



# Effective Health Care Program

## Imaging Tests for the Diagnosis and Staging of Pancreatic Adenocarcinoma

### Executive Summary

#### Background

##### Pancreatic Adenocarcinoma

Pancreatic cancer is the fourth most common cause of cancer death among men and women in the United States.<sup>1,2</sup> In 2013 in the United States, about 46,000 people received a diagnosis of pancreatic cancer and 40,000 died of the disease.<sup>3</sup> Risk factors for pancreatic cancer include tobacco use; personal history of chronic pancreatitis, diabetes, or obesity; and a family history of pancreatic cancer.<sup>1</sup> About 10 percent of patients with pancreatic cancer have a positive family history of the disease.<sup>4</sup> Pancreatic cancer incidence rates were reportedly highest among African American men (21.3 per 100,000) and women (17.6 per 100,000) during 2004 and 2008.<sup>1</sup> The second highest rates were reported for Caucasian men (16.8 per 100,000) and women (12.8 per 100,000).<sup>1</sup> The differences between these populations and burden of disease may be related to higher rates of cigarette smoking and diabetes mellitus among African American men than for Caucasian men and higher body mass indexes among African American women than for Caucasian women.<sup>1</sup>

##### Diagnosis and Staging

Patients with early pancreatic cancer are often asymptomatic or have only

#### Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

nonspecific symptoms such as malaise, fatigue, and appetite loss. As a result, patients often present with advanced stage disease when weight loss, jaundice, and severe abdominal pain often appear. Due to this delayed diagnosis, approximately 80 percent to 85 percent of cases are too



Agency for Healthcare Research and Quality

Advancing Excellence in Health Care • [www.ahrq.gov](http://www.ahrq.gov)

Effective  
Health Care

advanced to permit surgical resection,<sup>5</sup> and the median survival for patients with unresectable tumors is only 6–10 months.<sup>6</sup>

Given the poor prognosis of this disease, it is important for patients to understand differences in the ability of imaging modalities to diagnose pancreatic adenocarcinoma. Also, elucidating patients' experience and tolerance of various imaging modalities may help future patients weigh the benefits and harms of these tests and allow patients and their providers to incorporate individual values and priorities into the choice of imaging evaluation. Many patients are willing to experience some discomfort during an imaging test to ensure that their disease is appropriately diagnosed and staged. However, if two tests are equally accurate, test tolerance may be an important outcome.

Once pancreatic adenocarcinoma is diagnosed, staging the disease is critical and is the key determinant of clinical management, as well as a key predictor of survival. As noted, most cases are diagnosed at an advanced stage, precluding surgical resection.<sup>1</sup> When pancreatic adenocarcinoma is diagnosed at an advanced stage, the 5-year survival is approximately 2 percent. However, when pancreatic cancer is diagnosed at a localized stage, the 5-year survival is approximately 22 percent.<sup>1</sup>

Currently, there are no widely accepted clinical practice guidelines with strong recommendations on which imaging modalities to use in the diagnosis and staging of pancreatic cancer.

## Resectability

Surgical resection offers the only hope of cure and is typically determined via multidisciplinary consultation (e.g., surgeon, gastroenterologist, radiologist, medical oncologist, radiation oncologist) considering a variety of factors. The two key factors in assessing resectability are distant metastasis (which excludes resectability) and blood vessel involvement (which sometimes excludes unresectability, depending on the degree of involvement). The major blood vessels of focus are the celiac artery, common hepatic artery, superior mesenteric artery, superior mesenteric vein, and portal vein. The resectability criteria continue to evolve as surgical techniques advance and more tumors are resectable via reconstruction of blood vessels.<sup>7</sup>

Multidetector computed tomography (MDCT) scan is often the first imaging test in a patient whose symptoms suggest pancreatic adenocarcinoma. This widely available test provides three dimensional (3D) multiplanar reconstruction images within a single breath-hold, enabling

determination of tumor size, extent, and spread, with a standardized pancreas protocol.<sup>8,9</sup> However, MDCT does not always differentiate malignant from benign pancreatic lesions, and its ability to detect small tumors or small hepatic/peritoneal metastases is limited. Another concern about MDCT, particularly when used for screening, is that the procedure exposes the patient to potential harm through ionizing radiation.

MDCT can be performed using standard technique, whereby a single scan is obtained during delayed (i.e., venous) phase of enhancement. MDCT can also be performed using angiographic technique whereby images are obtained using at least two scans during arterial and venous phases of enhancement to permit more confident identification of arteries for surgical mapping.

Other commonly used imaging technologies for diagnosing and staging pancreatic adenocarcinoma include endoscopic ultrasound with fine-needle aspiration (EUS-FNA), magnetic resonance imaging (MRI), and positron emission tomography–computed tomography (PET/CT). EUS-FNA provides image-guided tissue sampling by placement of an endoscope into the upper gastrointestinal tract. MRI is noninvasive and provides detailed information about soft tissues, including the pancreas, in multiple planes. PET/CT provides information about tissue function through radiotracers that can be anatomically localized through CT. However, PET/CT exposes patients to radiation, mainly through the administration of the radiotracer.

## Screening

Screening for pancreatic adenocarcinoma is not recommended for the general population (e.g., the U.S. Preventive Services Task Force gives a D recommendation).<sup>10</sup> However, some recommend screening people at high risk of developing pancreatic cancer, such as those having two or more first-degree relatives with pancreatic cancer or those carrying specific genetic risk factors, such as Peutz-Jeghers syndrome or carriers of *BRCA2*, *PALB2*, *p16* gene mutations.<sup>11</sup> The most suitable imaging technology for screening high-risk populations is unclear.

## Objectives of This Review

Our objectives were to synthesize the available information on the diagnostic accuracy and clinical utility of various imaging tests for the diagnosis and staging of pancreatic adenocarcinoma, as well as screening for pancreatic adenocarcinoma. The availability of this information will assist clinicians in selecting imaging tests, may reduce variability across treatment centers in staging

protocols, and may improve patient outcomes. A secondary objective is to identify gaps in the evidence base to inform future research needs.

## Scope and Key Questions

The Key Questions (KQ) are listed below:

1. What is the comparative effectiveness of imaging techniques (e.g., MDCT, MDCT angiography, EUS-FNA, PET/CT, MRI) for **diagnosis** of pancreatic adenocarcinoma in adults with suspicious symptoms
  - a. What is the accuracy of each imaging technique for diagnosis and assessment of resectability?
  - b. What is the comparative accuracy of the different imaging techniques for diagnosis and assessment of resectability?
  - c. What is the comparative diagnostic accuracy of using a single imaging technique versus using multiple imaging techniques?
  - d. How is test experience (e.g., operative experience, assessor experience, center's annual case volume) related to comparative diagnostic accuracy of the different imaging strategies?
  - e. How are patient factors and tumor characteristics related to the comparative diagnostic accuracy of the different imaging strategies?
  - f. What is the comparative clinical management after the different imaging strategies when used for diagnosis?
  - g. What is the comparative impact of the different imaging strategies on long-term survival and quality of life when used for diagnosis?
2. What is the comparative effectiveness of imaging techniques (e.g., MDCT, MDCT angiography, EUS-FNA, PET/CT, MRI) for **staging** of pancreatic adenocarcinoma among adults with a diagnosis of pancreatic adenocarcinoma?
  - a. What is the staging accuracy of each imaging technique (for tumor size, lymph node status, vessel involvement, metastases, stage I–IV, and resectability)?
  - b. What is the comparative staging accuracy among the different imaging techniques?
  - c. What is the comparative staging accuracy of using a single imaging technique versus using multiple imaging techniques?
  - d. How is test experience (e.g., operative experience, assessor experience, center's annual case volume) related to comparative staging accuracy of the different imaging strategies?
  - e. How are patient factors and tumor characteristics related to the comparative staging accuracy of the different imaging strategies?
  - f. What is the comparative clinical management of the different imaging strategies when used for staging?
  - g. What is the comparative impact of the different imaging strategies on long-term survival and quality of life when used for staging?
3. What are the rates of harms of imaging techniques (e.g., MDCT, MDCT angiography, EUS-FNA, PET/CT, MRI) when used to diagnose and/or stage pancreatic adenocarcinoma?
  - a. How are patient factors related to the harms of different imaging techniques?
  - b. What are patient perspectives on the tolerance of different imaging techniques and the balance of benefits and harms of different imaging techniques?
4. What is the screening accuracy of imaging techniques (e.g., MDCT, MDCT angiography, EUS-FNA, PET/CT, MRI) for detecting precursor lesion(s) of pancreatic cancer or pancreatic adenocarcinoma in high-risk asymptomatic adults (i.e., those at genetic or familial risk of pancreatic adenocarcinoma)?

## PICOTS

### Populations

- Adult patients with symptoms in whom pancreatic adenocarcinoma is suspected
- Adult patients with symptoms with an established diagnosis of pancreatic adenocarcinoma
- Adult patients without symptoms who are at high risk of having or developing pancreatic adenocarcinoma (family history or genetic risk factor)

### Interventions

Imaging using one or more of the following tests:

- Multidetector computed tomography (MDCT)
- MDCT angiography (with or without 3D reconstruction)
- Endoscopic ultrasound with fine-needle aspiration (EUS-FNA)

- Magnetic resonance imaging (MRI)
- Positron emission tomography combined with computed tomography (PET/CT)

### Comparators

- Any direct comparisons of the imaging tests of interest
- Reference standards to assess test performance
  - Histopathological examination of tissue and/or biopsy
  - Intra-operative findings
  - Clinical followup

### Outcomes

- Accuracy
  - Test performance (sensitivity, specificity, under-, overstaging)
- Intermediate outcomes
  - Therapeutic management
- Clinical outcomes
  - Mortality
  - Quality of life
- Adverse effects and harms
  - Procedural harms of testing (e.g., radiation exposure, puncture from FNA)

### Timing

- Any time points will be considered

### Setting

- Any setting will be considered

## Methods

### Search Strategy

Medical librarians in the Evidence-based Practice Center (EPC) Information Center performed literature searches, following established systematic review protocols. We searched the following databases using controlled vocabulary and text words: Embase, MEDLINE, PubMed, and The Cochrane Library from 1980 through November 1, 2013. The literature searches will be updated during the peer review process, before finalization of the review.

### Study Selection

Our criteria are listed in five categories below: (1) publication criteria, (2) study design criteria, (3) patient criteria, (4) test criteria, and (5) data criteria.

### Publication Criteria

- Full-length articles: The article must have been published as a full-length, peer-reviewed study.
- Redundancy: To avoid double-counting patients, in instances in which several reports of the same or overlapping groups of patients were available, only outcome data based on the larger number of patients were included. However, we included data from publications with lower numbers of patients when either (a) a publication with lower patient enrollment reported an included outcome that was not reported by other publications of that study, or (b) a publication with lower patient enrollment reported longer followup data for an outcome.
- English language: Although we recognize that in some situations exclusion of non-English studies could lead to bias, we believe that the few instances in which this may occur do not justify the time and cost typically necessary for translating studies.
- Publication date: We included studies published since January 1, 2000. Older articles likely included outdated technologies. Studies of harms of imaging technologies that did not specifically involve pancreatic adenocarcinoma (i.e., any clinical indication), must have been published since January 1, 2009.

### Study Design Criteria

- For KQs on single-test accuracy: For KQs 1a and 1b, which address the performance of a single imaging test against a reference standard, we included only systematic reviews. EPC guidance by White et al. (2009)<sup>12</sup> states how existing systematic reviews can be used to replace de novo processes in comparative effectiveness reviews. We referred to the PICOTS-SD for the pertinent subquestion, and these seven components (Populations, Interventions, Comparisons, Outcomes, Time points, Setting, Study design) were the seven inclusion criteria. For quality, see section on risk of bias.
- For any KQs comparing two or more tests, the study must have compared both tests to a reference standard. The reference standard must not have been defined by either imaging test being assessed.
- For any KQs on single versus multiple tests, test experience, patient factors (e.g., age), or tumor characteristics (e.g., head or tail of pancreas), the study must have made a comparison of data to address the question.

- d. For any KQs involving comparative clinical management or long-term survival/quality of life, some patients must have received one of the imaging tests, and a separate group of patients must have received a different imaging test. This design permits a comparison of how the choice of test may influence management and/or survival and/or quality of life.
- e. For KQ3 on the rates of procedural harms, we included any reported harms data based on 50 or more patients, in the context of diagnosis or staging of pancreatic adenocarcinoma, on the harms of imaging procedures that contained a statement in the Methods section that the study planned in advance to capture harms/ complications data. Additionally, we included studies primarily of harms and adverse events associated with the use of each specific imaging modality, regardless of the type of cancer being detected, that were published in 2009 or later.
- f. For KQ3b on patient perspectives of imaging tests, any study design was accepted.
- g. For KQ4 on screening, we included any study that reported the performance of at least one included imaging test in the context of screening for either pancreatic adenocarcinoma itself or precursor lesions to pancreatic cancer.

### Patient Criteria

- a. To be included, the study must have reported data obtained from groups of patients in which at least 85 percent of the patients were from one of the patient populations of interest. If a study reported multiple populations, it must have reported data separately for one or more of the populations of interest.
- b. Adults. At least 85 percent of patients must have been aged 18 years or older, or data must have been reported separately for those aged 18 years or older.
- c. Studies of the screening/diagnosis/staging of *primary* pancreatic adenocarcinoma were included. Testing for *recurrent* pancreatic cancer was excluded.
- d. Data on imaging tests performed after any form of treatment (e.g., neoadjuvant chemotherapy) were excluded, but pretreatment imaging data were considered.

### Test Criteria

- a. Type of test. Only studies of the imaging tests of interest were included (listed in the key questions above). Studies of computed tomography (CT) that did not explicitly state that (or it could not be determined

that) CT was MDCT were assumed to be MDCT. Given our publication date criterion of 2000 and later, we believe it is safe to assume that CT performed in such studies was MDCT.

### Data Criteria

- a. The study must have reported data pertaining to one of the outcomes of interest (see the key questions section).
  - For accuracy outcomes (KQ1a through 1e, KQ2a through 2e, and KQ4), this means reporting enough information for one to calculate both sensitivity and specificity, along with corresponding confidence intervals (CIs).
  - For clinical management (KQ1f, KQ2f), this means reporting the percentage of patients who received a specific management strategy, after undergoing each imaging test (a separate group of patients corresponding to each imaging test).
  - For long-term survival (KQ1g, KQ2g), this means either reporting median survival after each imaging test (separate groups of patients), or mortality rates at a given time point (separate groups of patients), or other patient survival such as a hazard ratio.
  - For quality of life (KQ1g, KQ2g), this means reporting data on a previously validated quality-of-life instrument (such as the SF-36) after each imaging test (separate groups of patients).
  - For harms (KQ3), this means a statement appearing in the Methods section that harms/complications would be measured, reporting the occurrence of a procedure-related harm and number of patients at risk, or the reporting that no harms or complications occurred as a result of the procedure.
  - For patient perspectives (KQ3b), this means reporting the results of asking patients about their opinions or experience after having undergone one or more of the imaging tests.
- b. Regarding the minimum patient enrollment, for studies comparing imaging tests (KQ1b through 1g; KQ2b through 2g), we required data on at least 10 patients per imaging test. We also used a minimum of 10 for KQ3b on patient perspectives of imaging tests. We used a minimum of 50 patients for data on harms (KQ3) or screening (KQ4).
- c. For all KQs, the reported data must have included at least 50 percent of the patients who had initially enrolled in the study.

- d. Studies that reported data by tumor (e.g., x percent of pancreatic adenocarcinoma tumors were correctly detected) instead of by patient (e.g., x percent of enrolled patients were correctly given a diagnosis of pancreatic adenocarcinoma) were not excluded because of this difference. However, we separated the tumor-based data from the patient-based data because they measure different types of accuracy.

### Data Abstraction

Duplicate abstraction of comparative accuracy data was used to ensure accuracy. All discrepancies were resolved by consensus discussion. Elements abstracted included general study characteristics (e.g., country, setting, study design, number of subjects enrolled), patient characteristics (e.g., age, sex, comorbidities), details of the imaging methodology (e.g., radiotracer, timing of test), risk-of-bias items, and outcome data.

### Risk of Bias Evaluation

For systematic reviews of single-test accuracy, we used a revised AMSTAR (Assessment of Multiple Systematic Reviews) instrument. For each included review, two analysts independently answered 15 items on the AMSTAR instrument and independently assessed the systematic review as either high quality or not high quality. Discrepancies in the category assignment were resolved by consensus. For primary studies comparing two or more tests, we used a set of nine risk-of-bias items after considering the QUADAS-2,<sup>13</sup> as well as additional issues that specifically address bias in the comparison of diagnostic tests.

### Data Analysis and Synthesis

For comparing imaging tests, we synthesized the evidence using meta-analysis wherever appropriate and possible. When meta-analysis was not possible (because of clinical heterogeneity or limitations of reported data) or was judged to be inappropriate, the data were synthesized using a descriptive narrative review approach.

For each pair of imaging tests compared directly by a group of studies (e.g., MDCT and EUS-FNA) for a given clinical purpose (e.g., diagnosis), we performed bivariate meta-analysis<sup>14</sup> of each test's accuracy data using the "metandi" command in STATA,<sup>15</sup> or separate analyses of sensitivity and specificity using Meta-Disc.<sup>16</sup> Using the meta-analytic results, we used equation 39 in Trikalinos et al. (2013)<sup>17</sup> to compare the tests statistically (separately for sensitivity and specificity). For statistical tests, we set  $p=0.05$  two-tailed as the threshold for statistical significance. If a comparison was not statistically

significant, two reviewers independently judged whether the CI around the difference was sufficiently narrow to permit a conclusion of similar accuracy; disagreements were resolved by consensus.

### Strength of Evidence Grading

We used the EPC system for grading evidence on diagnostic tests as described in the EPC guidance chapter by Singh et al. (2012).<sup>18</sup> This system uses up to eight domains as inputs (study limitations, directness, consistency, precision, reporting bias, dose-response association, all plausible confounders would reduce the effect, strength of association). Reporting bias was addressed by considering unpublished trials listed in clinicaltrials.gov as well as trial funding sources. The output is a grade of the strength of evidence: high, moderate, low, or insufficient. This grade is made separately for each outcome of each comparison of each KQ. The grades are defined as follows:

- High: We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable—that is, another study would not change the conclusions
- Moderate: We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low: We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient. We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

We selected the most important outcomes to be graded. For this report, we graded evidence on comparative accuracy for diagnosis, resectability in patients with unstaged disease, staging (including its components T [tumor] staging, N [nodal] staging, metastases, vessel involvement, and precise stage), resectability in patients whose disease has been staged, and clinical outcomes (clinical management, survival, and quality of life). These were the most important outcomes, and the EPC guidance chapter by Singh et al. (2012)<sup>18</sup> can be applied. We did not

grade the strength of evidence from published systematic reviews on the accuracy of individual imaging tests, or the procedural harms of a single imaging test, or screening accuracy.

For each comparison and each outcome, we determined whether the evidence permitted an evidence-based conclusion. For comparative test accuracy, this meant determining whether the evidence was sufficient to permit one of the following three types of conclusions: (1) test A is more accurate than test B, (2) test B is more accurate than test A, or (3) tests A and B are similarly accurate. The first two types of conclusions required a statistically significant difference for either sensitivity or specificity (or both), whereas the third type of conclusion required a nonstatistically significant difference for both sensitivity and specificity, as well as independent judgments from two reviewers that the data were precise enough to indicate similar accuracy. If none of these three conclusions were appropriate, we graded the evidence insufficient. If the evidence was sufficient to permit a conclusion, then the grade was high, moderate, or low. The grade was provided by two independent raters, and discrepancies were resolved by consensus. When the evidence base consisted of a single study, the evidence was considered insufficient unless the study had all of the following characteristics: low risk of bias, the evidence was direct, there was nothing that raised concern about reporting bias, the finding was precise, and one of the three types of conclusions described above could be drawn.

### **Applicability**

The applicability of the evidence involves four key aspects: patients, tests/interventions, comparisons, and settings.<sup>19</sup> In considering the applicability of the findings to patients, we consulted large studies to ascertain the typical characteristics of patients newly given a diagnosis of pancreatic adenocarcinoma (e.g., age, sex) and then to assess whether the included studies enrolled similar patients. Some aspects of interventions may also affect applicability, for example, if a study uses an uncommonly used radiotracer. Settings of care were to be described, and if data permitted, subgroups of studies by setting were analyzed separately.

### **Peer Review and Publication**

Peer reviewers were invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. Peer review comments on the preliminary draft of the report were considered by the EPC in preparation of the final draft of the report. The

dispositions of the peer review comments are documented and will be published 3 months after the publication of the evidence report.

## **Results**

### **Evidence Base**

The literature searches identified 9,776 citations, and after duplicate review, we excluded 9,036 of them. The most common reason for exclusion was that the article did not involve diagnosis, staging, screening, or harms. We retrieved the other 740 articles, and after duplicate review, we excluded 610 of those. The most common reason was that the study reported data only on a single imaging test of interest and did not meet inclusion criteria for other KQs. See Appendix B for a list of the publications excluded at the full article level. We included the remaining 130 publications, which described 123 unique studies/reviews (seven publications reported overlapping patients). Of the 123, 15 were systematic reviews and 108 were primary studies.

### **KQ1: Comparative Effectiveness of Imaging Techniques for Diagnosis**

#### **KQ1a. What is the accuracy of each imaging technique for diagnosis and assessment of resectability?**

Thirteen systematic reviews met the inclusion criteria for this question, of which four were both recent (published 2009 or later) and of high quality (meeting all eight of the quality criteria on the revised AMSTAR instrument deemed most important). We did not grade the strength of the evidence from systematic reviews.

For EUS-FNA in diagnosing pancreatic cancer, we included eight reviews, and four were recent and high-quality. The four recent high-quality reviews reported summary sensitivity results ranging from 85 percent to 93 percent and summary specificity results ranging from 94 percent to 100 percent. CT was addressed in only one review, which was deemed not of high quality and is now outdated (2005). MRI was addressed in three reviews, none of which were high quality. The reviews agreed on MRI sensitivity, with meta-analysis results ranging from 84 percent to 86 percent, but differed on specificity, with the two reviews from one group reporting 91 percent specificity and the other review reporting 82 percent. PET/CT was addressed in three reviews, none of which were high quality.

The only review that included resectability as an outcome was outdated, of low quality, and analyzed only CT and MRI studies. For CT, the study estimated sensitivity at 81 percent and specificity at 82 percent; for MRI, the study estimated sensitivity at 82 percent and specificity at 78 percent.

The limitations of the evidence for KQ1a involve limitations of the available systematic reviews. A de novo analysis of single test diagnostic accuracy studies (which was outside the scope of this report) could may permit more estimates of single test diagnostic accuracy.

### **KQ1b. What is the comparative accuracy of the different imaging techniques for diagnosis and assessment of resectability?**

Eighteen included studies addressed this question. For diagnostic accuracy, three studies compared MDCT with EUS-FNA, seven studies compared MDCT with MRI, six studies compared MDCT with PET/CT, one study compared EUS-FNA with PET/CT, and one study compared MRI with PET/CT. For resectability in patients with unstaged disease, one study compared MDCT angiography with 3D reconstruction to MDCT angiography without 3D reconstruction, one study compared MDCT with EUS-FNA, and two studies compared MDCT with MRI. All studies had a low or moderate risk of bias.

In most cases, the combined evidence indicated neither a difference nor equivalence between two imaging technologies. The imprecision, therefore, often prevented any conclusions about comparative accuracy. For two cases, however, the evidence was sufficient to permit conclusions. One involved the comparison between MDCT and EUS-FNA with respect to the accuracy of resectability assessment in patients with unstaged disease. Based on one study, we found similar accuracy between the two modalities, with sensitivities of 64 percent to 68 percent and specificities of 88 percent to 92 percent. Another conclusion involved the comparison between MDCT and MRI with respect to diagnostic accuracy, which was performed in seven studies. These studies found consistently high sensitivity (89%) and specificity (90%) for both imaging modalities.

There were no included studies reporting pertinent data for all other subquestions for KQ1.

### **Conclusions for KQ1**

Thirteen included systematic reviews yielded the following conclusions regarding single-test accuracy for diagnosis and assessment of resectability in unstaged patients:

- Evidence was insufficient to permit accuracy estimates for MDCT angiography with or without 3D reconstruction.
- For diagnosis using MDCT, one systematic review yielded a sensitivity estimate of 91 percent (95% CI, 86% to 94%) and a specificity estimate of 85 percent (95% CI, 76% to 91%). (Strength of evidence from published systematic reviews was not graded.)
- For diagnosis using EUS-FNA, four high-quality and recent systematic reviews yielded sensitivity estimates ranging from 85 percent to 93 percent and specificity estimates ranging from 94 percent to 100 percent. (Strength of evidence from published systematic reviews was not graded.)
- For diagnosis using MRI, three systematic reviews yielded sensitivity estimates of 84 percent to 86 percent and specificity estimates of 82 percent to 91 percent.<sup>20-22</sup> (Strength of evidence from published systematic reviews was not graded.)
- For diagnosis using PET/CT, three systematic reviews yielded sensitivity estimates of 87 percent to 90 percent and specificity estimates of 80 percent to 85 percent.<sup>20,23,24</sup> (Strength of evidence from published systematic reviews was not graded.)
- For MDCT, in assessing the resectability of tumors in patients with unstaged disease, one systematic review yielded a sensitivity estimate of 81 percent (95% CI, 76% to 85 percent) and a specificity estimate of 82 percent (95 percent CI, 77 percent to 97%). (Strength of evidence from published systematic reviews was not graded.)
- For MRI, in assessing the resectability of tumors in patients with unstaged disease, one systematic review yielded a sensitivity estimate of 82 percent (95% CI, 69% to 91%) and a specificity estimate of 78 percent (95% CI, 63% to 87%).<sup>22</sup> (Strength of evidence from published systematic reviews was not graded.)

For comparative diagnostic accuracy and resectability in unstaged patients, we included 18 primary studies, and drew the following conclusions:

- Based on one study, MDCT and EUS-FNA have similar accuracy in the assessment of resectability of pancreatic adenocarcinoma in symptomatic adults with unstaged disease (Strength of evidence: low). Based on the study's prevalence of 53 percent, the results mean that those whose disease is deemed unresectable by either MDCT or EUS-FNA have about an 88 percent chance of their disease actually being unresectable, and those

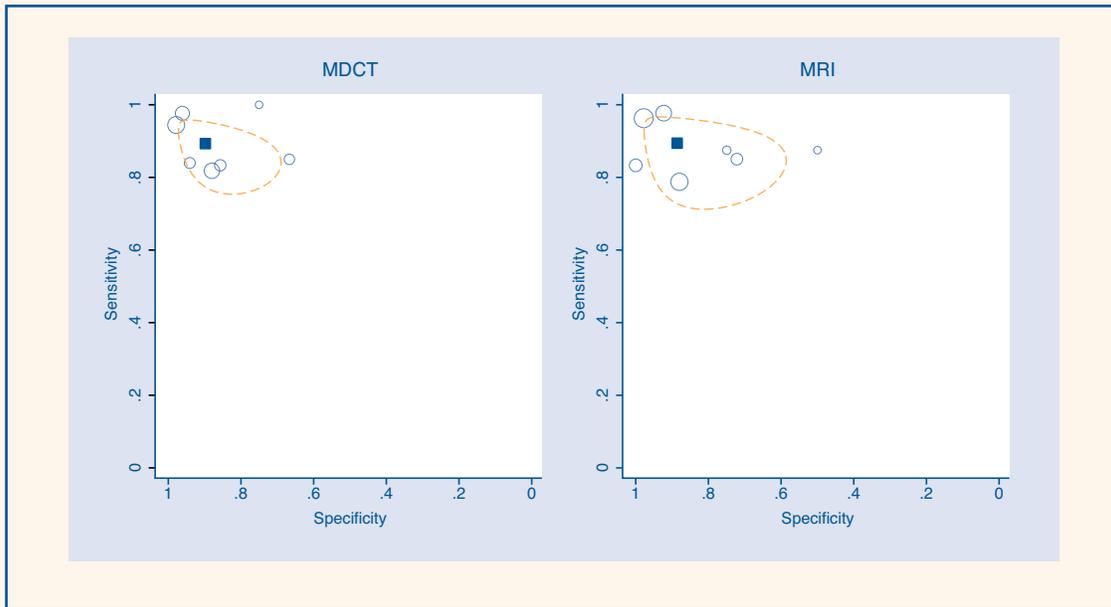
whose disease is deemed resectable by either test have about a 70 percent chance of their disease actually being resectable.

- Based on seven studies, MDCT and MRI have similar accuracy in the diagnosis of pancreatic adenocarcinoma in symptomatic adults (Strength of evidence: moderate). A figure of the meta-analysis appears in Figure A below. Based on the mean prevalence of 53 percent, the results mean that a patient with

a positive test result (on either MDCT or MRI) has approximately a 90 percent chance of having pancreatic adenocarcinoma, whereas a patient with a negative test result (on either MDCT or MRI) has only a 12 percent chance of having pancreatic adenocarcinoma.

No included studies addressed KQ1c-g, thus we drew no conclusions about those issues.

**Figure A. ROC plot of diagnostic accuracy, MDCT versus MRI**



The left side of the plot shows the multidetector computed tomography (MDCT) data in receiver operating characteristic (ROC) space; the right side shows the magnetic resonance imaging (MRI) data in ROC space. Each study contributed one circle to each side of the plot. The filled square shows the summary estimate, and the dashed region shows the 95% confidence interval range around the summary estimate.

For comparative test accuracy for KQ1, our strength of evidence assessments appear in Table A.

**Table A. Summary of evidence on comparative accuracy for KQ1**

<b>Comparison and Clinical Decisions</b>	<b># Studies, # Patients, and Overall Study Limitations</b>	<b>Directness</b>	<b>Consistency</b>	<b>Precision</b>	<b>Reporting Bias</b>	<b>Strength of Evidence</b>	<b>Conclusion</b>
MDCT angiography without 3D reconstruction vs. with 3D reconstruction; resectability in those with unstaged disease	1 Total N=57 Low	Direct	Unknown	Precise	Yes*	Insufficient	NA
MDCT vs. EUS FNA, diagnosis	3 Total N=302 Medium	Direct	Inconsistent	Imprecise	No	Insufficient	NA
MDCT vs. EUS FNA; resectability in those not staged	1 Total N=53 Low	Direct	Unknown	Precise	No	Low	Similar accuracy
MDCT vs. MRI, diagnosis	7 Total N=397 Medium	Direct	Consistent	Precise	No	Moderate	Similar accuracy
MDCT vs. MRI; resectability in those not staged	2 Total N=79 Low	Direct	Consistent	Imprecise	No	Insufficient	NA
MDCT vs. PET/CT, diagnosis	6 Total N=278 Medium	Direct	Consistent	Imprecise	No	Insufficient	NA
EUS-FNA vs. PET/CT, diagnosis	1 Total N=45 Medium	Direct	Unknown	Imprecise	No	Insufficient	NA
MRI vs. PET/CT, diagnosis	1 Total N=38 Medium	Direct	Unknown	Imprecise	No	Insufficient	NA

EUS-FNA = Endoscopic ultrasound with fine-needle aspiration; MDCT = multidetector computed tomography; MRI = magnetic

resonance imaging; NA = not applicable; PET/CT = positron emission tomography–computed tomography.

Note: In addition to considering study limitations and the domains of directness, consistency, precision and reporting bias, we considered whether a conclusion could be drawn about difference or equivalence in determining the strength of evidence.

\*Possible reporting bias in the single study involving MDCT angiography because the study was performed by the developers of the 3D reconstruction software under consideration.

## **KQ2: Comparative Effectiveness of Imaging Techniques for Staging**

### **KQ2a. What is the staging accuracy of each imaging technique (for tumor size, lymph node status, vessel involvement, metastases, stage I–IV, and resectability)?**

Three systematic reviews were included for this question. There were two reviews for the diagnosis of vascular invasion: both included CT results and one also reviewed MRI data. One review analyzed both the full set of CT studies and a subset of studies using multi-slice scanners. Sensitivity was considerably higher for the more recent studies than for the older ones (Evidence Table C-6), with no corresponding loss of specificity. A review of MRI found only four studies and thus had a large uncertainty in its results. A review of PET for pancreatic cancer staging also tabulated a subset of studies using integrated PET/CT scanners. It found only one such study, which reported on only 50 patients.

As with KQ1a, the limitations of the evidence for KQ2a involve limitations of the available systematic reviews. A de novo analysis of single test staging accuracy studies (which was outside the scope of this report) may permit more estimates of single test staging accuracy.

### **KQ2b. What is the comparative staging accuracy of the different imaging techniques?**

Twelve included studies (low or moderate risk of bias) addressed this question. For the accuracy of the assessment of metastases, five studies compared MDCT with MRI, and two compared MDCT and PET/CT. Three studies also compared MDCT and MRI with respect to the assessment of vessel involvement. All other test comparisons and aspects of staging were analyzed by no more than one study apiece.

In most cases, the combined evidence indicated neither a difference nor equivalence between two imaging technologies. The imprecision, therefore, often prevented any conclusions about comparative accuracy. For three cases, however, the evidence was sufficient to permit conclusions. One conclusion, based on one study, involved the superiority in T-stage accuracy of EUS-FNA over MDCT (~67% of cases were accurately T-staged by EUS-FNA as compared with only 41% by MDCT; this was due to a lower rate of undersizing the tumor by EUS-FNA). Another conclusion, based on three studies, was the similarity in the accuracy of the assessment of vessel involvement by MDCT and MRI (sensitivities of 62% to 68%, specificities of 96% to 97%). The third conclusion,

based on two studies, was that PET/CT is more accurate in assessing distant metastases than MDCT (67% vs. 57% for sensitivity, and 100% vs. 91% for specificity).

No included studies addressed KQ2c-g, thus we drew no conclusions about those issues.

## **Conclusions for KQ2**

Three included systematic reviews yielded the following conclusion about single-test accuracy of imaging tests for staging and tumor resectability in patients whose disease is staged:

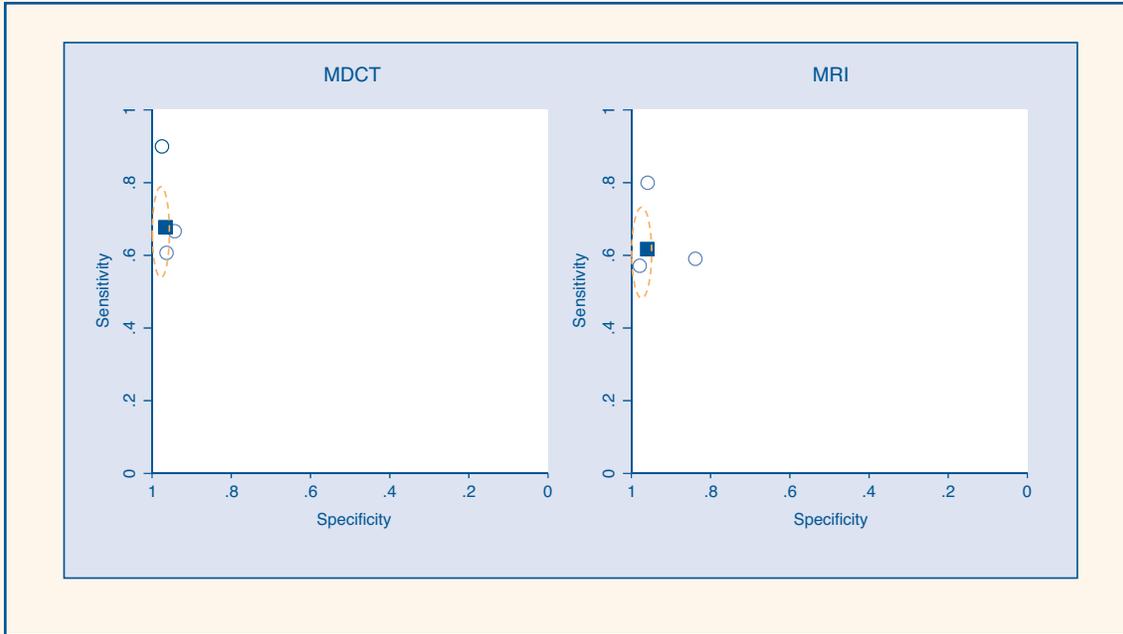
- Two systematic reviews that were not high quality reported on CT for assessing vascular invasion. Both concluded that sensitivity and specificity were worse for the subset of studies using older or single-slice CT scanners than for the studies using newer multi-slice CT. Summary sensitivity values for the newer scanners ranged from 80 percent to 85 percent while summary specificity ranged from 82 percent to 97 percent. The evidence base in both reviews was small: four or five studies each.
- One low-quality systematic review reported on MR for assessing vascular invasion, concluding it had sensitivity of 63 percent and specificity of 93 percent. The evidence base was only four studies.
- One review of PET/CT included only a single study, which had reported 82 percent sensitivity and 97 percent specificity for detecting liver metastasis.

For comparative staging accuracy, we included a total of 12 primary studies, and we drew the following conclusions:

- Based on one study, EUS-FNA is more accurate than MDCT in assessing the T stage of pancreatic adenocarcinoma in symptomatic adults (Strength of evidence: low)
- Based on three studies, MDCT and MRI have similar accuracy in assessing the vessel involvement of pancreatic adenocarcinoma in symptomatic adults (Strength of evidence: moderate). A figure of the meta-analysis appears in Figure B.
- Based on two studies, PET/CT is more accurate than MDCT in assessing distant metastases of pancreatic adenocarcinoma in symptomatic adults (Strength of evidence: low). A figure of the meta-analysis appears in Figure C.

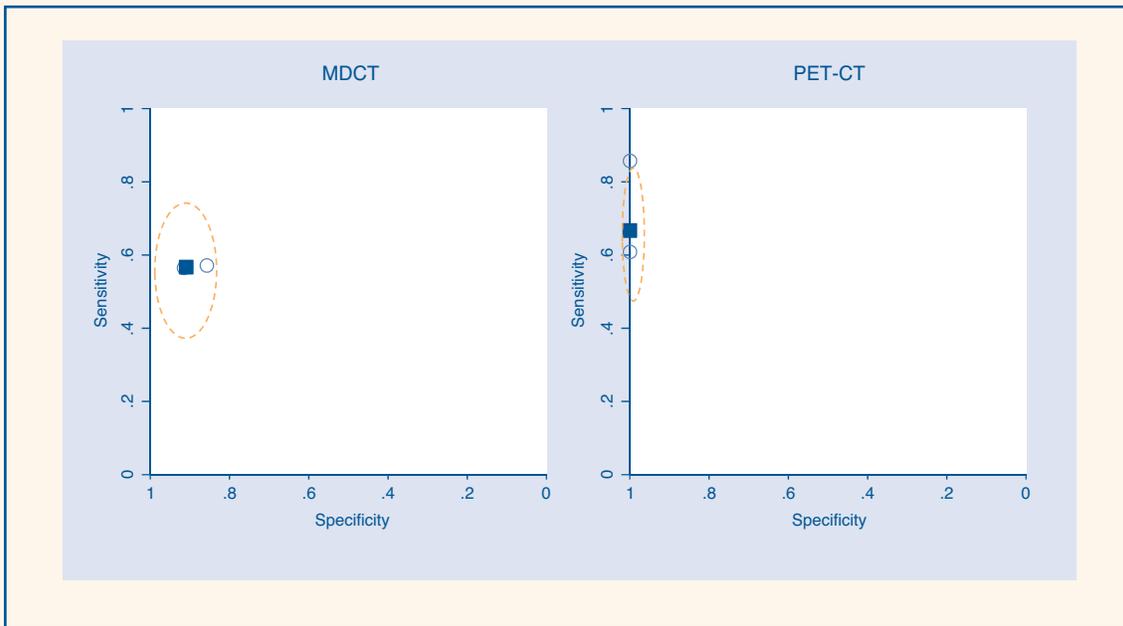
No included studies addressed KQ2c-g, thus we drew no conclusions about those issues.

**Figure B. ROC plot of vessel involvement, MDCT versus MRI**



The left side of the plot shows the multidetector computed tomography (MDCT) data in receiver operating characteristic (ROC) space; the right side shows the magnetic resonance imaging (MRI) data in ROC space. Each study contributed one circle to each side of the plot. The filled square shows the summary estimate, and the dashed region shows the 95% confidence interval range around the summary estimate.

**Figure C. ROC plot of metastases, MDCT versus PET-CT**



The left side of the plot shows the multidetector computed tomography (MDCT) data in receiver operating characteristic (ROC) space; the right side shows the positron emission tomography–computed tomography (PET-CT) data in ROC space. Each study contributed one circle to each side of the plot. The filled square shows the summary estimate, and the dashed region shows the 95% confidence interval range around the summary estimate.

For comparative test accuracy for KQ2, our strength-of-evidence assessments appear in Table B below.

**Table B. Summary of evidence on comparative accuracy for KQ2**

<b>Comparison and Staging Judgment</b>	<b># Studies, # Patients, and Overall Study Limitations</b>	<b>Directness</b>	<b>Consistency</b>	<b>Precision</b>	<b>Reporting Bias</b>	<b>Strength of Evidence</b>	<b>Conclusion</b>
MDCT vs. EUS-FNA, T staging	1 Total N=49 Low	Direct	Unknown	Precise	No	Low	Evidence favors EUS FNA
MDCT vs. EUS-FNA, vessel involvement	1 Total N=50 Medium	Direct	Unknown	Imprecise	No	Insufficient	NA
MDCT vs. MRI, T staging	1 Total N=59 Low	Direct	Unknown	Imprecise	No	Insufficient	NA
MDCT vs. MRI, N staging	1 Total N=58 Low	Direct	Unknown	Imprecise	No	Insufficient	NA
MDCT vs. MRI, Metastases	5 Total N=232 Low	Direct	Inconsistent	Imprecise	No	Insufficient	NA
MDCT vs. MRI, precise stage	1 Total N=59 Low	Direct	Unknown	Imprecise	No	Insufficient	NA
MDCT vs. MRI, vessel involvement	3 Total N=213 Low	Direct	Consistent	Imprecise	No	Moderate	Similar accuracy
MDCT vs. MRI; resectability in those staged	1 Total N=59 Low	Direct	Unknown	Imprecise	No	Insufficient	NA
MDCT vs. PET/CT, N staging	1 Total N=47 Medium	Direct	Unknown	Precise	No	Insufficient	NA
MDCT vs. PET/CT, metastases	2 Total N=96 Medium	Direct	Consistent	Precise	No	Low*	Evidence favors PET/CT

**Table B. Summary of evidence on comparative accuracy for KQ2 (continued)**

Comparison and Staging Judgment	# Studies, # Patients, and Overall Study Limitations	Directness	Consistency	Precision	Reporting Bias	Strength of Evidence	Conclusion
EUS-FNA vs. MRI, precise stage	1 Total N=48 Medium	Direct	Unknown	Precise	No	Insufficient	NA
MRI vs. PET/CT, metastases	1 Total N=14 Low	Direct	Unknown	Imprecise	No	Insufficient	NA

EUS-FNA = Endoscopic ultrasound with fine-needle aspiration; M = metastasis; MDCT = multidetector computed tomography; MRI = magnetic resonance imaging; N = nodal; NA = not applicable; PET/CT = positron emission tomography–computed tomography; T = tumor.

Note: In addition to considering study limitations and the domains of directness, consistency, precision and reporting bias, we considered whether a conclusion could be drawn about difference or equivalence in determining the strength of evidence.

\*We graded the evidence as Low due to medium risk of bias, the fact that there were only two studies, and the advantage of PET-CT over MDCT was statistically significant for specificity but not for sensitivity.

### **KQ3: Harms of Imaging Techniques for Diagnosis and/or Staging**

We included a total of 78 studies for this KQ: 50 described harms due to imaging tests for the diagnosis/staging of pancreatic cancer and were published in the year 2000 or later, and the other 28 were not specific to pancreatic cancer and were published in the year 2009 or later. The large majority of pancreas-specific studies reported the procedural harms of EUS-FNA. The most commonly reported harms in such studies were pancreatitis (occurring in 0% to 3.7% of patients), postprocedural pain (occurring in 0.1% to 2.0% of patients), and bleeding/puncture/perforation (occurring in 0% to 4.3% of patients).

#### **KQ3a. How are patient factors related to the harms of different imaging techniques?**

No included studies addressed this question.

#### **KQ3b. What are patient perspectives on the tolerance of different imaging techniques and the balance of benefits and harms of different imaging techniques?**

One included study found that about 10 percent of patients state that EUS-FNA is very uncomfortable, and 11 percent of patients state that MRI is very uncomfortable. For EUS-FNA, the stated reason for lack of comfort involved either

inadequate sedation or oversedation, whereas for MRI the stated reason involved claustrophobia.

### **Conclusions for KQ3**

In the diagnosis and staging of pancreatic adenocarcinoma, different imaging tests are associated with different types of harms. We did not grade the strength of evidence for harms.

- MDCT and PET/CT use radiation and therefore can cause cancer, but the size of the risk is not possible to estimate specifically when used for diagnosis/staging of pancreatic adenocarcinoma.
- EUS-FNA risks are due to the physical invasiveness of the procedure and primarily involve pancreatitis, postprocedural pain, and puncture, perforation, and bleeding.
- MRI risks mainly involve adverse reactions to contrast media.
- Regarding patient tolerance, one study of screening found that about 10 percent of patients stated that EUS-FNA and MRI are very uncomfortable.

## **KQ4: Imaging Techniques for Screening Asymptomatic People**

We included a total of six studies for this KQ, five of which were published in the year 2009 or later. The group of studies was heterogeneous in the populations studied, imaging tests examined, the design of study, and reporting of results, which limits generation of conclusions. Studies defined high-risk individuals (HRIs) differently, with most based on a combination of personal and family history of pancreatic cancer and/or a familial cancer syndrome (i.e., familial pancreatic cancer) and/or a hereditary predisposition to tumors (i.e., Peutz-Jeghers syndrome). One study had a control arm of non-HRIs, however, we examined only the data on HRIs. Two studies looked at one-time-only initial screening of HRIs, whereas four studies had followup screening annually or more frequently for individuals for whom it was indicated. Followup times ranged from 5 to 50.4 months.

One study examined the use of MRI only for screening HRIs, whereas the others looked at a combination of MRI/magnetic resonance cholangiopancreatography (MRCP) with EUS with or without FNA, some with the addition of MDCT and endoscopic retrograde cholangiopancreatography. Most of these studies were not designed to assess accuracy of individual imaging modalities for screening HRIs, but rather the accuracy of screening HRIs with a combination of imaging modalities as deemed clinically appropriate. Similarly, they were not designed to assess comparative accuracy of imaging modalities. Therefore, studies did not uniformly or comprehensively report results for each imaging modality performed, which prevents conclusions about accuracy of any particular imaging tests or comparative accuracy.

### **Conclusions for KQ4**

Six included studies involved screening high-risk asymptomatic adults for detecting precursor lesion(s) of pancreatic cancer or pancreatic adenocarcinoma.

- No accuracy estimates are possible for any single imaging modality, because the six included screening studies provided accuracy data only for a joint set of imaging tests.
- Two percent to 18 percent of HRIs screened received either a biopsy or surgery based on imaging findings on any imaging modality— MDCT, EUS-FNA, or MRI— amounting to a total of 43 of 665 HRIs (7%) from all six studies.

- Of 46 patients with a pathological specimen from either biopsy or surgery in the six screening studies, 17 total (1.1% to 9.0% of HRIs screened) had true-positive findings (i.e., pathology-confirmed precursor lesions or pancreatic adenocarcinoma); 19 total (0% to 9.8% of HRIs screened) had major false-positive findings (i.e., patient had surgical resection based on imaging and pathology that showed a benign lesion, e.g., branch duct intraductal papillary mucinous neoplasia [BD-IPMN] with low-grade dysplasia); seven total (0% to 9.2% of HRIs screened) had a minor false-positive finding (i.e., patient had a FNA biopsy based on imaging but pathology was normal, so no surgery was performed). An additional three patients (0% to 1.5% of HRIs screened) had false-negative findings (i.e., patient's cancer was missed on image screening but found on later screening with pathology confirmation).

## **Discussion**

### **Key Findings and Strength of Evidence**

This comparative review summarizes evidence on imaging tests (MDCT, MDCT angiography, EUS-FNA, MRI, and PET/CT) for pancreatic adenocarcinoma with respect to four areas: diagnosis, staging, harms, and screening. Diagnostic and staging accuracy are reasonable for several tests, but direct comparative evidence was generally not precise enough to demonstrate clear advantages of one test over another or to demonstrate similar accuracy among tests. We conclude that MDCT and EUS-FNA have similar accuracy in assessing resectability of tumors in patients with unstaged disease and that EUS-FNA is less likely to undersize tumors than MDCT with respect to T staging. Further, we conclude that MDCT and MRI are similarly accurate with respect to both diagnosis and assessment of vessel involvement. For PET/CT, evidence was generally inconclusive, but we found sufficient evidence to conclude that PET/CT is more accurate in assessing distant metastases than MDCT (67% vs. 57% for sensitivity, and 100% vs. 91% for specificity).

Regarding the procedural harms of imaging tests in the diagnosis and staging of pancreatic adenocarcinoma, the harms of concern are different for different tests. MDCT and PET/CT use radiation and, therefore, could theoretically increase the risk of developing cancer over time. However, the size of the risk is not possible to estimate, specifically when used for diagnosis and staging of pancreatic adenocarcinoma. Furthermore data on the importance of this risk relative to the data obtained from these tests is unknown. EUS-FNA risks are due to the physical invasiveness of this procedure and primarily

involve pancreatitis, postprocedural pain, and puncture, perforation, and bleeding. Regarding patient tolerance, one study of screening found that about 10 percent of patients state that EUS-FNA and MRI are very uncomfortable.

One of the practical challenges of this review is that while our KQs looked separately at the comparative effectiveness of imaging procedures for diagnosis, staging, and resectability, generally speaking these determinations occur simultaneously or in rapid succession. So, the question naturally arises, do our findings mean that all four imaging modalities should be used in the evaluation of patient's with suspected pancreatic adenocarcinoma? Specifically, should an individual have an MDCT or MRI for diagnosis, assessment of vessel involvement, and potential resectability determination, followed by an EUS-FNA for tumor staging, followed by a PET/CT for metastatic staging? Although our results did not permit determination of the optimal sequencing of imaging tests, they suggest that MDCT or MRI, plus EUS-FNA, plus PET/CT may all be appropriate for the diagnosing, staging, and resectability determination of suspected pancreatic adenocarcinoma. However, it should be noted that these four imaging studies are not equally available at all institutions, and each study has its associated risks of harms as well as patient preferences and tolerances.

Existing practice follows a multi-modality paradigm for diagnosis and staging of pancreatic cancer that is largely institution-specific based on technology and resource availability and institution and provider preference, an approach that allows for potential inappropriate variation and disparities in care. This report sheds additional light on which imaging modalities are more accurate or roughly equivalent for some aspects of diagnosis and staging of pancreatic adenocarcinoma, and could be incorporated into additional guidance developed for clinicians.

When screening individuals at high risk of developing pancreatic adenocarcinoma, 2 percent to 18 percent of HRIs screened, by any imaging modality (MDCT, EUS-FNA, MRI) had a result that warranted either a biopsy or surgery based on imaging findings. Available studies do not report results for individual imaging modalities, therefore, one cannot determine the screening accuracy of any given imaging test. At this time, further research is needed to elucidate the preferred imaging modalities (and other tests) for screening HRIs.

## Applicability

The applicability of the existing evidence to current practice is mixed:

- Regarding patients, the typical age of patients in the included studies was 60–65 years, which is slightly younger than the median age at diagnosis of pancreatic cancer (71 years). Resection may be more appropriate for younger patients (because of fewer comorbidities), but the comparative accuracy of different tests (e.g., MDCT vs. EUS-FNA) may not vary by age. The gender ratio in the included studies was representative (slightly fewer than half the patients were women).
- Regarding tests and comparisons, we attempted to ensure applicability by including only studies of imaging technologies that are currently in wide use for the diagnosis and staging of pancreatic adenocarcinoma. Specific test protocols, however, may differ between the studies we included and the typical test parameters used outside the context of a research study.
- Regarding settings, academic settings were overrepresented in the evidence we reviewed. The implication of this is unclear, but the test readers or practitioners in these publications may be more experienced than at nonacademic centers.

## Research Gaps

We identified four important gaps. The first important gap concerns the general lack of specific evidence on MDCT angiography, which is a newer technology that has not been sufficiently studied.

The second important gap concerns the lack of evidence on comparative longer-term outcomes such as how patients were managed differently after different tests, the length of survival after undergoing different imaging tests, and the quality of patients' lives after different tests. No studies have provided comparative management or health outcome information in the context of diagnosis and staging of pancreatic adenocarcinoma.

The third important gap concerns the lack of evidence on important factors that could influence comparative accuracy, such as the prior experience of test readers (e.g., two tests may have similar accuracy if readers are very experienced, but one may be much better if readers are less experienced), patient factors (e.g., for patients with jaundice, one test may be better, but for patients without jaundice that same test is worse), and tumor characteristics

(e.g., for staging small tumors, one test is best, but for large tumors, another test is best). Again, no studies provided pertinent data, so the reason for this gap also is insufficient or imprecise information.

The fourth important gap concerns the screening of asymptomatic high-risk people. No studies have reported test-specific screening accuracy (insufficient or imprecise information). This is an important gap in the literature because there is little evidence to justify the choice of one screening test over another.

## Conclusions

We have comprehensively reviewed the evidence on commonly used imaging tests for the diagnosis and staging of pancreatic adenocarcinoma, as well as screening for pancreatic adenocarcinoma in high risk individuals. Some conclusions are possible at this time, specifically regarding relative test accuracy for different clinical purposes, but many uncertainties remain. Chief among these are the impact of imaging tests on patient management and long-term survival, the influence of patient factors and tumor characteristics on comparative accuracy, and test accuracy when used for screening high risk individuals.

## References

1. American Cancer Society. Cancer facts & figures 2012. Atlanta, GA: American Cancer Society; 2012. 65 p. [www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-031941.pdf](http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-031941.pdf).
2. Sharma C, Eltawil KM, Renfrew PD, et al. Advances in diagnosis, treatment and palliation of pancreatic carcinoma: 1990-2010. *World J Gastroenterol*. 2011 Feb 21;17(7):867-97. PMID: 21412497.
3. Siegel R, Ma J, Zou Z, et al. Cancer Statistics, 2014. *CA Cancer J Clin*. 2014 Jan-Feb;64(1):9-29. PMID: 24399786.
4. Dabizzi E, Assef MS, Raimondo M. Diagnostic management of pancreatic cancer. *Cancers*. 2011 Mar;3(1):494-509.
5. Pancreatic cancer treatment (PDQ). Health professional version. Bethesda, MD: National Cancer Institute; 2012 Jul 17. [www.cancer.gov/cancertopics/pdq/treatment/pancreatic/HealthProfessional/page1](http://www.cancer.gov/cancertopics/pdq/treatment/pancreatic/HealthProfessional/page1). Accessed 2012 Nov 8.
6. Benson AB III, Myerson RJ, Sasson AR. Pancreatic, neuroendocrine GI, and adrenal cancers. In: *Cancer management: a multidisciplinary approach*. 14th ed. New York, NY: UBM Medica LLC; 2011 Oct 28. <http://www.cancernetwork.com/cancer-management/pancreatic/article/10165/1802606>. Accessed 2013 Feb 28.
7. Tamm EP, Balachandran A, Bhosale PR, et al. Imaging of pancreatic adenocarcinoma: update on staging/resectability. *Radiol Clin North Am*. 2012 May;50(3):407-28. PMID: 22560689.
8. Callery MP, Chang KJ, Fishman EK, et al. Pretreatment assessment of resectable and borderline resectable pancreatic cancer: expert consensus statement. *Ann Surg Oncol*. 2009 Jul;16(7):1727-33. PMID: 19396496.
9. Wong JC, Lu DS. Staging of pancreatic adenocarcinoma by imaging studies. *Clin Gastroenterol Hepatol*. 2008 Dec;6(12):1301-8. PMID: 18948228.
10. U.S. Preventive Services Task Force. Screening for pancreatic cancer: recommendation statement. Rockville, MD: Agency for Healthcare Research and Quality; 2004 Feb. 3.
11. Canto MI, Harinck F, Hruban RH, et al. International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. *Gut*. 2013 Mar;62(3):339-47. Epub 2012 Nov 7. PMID: 23135763.
12. White CM, Ip S, McPheeters M, et al. Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: *Agency for Healthcare Research and Quality. Methods guide for comparative effectiveness reviews*. Rockville, MD: Agency for Healthcare Research and Quality; September 2009. <http://effectivehealthcare.ahrq.gov/healthInfo.cfm?infotype=rr&ProcessID=60>. PMID: 21433402.
13. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003 Nov 10;3(1):25. PMID: 14606960.
14. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005 Oct;58(10):982-90. PMID: 16168343.
15. STATA statistics/data analysis. MP parallel edition. College Station, TX: StataCorp; 1984-2007. Single user Stata for Windows. <http://www.stata.com>.
16. Zamora J, Abraira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31. PMID: 16836745.
17. Trikalinos TA, Hoaglin DC, Small KM, et al. Evaluating practices and developing tools for comparative effectiveness reviews of diagnostic test accuracy: methods for the joint meta-analysis of multiple tests. Methods research report. (Prepared by the Tufts Evidence-based Practice Center, under Contract No. 290-2007-10055-I.) AHRQ publication no. 12(13)-EHC151-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013 Jan. <http://effectivehealthcare.ahrq.gov>.
18. Singh S, Chang S, Matchar DB, et al. Grading a body of evidence on diagnostic tests. AHRQ Publication No. 12-EHC079-EF. Chapter 7 of *Methods Guide for Medical Test Reviews* (AHRQ Publication No. 12-EHC017). Rockville, MD: Agency for Healthcare Research and Quality; June 2012. <http://www.effectivehealthcare.ahrq.gov/reports/final.cfm>. Also published as a special supplement to the *Journal of General Internal Medicine*, July 2012.

19. Atkins D, Chang S, Gartlehner G, et al. Assessing applicability when comparing medical interventions: Agency for Healthcare Research and Quality and the Effective Health Care Program. *J Clin Epidemiol*. 2011 Nov;64(11):1198-207. Epub 2011 Apr 3. PMID: 21463926.
20. Wu LM, Hu JN, Hua J, et al. Diagnostic value of diffusion-weighted magnetic resonance imaging compared with Fluorodeoxyglucose positron emission tomography/computed tomography for pancreatic malignancy: a meta-analysis using a hierarchical regression model. *J Gastroenterol Hepatol*. 2012 Jun;27(6):1027-35. PMID: 22414092.
21. Wu LM, Xu JR, Hua J, et al. Value of diffusion-weighted imaging for the discrimination of pancreatic lesions: a meta-analysis. *Eur J Gastroenterol Hepatol*. 2012 Feb;24(2):134-42. PMID: 22241215.
22. Bipat S, Phoa SS, van Delden OM, et al. Ultrasonography, computed tomography and magnetic resonance imaging for diagnosis and determining resectability of pancreatic adenocarcinoma: a meta-analysis. *J Comput Assist Tomogr*. 2005 Jul-Aug;29(4):438-45. PMID: 16012297.
23. Tang S, Huang G, Liu J, et al. Usefulness of 18F-FDG PET, combined FDG-PET/CT and EUS in diagnosing primary pancreatic carcinoma: a meta-analysis. *Eur J Radiol*. 2011 Apr;78(1):142-50. PMID: 19854016.
24. Wang Z, Chen JQ, Liu JL, et al. FDG-PET in diagnosis, staging and prognosis of pancreatic carcinoma: a meta-analysis. *World J Gastroenterol*. 2013 Aug 7;19(29):4808-17.

## Full Report

This executive summary is part of the following document: Treadwell JR, Mitchell MD, Eatmon K, Jue J, Zafar H, Teitelbaum U, Schoelles K. Imaging Tests for the Diagnosis and Staging of Pancreatic Adenocarcinoma. Comparative Effectiveness Review No. 141. (Prepared by the ECRI Institute-Penn Medicine Evidence-based Practice Center under Contract No. 290-2012-00011-I.) AHRQ Publication No.14-EHC045-EF. Rockville, MD: Agency for Healthcare Research and Quality. September 2014. [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

