



Systematic appraisal of guidelines for the diagnosis of pancreatic cancer

Xin-Jie Liu^{1#}, Ting Yang^{1#}, Xin Shi^{2#}, Bing-He Xiao¹, Li-Ya An¹, Su-Yun Zheng², Yu-Xing Qi¹, Da-Li Sun¹

¹Department of Gastrointestinal Surgery, Second Affiliated Hospital of Kunming Medical University/The Second Faculty of Clinical Medicine, Kunming Medical University, Kunming, China; ²Department of Digestive Endoscopy Center, Qujing First Hospital/Qujing Affiliated Hospital of Kunming Medical University, Qujing, China

Contributions: (I) Conception and design: None; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Su-Yun Zheng. Department of Digestive Endoscopy Center, Qujing First Hospital/Qujing Affiliated Hospital of Kunming Medical University, Qujing, China. Email: 1227920977@qq.com; Yu-Xing Qi; Da-Li Sun. Department of Gastrointestinal Surgery, Second Affiliated Hospital of Kunming Medical University/The Second Faculty of Clinical Medicine, Kunming Medical University, Kunming, China. Email: 1154180827@qq.com; sundali2018@126.com.

Abstract: This review aimed at assessing current guidelines' methodological quality systematically for pancreatic cancer's diagnosis and to reveal the heterogeneity of the recommendations among the evaluated guidelines. A systematic search was conducted to find the latest guidelines for pancreatic cancer's diagnosis. The Appraisal of Guidelines for Research and Evaluation (AGREE II) tool was used to assess the qualified guidelines' feature. We extracted the main recommendations for the diagnosis of pancreatic cancer from the guidelines and performed a heterogeneity evaluation. The highest-level evidence that supported these recommendations was further extracted and analysed. Nine guidelines for the diagnosis of pancreatic cancer were included in this study. Four of the guidelines had an overall score of more than 60% and thus are recommended for clinical use. Further analysis of the heterogeneity of the main recommendations for the diagnosis of pancreatic cancer in the guidelines revealed that the recommendations vary greatly among the different guidelines. The main reasons for the great differences include the neglect of symptoms and signs, great differences in the items involved in recommendations for the diagnosis of pancreatic cancer, inconsistent recommendations for some indicators (carbohydrate antigen 19-9 and ERCP), the unreasonable citation of evidence, and the failure of some recommendations to provide evidence supporting the recommendations. For most recommendations, there was a low level of evidence and a dearth of high-quality study evidence. Recommendations for pancreatic cancer diagnosis have been significantly inconsistent over the past five years. The quality of the guidelines for diagnosing pancreatic cancer also varies. The improvement by the guideline creators of the factors that contribute to the differences mentioned above will be a shortcut to update the guidelines for the diagnosis of pancreatic cancer.

Keywords: Pancreatic cancer; diagnosis; systematic review; AGREE II; guidelines

Submitted Aug 25, 2020. Accepted for publication Mar 10, 2021.

doi: 10.21037/gs-20-676

View this article at: <http://dx.doi.org/10.21037/gs-20-676>

Introduction

Pancreatic cancer is a malignant tumour of the digestive tract that is difficult to diagnose and treat and is now the third most common cause of cancer-related death (1). The incidence and mortality of pancreatic cancer have increased significantly in recent years. Pancreatic cancer lacks specific signs and symptoms in the early stage (2), has insidious clinical manifestations, develops rapidly, and has a poor prognosis. In the vast majority of patients, the cancer has metastasized when it is initially diagnosed, and only a few patients are diagnosed in the early stages (3). Because of the difficulty of early diagnosis, most patients lose the opportunity for surgical treatment, so the early diagnosis and effective treatment of pancreatic cancer are of great importance. Obviously, this situation has received much attention from researchers around the world. In recent years, a number of guidelines on how to diagnose pancreatic cancer have been developed (4-12), but users of the guidelines will be confused about how to use them to diagnose pancreatic cancer, as the guidelines vary widely in terms of quality and rigour as well as in terms of the diagnosis of pancreatic cancer. Therefore, our goal is to analyse current guidelines' methodological quality for pancreatic cancer's diagnosis, reveal the heterogeneity of the guidelines' recommendations for pancreatic cancer's diagnosis, and discuss the underlying causes of this heterogeneity to help clinicians choose the most appropriate guidelines and recommendations and to help guideline developers update and improve the guidelines.

Methods

Study design

This study used the Appraisal of Guidelines for Research and Evaluation (AGREE II) tool to comprehensively evaluate and analyse the diagnostic guidelines for pancreatic cancer. This review was conducted in the light of the Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocols (PRISMA-P) guidelines (13).

Retrieval strategy

The following databases were searched comprehensively in this study: PubMed, Web of Science, Ovid, and ScienceDirect. To take into account the impact of a large time span on the currency of evidence in analysing the heterogeneity of the evidence and recommendations

among the guidelines, we limited the database searches to a period of 5 years (January 2014 to September 2019). The search terms included "pancreatic cancer" ("pancreatic cancer", "pancreatic neoplasm", "pancreatic carcinoma", "pancreatic adenoma", "pancreatic tumour"), "guidelines" ("statement, guideline", "recommendation", "consensus"), and "diagnosis", and the search conditions were limited to titles. At the same time, guideline development organizations' official websites were searched in this study, including NCCN (https://www.nccn.org/professionals/physician_gls/f_guidelines.asp), NICE (<https://www.nice.org.uk/guidance>) and SIGN (<http://www.sign.ac.uk/>), and guideline databases such as GIN (<http://www.gin.net/>) and NGC (<https://www.guideline.gov/>). References included in the guidelines were manually retrieved to screen possible guidelines.

Guidelines' selection

The inclusion criteria were as follows: (I) recent guidelines for pancreatic cancer; (II) guidelines with recommendations for pancreatic cancer's diagnosis; (III) the full text of the article; and (IV) published in English. The exclusion criteria were as follows: (I) older guidelines issued by the same organization; (II) repeated guidelines; (III) only the abstract of the article; and (IV) a review of pancreatic cancer diagnosis. Reviewers (LXJ and YT) independently reviewed the inclusion guidelines. The guidelines were screened, and those that met the inclusion criteria were identified; then, basic information, including the guideline name, author, year of publication, and number of recommendations, was extracted.

Quality evaluation of guidelines

The quality of the guidelines was evaluated according to the latest edition of the AGREE II tool (2017 version) (14). The AGREE II tool is a tested assessment tool that is designed to measure and quantify clinical guidelines. The AGREE II tool includes 23 items in 6 fields: (I) Scope and purpose, covering the guidelines' overall goals, particular health issues and target groups (items 1 to 3); (II) Stakeholder involvement, meaning the guidelines were developed by suitable stakeholders and represent the intended users' views (items 4 to 6); (III) Rigour of development, involving the procedure of collecting and screening evidence, the method of producing recommendations and the process of updating

recommendations (items 7 to 14); (IV) Clarity and presentation concerning the structure, format and language of the guidelines (items 15 to 17); (V) Applicability, including obstacles and facilitators that may be encountered during implementation, strategies for improving understanding, and resource issues involved in application (items 18 to 21); and (VI) Editorial independence, referring to the formulation of recommendations without expressing a preference among competing interests (items 22 to 23). Each field was independently evaluated with the help of four assessors (LXJ, YT, QYX, XBH) that were coached to apply the AGREE II tool by learning the AGREE on-line tutorials and were blinded to the other assessors' ratings. Each item was rated on a 7-point scale as follows: 1 point indicated strong disagreement, and 7 points indicated strong agreement. 1 point was given when little or no relevant information was provided. While a statement did not satisfy the criteria completely or satisfied just one item of the standards, the score was 2 to 6. The nearer the standards had been or the greater they were considered, the better the score. 7 point was given while the statement met or considered all the standards completely. All items with a score difference of 3 or more points were discussed further. Eventually, a assessor aggregated all the scores for each item and worked out the score for each field by practicing the following formula: (acquired score - minimum possible score) / (maximum possible score - minimum possible score) × 100%. After reviewing 23 items and based on the assessors' comprehensive judgement, the assessed guidelines were separated into three categories according to the AGREE II score, including: recommended, revised recommended, and not recommended. Guidance is not provided by the AGREE II manual on how to interpret scores. Thus, to ensure the consistence of the AGREE II tool evaluation of the included guidelines and the recommendations for evidence's level, the following method was used in this study: guidelines with an overall score of >60% are recommended, guidelines with an overall score of 30% to 60% are recommended after revision, and guidelines with an overall score of <30% are not recommended.

Heterogeneity assessment of guideline diagnostic entries

The same diagnostic item was recommended by at least 4 guidelines, and the heterogeneity of this diagnostic item's recommendations was compared in different guidelines. The scoring applying a consistent rate measurement scale (15):

- ❖ 0–20%: Radically different
- ❖ 20–40%: Numerous major differences
- ❖ 40–60%: Some major differences
- ❖ 60–80%: Only minor differences
- ❖ 80–100%: Essentially identical

We extracted the main recommendations in pancreatic cancer diagnosis from the guidelines and evaluated the heterogeneity; then, we further extracted and analysed the highest-level evidence supporting these recommendations. We reclassified this evidence by using the Oxford Centre for Evidence-Based Medicine (OCEBM) grading system (16).

Statistical analysis

The standardized score for each field was calculated applying descriptive statistical analysis and presented as a percentage. The range and median were also listed for each field. Two-way analysis of variance was used to calculate intragroup correlation coefficients (ICCs) to test whether the scores of the four evaluators were consistent. An ICC between 0.01 and 0.20 is considered to be a minor consistency, 0.21 to 0.40 is fair, 0.41 to 0.60 is moderate, 0.61 to 0.80 is substantial, and 0.81 to 1.00 is very good (17). $P < 0.05$ is considered statistically to be significant. Statistical analysis was carried out applying SPSS version 19.0 (SPSS Inc., Chicago, IL, USA).

Results

Guideline characteristics

Through the title screening, a total of 47 articles were retrieved. After we reviewed these articles, we excluded 4 articles because of duplication. Twelve of the remaining documents were excluded because they met the exclusion criteria. We identified 35 studies for further evaluation, and 9 guidelines ultimately met the inclusion criteria for our evaluation (*Figure 1*). The prime characteristics of the 9 guidelines that are included in this research are demonstrated in *Table 1*. One was drafted by the European Society for Medical Oncology, two were drafted in Spain, two were drafted in China, and the remaining guidelines were drafted in Canada, the United Kingdom, Italy, and France. See *Table 1* for other characteristics.

The guidelines' quality evaluation

The outcomes of the evaluation of the quality of the included

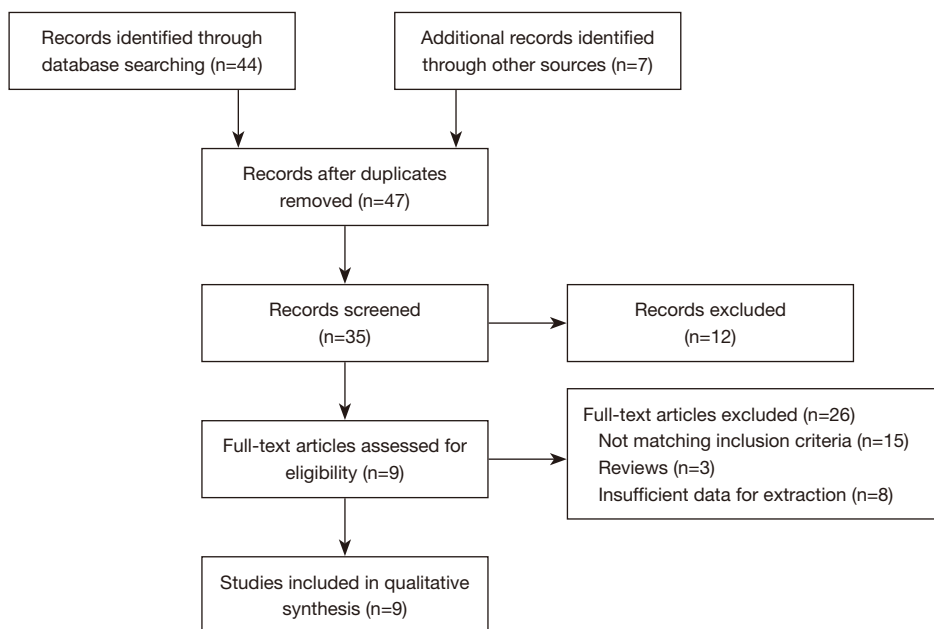


Figure 1 Flow chart of study selection.

Table 1 Characteristics of the identified guidelines for the diagnosis of pancreatic cancer

Guideline ID	Short Name	Country applied	Version	Topic	Funding	Grading system used	Development method
AIGO AISP, 2014 (4)	AI	Italian	Original version	The diagnostic work-up and follow-up of cystic pancreatic neoplasms	Wilson Cook Inc, Pentax, Olympus, and Abbott Italia SpA	Oxford criteria	EB
NHCPRC, 2019 (5)	NH	China	Original version	Diagnosis and treatment of pancreatic cancer	Not reported	None	CB
Hidalgo et al., 2017 (6)	Hi	Spain	Original version	Diagnosis, treatment and follow-up of patients with pancreatic cancer in Spain	Fundacio'n ECO	IDSAUSPHSGS	CB
Singh et al., 2015 (7)	Si	Canadian	Original version	Diagnosis and Management of Pancreatic Neuroendocrine Tumors	Novartis Pharmaceuticals Canada	Unclear	CB
O'Reilly et al., 2018 (8)	Or	UK	Original version	Diagnosis and management of pancreatic cancer in adults	Not reported	GRADE	EB
Li et al., 2019 (9)	Li	China	Original version	Intraoperative Radiotherapy for Pancreatic Cancer	CIFMS	None	CB
Neuzille et al., 2018 (10)	Ne	French	Original version	Pancreatic cancer: diagnosis, treatment and follow-up	Not reported	GRADE	CB
Benavides et al., 2014 (11)	Be	Spain	Original version	Diagnosis and management of exocrine pancreatic cancer	Not reported	None	CB
Ducreux et al., 2015 (12)	Du	Europe	Update	Pancreatic cancer: diagnosis, treatment and follow-up	Celgene	Adapted from IDSAUSPHSGS	CB

EB, evidence-based guideline; CB, consensus-based guideline.

Table 2 Assessment of the quality of the included guidelines using AGREE II instrument

Guideline	Scope and purpose	Stakeholder involvement	Rigour of development	Clarity and presentation	Applicability	Editorial independence	ICC	Overall assessment	
AIGO AISP, 2014 (4)	95.8%	73.6%	85.4%	100.0%	85.4%	58.3%	0.899	83.7%	R
NHCPRC, 2019 (5)	83.3%	16.6%	18.8%	97.2%	66.7%	0.0%	0.993	46.0%	RM
Hidalgo <i>et al.</i> , 2017 (6)	94.4%	41.6%	39.1%	97.2%	79.2%	70.8%	0.992	67.6%	R
Singh <i>et al.</i> , 2015 (7)	88.8%	44.4%	39.1%	97.2%	51.0%	62.5%	0.990	59.2%	RM
O'Reilly <i>et al.</i> , 2018 (8)	95.8%	55.5%	67.7%	97.2%	66.7%	85.4%	0.985	75.4%	R
Li <i>et al.</i> , 2019 (9)	91.6%	52.7%	21.3%	91.7%	64.6%	45.8%	0.989	56.7%	RM
Neuzille <i>et al.</i> , 2018 (10)	86.1%	38.8%	56.2%	100.0%	64.6%	45.8%	0.982	64.7%	R
Benavides <i>et al.</i> , 2014 (11)	97.2%	37.5%	22.4%	97.2%	49.0%	47.9%	0.994	52.8%	RM
Ducreux <i>et al.</i> , 2015 (12)	83.3%	33.3%	40.1%	97.2%	49.0%	50.0%	0.986	55.3%	RM
Median score (range)	90.7% (83.3–97.2%)	43.8% (16.7–73.6%)	43.3% (18.8–85.4%)	97.2% (91.7–100.0%)	64.0% (49.0–85.4%)	51.9% (0.0–85.4%)	–	–	–

R, recommended; RM, recommended with modifications; NR, not recommended.

guidelines that use the AGREE II tool are shown in *Table 2*. Scope and purpose (90.7%; range, 83.3–97.2%) and clarity and presentation (97.2%; range, 91.6–100.0%) had relatively high median values. Stakeholder involvement (43.8%; range, 16.7–73.6%) and rigour of development (43.3%; range, 18.8–85.4%) had relatively low median values. The median value of editorial independence was 51.9% (range, 0.0–85.4%). Finally, we provided a general suggestion according to the score. Each guideline's overall score is listed in *Table 2*. Four of the guidelines had an overall score greater than 60% and are recommended (4,6,8,10). Five guidelines had an overall rating between 30% and 60%, which are fall into the recommended category but still need to be improved (5,7,9,11,12). Four evaluators took part in the guidelines' evaluation for diagnosing pancreatic cancer. In this research, the ICCs of the AGREE II evaluation that was conducted by the 4 evaluators were all more outstanding than 0.8, suggesting a high level of the project scores' consistency among the evaluators.

Heterogeneity of recommendations and evidence in guidelines for pancreatic cancer's diagnosis

We referred to the key recommendation items for pancreatic cancer diagnosis in the high-quality guidelines (4) to further analyse the reasons for the heterogeneity of pancreatic cancer recommendations among the different guidelines, and extracted key recommendations from the 9 included guidelines (*Table 3*). MSRA was introduced (15) to further analyse the differences in the key recommendations among the different guidelines, and the Italian Association of Hospital Gastroenterologists and Endoscopists (AIGO) Guideline (4) was used as a reference to compare the similarity of key recommendations (*Table 4*). As long as the similarity went beyond 60% and the same recommendation item was made by more than 4 guidelines, the strongest evidence that supported this recommendation was extracted and reclassified using the OCEBM grading system (*Figure 2*)

Table 3 Recommendations for diagnosis of pancreatic cancer in the included guidelines

	Al (4)	NH (5)	Hi (6)	Si (7)	Or (8)	Li (9)	Ne (10)	Be(11)	Du (12)
History and examination									
Signs/symptoms	●	●	●	●	●	—	●	—	●
Personal or familial history	●	—	●	—	●	—	—	—	—
Imaging examination									
MRI/MRCP	●	●	●	●	●	●	●	●	●
Ultrasound	○	●	●	—	—	—	—	—	—
EUS	●	●	●	●	●	●	●	●	●
CT/ MDCT	●	●	●	●	●	●	●	●	●
PET-CT	●	●	—	—	●	—	—	—	—
S-MRCP	○	—	—	—	—	—	—	—	—
Laparoscopy	—	—	—	—	—	—	●	—	—
PET	—	—	—	●	—	●	—	○	—
SRS	—	—	—	●	—	—	—	—	—
Bone scans	—	●	—	—	—	—	—	—	—
ERCP	●	●	●	—	●	—	●	●	—
IDUS under ERCP	—	●	—	—	—	—	—	—	—
Laboratory test									
CgA	—	—	—	●	—	—	—	—	—
Pancreatic Polypeptide	—	—	—	●	—	—	—	—	—
Pancreatstatin	—	—	—	●	—	—	—	—	—
Percutaneous guided sampling	○	—	—	—	—	—	—	—	—
CA 19-9	●	●	●	—	—	—	○	—	○
CA125	—	●	—	—	—	—	—	—	—
CA242	—	●	—	—	—	—	—	—	—
Ki67	—	—	—	●	—	—	—	—	—
Immunohistochemistry	—	●	—	●	—	—	—	—	—
Glucose levels	●	—	—	—	—	—	—	—	—
Biliary brushing for cytology	—	—	—	—	●	—	—	—	—
EUS-FNA/FNA	●	●	●	●	●	●	●	●	●
Biopsy	●	●	●	●	—	—	●	—	●
Rapid on-site evaluation	○	—	—	—	—	—	—	—	—
CEA assay	●	●	—	—	●	—	—	—	—
Pancreatic enzymes	○	—	—	—	—	—	—	—	—
Amylase	●	—	—	—	—	—	—	—	—

● Indicates being recommended definitely; ● indicates being mentioned; ○ indicates being not recommended; — indicates being not mentioned. MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography; EUS, endoscopic ultrasonography; CT, computed tomography; MDCT, multi-detector computed tomography; PET-CT, positron emission tomography-computed tomography; SRS, somatostatin-receptor scintigraphy; IDUS, intraductal ultrasound; CgA, chromogranin A; CA 19-9, carbohydrate antigen 19-9; CA 125, carbohydrate antigen 125; CA 242, carbohydrate antigen 242; EUS-FNA, endoscopic ultrasound-guided fine-needle aspiration; FNA, fine-needle aspiration; CEA, carcinoembryonic antigen.

Table 4 Scientific agreement of formulated recommendations for diagnosis of pancreatic cancer in the included guidelines

	AI (4)	NH (5)	Hi (6)	Si (7)	Or (8)	Li (9)	Ne (10)	Be (11)	Du (12)
Signs/symptoms	–	80–100%	80–100%	80–100%	60–80%	–	80–100%	–	80–100%
MRI/MRCP	–	80–100%	80–100%	80–100%	80–100%	80–100%	80–100%	80–100%	80–100%
EUS	–	80–100%	80–100%	80–100%	80–100%	80–100%	80–100%	60–80%	60–80%
CT/MDCT	–	80–100%	80–100%	80–100%	80–100%	80–100%	80–100%	60–80%	60–80%
ERCP	–	80–100%	80–100%	–	80–100%	–	80–100%	20–40%	–
CA 19-9	–	80–100%	80–100%	–	–	–	0%–20%	–	0%–20%
EUS-FNA/FNA	–	80–100%	80–100%	80–100%	80–100%	80–100%	60–80%	40–60%	80–100%
Biopsy	–	60–80%	60–80%	80–100%	–	–	80–100%	–	80–100%

measurement scale of rate of agreement: 0–20%: radically different; 20–40%: numerous major scientific disagreements present; 40–60%: few major scientific disagreements present; 60–80%: only minor scientific disagreements present; 80–100%: absolute scientific agreement. In blank fields, no information is available. MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography; EUS, endoscopic ultrasonography; CT, computed tomography; MDCT, multi-detector computed tomography; ERCP, endoscopic retrograde cholangiopancreatography; EUS-FNA, endoscopic ultrasound-guided fine-needle aspiration; FNA, fine-needle aspiration.

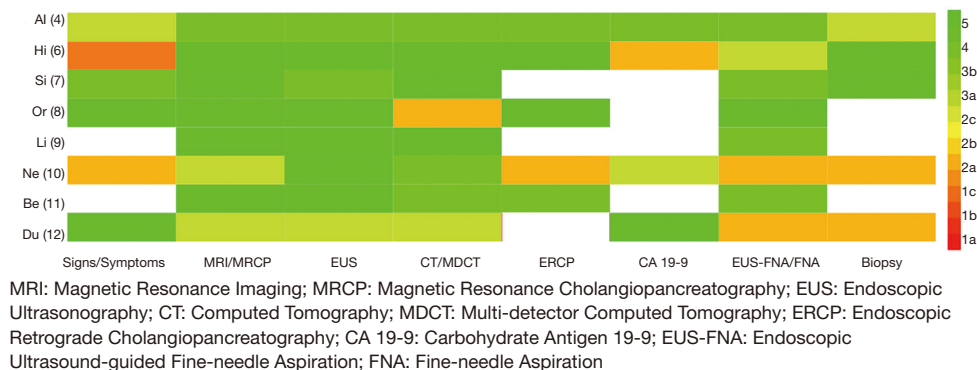


Figure 2 Distribution of the highest level of evidence to support similar recommendations for diagnosis of pancreatic cancer among included guidelines.

to determine the selected evidence’s impact on the strength of the recommendation. Guideline (5) is not included in *Figure 2* because it did not provide references.

The details of main recommendations and supporting evidence for pancreatic cancer diagnosis are shown in *Table 3* and *Table S1* (18). In terms of history and examination, 7 guidelines mentioned signs/symptoms (4-8,10,12), and 3 guidelines mentioned personal or familial history (4,6,8). The rate of agreement for signs/symptoms was 80–100% in guidelines (5-7,10,12) and 60–80% in guideline (8) (*Table 4*). The highest level of evidence supporting this recommendation was 2a (*Figure 2*). In terms of imaging examination, all the guidelines refer to EUS, CT/MDCT and

MRI/MRCP. The rate of agreement for EUS was 80–100% in 6 guidelines (5-10) and 60–80% in 2 guidelines (11,12) (*Table 4*). The highest level of evidence supporting this recommendation was 3a (*Figure 2*). The rate of agreement for CT/MDCT was 80–100% in 6 guidelines (5-10) and 60–80% in 2 guidelines (11,12). The highest level of evidence that supported this recommendation was 2b (*Figure 2*). All the guidelines had an MRI/MRCP agreement rate of 80–100% (*Table 4*), and the highest level of evidence that supported this recommendation was 3a (*Figure 2*). Six guidelines mentioned ERCP; the agreement rate was 80–100% for 4 guidelines (5-6,8,10) and 20–40% for 1 guideline (11) (*Table 4*). The highest level of evidence supporting this recommendation

was 2b (*Figure 2*). In terms of laboratory tests, 5 guidelines mentioned carbohydrate antigen (CA) 19-9 (4-6,10,12); of these guidelines, 2 had an agreement rate of 80–100% (5,6), and 2 are not recommended (10,12) (*Table 4*). The highest level of evidence supporting this recommendation was 2b (*Figure 2*). All the guidelines mentioned fine-needle aspiration (FNA), and the agreement rate was 80–100% for 6 guidelines (5-9,12), 60–80% for 1 guideline (10), and 40–60% for 1 guideline (11) (*Table 4*). The highest level of evidence supporting this recommendation was 2b (*Figure 2*). Six guidelines mentioned biopsy; the agreement rate was 80–100% for 3 guidelines (7,10,12) and 60–80% for 2 guidelines (5,6) (*Table 4*). The highest level of evidence supporting this recommendation was 2b (*Figure 2*).

Discussion

Principal findings

In this review, we evaluated the diagnostic guidelines for pancreatic cancer using the AGREE II tool and found that the diagnosis of pancreatic cancer differed greatly among different guidelines. The main reasons for the large differences include lack of attention to symptoms and signs, large differences in the items involved in the recommendations for the diagnosis of pancreatic cancer, contradictory recommendations for some indicators (CA19-9 and ERCP), the unreasonable citation of evidence, and the absence of some recommendations and evidence to support the recommendations. In addition, the quality of the pancreatic cancer diagnostic guidelines varied significantly among the guidelines and even among the same guideline's different domains.

Quality evaluation of guidelines by AGREE II

Apart from the cooperation and contributions of different professional medical teams, the intentions, opinions, and preferences of stakeholders (including patients) are an important part of developing guidelines. In this research, stakeholder participation's median score was 43.8%, largely for the reason that most of the associations developing the guidelines failed to involve patients and guideline users. Supposing that patients and users are invited to take part in the process of developing guidelines, the guidelines' applicability can be improved to some degree, and the drive of clinicians to practice them can be enhanced.

Rigour is an important component in the guidelines'

development, and it is also a significant criterion for determining whether the guidelines are believable and whether users should adopt them. Three guidelines had a rigour score of more than 50%, and only one of these guidelines scored more than 80%; most of the guidelines scored very low. The results regarding rigour were as follows: only 2 guidelines explicitly stated that a systematic approach had been used to gather evidence (4,8), and the other 7 guidelines did not specify the standards used to choose evidence. Only three guidelines had been externally reviewed by experts before publication (4,10,12). None of the guidelines provided an update process.

Regarding applicability and editorial independence, the median scores of the included guidelines were 64.0% and 51.8%, respectively. The low applicability of the guidelines is related to the dearth of consideration of obstacles and facilitators in the potential resource investment and application process (7,11,12). Only two of the guidelines clearly stated that the interests or views of funding agencies would take no impact on the development of the guidelines (4,8). Competing interests are a usual source of bias and are often overlooked. Full account of conflicts of interest in the guideline development process should be taken by the guideline development committee to increase editorial independence.

Causes of heterogeneity in recommendations and evidence for diagnostic guidelines for pancreatic cancer

- (I) Recommendations for pancreatic cancer-related symptoms and signs did not receive enough attention, although symptoms and signs are one of the important bases for the clinical diagnosis of diseases. Six guidelines mentioned symptoms and signs, and all agreed that the signs and symptoms of pancreatic cystic lesions involve abdominal pain, acute pancreatitis, nausea, vomiting, weight loss, steatorrhea, anorexia, recent onset or exacerbation of diabetes, obstructive jaundice and palpable mass (4-7,10,12). The six guidelines had relatively high levels of consensus on symptoms and signs. However, two guidelines had no recommendations for related symptoms and signs (9,11). One of the guidelines (8) addressed only the psychological impact of symptoms on pancreatic cancer patients but did not emphasize the value of signs and symptoms in the diagnosis of pancreatic cancer. Obviously, guideline developers have not paid enough attention to symptoms and signs

as a basis for diagnosing pancreatic cancer.

- (II) The items involved in the recommendations for diagnosing pancreatic cancer are significantly different: Although there is no strict requirement that every pancreatic cancer diagnostic guideline should include all known diagnostic methods, comprehensive diagnostic items will provide guideline users with a more accurate reference. Of the 9 guidelines that we evaluated, only MRI/MRCP, EUS, CT/MDCT, and EUS-FNA/FNA were recommended for pancreatic cancer. The diagnostic items recommended by 5 guidelines were symptoms and signs, ERCP, CA19-9; other diagnostic items were recommended by or a concern of fewer than 4 guidelines (*Table 3*). The analysis indicated the following main reasons: (I) Most of the included guidelines did not use a systematic search method for evidence retrieval (5-7,9,11,12); thus, they overlooked some relevant evidence for pancreatic cancer diagnosis methods, such as laparoscopy (10), somatostatin-receptor scintigraphy (SRS) (7), bone scan (5), IDUS under ERCP (5), CgA (7), pancreatic polypeptide (7), and pancreatitis (7). (II) The guidelines had different concerns: in one guideline (9), because the author focused on intraoperative radiotherapy to more accurately assess tumour staging and resectability, the preoperative diagnosis of pancreatic cancer mentioned only the use of CT and radiological methods such as MRCP, EUS and PET. Another guideline (11) focused on the imaging and marker diagnosis of exocrine pancreatic cancer and lacked recommendations of other reference standards, such as symptoms and physical examination.
- (III) The recommendations of CA19-9 and ERCP in the different guidelines were inconsistent, and the evidence cited was unreasonable: there was a great difference between the guidelines in terms of recommending the test for CA19-9 (*Figure 2*). In one guideline (4), the test for CA19-9 was definitely recommended, and the highest-level evidence cited was a case-control study (19). Two guidelines mentioned that CA19-9 can be recommended and is the most commonly used marker for the diagnosis of pancreatic cancer, but the false positive rate is relatively high (5,6). One of these guidelines (6) cited the results of a diagnostic study (20), but the other guideline (5) did not provide specific evidence, and we could not judge the source of the evidence. Two guidelines did not recommend CA19-

9 as a reference standard for the diagnosis of pancreatic cancer (10,12); the highest-level evidence used in these guidelines (10) was the results of a meta-analysis of diagnostic studies (21), which supported the usefulness of CA19-9 in the diagnosis of pancreatic cancer. The other guideline (12) did not provide specific evidence. Six guidelines referred to ERCP, and one guideline suggested that ERCP can be used to obtain tissue for pathological diagnosis with effects similar to those of EUS (11). The highest-level evidence cited was case analysis (22). Other guidelines suggested that ERCP was less sensitive than EUS-FNA for histological examinations and had lower benefits for histological diagnosis (4-6,8,10). Among them, one guideline did not provide specific evidence, and we could not judge the source of the evidence (5). The highest-level evidence cited in guideline (6) was guideline (12). The highest-level evidence cited in guideline (8) was guideline (23). The highest-level evidence cited in guideline (4) was a case analysis (24). The highest-level evidence cited in guideline (10) was a randomized controlled trial (25).

- (IV) The pancreatic cancer diagnostic guidelines did not provide evidence to support recommendations: evidence-based recommendations can be more convincing, and they can determine the basis on which guideline developers develop recommendations. One of the guidelines that we included (5), although providing recommendations, did not provide the source of evidence. In another guideline (8), the guideline authors explicitly recommended biliary brushing for cytology in the absence of histological diagnosis and ERCP examination to relieve biliary obstruction, but the author did not provide the corresponding evidence; thus, we could not evaluate the reliability and accuracy of this method. Two guidelines with CA19-9 recommendations also did not provide the source of evidence (5,12).
- (V) The target population of the guideline involves different types of pancreatic malignant tumours; although the guidelines included are all related to the diagnosis of pancreatic cancer, the focus of the target population differs among them. One of the nine guidelines concerns cystic pancreatic neoplasms (4), 4 guidelines concern pancreatic cancer (5,8,9,12), 1 guideline concerns pancreatic ductal adenocarcinoma (6), 1 guideline concerns pancreatic neuroendocrine tumours (7), 1 guideline concerns pancreatic adenocarcinoma (10), and 1 guideline concerns exocrine pancreatic cancer (11).

This difference has a certain impact on the use of specific recommendations and evidence in the guidelines.

Besides, the highest level of evidence to support the recommendations for pancreatic cancer's diagnosis in the guidelines was extracted by us. Our preliminary review of the literature showed that most of the recommendations cited had a low level of evidence and lacked high-quality research evidence. The evidence cited was mainly case-control studies and expert opinion. This lack of evidence means that the opinions of the guideline development panel would have greatly influenced the recommendations of the guideline. Therefore, the method of developing recommendations and the process of making the final decision in the guideline development process should be described in detail. Only two guidelines performed better in this area (4,8).

Our study's results provide future guidelines' development with some references for pancreatic cancer's diagnosis: (I) Pancreatic cancer-related symptoms and signs should be taken seriously as recommendations. (II) A systematic retrieval method should be used to gather evidence to show the diagnosis of pancreatic cancer as comprehensively as possible for the reference of the guideline users. (III) Before the development of new guidelines, it is recommended that previously published guidelines for the diagnosis of pancreatic cancer be reviewed and that more detailed evidence be provided for contradictory recommendations such as CA19-9 and ERCP. (IV) A detailed source of evidence for any recommendation should be provided so that users of the guidelines can trace the evidence. (V) Before implementation, the existent guidelines should be closely reviewed for compliance with quality criteria so as to make the clinical application of these guidelines more standardized. (VI) Guideline developers should be familiar with guideline development criteria, like the AGREE II instrument. (VII) Clinical guidelines' development should take into account the intentions of the target group (patients, the public, etc.). (VIII) The guidelines should be reviewed externally by experts before they are published. (IX) Wherever possible, high-level evidence should be picked, and consensus conferences should be held to develop recommendations. (X) Rigorous investigations of interest's possible conflicts should be organized for the developers, and the development procedure should be made transparent.

Strengths and limitations

This research's advantages are as follows: (I) Since

appropriate weights were used in each area evaluated in this study, the reliability of the guidance assessments and recommendations is high, and (II) our detailed analysis of recommendations and evidence for the diagnosis of pancreatic cancer can help guideline developers identify gaps in practice and help users decide which recommendations to follow. This research's limitations are as follows: (I) In this research, we assessed only guidelines that were written in English and excluded guidelines that were issued in other languages, and (II) the AGREE II tool focuses only on the methodology for developing guidelines, not on assessing the impact of recommendations on patient clinical outcomes.

Conclusion

This study finds significant inconsistencies in the recommendations for pancreatic cancer diagnosis over a 5-year period. The quality of the diagnostic guidelines for pancreatic cancer varies widely, especially in the areas of the participants in making the guidelines, rigorous development, and editorial independence, and there is still much room for improvement. The pancreatic cancer diagnosis recommendations vary widely among the guidelines. The main reasons include the lack of attention to symptoms and signs, great differences in the items included in the recommendations for diagnosing pancreatic cancer, contradictory recommendations for some indicators (CA19-9 and ERCP), the unreasonable citation of evidence, and the lack of supporting evidence for some recommendations. When the guideline developers update their diagnostic guidelines for pancreatic cancer, they would be well advised to solve the above problems, which will promote the clinical use of the guidelines.

Acknowledgments

Funding: This study was supported by Yunnan health training project of high level talents (No. H2017038) and Kunming Medical University training project of 100 talents (No. 60117190466) to DLS.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/gs-20-676>). The authors declare that they have no competing interests.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66:7-30.
2. Oldfield LE, Connor AA, Gallinger S. Molecular Events in the Natural History of Pancreatic Cancer. *Trends Cancer* 2017;3:336-46.
3. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
4. Italian Association of Hospital Gastroenterologists and Endoscopists, Italian Association for the Study of the Pancreas, Buscarini E, et al. Italian consensus guidelines for the diagnostic work-up and follow-up of cystic pancreatic neoplasms. *Dig Liver Dis* 2014;46:479-93.
5. National Health Commission of the People's Republic of China. Chinese guidelines for diagnosis and treatment of pancreatic cancer 2018. *Chin J Cancer Res* 2019;31:278-94.
6. Hidalgo M, Álvarez R, Gallego J, et al. Consensus guidelines for diagnosis, treatment and follow-up of patients with pancreatic cancer in Spain. *Clin Transl Oncol* 2017;19:667-81.
7. Singh S, Dey C, Kennecke H, et al. Consensus Recommendations for the Diagnosis and Management of Pancreatic Neuroendocrine Tumors: Guidelines from a Canadian National Expert Group. *Ann Surg Oncol* 2015;22:2685-99.
8. O'Reilly D, Fou L, Hasler E, et al. Diagnosis and management of pancreatic cancer in adults: A summary of guidelines from the UK National Institute for Health and Care Excellence. *Pancreatol* 2018;18:962-70.
9. Li Y, Feng Q, Jin J, et al. Experts' consensus on intraoperative radiotherapy for pancreatic cancer. *Cancer Lett* 2019;449:1-7.
10. Neuzillet C, Gaujoux S, Williet N, et al. Pancreatic cancer: French clinical practice guidelines for diagnosis, treatment and follow-up (SNFGE, FFCD, GERCOR, UNICANCER, SFCD, SFED, SFRO, ACHBT, AFC). *Dig Liver Dis* 2018;50:1257-71.
11. Benavides M, Abad A, Ales I, et al. TTD consensus document on the diagnosis and management of exocrine pancreatic cancer. *Clin Transl Oncol* 2014;16:865-78.
12. Ducreux M, Cuhna AS, Caramella C, et al. Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015;26:v56-v68.
13. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;350:g7647.
14. The AGREE Next Steps Consortium: Appraisal of guidelines for research & evaluation II 2017. Available online: <http://www.agreetrust.org>. Accessed 14 February 2020
15. Pentheroudakis G, Stahel R, Hansen H, et al. Heterogeneity in cancer guidelines: should we eradicate or tolerate? *Ann Oncol* 2008;19:2067-78.
16. The Centre for Evidence-Based Medicine develops: Promotes and disseminates better evidence for healthcare. Available online: <http://www.cebm.net/index.aspx?o=5653>. Accessed 14 February 2020
17. Shen WQ, Yao L, Wang XQ, et al. Quality assessment of cancer cachexia clinical practice guidelines. *Cancer Treat Rev* 2018;70:9-15.
18. Oxford Centre for Evidence-based Medicine – levels of evidence (March 2009). Centre for Evidence-based Medicine Web site. Available online: <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009> Updated March 2009.
19. Lee CJ, Scheiman J, Anderson MA, et al. Risk of malignancy in resected cystic tumors of the pancreas < or =3 cm in size: is it safe to observe asymptomatic patients? A multi-institutional report. *J Gastrointest Surg* 2008;12:234-42.
20. Cwik G, Wallner G, Skoczylas T, et al. Cancer antigens 19-9 and 125 in the differential diagnosis of pancreatic mass lesions. *Arch Surg* 2006;141:968-73.
21. Huang Z, Liu F. Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a meta-analysis. *Tumour Biol* 2014;35:7459-65.
22. Matsuyama M, Ishii H, Kuraoka K, et al. Ultrasound-

- guided vs endoscopic ultrasound-guided fine-needle aspiration for pancreatic cancer diagnosis. *World J Gastroenterol* 2013;19:2368-73.
23. National Institute for Health and Care Excellence (NICE): Pancreatic cancer: diagnosis and management in adults. NICE Clinical Guideline CG32 [Electronic version]. Available online: <https://www.nice.org.uk/guidance/cg32>, 2018. Accessed 14 February 2020
24. Hara T, Yamaguchi T, Ishihara T, et al. Diagnosis and

- patient management of intraductal papillary-mucinous tumor of the pancreas by using peroral pancreatoscopy and intraductal ultrasonography. *Gastroenterology* 2002;122:34-43.
25. Weilert F, Bhat YM, Binmoeller KF, et al. EUS-FNA is superior to ERCP-based tissue sampling in suspected malignant biliary obstruction: results of a prospective, single-blind, comparative study. *Gastrointest Endosc* 2014;80:97-104.

Cite this article as: Liu XJ, Yang T, Shi X, Xiao BH, An LY, Zheng SY, Qi YX, Sun DL. Systematic appraisal of guidelines for the diagnosis of pancreatic cancer. *Gland Surg* 2021;10(4):1487-1498. doi: 10.21037/gs-20-676

Table S1 Levels of evidence and grades of the recommendations based on the Oxford Centre for Evidence-Based Medicine (18)

Strength of recommendation	Quality of evidence	Description
A	1a	Systematic review (SR) with homogeneity of randomized controlled trials (RCTs)
	1b	Individual RCT (with narrow confidence interval)
	1c	All or none
B	2a	SR (with homogeneity) of cohort studies
	2b	Individual cohort study (including low quality RCT; for example, <80% follow-up)
	2c	“Outcomes” research; ecological studies
	3a	SR with homogeneity of case-control studies
	3b	Individual case-control study
C	4	Case series (and poor quality cohort and case-control studies)
D	5	Expert opinion without explicit critical appraisal, or based on physiology, bench research, or “first principles”