evaluation. Oncology 2006;70(5):339-44.

68. Jadvar H, Kherbache HM, Pinski JK, et al. Diagnostic role of [F-18]-FDG positron emission tomography in restaging renal cell carcinoma. Clin Nephrol 2003;60(6):395-400.

69. Kang DE, White RLJr, Zuger JH, et al. Clinical use of fluorodeoxyglucose F18 positron emission tomography for detection of renal cell carcinoma. J Urol 2004;171(5):1806-9.

70. Kumar R, Chauhan A, Lakhani P, et al. 2-Deoxy-2-[F-18]fluoro-D-glucose-positron emission tomography in characterization of solid renal masses. Mol Imaging Biol 2005;7(6):431-9.

71. Majhail NS, Urbain JL, Albani JM, et al. F-18 fluorodeoxyglucose positron emission tomography in the evaluation of distant metastases from renal cell carcinoma. J Clin Oncol 2003;21(21):3995-4000.

72. Bristow RE, Giuntoli RL2nd, Pannu HK, et al. Combined PET/CT for detecting recurrent ovarian cancer limited to retroperitoneal lymph nodes. Gynecol Oncol 2005;99(2):294-300.

73. Bristow RE, del Carmen MG, Pannu HK, et al. Clinically occult recurrent ovarian cancer: patient selection for secondary cytoreductive surgery using combined PET/CT. Gynecol Oncol 2003;90(3):519-28.

74. Castellucci P, Perrone AM, Picchio M, et al. Diagnostic accuracy of 18F-FDG PET/CT in characterizing ovarian lesions and staging ovarian cancer: correlation with transvaginal ultrasonography, computed tomography, and histology. Nucl Med Commun 2007;28(8):589-95.

75. Chung HH, Kang WJ, Kim JW, et al. Role of [18F]FDG PET/CT in the assessment of suspected recurrent ovarian cancer: correlation with clinical or histological findings. Eur J Nucl Med Mol Imaging 2007;34(4):480-6.

76. Drieskens O, Stroobants S, Gysen M, et al. Positron emission tomography with FDG in the detection of peritoneal and retroperitoneal metastases of ovarian cancer. Gynecol Obstet Invest 2003;55(3):130-4.

77. Garcia-Velloso MJ, Jurado M, Ceamanos C, et al. Diagnostic accuracy of FDG PET in the follow-up of platinum-sensitive epithelial ovarian carcinoma. Eur J Nucl Med Mol Imaging 2007;34(9):1396-405.

78. Hauth EA, Antoch G, Stattaus J, et al. Evaluation of integrated whole-body PET/CT in the detection of recurrent ovarian cancer. Eur J Radiol 2005;56(2):263-8.

79. Kawahara K, Yoshida Y, Kurokawa T, et al. Evaluation of positron emission tomography with tracer 18-fluorodeoxyglucose in addition to magnetic resonance imaging in the diagnosis of ovarian cancer in selected women after ultrasonography. J Comput Assist Tomogr 2004;28(4):505-16.

80. Kim CK, Park BK, Choi JY, et al. Detection of recurrent ovarian cancer at MRI: comparison with integrated PET/CT. J Comput Assist Tomogr 2007;31(6):868-75.

81. Murakami M, Miyamoto T, Iida T, et al. Whole-body positron emission tomography and tumor marker CA125 for detection of recurrence in epithelial ovarian cancer. Int J Gynecological Cancer 2006;16(Suppl 1):99-107.

82. Nanni C, Rubello D, Farsad M, et al. (18)F-FDG PET/CT in the evaluation of recurrent ovarian cancer: a prospective study on forty-one patients. Eur J Surg Oncol 2005;31(7):792-7.

83. Pannu HK, Cohade C, Bristow RE, et al. PET-CT detection of abdominal recurrence of ovarian cancer: radiologic-surgical correlation. Abdom Imaginging 2004;29(3):398-403.

84. Picchio M, Sironi S, Messa C, et al. Advanced ovarian carcinoma: usefulness of (18)F-FDG-PET in combination with CT for lesion detection after primary treatment. Q J Nucl Med 2003;47(2):77-84.

85. Risum S, Hogdall C, Loft A, et al. The diagnostic value of PET/CT for primary ovarian cancer--a prospective study. Gynecol Oncol 2007;105(1):145-9.

86. Sebastian S, Lee SI, Horowitz NS, et al. PET-CT vs. CT alone in ovarian cancer recurrence. Abdom Imaging 2008;33(1):112-8.

87. Sironi S, Messa C, Mangili G, et al. Integrated FDG PET/CT in patients with persistent ovarian cancer: correlation with histologic findings. Radiology 2004;233(2):433-40.

88. Takekuma M, Maeda M, Ozawa T, et al. Positron emission tomography with 18F-fluoro-2-deoxyglucose for the detection of recurrent ovarian cancer. Int J Clin Oncol 2005;10(3):177-81.

89. Thrall MM, DeLoia JA, Gallion H, et al. Clinical use of combined positron emission tomography and computed tomography (FDG-PET/CT) in recurrent ovarian cancer. Gynecol Oncol 2007;105(1):17-22.

90. Yoshida Y, Kurokawa T, Kawahara K, et al. Incremental benefits of FDG positron emission tomography over CT alone for the preoperative staging of ovarian cancer. AJR Am J Roentgenol 2004;182(1):227-33.

91. Bang S, Chung HW, Park SW, et al. The clinical usefulness of 18-fluorodeoxyglucose positron emission tomography in the differential diagnosis, staging, and response evaluation after concurrent chemoradiotherapy for pancreatic cancer. J Clin Gastroenterol 2006;40(10):923-9.

92. Borbath I, Van Beers BE, Lonneux M, et al. Preoperative assessment of pancreatic tumors using magnetic resonance imaging, endoscopic ultrasonography, positron emission tomography and laparoscopy. Pancreatology 2005;5(6): 553-61.

93. Casneuf V, Delrue L, Kelles A, et al. Is combined [18]F-fluorodeoxyglucose-positron emission tomography/computed tomography superior to positron emission tomography or computed tomography alone for diagnosis, staging and restaging of pancreatic lesions? Acta Gastro-Ent Belg 2007;70(4):331-8.

94. Giorgi MC, Cunha RM, Soares J Jr, et al. Dual-head gamma camera coincidence imaging in pancreatic cancer. Rev Esp Med Nucl 2004;23(2):90-4.

95. Heinrich S, Goerres GW, Schafer M, et al. Positron emission tomography/computed tomography influences on the management of resectable pancreatic cancer and its cost-effectiveness. Ann Surg 2005;242(2):235-43.

96. Lemke AJ, Niehues SM, Hosten N, et al. Retrospective digital image fusion of multidetector CT and 18F-FDG PET: clinical value in pancreatic lesions--a prospective study with 104 patients. J Nucl Med 2004;45(8):1279-86.

97. Lytras D, Connor S, Bosonnet L, et al. Positron emission tomography does not add to computed tomography for the diagnosis and staging of pancreatic cancer. Dig Surg 2005;22(1-2):55-62.

98. Maemura K, Takao S, Shinchi H, et al. Role of positron emission tomography in decisions on treatment strategies for pancreatic cancer. J Hepatobil Pancreat Surg 2006;13(5):435-41.

99. Mansour JC, Schwartz L, Pandit-Taskar N, et al. The utility of F-18 fluorodeoxyglucose whole body PET imaging for determining malignancy in cystic lesions of the pancreas. J Gastrointest Surg 2006;10(10):1354-60.

100. Nishiyama Y, Yamamoto Y, Monden T, et al. Evaluation of delayed additional FDG PET imaging in patients with pancreatic tumour. Nucl Med Commun 2005;26(10):895-901.

101. Nishiyama Y, Yamamoto Y, Yokoe K, et al. Contribution of whole body FDG-PET to the detection of distant metastasis in pancreatic cancer. Ann Nucl Med 2005;19(6):491-7.

102. Rasmussen I , Sorensen J, Langstrom B, et al. Is positron emission tomography using 18F-fluorodeoxyglucose and 11C-acetate valuable in diagnosing indeterminate pancreatic masses? Scand J Surg 2004;93(3):191-7.

103. Ruf J, Lopez Hanninen E, Bohmig M, et al. Impact of FDG-PET/MRI image fusion on the detection of pancreatic cancer. Pancreatology 2006;6(6):512-9.

104. Ruf J, Lopez Hanninen E, Oettle H, et al. Detection of recurrent pancreatic cancer: comparison of FDG-PET with CT/MRI. Pancreatology 2005;5(2-3):266-72.

105. Sperti C, Bissoli S, Pasquali C, et al. 18-fluorodeoxyglucose positron emission tomography enhances computed tomography diagnosis of malignant intraductal papillary mucinous neoplasms of the pancreas. Ann Surg 2007;246(6):932-9.

106. van Kouwen MC, Jansen JB, van Goor H, et al. FDG-PET is able to detect pancreatic carcinoma in chronic pancreatitis. Eur J Nucl Med Mol Imaging 2005;32(4):399-404.

107. Wakabayashi H, Nishiyama Y, Otani T, et al. Role of 18F-fluorodeoxyglucose positron emission tomography imaging in surgery for pancreatic cancer. World J Gastroenterol 2008;14(1):64-9.

108. Chang CH, Wu HC, Tsai JJ, et al. Detecting metastatic pelvic lymph nodes by 18F-2-deoxyglucose positron emission tomography in patients with prostate-specific antigen relapse after treatment for localized prostate cancer. Urol Int 2003;70(4):311-5.

109. Jadvar H, Pinski JK, Conti PS . FDG PET in suspected recurrent and metastatic prostate cancer. Oncol Reports 2003;10(5):1485-8.

110. Oyama N, Miller TR, Dehdashti F, et al. 11C-acetate PET imaging of prostate cancer: detection of recurrent disease at PSA relapse. J Nucl Med 2003;44(4):549-55.

111. Schoder H, Herrmann K, Gonen M, et al. 2-[18F]fluoro-2-deoxyglucose positron emission tomography for the detection of disease in patients with prostate-specific antigen relapse after radical prostatectomy. Clin Cancer Res 2005;11(13):4761-9.

112. Blum R, MacManus MP, Rischin D, et al. Impact of positron emission tomography on the management of patients with small-cell lung cancer: preliminary experience. Am J Clin Oncol 2004;27(2):164-71.

113. Bradley JD, Dehdashti F, Mintun MA, et al. Positron emission tomography in limited-stage small-cell lung cancer: a prospective study. J Clin Oncol 2004;22(16):3248-54.

114. Brink I, Schumacher T, Mix M, et al. Impact of [18F]FDG-PET on the primary staging of small-cell lung cancer. Eur J Nucl Med Mol Imaging 2004;31(12):1614-20.

115. Fischer BM, Mortensen J, Langer SW, et al. A prospective study of PET/CT in initial staging of small-cell lung cancer: comparison with CT, bone scintigraphy and bone marrow analysis. Ann Oncol 2007;18(2):338-45.

116. Fischer BM, Mortensen J, Langer SW, et al. PET/CT imaging in response evaluation of patients with small cell lung cancer. Lung Cancer 2006;54(1):41-9.

117. Kamel EM, Zwahlen D, Wyss MT, et al. Whole-body (18)F-FDG PET improves the management of patients with small cell lung cancer. J Nucl Med 2003;44(12):1911-7.

118. Kut V, Spies W, Spies S, et al. Staging and monitoring of small cell lung cancer using [18F]fluoro-2-deoxy-D-glucose-positron emission tomography (FDG-PET). Am J Clin Oncol 2007;30(1):45-50.

119. Niho S, Fujii H, Murakami K, et al. Detection of unsuspected distant metastases and/or regional nodes by FDG-PET [corrected] scan in apparent limited-disease small-cell lung cancer.[erratum appears in Lung Cancer. 2007 Dec;58(3):432]. Lung Cancer 2007;57(3):328-33.

120. Pandit N, Gonen M, Krug L, et al. Prognostic value of [18F]FDG-PET imaging in small cell lung cancer. Eur J Nucl Med Mol Imaging 2003;30(1):78-84.

121. Vinjamuri M , Craig M, Campbell-Fontaine A, et al. Can positron emission tomography be used as a staging tool for small-cell lung cancer? Clin Lung Cancer 2008;9(1):30-4.

122. Becherer A, De Santis M, Karanikas G, et al. FDG PET is superior to CT in the prediction of viable tumour in post-chemotherapy seminoma residuals. Eur J Radiol 2005;54(2):284-8.

123. Hinz S, Schrader M, Kempkensteffen C, et al. The role of positron emission tomography in the evaluation of residual masses after chemotherapy for advanced stage seminoma. J Urol 2008;179(3):936-40.

124. Karapetis CS, Strickland AH, Yip D, et al. Use of fluorodeoxyglucose positron emission tomography scans in patients with advanced germ cell tumour following chemotherapy: single-centre experience with long-term follow up. Intern Med J 2003;33(9-10):427-35.

125. Lassen U, Daugaard G, Eigtved A, et al. Whole-body FDG-PET in patients with stage I non-seminomatous germ cell tumours. Eur J Nucl Med Mol Imaging 2003 ;30(3):396-402.

126. Mangili G, Picchio M, Sironi S, et al. Integrated PET/CT as a first-line re-staging modality in patients with suspected recurrence of ovarian cancer. Eur J Nucl Med Mol Imaging 2007;34(5):658-66.

127. B, Neesham D, Quinn M, et al. The impact of PET/CT in the management of recurrent ovarian cancer. Gynecol Oncol 2006;103(1):271-6.

128. Soussan M, Wartski M, Cherel P, et al. Impact of FDG PET-CT imaging on the decision making in the biologic suspicion of ovarian carcinoma recurrence. Gynecol Oncol 2008;108(1):160-5.

129. Padma MV, Said S, Jacobs M, et al. Prediction of pathology and survival by FDG PET in gliomas. J Neuro-Oncol 2003;64(3):227-37.

130. Kim S, Chung JK, Kang SB, et al. [18F]FDG PET as a substitute for second-look laparotomy in patients with advanced ovarian carcinoma. Eur J Nucl Med Mol Imaging 2004;31(2):196-201.

131. Kirkali Z, Chan T, Manoharan M, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. Urology 2005;66(suppl 6A):4-34.

132. Babjuk M, Oosterlinck W, Sylvester R, et al. EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder. Eur Urol 2008;Apr 30. Epub ahead of print.

133. Ries LAG, Melbert D, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2005. Bethesda (MD): National Cancer Institute, 2008.

134. Kitamura H, Tsukamoto T. Early bladder cancer: concept, diagnosis, and management. Int J Clin Oncol 2006;11(1):28-37.

135. Sylvester RJ, van der Meijden A, Witjes JA, et al. High-grade Ta urothelial carcinoma and carcinoma in situ of the bladder. Urology 2005;66(suppl 6a):90-107.

136. Troung MT. Current role of radiation therapy in the mangement of malignant brain tumors. Hematol Oncol Clin N Am 2006;20(2):431-53.

137. Fisher JL, Schwartzbaum JA, Wrensch M, et al. Epidemology of brain tumors. Neurol Clin 2007;25(4):867-90.

138. Chandana SR, Movva S, Arora M ST. Primary brain tumors in adults. Am Fam Physician 2008;77(10):1423-30.

139. Safaeian M, Solomon D, Castle PE. Cervical cancer prevention - cervical screening: science in evolution. Obstet Gynecol Clin North Am 2007;34(4):739-60.

140. Cannistra SA, Niloff JM. Cancer of the uterine cervix. NEJM 1996;334(16):1030-8.

141. Whitcomb BP. Gynecologic malignancies. Surg Clin North Am 2008;88(2):301-17.

142. Rose PG. Chemoradiotherapy for cervical cancer. Eur J Cancer 2002;38(2):270-8.

143. Petignat P, Roy M. Diagnosis and management of cervical cancer. BMJ 2007;335(7623):765-8.

144. Jadvar H, Conti P. The reproductive tract. Semin Nucl Med 2004;34(4):262-73.

145. Singh AK, Grigsby PW, Dehdashti F, et al. FDG-PET lymph node staging and survival of patients with FIGO stage IIIb cervical carcinoma. Int J Radiat Oncol Biol Phys 2003;56(2):489-93.

146. Rini BI, Rathmell, WK, et al. Renal cell carcinoma. Curr Opin Oncol 2008;20(3):300-6.

147. Corgna E, Betti M, Gatta G, et al. Renal Cancer. Crit Rev Oncol 2007;64(3):247-62.

148. Holshneider CH, Berek JS. Ovarian cancer: epidemiology, biology and prognostic factors. Semin Surg Oncol 2000;19(1):3-10.

149. Modugno F, Ovarian Cancer and High-Risk Woman Symposium Presenters. Ovarian cancer and high-risk women: implications for prevention, screening, and early detection. Gynecol Oncol 2003;91(1):15-31.

150. Scully RE. Classification of human ovarian tumors. Environ Health Perspect 1987;73:15-24.

151. Abei S, Castiglione M. Epithelial ovarian carcinoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2008;19(suppl2):ii14-6.

152. Freelove R, Walling AD. Pancreatic cancer: diagnosis and management. Am Fam Physician 2006;73(3):485-92.

153. Goonetilleke KS, Siriwardena AK. Current status of gene expression profiling of pancreatic cancer. Int J Surg 2008;6(1):81-3.

154. Li D, Xie K, Wolff R, et al. Pancreatic cancer. Lancet 2004;363:1049-57.

155. Michl P, Pauls S, Gress TM. Evidence-based diagnosis and staging of pancreatic cancer. Best Pract Res Clin Gastroenterol 2006;20(2):227-51.

156. Horwich A, Parker C, Kataja V, et al. Prostate cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2008;19(suppl2):ii45-6.

157. U.S. Cancer Statistics Working Group. United States cancer statistics: 2004 incidence and mortality. 2007: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention *and* National Cancer Institute, Atlanta (GA).

158. National Cancer Institute: Bethesda M. Cancer advances in focus: prostate cancer. [web page]. Available at: http://www.cancer.gov/cancertopics/cancer-advances-in-focus/prostate.pdf. Accessed June 16, 2008.

159. National Cancer Institute: Bethesda M. Prostate Cancer Screening (PDQ®) [web page]. Available at: http://www.nci.nih.gov/cancertopics/pdq/screening/prostate/healthprofessional/allpages#Section_20. Accessed June 16, 2008.

160. Damber JE AG. Prostate cancer. Lancet 2008;371(9625):1710-21.

161. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol 1974;111(1):58-64.

162. Jewett HJ. The present status of radical prostatectomy for stages A and B prostatic cancer. Urol Clin North Am 1975;2(1):105-24.

163. Prostate. American Joint Committee on Cancer AJCC Cancer Staging Manual. 6th ed ed. New York (NY): Springer; 2002. p. 309-16.

164. Dahm P, Yeung LL, Chang SS, et al. A critical review of clinical practice guidelines for the management of clinically localized prostate cancer. J Urol 2008;180:451-60.

165. Wilt TJ, MacDonald R, Rutks I, et al. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. Ann Intern Med 2008;148(6):435-48.

166. Lin AM, Small EJ. Prostate cancer update: 2007. Curr Opin Oncol 2008;20(3):294-9.

167. Robinson JW , Saliken JC, Donnelly BJ, et al. Quality-of-life outcomes for men treated with cryosurgery for localized prostate carcinoma. Cancer 1999;86(9):1793-801.

168. Alibhai SM, Klotz LH. A systematic review of randomized trials in localized prostate cancer. Can J Urol 2004;11(1):2110-7.

169. Ferraldeschi R, Baka S, Jyoti B, et al. Modern management of small-cell lung cancer. Drugs 2007;67(15):2135-52.

170. Sher T, Dy Gk, Adjei AA. Small cell lung cancer. Mayo Clin Proc 2008;83(3):355-67.

171. Govindan R, Page N, Morgensztern D, et al. Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic and end results database. J Clin Oncol 2006;24(28):4539-44.

172. Chua YJ, Steer C, Yip D. Recent advances in management of small-cell lung cancer. Cancer Treat Rev 2004;30(6):521-43.

173. Leo F, Pastorino U. Surgery in small-cell lung carcinoma: where is the rationale? Sem Surg Oncol 2003;21(3):176-81.

174. Shepherd FA , Crowly J, Van Houtte P, et al. International association for the study of lung cancer lung cancer staging project: proposals regarding the clinical staging of small cell lung cancer in the forthcoming (seventh) edition of the tumor, node, metastasis classification for lung cancer. J Thor Oncol 2007;2(12):1067-77.

175. Feldman DR, Bosl George J, Sheinfeld J, et al. Medical treatment of advanced testicular cancer. JAMA 2008;299(6):672-84..

176. National Cancer Institute: Bethesda M. Testicular cancer: questions and answers. [web page]. Available at: http://www.cancer.gov/cancertopics/factsheet/Sites-Types/testicular

177. Shaw J. Diagnosis and treatment of testicular cancer. Am Fam Physician 2008;77(4):469-74.

178. Krege S, Beyer J, Souchon R, et al. European consensus conference on diagnosis and treatment of germ cell cancer: a report of the second meeting of the European Germ Cell Cancer Consensus Group (EGCCCG): part I. Eur Urol 2008;53(3):478-96.

179. Machtens S, Serth J, Meyer A, et al. World J Urol 2007;25(4):341-9.

180. Jafri SZ, Dinan D, Francis IR, et al. Pretreatment staging of invasive bladder cancer. Reston (VA): American College of Radiology, 2007.

181. Ascher SM, Brown JJ, Walsh JW. Diagnostic Imaging techniques in gynecologic oncology. Principles and Practice of Gynecologic Oncology. 2 ed. Philidelphia: Lippincott-Raven Publishers, 1997.

182. Dose J, Kotzerke J, Schicha H, et al. Clinical value of FDG PET in the diagnosis of gynecological tumors: recommendations of the 3rd interdisciplinary german Consensus conference PET in Oncology III. Geburtsh Frauenheilk 2002;62(7):644-51.

183. Javitt MC, Fleischer AC, Andreotti RF, et al.

Staging and follow-up of ovarian cancer. Reston (VA): American College of Radiology, 2007.

184. FDG positron emission tomography in pancreas. Chicago, IL: Blue Cross and Blue Shield Association, 2000.

185. Higashi T, Saga T, Nakamoto Y, et al. Diagnosis of pancreatic cancer using fluorine-18 fluorodeoxyglucose positron emission tomography (FDG PET): usefulness and limitations in "clinical reality". Ann Nuclear Med 2003;17(4):261-79.

186. Pound CR, Partin AW, Eisenberger MA, et al. Natural history of progression after PSA elevation following radical prostatectomy. JAMA 1999;281:1591-7.

187. Anscher MS. Salvage therapy for recurrent prostate cancer: the earlier the better. JAMA 2004;291:1380-2.

188. Pignon JP, Arriagada R, Ihde DC et al. Effect of thoracic radiotherapy on mortality in limited small cell lung cancer. A meta-analysis of 13 randomized trials among 2,140 patients. N Engl J Med 1992;327:1618-24.

189. Papanicolaou N, Francis IR, Baumgarten DA, et al. taging of testicular malignancy. Reston (VA): American College of Radiology, 2007.

190. Mallet S, Deeks JJ, Halligan S, et al. Systematic reviews of diagnostic tests in cancer: review of methods and reporting. BMJ 2006; 333:413.