Journal of Digestive Diseases

Journal of Digestive Diseases 2016; **17;** 95–103 doi: 10.1111/1751-2980.12310

Meta analysis

How reliable is the Ki-67 cytological index in grading pancreatic neuroendocrine tumors? A meta-analysis

Jun LI,*,† Jin Ping LIN,*,† Liu Hong SHI,‡ Wei Jia WANG,§ Ai Qing LI,*,† Jian Min SI*,† & Shu Jie CHEN*,†

*Department of Gastroenterology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, †Institute of Gastroenterology, Zhejiang University, ‡Department of Ultrasound, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, China, and §Department of Internal Medicine, Tufts Medical Center, Boston, Massachusetts, USA

OBJECTIVE: To investigate the accuracy of the cytological Ki-67 index in distinguishing intermediate and high-grade (G2 + G3) from low-grade (G1) pancreatic neuroendocrine tumors (PNETs).

METHODS: Two investigators independently searched databases to identify eligible studies using the following term: ('Ki-67') AND ('pancreatic endocrine tumor' OR 'pancreatic neuroendocrine tumor' OR 'pancreatic endocrine tumour' OR 'pancreatic neuroendocrine tumour' OR 'pancreatic endocrine tumors' OR 'pancreatic neuroendocrine tumors' OR 'pancreatic endocrine tumours' OR 'pancreatic neuroendocrine tumours'), and meta-analysis was performed to calculate the pooled sensitivity, specificity, positive (PLR) and negative likelihood ratio (NLR), and diagnostic odds ratio (DOR).

RESULTS: A total of 263 lesions from 13 studies were included in the study. The pooled sensitivity and specificity of Ki-67 (cut-off value: 2%) in the differential diagnosis of G2 + G3 from G1 PNETs were 64% and 87%, respectively. The pooled PLR, NLR and DOR were 3.96, 0.42 and 11.21, respectively. The area under the summary receiver operating characteristic curve (AUROC) was 0.8397. While the cut-off value of Ki-67 index was set as 5%, the sensitivity and specificity were increased up to 69% and 93%, respectively, and the AUROC was increased to 0.955.

CONCLUSION: The cytological Ki-67 index is very useful in distinguishing intermediate and high-grade from low-grade PNETs, and a cut-off value of 5% had a better predictive value compared with that of 2%.

KEY WORDS: Endoscopic ultrasound-guided fine needle aspiration, Ki-67, meta-analysis, neoplasm grading, neuroendocrine tumor, pancreatic neoplasms.

INTRODUCTION

Pancreatic neuroendocrine tumor (PNET) is a rare malignancy, characterized by low mitotic rates based on histopathology.[1] Since Nicholls[2] first described a case of adenoma originating from the pancreatic islet in 1902, the incidence of PNETs has been reported to be steadily increasing.[3] Various nomenclature systems had been employed to classify PNETs, causing much confusion, until the currently widely used World Health Organization (WHO) 2010 classification that integrates these nomenclatures was published. Based on the Ki-67 index and mitotic count, PNETs are divided into low-grade (G1) and intermediate-grade (G2) NETs, and neuroendocrine carcinomas (G3).[4]

Correspondence to: Jian Min SI, Department of Gastroenterology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, No. 3 East Qingchun Road, Hangzhou, Zhejiang Province 310016, China. Email: jianmin_si@zju.edu.cn

Conflict of interest: None.

The abovementioned grading system (G1–3) is intimately associated with the prognosis of the patients. A 5-year survival rate of patients with G3 NET has been reported to be less than 30%, while that for patients with G1 NET was 95.7%; the prognosis of those with G2 NET has been unpredictable.[5] Therefore, surgery and other aggressive therapies are recommended for G3 PNETs, and therapeutic management of G2 tumors remains controversial, while G1 tumors are treated in a "wait-and-see" policy.[6–8] Most G1 PNETs are indolent, while quite a number of PNETs are incidentally detected by computed tomography (CT), magnetic resonance imaging (MRI) or at autopsy.[9] Therefore, obtaining specimens before operation in order to distinguish among G1, G2 and G3 tumors is one of the top priorities in managing PNETs, which will help to avoid unnecessary aggressive therapies.

Endoscopic ultrasound fine-needle aspiration (EUS-FNA) has been used to preoperatively diagnose PNETs for over 10 years in clinical setting, which allows preoperative pathological diagnosis of the tumors. Ki-67 is a nuclear antigen presenting in the S (synthesis), G2 (second gap), M (mitotic) and late G1 (first gap) phases, but is absent at the G0 and early G1 phase.[10] Ki-67 index has been regarded as a pivotal variable in grading of PNETs.[4] A recent study showed that a Ki-67 index of over 10% was more sensitive for indicating the malignant phenotype PNETs compared with mitotic count and tumor size.[11] A number of studies[12–24] have evaluated a possibility of grading PNETs using Ki-67 stain of the specimens obtained by EUS-FNA or surgery with a cut-off value of 2%, among which seven have also reported the diagnostic accuracy using a cut-off value of 5%.[12,15,19–22,24] In this meta-analysis we aimed to evaluate the efficacy of Ki-67 index in grading PNETs based on the histopathology of the specimens obtained from EUS-FNA in comparison with surgically resected specimens.

MATERIALS AND METHODS

Literature search

A comprehensive literature search was conducted on the PubMed, Web of Science, EMBASE, Ovid, Scopus, MEDLINE and the Cochrane Library databases covering all articles published from 1 January 2000 to 4 December 2015. The following terms were used for the search: ('Ki-67') AND ('pancreatic endocrine tumor' OR 'pancreatic neuroendocrine tumor' OR 'pancreatic endocrine tumour' OR 'pancreatic neuroendocrine tumour' OR 'pancreatic endocrine tumors' OR 'pancreatic neuroendocrine tumors' OR 'pancreatic endocrine tumours' OR 'pancreatic neuroendocrine tumours'). Additionally, abstracts presented at the conferences were also searched.

Inclusion and exclusion criteria

Studies provided Ki-67 index and/or histological grades of the cell blocks obtained from EUS-FNA and surgical resection were included in the meta-analysis. Studies met any of the following criteria were excluded: (i) reviews, consensus, guidelines, meta-analyses, case reports/case series and editorials; (ii) articles not in English; (iii) studies on animals or cell lines; (iv) articles did not provide the data on Ki-67 index and/or histopathological evidence; (v) those with data that was insufficient to construct a 2×2 table for the calculation of true positive (TP), false positive (FP), false negative (FN) and true negative (TN) rates (positive, G2 and G3; negative, G1; using a cut-off value of 2%). When there was any duplicate, the study containing more sufficient data was included.

Data extraction

The following information was extracted from each study: first author, year of publication, country, number of centers, publication type (full-text or abstract), study design (prospective or retrospective), total number, age and sex of patients, total number and location of the lesions (pancreatic head/body/tail), diameter of the lesions, needle size of EUS-FNA, number of the lesions with both biopsy cytopathology and surgical histopathological results and counted cell numbers. Moreover, TP, FP, FN and TN were recorded. The sensitivity and specificity of Ki-67 index were defined as its ability to differentiate G1 lesions from G2 and G3 lesions, in which histopathology (including Ki-67 index) of the surgically resected specimens was set as the gold standard. Data extraction was conducted by two authors (Jun LI and Jin Ping LIN) independently. Any disagreement was resolved by consensus.

Quality assessment

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) questionnaire was used to assess the quality of the selected studies.[25] Answer to each question was marked as yes, no or unclear.

Statistical analysis

Statistical analyses were performed using Meta-Disc 1.4 (Ramony Cajal Hospital, Madrid, Spain) and the Metabias package of STATA 12.0 (StataCorp LP, College Station, TX, USA).The Cochrane Q statistic of I^2 (inconsistency) was used to test heterogeneity among various studies. $I^2 > 50\%$ was considered significant for heterogeneity, indicating the use of a random-effect model to derive pooled results. A meta-analysis of sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR) was conducted by pooling data from all data series. A summary receiver operating characteristic (sROC) curve was drawn and the area under the sROC curve (AUROC) was calculated to estimate the pooled sensitivity and specificity.[26] Spearman's correlation coefficient between sensitivity and specificity was calculated to determine the cut-off value effect between the studies.[27] A meta-regression analysis was performed to investigate the possible sources of heterogeneity by assessing the effects of publication type (full-text *vs* abstract), number of centers (multicenter *vs* single center), study design (prospective *vs* retrospective) and sample size (>20 lesions *vs* ≤20 lesions). Publication bias was assessed by Begg's asymmetry test and a funnel plot was constructed based on lnDOR *vs* standard error (SE) of lnDOR.[28]

RESULTS

Characteristics of the studies

A total of 636 articles were identified from the primary literature search, among them 264 articles were excluded due to the publication type such as reviews, consensus, guidelines, systematic review, editorials, conference summaries, erratum, commentaries, case reports/case series (n = 248), not published in English (n = 13) or laboratory studies (n = 3). The remaining potentially relevant 372 articles were evaluated, and 359 articles were further excluded because Ki-67 index was not mentioned (n = 301), in the absence of histopathological evidence of resected specimens (n = 18), unable to construct 2 × 2 table (n = 34) or duplicates (n = 6). Finally, a total of 13 studies[12–24] with 263 lesions were included in the meta-analysis. The flowchart of study selection is shown in Fig. 1.

The main characteristics of eligible studies are outlined in Table 1. Three studies were conducted in Italy, another five in Japan, two each in the USA and Greece and one in Belgium. The full-text was available from 11 studies and the other two studies were published as meeting abstracts. Only one study was prospectively designed and the remaining 12 were retrospective.

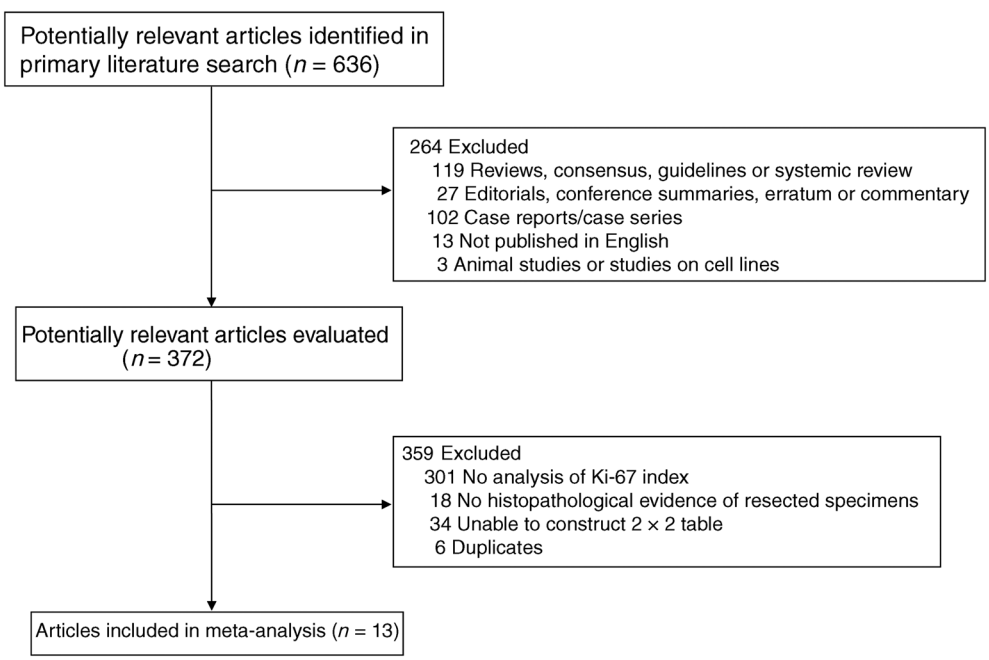

Figure 1. Flowchart demonstrating the process for selecting eligible studies.

Table 1. Characteristics of the eligible studies

Study	Country/ centers	Manuscript type	Study design	Counted cells (*n*)	Total patients /lesions (*n*)	Age, mean (Range)	Sex, *n* (male /female)	Location, *n* (head /body /tail)	Functional /non-functional tumor (*n*)	Diameter, mm (mean)	Needle size (gauge)	Lesions with both cytological and histopathological evidence	TP	FP	FN	TN
Sugimoto *et al.*[12] (2015)	Japan/ Single	Full-text	Retrospective	Unclear	10/10	61(44–79)	7/3	52/2/3	NA	28	19/22/25	8	2	0	1	5
Fujimori *et al.*[13] (2016)	Japan/ Single	Full-text	Retrospective	1000–2000	51/61	57 (23–81)	23/28	16/45 (body + tail)	8/53	18	NA	13	2	0	4	7
Sano *et al.*[14] (2014)	Japan/ Single	Abstract	Retrospective	Unclear	34/35	NA	NA	11/23 (body + tail)	NA	21.9	NA	16	2	1	3	10
Carlinfante *et al.*[15] (2014)	Italy/ Multiple	Full-text	Retrospective	800–2000	48/53	59 (32–82)	29/19	20/15/18	NA	17.0	19/22/25	53	14	2	5	32
Hasegawa *et al.*[16] (2014)	Japan/ Single	Full-text	Retrospective	400 (approximate)	58/60	54 (23–81)	28/30	15/30/15	6/52	28.1 (median)	22/25	27	6	1	5	15
Wobker *et al.*[17] (2014)	USA/ Single	Abstract	Retrospective	Up to 500	NA/10	NA	NA	NA	NA	NA	NA	10	5	4	0	1
Weynand *et al.*[18] (2014)	Belgium/ Single	Full-text	Retrospective	≥200	45/46	56 (24–85)	21/24	26/20 (body + tail)	NA	NA	22	27	6	1	10	10
Farrell *et al.*[19] (2014)	USA/ Single	Full-text	Retrospective	>100	22/22	54 (23–77)	9/13	8/7/7	5/16 (1 NA)	30.3	19/25	22	8	1	1	12
Tatsumoto *et al.*[20] (2013)	Japan/ Single	Full-text	Retrospective	Unclear	22/25	NA	12/10	NA	2/20	NA	22/25	9	3	2	1	3
Larghi *et al.*[21] (2012)	Italy/ Multiple	Full-text	Prospective	Unclear	30/30	56	13/17	12/18 (body ± tail)	0	16.9	19	12	3	1	1	7
Kaklamanos *et al.*[22] (2011)	Greece/ Single	Single	Retrospective	Unclear	26/26	NA	NA	NA	NA	NA	NA	26	7	4	8	7
Chatzipantelis *et al.*[23] (2009)	Greece/ Single	Full-text	Retrospective	Unclear	35/NA	NA	NA	NA	NA	NA	22	22	12	0	4	6
Piani *et al.*[24] (2008)	Italy/ Single	Full-text	Retrospective	Unclear	24/NA	54 (24–82)	9/9	7/7/4	7/11	30.1	22/25	18	8	2	0	8

The cut-off value of Ki-67 index is 2% for differentiating G1 from G2 and G3. FN, false negative; FP, false positive; NA, not available; TN, true negative; TP, true positive.

Quality of the studies

The quality of each selected study was evaluated with 14 items using QUADAS (Table 2). The studies conducted by Sugimoto *et al.*,[12] Carlinfante *et al.*[15] and Piani *et al.*[24] selected patients who were referred for EUS-FNA; therefore, question 1 was answered as "no". Follow-up data was not available in nine of the studies,[12–15,17,19,20,22,24] and question 3 was rated as no. None of the included studies mentioned whether the histopathological results were interpreted with the knowledge of Ki-67 index or not, so question 11 was labeled as unclear for all. Collectively, 11 to 13 of the 14 quality questions for the 13 included studies were marked yes, suggesting they were of moderate to good quality.

Diagnostic accuracy

Significant heterogeneities were found in sensitivity (I^2 = 55.6%), specificity (I^2 = 54.6%), PLR (I^2 = 63.9%). Calculated by a random-effect model, the cut-off value of Ki-67 index of 2% had a pooled sensitivity of 64% [95% confidence interval (CI) 55–73%] in the differential diagnosis of G2 + G3 and G1 PNETs (Fig. 2a). The pooled specificity and PLR were 87% (95% CI 80–92%, Fig. 2b) and 3.96 (95% CI 2.00–7.86, Fig. 2c), respectively. The pooled positive predictive value (PPV) and negative predictive value (NPV) were 80% and 74%, respectively. There was no significant heterogeneity in NLR (I^2 = 38.9%) and DOR (I^2 = 12.4%); therefore, the pooled NLR and DOR were analyzed by a fixed-effect model, with the value of 42% (95% CI 32–54%) and 11.21 (95% CI 5.98–21.05) (Fig. 2d,e), respectively.

The AUROC was 0.8397, suggesting a good overall accuracy for distinguishing intermediate and high-grade PNETs from low-grade ones using Ki-67 index (Fig. 3). The pattern of the points in this plot suggests a shoulder-arm shape, implying the possibility of a cut-off point effect. A Spearman's rank correlation test was then performed. The correlation coefficient between the log of sensitivity and log of 1-specificity was 0.397 (P = 0.179), indicating that there was no significant cut-off value effect. A meta-regression analysis was further performed to explore the possible sources of heterogeneity that were not induced by a cut-off value effect. However, none of the categories of publication type, the number of centers, study design and sample size significantly affected the DOR (Table 3). When 5% was set as the cut-off value for Ki-67 index, heterogeneities were reduced for pooled estimates of sensitivity, specificity, PLR and DOR, and they were amplified for NLR. The pooled sensitivity, specificity and DOR were increased up to 69% (95% CI 48–86%), 93% (95% CI 88–97%) and 21.18 (95% CI 7.13–62.93), respectively (Table 4). The AUROC was enlarged up to 0.955 (SE = 0.042).

Begg's funnel plot of lnDOR *vs* SE of lnDOR did not show significant asymmetry (P = 0.343 for bias), indicating that there was no significant publication bias in this meta-analysis (Fig. 4).

DISCUSSION

Owing to the development of cross-sectional imaging and endoscopic examinations, PNETs have been increasingly detected and diagnosed during the past few decades.[3] An ensuing issue is whether all these

Table 2. Quality of studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)[25]

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Sugimoto *et al.*[12] (2015)	N	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Fujimori *et al.*[13] (2016)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Sano *et al.*[14] (2014)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Carlinfante *et al.*[15] (2014)	N	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Hasegawa *et al.*[16] (2014)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Wobker *et al.*[17] (2014)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Weynand *et al.*[18] (2014)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Farrell *et al.*[19] (2014)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Tatsumoto *et al.*[20] (2013)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Larghi *et al.*[21] (2012)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Kaklamanos *et al.*[22] (2011)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Chatzipantelis *et al.*[23] (2009)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Piani *et al.*[24] (2008)	N	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y

Y, yes; N, no; U, unclear.

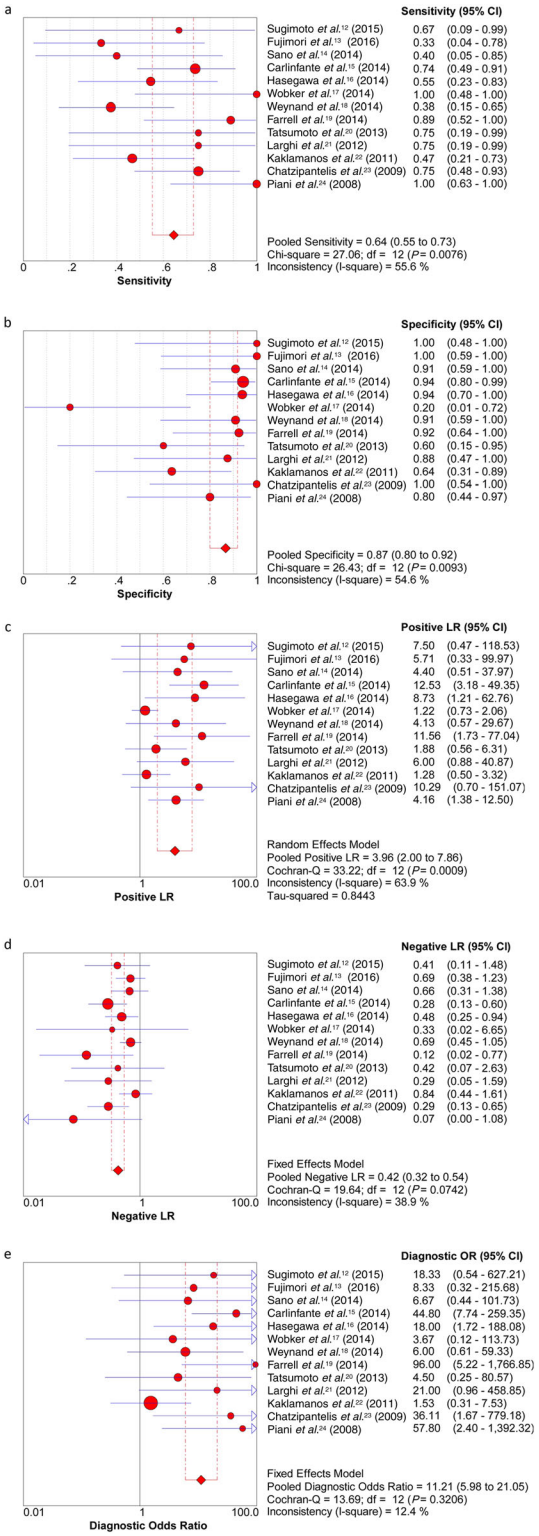

Figure 2. Forest plot showing (a) pooled sensitivity, (b) specificity, (c) positive likelihood ratio (LR), (d) negative LR and (e) diagnostic odds ratio (OR) of Ki–67 index in distinguishing intermediate and high–grade pancreatic neuroendocrine tumors from low–grade tumors. CI, confidence interval.

lesions need to be treated aggressively with surgery or be managed in a "wait-and-see" policy. Currently, the managements of the tumors are mostly based on the accurate histopathological staging, making it appealing to obtain the staging information before treatment.[21] Our results indicated that with a cut-off value of cytological Ki-67 index of 2%, the sensitivity and specificity were 64% and 87%, respectively, in distinguishing intermediate and high-grade (G2 + G3) PNETs from low-grade (G1) tumors, with the PPV and NPV being 80% and 74% (data not shown). The AUROC was 0.8397, indicating that the cytological Ki-67 index was a valuable parameter in the diagnostic work-up of PNETs.

Several studies have set different cut-off values for the Ki-67 index of resected PNET specimens. The follow-up data showed that a prognostic capability of the grading system might be improved by increasing the Ki-67 threshold value from 2% to 5% for distinguishing G1 from G2 tumors.[29] In this meta-analysis, the cytological Ki-67 cut-off value of 5% achieved a higher sensitivity and specificity and the AUROC was increased to 0.955, showing even a better diagnostic accuracy. However, data with a cut-off value of 5% were analyzed in only 148 lesions, thus a direct comparison of the value between 2% and 5% in such a small group of lesions was not applicable. Large studies on the diagnostic value of the cytological Ki-67 index in the diagnosis of PNETs regarding the cut-off values of both 2% and 5% are needed in the future.

Intratumoral heterogeneity of Ki-67 index can be found in PNET specimens; therefore, the European Neuroendocrine Tumor Society (ENETS) Consensus Guidelines have recommended the cytological Ki-67 index be evaluated by manual counting as a percentage of at least 500 cells.[30] However, inadequate tissues obtained lead to a possibility that less than 500 cells can be found within the entire field of view, and accordingly, the sensitivity of the index might be greatly impaired. In the study by Weynand *et al.*,[18] with a minimum of 200 counted cells, the sensitivity of the cytological Ki-67 index was only 37.5%, which is much lower than the pooled sensitivity in this meta-analysis. If cell blocks under 200 cells were also included, the sensitivity was even lower (35.3%). To solve this problem, several promising techniques including optical biopsies with needle-based confocal laser endomicroscopy, contrast-enhanced EUS and elastography have been developed to help endoscopists with target aspiration. In addition, an increased amount of materials sampled in fewer passes

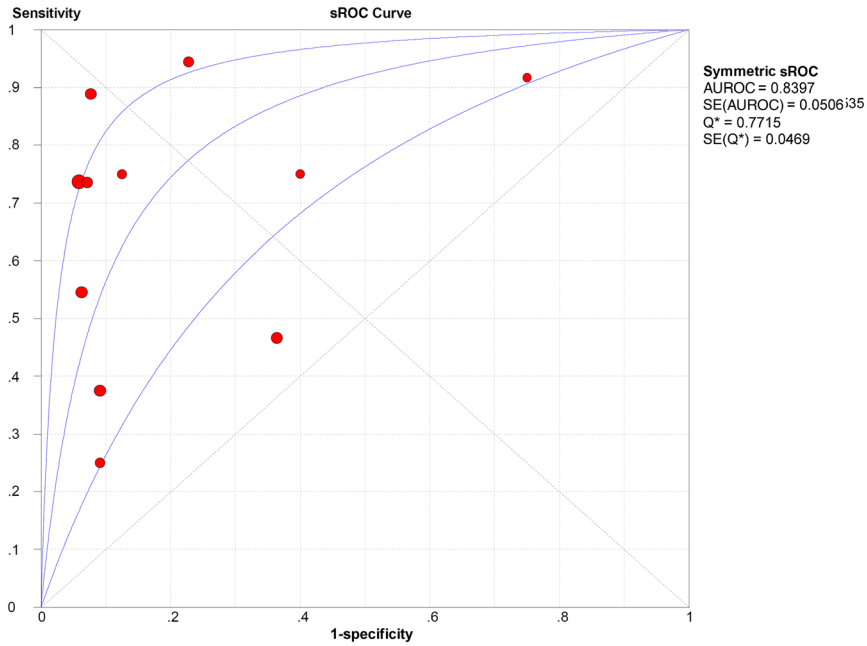

Figure 3. Summary receiver operating characteristic (sROC) curve for Ki-67 cytological index in identifying intermediate and high-grade pancreatic neuroendocrine tumors from low-grade tumors. AUROC, area under the sROC curve; SE, standard error.

Table 3. Meta-regression analysis for possible sources of heterogeneity

Variances	Coefficient standard	Standard error	*P* value	RDOR	95% CI
Inverse variance weights 1					
Cte	1.648	1.2854	0.2406	NA	NA
S	0	0.2888	0.9996	NA	NA
Sample size	−0.327	1.1986	0.7929	0.72	0.04–12.27
Design of study	−1.085	2.5352	0.6816	0.34	0.00–135.64
Number of Centers	1.531	1.4727	0.333	4.62	0.14–150.44
Publication type	0.95	1.5996	0.5713	2.59	0.06–113.57
Inverse variance weights 2					
Cte	1.653	1.2554	0.2244	NA	NA
S	0.016	0.2739	0.956	NA	NA
Design of study	−0.772	2.198	0.7346	0.46	0.00–73.47
Number of Centers	1.448	1.3623	0.3187	4.26	0.18–98.49
Publication type	0.728	1.3971	0.6164	2.07	0.08–51.93
Inverse variance weights 3					
Cte	1.653	1.223	0.2094	NA	NA
S	0.001	0.2649	0.9977	NA	NA
Number of Centers	1.232	1.1089	0.2954	3.43	0.28–42.12
Publication type	0.682	1.3571	0.6276	1.98	0.09–42.59
Inverse variance weights 4					
Cte	2.167	0.5324	0.0023	NA	NA
S	−0.030	0.2510	0.9071	NA	NA
Number of Centers	1.367	1.0363	0.2167	3.92	0.39–39.48
Inverse variance weights 5					
Cte	1.648	1.2690	0.2231	NA	NA
S	−0.007	0.2749	0.9799	NA	NA
Publication type	0.968	1.3932	0.5029	2.63	0.12–58.70

CI, confidence interval; Cte, constant term in the equation; NA, not available; RDOR, relative diagnostic odds ratio; S, indicator of cut-off point.

Table 4. Meta-analysis with 5% as the cut-off point for the Ki-67 index (cytology)

Pooled results	Value	95% CI	*P* value	I^2 (%)
Sensitivity	0.69	0.48−0.86	0.097	44.1
Specificity	0.93	0.88−0.97	0.125	40.0
PLR	8.84	4.38−17.82	0.584	0.0
NLR	0.44	0.21−0.91	0.016	61.4
DOR	21.18	7.13−62.93	0.388	5.1

CI, confidence interval; DOR, diagnostic odds ratio; I^2, inconsistency, NLR, negative likelihood ratio; PLR, positive likelihood ratio.

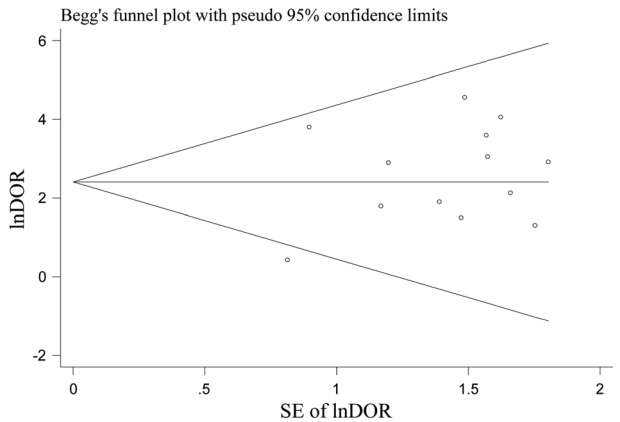

Figure 4. Begg's funnel plot comparing ln diagnostic odds ratio (lnDOR) with standard error (SE) of lnDOR.

with optimal safety for patients are available with newly designed needles.[31] Another issue for the evaluation of the cytological Ki-67 index is that counting hundreds of thousands of cells manually is labor-intensive and time-consuming. An "eyeballed" estimate of labeling percentage by pathologists is therefore a widely used technique in daily clinical practice, which adds subjective factors to the results.[32] Another noteworthy fact is that the difference in the labeling rate between G1 and G2 PNETs is subtle. As a result, it is sometimes hard to guarantee consistency in quantifying Ki-67. A recent study compared the diagnostic accuracy between digital Ki-67 calculation, "eyeballed" estimate of labeling percentage and manual counting, which achieved a perfect agreement between the digital and manual Ki-67 calculation,[32] suggesting that the relatively objective method using digital pathology and software analysis might lead to reliable and reproducible results conveniently.

Mitotic count is another evaluating index in the WHO grading system for PNETs.[4] Phosphohistone-H3 (PHH-3) is a core histone protein, the level of which increases sharply during mitosis. A retrospective study assessed the potential value of PHH-3 mitotic index of the cytological specimens obtained by EUS-FNA in grading PNETs, where a reliable correlation between the Ki-67 proliferation index and the PHH-3 mitotic index was observed.[33] According to the WHO grading system, if the grade differs for Ki-67 and mitotic parameters, the higher grade should be used.[4] Although a single PHH-3 cytological index is not as valuable as Ki-67, a combination of both might be helpful in improving the sensitivity for distinguishing between G2 + G3 and G1 PNETs.

This meta-analysis had some limitations. First, there were significant inter-study heterogeneities in pooled sensitivity, specificity and PLR. However, the Spearman's rank correlation test showed no significant cut-off value effect and the meta-regression analysis failed to identify the possible sources of heterogeneity from the categories of publication type, number of centers, study design and sample size. Second, an adequate number of counted cells was a prerequisite for an accurate Ki-67 index, but only six of the 13 included studies recorded the counted cell numbers and their lower limits were not the same among those six studies. This might be a source of heterogeneity, but subgroup analysis was not applicable due to the limited sample size. Therefore, it is necessary to establish a standard method to aspirate enough PNET cell blocks by EUS-FNA. Third, most of the included studies were retrospectively designed, and there might be errors associated with the retrospective retrieval of information. In view of these limitations, prospective and multicenter studies are still needed in order to set an optimal cut-off value for the cytological Ki-67 index in grading PNETs.

In conclusion, the cytological Ki-67 index is a valuable parameter for grading PNETs, and a cut-off value of 5% has an incremental benefit over the use of 2% in distinguishing intermediate and high-grade (G2 + G3) from low-grade (G1) PNETs.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (nos. 81472214, 81302070, 81101836), the Key Innovation Team of Zhejiang Province (no. 2013TD13) and the National Basic Research Program of China (973 Program) (no. 2012CB945004).

REFERENCES

1 Klöppel G. Tumour biology and histopathology of neuroendocrine tumours. *Best Pract Res Clin Endocrinol Metab* 2007; **21**: 15–31.
2 Nicholls AG. Simple adenoma of the pancreas arising from an island of langerhans. *J Med Res* 1902; **8**: 385–95.
3 Yao JC, Hassan M, Phan A *et al.* One hundred years after "carcinoid": epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. *J Clin Oncol* 2008; **26**: 3063–72.
4 Klimstra DS. Tumours of the pancreas. In: Bosman FT, Carneiro F, Hruban RH, eds. *WHO Classification of Tumours of the Digestive System, Fourth Edition*. Lyons: International Agency for Research on Cancer, 2010: 279–337.
5 Pape UF, Jann H, Müller-Nordhorn J *et al.* Prognostic relevance of a novel TNM classification system for upper gastroenteropancreatic neuroendocrine tumors. *Cancer* 2008; **113**: 256–65.
6 Partelli S, Maurizi A, Tamburrino D, Crippa S, Pandolfi S, Falconi M. Surgical management of pancreatic neuroendocrine neoplasms. *Ann Saudi Med* 2014; **34**: 1–5.
7 Artale S, Giannetta L, Cerea G *et al*. Treatment of metastatic neuroendocrine carcinomas based on WHO classification. *Anticancer Res* 2005; **25**: 4463–9.
8 Gaujoux S, Partelli S, Maire F *et al.* Observational study of natural history of small sporadic nonfunctioning pancreatic neuroendocrine tumors. *J Clin Endocrinol Metab* 2013; **98**: 4784–9.
9 Oberstein PE, Saif MW. Safety and efficacy of everolimus in adult patients with neuroendocrine tumors. *Clin Med Insights Oncol* 2012; **6**: 41–51.
10 Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, Stein H. Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J Immunol* 1984; **133**: 1710–5.
11 Lowe K, Khithani A, Liu E *et al.* Ki-67 labeling: a more sensitive indicator of malignant phenotype than mitotic count or tumor size? *J Surg Oncol* 2012; **106**: 724–7.
12 Sugimoto M, Takagi T, Hikichi T *et al.* Efficacy of endoscopic ultrasonography-guided fine needle aspiration for pancreatic neuroendocrine tumor grading. *World J Gastroenterol* 2015; **21**: 8118–24.
13 Fujimori N, Osoegawa T, Lee L *et al.* Efficacy of endoscopic ultrasonography and endoscopic ultrasonography-guided fine-needle aspiration for the diagnosis and grading of pancreatic neuroendocrine tumors. *Scand J Gastroenterol* 2016; **51**: 245–52.
14 Sano I, Katanuma A, Maguchi H *et al.* Diagnostic accuracy of EUS-FNA for WHO-grading of pancreatic neuroendocrine tumors. *J Gastroenterol Hepatol* 2014; **29 Suppl 3**: 38.
15 Carlinfante G, Baccarini P, Berretti D *et al.* Ki-67 cytological index can distinguish well-differentiated from poorly differentiated pancreatic neuroendocrine tumors: a comparative cytohistological study of 53 cases. *Virchows Arch* 2014; **465**: 49–55.
16 Hasegawa T, Yamao K, Hijioka S *et al.* Evaluation of Ki-67 index in EUS-FNA specimens for the assessment of malignancy risk in pancreatic neuroendocrine tumors. *Endoscopy* 2014; **46**: 32–8.
17 Wobker SE, Greene KG, Dodd LG. Correlation between the Ki-67 labeling index of paired cytology and resection specimens from pancreatic neuroendocrine tumors. *Lab Invest* 2014; **94 Suppl**: 457A.
18 Weynand B, Borbath I, Bernard V *et al.* Pancreatic neuroendocrine tumour grading on endoscopic ultrasound-guided fine needle aspiration: high reproducibility and inter-observer agreement of the Ki-67 labelling index. *Cytopathology* 2014; **25**: 389–95.
19 Farrell JM, Pang JC, Kim GE, Tabatabai ZL. Pancreatic neuroendocrine tumors: accurate grading with Ki-67 index on fine-needle aspiration specimens using the WHO 2010/ENETS criteria. *Cancer Cytopathol* 2014; **122**: 770–8.
20 Tatsumoto S, Kodama Y, Sakurai Y, Shinohara T, Katanuma A, Maguchi H. Pancreatic neuroendocrine neoplasm: correlation between computed tomography enhancement patterns and prognostic factors of surgical and endoscopic ultrasound-guided fine-needle aspiration biopsy specimens. *Abdom Imaging* 2013; **38**: 358–66.
21 Larghi A, Capurso G, Carnuccio A *et al.* Ki-67 grading of nonfunctioning pancreatic neuroendocrine tumors on histologic samples obtained by EUS-guided fine-needle tissue acquisition: a prospective study. *Gastrointest Endosc* 2012; **76**: 570–7.
22 Kaklamanos M, Karoumpalis I, Salla C *et al.* Diagnostic accuracy and clinical significance of the fine needle aspiration Ki-67 labelling index in pancreatic endocrine tumours. *Endocr Relat Cancer* 2011; **18**: L1–3.
23 Chatzipantelis P, Konstantinou P, Kaklamanos M, Apostolou G, Salla C. The role of cytomorphology and proliferative activity in predicting biologic behavior of pancreatic neuroendocrine tumors: a study by endoscopic ultrasound-guided fine-needle aspiration cytology. *Cancer* 2009; **117**: 211–6.
24 Piani C, Franchi GM, Cappelletti C *et al.* Cytological Ki-67 in pancreatic endocrine tumours: an opportunity for pre-operative grading. *Endocr Relat Cancer* 2008; **15**: 175–81.
25 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25.
26 Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993; **12**: 1293–316.
27 Devillé WL, Buntinx F, Bouter LM *et al.* Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002; **2**: 9.
28 Hewitt MJ, McPhail MJ, Possamai L, Dhar A, Vlavianos P, Monahan KJ. EUS-guided FNA for diagnosis of solid pancreatic neoplasms: a meta-analysis. *Gastrointest Endosc* 2012; **75**: 319–31.
29 Boninsegna L, Panzuto F, Partelli S *et al.* Malignant pancreatic neuroendocrine tumour: lymph node ratio and Ki67 are predictors of recurrence after curative resections. *Eur J Cancer* 2012; **48**: 1608–15.
30 Klöppel G, Couvelard A, Perren A *et al.*; Mallorca Consensus Conference participants; European Neuroendocrine Tumor Society. ENETS Consensus Guidelines for the Standards of Care in Neuroendocrine Tumors: towards a standardized approach to the diagnosis of gastroenteropancreatic neuroendocrine tumors and their prognostic stratification. *Neuroendocrinology* 2009; **90**: 162–6.
31 Deprez PH. Future directions in EUS-guided tissue acquisition. *Gastrointest Endosc Clin N Am* 2014; **24**: 143–9.
32 Tang LH, Gonen M, Hedvat C, Modlin IM, Klimstra DS. Objective quantification of the Ki67 proliferative index in neuroendocrine tumors of the gastroenteropancreatic system: a comparison of digital image analysis with manual methods. *Am J Surg Pathol* 2012; **36**: 1761–70.
33 Draganova-Tacheva R, Bibbo M, Birbe R, Daskalakis C, Solomides C. The potential value of phosphohistone-h3 mitotic index determined by digital image analysis in the assessment of pancreatic endocrine tumors in fine-needle aspiration cytology specimens. *Acta Cytol* 2013; **57**: 291–5.