Ther Adv Gastroenterol
2019, Vol. 12: 1–17
DOI: 10.1177/1756284819890537

Assessing different diagnostic tests for gastroesophageal reflux disease: a systematic review and network meta-analysis

Mengyu Zhang, John E. Pandolfino, Xuyu Zhou, Niandi Tan, Yuwen Li, Minhu Chen and Yinglian Xiao

Abstract

Background: The aim of the current systematic review and network meta-analysis (NMA) was to assess the diagnostic characteristics of the gastroesophageal reflux disease questionnaire (GERDQ), proton-pump inhibitor (PPI) test, baseline impedance, mucosal impedance, dilated intercellular spaces (DIS), salivary pepsin, esophageal pH/pH impedance monitoring and endoscopy for gastroesophageal reflux disease (GERD).

Methods: We searched PubMed and the Cochrane Controlled Trial Register database (from inception to 10 April 2018) for studies assessing the diagnostic characteristics of the GERDQ, PPI test, baseline impedance, mucosal impedance, DIS, or salivary pepsin and esophageal pH/pH impedance monitoring/endoscopy in patients with GERD. Direct pairwise comparison and a NMA using Bayesian methods under random effects were performed. We also assessed the ranking probability.

Results: A total of 40 studies were identified. The NMA found no significant difference among the baseline impedance, mucosal impedance, and esophageal pH/pH impedance monitoring and endoscopy in terms of both sensitivity and specificity. It was also demonstrated that the salivary pepsin detected by the Peptest device had comparable specificity to esophageal pH/pH impedance monitoring and endoscopy. Results of ranking probability indicated that esophageal pH/pH impedance monitoring and endoscopy had highest sensitivity and specificity, followed by mucosal impedance and baseline impedance, whereas GERDQ had the lowest sensitivity and PPI test had the lowest specificity.

Conclusions: In a systematic review and NMA of studies of patients with GERD, we found that baseline impedance and mucosal impedance have relatively high diagnostic performance, similar to esophageal pH/pH impedance monitoring and endoscopy.

Keywords: baseline impedance, dilated intercellular space, GERDQ, mucosal impedance, network meta-analysis, proton-pump inhibitor test, salivary pepsin

Received: 14 June 2019; revised manuscript accepted: 15 October 2019.

Correspondence to:
Yinglian Xiao
Department of Gastroenterology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China
xyingl@mail.sysu.edu.cn

Mengyu Zhang
Niandi Tan
Yuwen Li
Minhu Chen
Department of Gastroenterology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

John E. Pandolfino
Department of Medicine, Northwestern University, Chicago, IL, USA

Xuyu Zhou
Medical Information Research Institute, Sun Yat-sen University, Guangzhou, China

Introduction

Gastroesophageal reflux disease (GERD) is one of the most common healthcare issues, with an estimated worldwide prevalence of up to 33%[1,2] resulting in a heavy economic burden of approximately $13.0 billion/year to the healthcare system in the USA alone, due to the different diagnostic testing and overuse of proton-pump inhibitors (PPIs) to a great extent.[3] Albeit with the advances of diagnostic tests for GERD, the lack of a 'gold standard' has made identifying patients with GERD one of the biggest dilemmas in clinical practice.

Although GERD is generally empirically diagnosed based on typical reflux symptoms (heartburn and regurgitation),[4,5] the sensitivity and specificity of the symptom-based diagnosis of GERD is limited due to the complex symptom spectrum for GERD.[6] Patients with suspected GERD symptoms are often first tested for a response to PPI therapy, which definitely results in unnecessary overuse of PPIs because of its high placebo effect and low specificity.[7] So far, upper endoscopy and esophageal pH/pH impedance testing are usually performed to detect GERD complications, as well as documentation of the presence of reflux for an objective GERD diagnosis.[8] In order to develop a better understanding of the pathophysiology and improve appropriate GERD diagnosis, several new diagnostic tests, such as baseline impedance, esophageal mucosal impedance, salivary pepsin, and histopathology have been developed in recent years.[9]

To the best of our knowledge, there has been little published information regarding the comparison of diagnostic performance among individual tests. Therefore, we performed a systematic review and network meta-analysis (NMA) to assess the diagnostic characteristics of the GERDQ questionnaire, PPI test, baseline impedance, mucosal impedance, dilated intercellular spaces (DIS), salivary pepsin, esophageal pH/pH impedance monitoring and endoscopy for GERD.

Methods

Search strategy

An electronic and manual search of PubMed and the Cochrane Controlled Trial Register database for relevant articles from inception to April 2018 was performed by two authors independently. The combination of keywords and free text including: GERD Questionnaires; omeprazole, lansoprazole, pantoprazole, rabeprazole, esomeprazole, and PPI test; baseline impedance; mucosal impedance; dilated intercellular spaces; pepsin; esophageal pH/pH impedance monitoring; and endoscopy were used as search terms of different diagnostic tests for GERD. Additional search terms were: GERD or GORD, expanded to diagnosis, screening, reproducibility of results, sensitivity and specificity, false-negative reactions, false-positive reactions, predictive value, accuracy, and likelihood ratio.[10]

The searches were limited to English- or Chinese-language studies performed in adults (age $>$ 18 years). Only data accessible in peer-reviewed journals were included to minimize potential sources of bias and inaccuracy.[11]

Inclusion criteria and exclusion criteria

Studies were screened for inclusion and final decisions on exclusion were made by two authors independently. Studies assessing the diagnostic characteristics of the GERDQ, PPI test, baseline impedance, mucosal impedance, DIS, salivary pepsin, esophageal pH/pH impedance monitoring, or endoscopy in adults with presumptive GERD were screened to be included. Studies were excluded if they focused only on children (age $<$ 18 years) or patients who have specific diseases (such as cardiovascular disease) or who have had operations, if they focused exclusively on patients with extraesophageal GERD symptoms (such as asthma or laryngitis).

Data extraction

Data from the included studies were extracted by two authors independently. Information extracted included patient characteristics, study design, setting, gold standard and diagnostic modalities, and definitions of outcomes. Numbers were extracted directly from the tables or derived from percentages if only the total number of patients was available. Discrepancies were resolved by discussion until consensus was achieved for all data.

Quality assessment of studies

The quality of all included studies was assessed by researchers, according to the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool.[12] The QUADAS-2 tool included the following four key domains: patient selection, index test, reference standard, and flow of patients through the study, and timing of the index tests and reference standard (flow and timing). Review Manager 5 (RevMan 5.2.3, Cochrane Collaboration, Oxford, UK) statistical computing software was used to carry out quality assessment and investigation of publication bias.

The positive standard of diagnostic tests for GERD

The following standards for GERD were used in the current study, all of which were based on commonly accepted measures (Table 1).

Table 1. The information of included studies.

Comparison (author)	Evaluable patients (*n*)	Men (%)	Average age	Prevalence (%) of esophagitis	Reference tests					Study test	Outcome measure extracted
					Upper endoscopy	pH/pH impedance monitoring					
						AET	Number of reflux events	DeMeester score	Symptom reflux association		
A *versus* B											
Frazzoni *et al.*[13]	121	51.89	NA	42.33	Erosion	Total AET > 3.2%	>48	—	SAP ⩾ 95% or SI ⩾ 50%	MNBI	A cut-off value of 2292 Ω
Ravi *et al.*[14]	55	41.38	60	22.73	—	Total AET ⩾ 5.0%	—	—	—	Baseline impedance	A cut-off value of 2268 Ω
Frazzoni *et al.*[15]	289	45.67	50	23.53	Erosion	Total AET > 3.2%	—	—	SAP ⩾ 95% or SI ⩾ 50%	MNBI	A cut-off value of 2292 Ω
Kandulski *et al.*[16]	36	30.77	55	30.77	Erosion	Total AET > 4.2%	—	—	SAP ⩾ 95% or SI ⩾ 50%	Baseline impedance	A cut-off value of 2100 Ω
A *versus* C											
Hayat *et al.*[17]	198	44.14	49.7	NA	—	Total AET > 4.2%	—	—	SAP ⩾ 95%	Salivary pepsin	At least one sample +ve > 16 ng/ml
Saritas Yuksel *et al.*[18]	47	65.52	50	23.4	Erosion	Total AET > 4.2%	—	—	—	Salivary pepsin	A cut-off value of +1.0 or greater
A *versus* D											
Cui *et al.*[19]	565	40.7	49.4	NA	Erosion	—	—	⩾14.7	SAP ⩾ 95%	DIS	A cut-off value of 0.9 μm
Zhou *et al.*[20]	576	41.67	49.3	27.04	Erosion	—	—	⩾14.7	SAP ⩾ 95%	DIS	A cut-off value of 0.9 μm
Cui *et al.*[21]	161	59.66	NA	48.74	Erosion	—	—	⩾14.7	—	DIS	A cut-off value of 0.85 μm
Mastracci *et al.*[22]	139	57.14	52	40.34	Erosion	Total AET > 5.5%	—	—	—	DIS	Mild and marked DIS
Vela *et al.*[23]	26	NA	NA	11.54	Erosion	Total AET > 5.3%	—	—	SAP ⩾ 95%	DIS	A cut-off value of 0.68 μm
A *versus* E											
Zhou *et al.*[20]	636	41.67	49.3	27.04	Erosion	—	—	⩾14.7	SAP ⩾ 95%	GERDQ	A cut-off score of 8

(Continued)

Table 1. (Continued)

Comparison (author)	Evaluable patients (*n*)	Men (%)	Average age	Prevalence (%) of esophagitis	Reference tests					Study test	Outcome measure extracted
					Upper endoscopy	pH/pH impedance monitoring					
						AET	Number of reflux events	DeMeester score	Symptom reflux association		
Zavala-Gonzales *et al.*[24]	252	37	49.49	43.65	Erosion	Total AET > 4.2%	—	—	SI ⩾ 50%	GERDQ	A cut-off score of 8
Jonasson *et al.*[25]	169	53	47.4	81.07	Erosion	Total AET > 5.5%, or upright AET > 6.7%, or supine AET > 6.9%	—	—	SAP ⩾ 95%	GERDQ	A cut-off score of 8
Lacy *et al.*[26]	177	28.2	51.00	16.87	—	Total AET > 5.3%	—	—	—	GERDQ	A cut-off score of 8
Jones *et al.*[27]	308	46.43	47	37.66	Erosion	Total AET > 5.5%	—	—	SAP ⩾ 95%	GERDQ	A cut-off score of 8
Bai *et al.*[28]	8065	50.1	46.3	17.79	Erosion	—	—	—	—	GERDQ	A cut-off score of 8
A *versus* F											
Zhou *et al.*[20]	636	41.67	49.3	27.04	Erosion	—	—	⩾14.7	SAP ⩾ 95%	Esomeprazole, 20 mg BID for 2 weeks	Complete symptom relief
Jonasson *et al.*[25]	124	53	47.4	81.07	Erosion	Total AET > 5.5%, or upright AET > 6.7%, or supine AET > 6.9%	—	—	SAP ⩾95%	PPI for 4 weeks	Symptoms decrease to 1 day/week
Lee *et al.*[29]	188	41.49	49.9	37.23	Erosion	—	—	—	—	Lansoprazole, 15 mg/30 mg/60 mg once daily for 2 weeks	Symptom score improved by more than 50%
Dent *et al.*[30]	296	46.43	47	37.66	Erosion	Total AET > 5.5%	—	—	SAP ⩾ 95%	Esomeprazole, 40 mg daily for 14 ± 3 days	Complete symptom relief
Cho *et al.*[31]	73	57.53	47	63.01	Erosion	Total AET > 4.4%	—	—	SAP ⩾ 95%	Lansoprazole, 30 mg BID for 2 weeks	Symptom score improved by more than 50%
Zheng *et al.*[32]	27	29.63	57	3.7	Erosion	Total AET > 4.0%	—	⩾ 14.0	SI ⩾ 50%	Esomeprazole, 20 mg BID for 2 weeks	NA

(Continued)

Table 1. (Continued)

Comparison (author)	Evaluable patients (*n*)	Men (%)	Average age	Prevalence (%) of esophagitis	Reference tests					Study test	Outcome measure extracted
					Upper endoscopy	pH/pH impedance monitoring					
						AET	Number of reflux events	DeMeester score	Symptom reflux association		
Lee *et al.*[33]	164	43.29	NA	42.07	Erosion	—	—	—	—	Rabeprazole, 20 mg/40 mg for 2 weeks	50% symptom reduction
Fan *et al.*[34]	32	37.5	47.3	0	—	—	—	⩾14.7	SI ⩾ 50%	Rabeprazole, 10 mg BID for 2 weeks	Symptom score reduced by more than a third
des Varannes *et al.*[35]	29	40.28	NA	31.94	Erosion	Total AET > 4.2%	—	—	SAP ⩾ 95% or SI ⩾ 50%	Rabeprazole, 20 mg BID for 1 week	The cut-off descriptor of at least a 'clear improvement'
Aanen *et al.*[36]	67	62.16	51	NA	—	Total AET > 4.2%	—	—	SAP ⩾ 95% or SI ⩾ 50%	Esomeprazole, 40 mg daily for 2 weeks	Adequate symptom suppression
Dickman *et al.*[37]	35	65.71	55.6	34.29	Erosion	Total AET > 4.2%, or upright AET > 6.0%, or supine AET > 1.2%				Rabeprazole, 20 mg BID for 1 week	Symptom score improved by more than 50%
Juul-Hansen *et al.*[38]	52	33.85	47	0	—	Total AET > 4.0%	—	—	—	Lansoprazole, 60 mg for 1 week	Symptom relief: yes/no
Pandak *et al.*[39]	38	42.86	NA	26.19	Erosion	Total AET > 4.2%	—	—	—	Omeprazole, 40 mg/d orally BID for 2 weeks	Symptom score improved by more than two points
Juul-Hansen *et al.*[40]	56	31.25	54	0	—	Total AET > 4.2%	—	—	—	Lansoprazole 60 mg daily for 1 week	Antacid use decreased by more than 75%
Fass *et al.*[41]	35	94.29	55	100	Erosion	Total AET > 4.2%	—	—	—	Omeprazole (40 mg a.m. plus 20 mg p.m.) for 1 week	Heartburn score improved by more than 50%
Fass *et al.*[42]	42	76.19	55.2	50	Erosion	Total AET > 4.2%	—	—	—	Omeprazole (40 mg a.m. plus 20 mg p.m.) for 1 week	Heartburn score improved by more than 50%

(Continued)

Table 1. (Continued)

Comparison (author)	Evaluable patients (*n*)	Men (%)	Average age	Prevalence (%) of esophagitis	Reference tests					Study test	Outcome measure extracted
					Upper endoscopy	pH/pH impedance monitoring					
						AET	Number of reflux events	DeMeester score	Symptom reflux association		
Bate *et al.*[43]	58	55.07	47.4	49.28	—	Total AET > 4.0%	—	—	—	Omeprazole, 40 mg o.m. for 2 weeks	Symptoms improved by at least two grades (or from mild to none)
Johnsson *et al.*[44]	160	NA	NA	57.5	Erosion	Total AET > 4.0%	—	—	—	Omeprazole, 20 mg BID for 1 week	Symptoms decreased by at least one grade
Fass *et al.*[45]	37	97.44	60.1	43.24	Erosion	Total AET > 4.2%	—	—	—	Omeprazole (40 mg a.m. plus 20 mg p.m.) for 1 week	More than 50% reduction in symptoms
Carlsson *et al.*[46]	225	50.1	50	61	Erosion	—	—	—	—	Omeprazole, 20 mg for 4 weeks	Complete symptom relief
Galmiche, Barthelemy and Hamelin[47]	141	44.7	50	26	Erosion	—	—	—	—	Omeprazole, 20 mg for 4 weeks	Symptoms decrease to 1 day/week, less than mild
Hatlebakk *et al.*[48]	161	57	49	48	Erosion	—	—	—	—	Omeprazole, 20 mg for 4 weeks	Symptoms decrease to 1 day/week, less than mild
Venables *et al.*[49]	330	52	52	31	Erosion	—	—	—	—	Omeprazole, 20 mg for 4 weeks	Symptoms decrease to 1 day/week, less than mild
A *versus* G											
Ates *et al.*[50]	268	NA	NA	22.76	Erosion	Total AET > 5.3%	—	—	—	Mucosal impedance	A cut-off value of 2019 Ω
Saritas Yuksel *et al.*[51]	69	NA	NA	27.54	Erosion	Total AET > 5.3%				Mucosal impedance	A cut-off value of 3200 Ω
F *versus* E											
Xu *et al.*[52]	2014	126	46.83	45.6	NA	8-week PPI treatment				GERDQ	A cut-off score of 8

A: esophageal pH/pH impedance monitoring and/or endoscopy; B: baseline impedance; C: salivary pepsin; D: DIS; E: GERDQ; F: PPI test; G: mucosal impedance.
AET, acid exposure time; BID, twice daily; DIS, dilated intercellular space; GERD, gastroesophageal reflux disease; GERDQ, GERD questionnaire; MNBI, mean nocturnal baseline impedance; NA, not applicable; o.m., every morning; SAP, symptom association probability; SI, symptom index; PPI, proton-pump inhibitor.

Upper endoscopy or esophageal pH/pH impedance monitoring. Esophageal mucosal breaks on upper endoscopy suggest the presence of GERD. GERD was diagnosed in studies when patients had esophagitis of any grade in one of the commonly used classification systems, such as the Los Angeles or Hetzel–Dent grading systems.

Ambulatory esophageal pH/pH impedance monitoring is generally considered to provide the most objective evidence for pathologic reflux. The results were considered abnormal based on criteria defined in the individual studies. Whenever possible, we chose definitions that would reasonably be interpreted as abnormal in clinical practice, including: (a) acid exposure time (AET) $\geqslant$ 3.2%–5.5% of the monitoring time; (b) DeMeester score $\geqslant$ 14; (c) positive symptom reflux association, such as symptom-associated probability (SAP) $\geqslant$ 95% or symptom index (SI) $\geqslant$ 50%.

Baseline impedance. The baseline impedance was assessed on the pH/impedance system, and the cut-off value ranging from 2100 Ω to 2292 Ω was established and used in individual studies.

Salivary pepsin. Salivary pepsin was detected and quantitatively/semiquantitatively measured by non-invasive rapid salivary pepsin lateral flow device (LFD) (Peptest, RDBiomed, Hull, UK). The cut-off value was used based on criteria defined in individual studies.

Dilated intercellular space. The quantitative or semiquantitative measurement of DIS under light microscopy was performed, and the cut-off value was used based on criteria defined in individual studies.

GERDQ. GERDQ was considered positive if the score was >8.

Proton-pump inhibitor test. The definition of 'a positive PPI test' was based on criteria defined in individual studies. Whenever possible, we chose definitions that would reasonably be interpreted as representing success in clinical practice, and 'complete relief of heartburn' is the most commonly adopted criteria.

Mucosal impedance. The cut-off value of mucosal impedance was also established and used in the individual studies.

Statistical analysis

Firstly, traditional pairwise meta-analyses were performed for studies to compare different diagnostic modalities using the Stata version 12.0 software (StataCorp, College Station, TX, USA). The pooled estimates of odd ratios (ORs) and 95% confidence intervals (CIs) for sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and diagnostic odds ratio of GERD were calculated if there were at least four studies included that had no threshold effect. Area under receiver operating characteristic (AUROC) was also calculated. When the AUROC is closer to 1, the clinical value is greater. When the AUROC is between 0.5 and 0.7, the clinical value is lower. An AUROC value > 0.7 indicates that the clinical value is good. Heterogeneity among studies was tested using the I^2 and Chi-square tests.[53] Secondly, the evidence network structure was drawn using the R version 3.5.1 statistical computing software and network package. Each node represents different diagnostic tests, with the node size reflecting the number of patients, and the thickness of lines between nodes indicating the number of included studies. Thirdly, Bayesian network meta-analyses were performed to combine the effective sizes of direct and indirect comparisons. Lack of autocorrelation and convergence were checked and confirmed by four chains and a 20,000-simulation burn-in phase; finally, direct probability statements were derived from an additional 50,000-simulation phase.[54] The consistency between direct and indirect evidence was assessed with the node-splitting method, and the consistency or inconsistency model was selected accordingly.[55] The ranking probability was then used to calculate the probability of each diagnostic test being the most effective diagnostic method based using a Bayesian approach, and the bar charts of the ranking probability were also produced; the larger the value is, the better the rank of the diagnostic test.[56,57] Comparison-adjusted funnel plots were performed to detect the small study effects on data.[56,58] R (version 3.5.1) package GeMTC was used for this network meta-analysis.

Results

Study search flow

A total of 6223 potentially relevant studies were initially retrieved and identified, of which 705 duplicate studies were excluded. Of the 5518 citations, 5450 citations were ruled out during

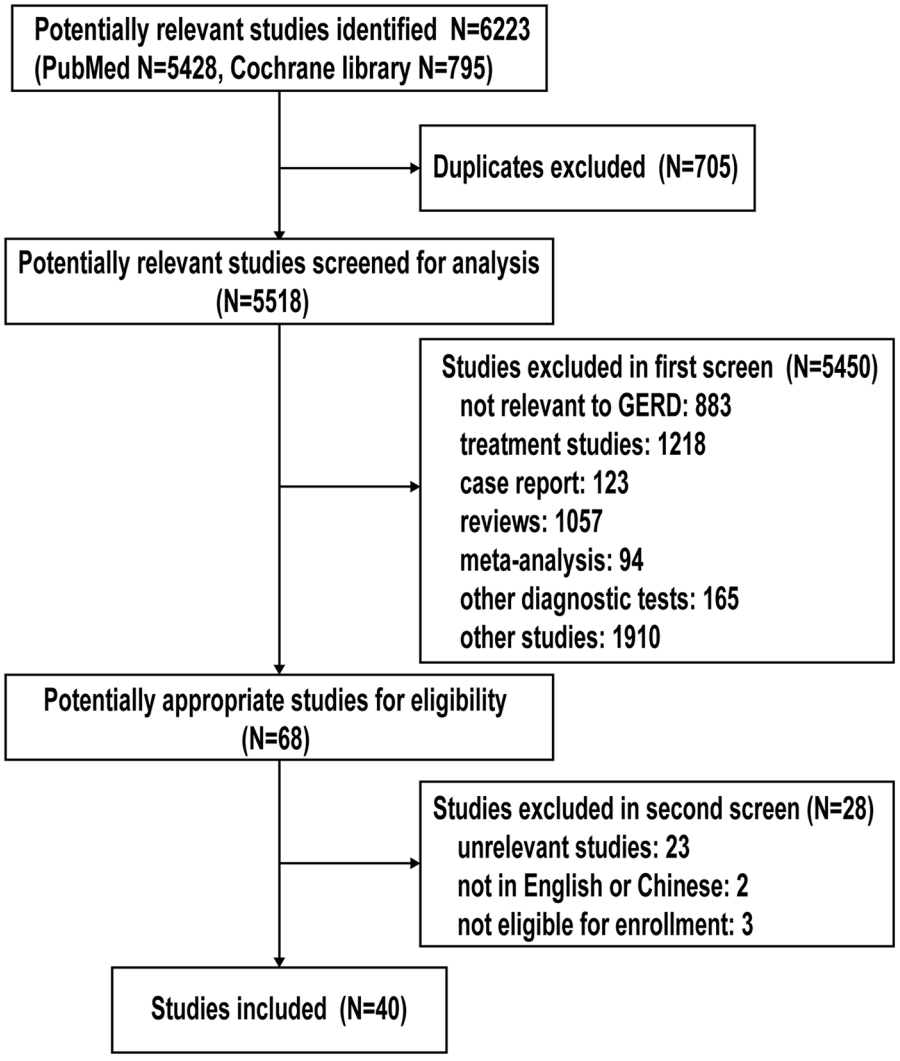

Figure 1. The PRISMA study search flow.
PRISMA, Preferred Reporting Items for Systematic Review and Meta-Analysis.
GERD, gastroesophageal reflux disease.

the first screen by abstract review. A total of 68 studies were then evaluated for eligibility by full-text review. After full-text review, studies unrelated to diagnostic tests for GERD ($n=23$), studies not in English/Chinese language ($n=2$), and studies not eligible for enrollment ($n=3$) were excluded. Altogether, a total of 40 published studies meeting the predetermined inclusion criteria were identified [13–52] (Figure 1).

Study characteristics and qualities

Of 40 included studies, the evaluation of esophageal impedance for GERD diagnosis was performed in 4 studies, salivary pepsin by Peptest in 2 studies, DIS in 5 studies, GERDQ in 6 studies, and PPI test in 23 studies when compared with esophageal pH/pH impedance monitoring or endoscopy. One study compared the diagnostic accuracy of GERDQ and PPI test. The clinical information of the included studies is shown in Table 1. The diagnostic characteristic of diagnostic tests for GERD varied across studies (Table 2). The evaluation of the risk of bias and applicability concerns using the QUADAS-2 was shown in Figure 2 and Supplementary Figure 1.

Pairwise meta-analysis for diagnostic tests for GERD

A direct pairwise meta-analysis of the diagnostic performance of six different tests for GERD

Table 2. Diagnostic evaluation of seven diagnostic tests for gastroesophageal reflux disease.

Comparison (author)	Patients					Prevalence of GERD according to reference test	SEN	SPE	PPV	NPV	LR+	LR–
	TP	FP	FN	TN	Total							
A *versus* B												
Frazzoni *et al.*[13]	60	3	20	38	121	0.66	0.75	0.93	0.95	0.66	10.25	0.27
Ravi *et al.*[14]	25	5	4	21	55	0.53	0.86	0.81	0.83	0.84	4.48	0.17
Frazzoni *et al.*[15]	62	31	6	190	289	0.24	0.91	0.86	0.67	0.97	6.50	0.10
Kandulski *et al.*[16]	15	5	4	12	36	0.53	0.79	0.71	0.75	0.75	2.68	0.30
A *versus* C												
Hayat *et al.*[17]	66	40	18	74	198	0.42	0.79	0.65	0.62	0.80	2.24	0.33
Saritas Yuksel *et al.*[18]	11	2	11	23	47	0.47	0.50	0.92	0.85	0.68	6.25	0.54
A *versus* D												
Cui *et al.*[19]	186	123	111	145	565	0.53	0.63	0.54	0.60	0.57	1.36	0.69
Zhou *et al.*[20]	188	118	119	151	576	0.53	0.61	0.56	0.61	0.56	1.40	0.69
Cui *et al.*[21]	111	0	8	42	161	0.74	0.93	1.00	1.00	0.84	NA	0.07
Mastracci *et al.*[22]	102	6	17	14	139	0.86	0.86	0.70	0.94	0.45	2.86	0.20
Vela *et al.*[23]	9	1	6	10	26	0.58	0.60	0.91	0.90	0.63	6.60	0.44
A *versus* E												
Zhou *et al.*[20]	203	145	149	139	636	0.55	0.58	0.49	0.58	0.48	1.13	0.86
Zavala-Gonzales *et al.*[24]	129	20	51	52	252	0.71	0.72	0.72	0.87	0.50	2.58	0.39
Jonasson *et al.*[25]	115	11	32	11	169	0.87	0.78	0.50	0.91	0.26	1.56	0.44
Lacy *et al.*[26]	77	41	31	28	177	0.61	0.71	0.41	0.65	0.47	1.20	0.71
Jones *et al.*[27]	125	33	69	81	308	0.63	0.64	0.71	0.79	0.54	2.23	0.50
Bai *et al.*[28]	620	1405	815	5225	8065	0.18	0.43	0.79	0.31	0.87	2.04	0.72
A *versus* F												
Zhou *et al.*[20]	248	158	104	126	636	0.55	0.70	0.44	0.61	0.55	1.27	0.67
Jonasson *et al.*[25]	90	4	27	3	124	0.94	0.77	0.43	0.96	0.10	1.35	0.54
Lee *et al.*[29]	55	82	15	36	188	0.37	0.79	0.31	0.40	0.71	1.13	0.70
Dent *et al.*[30]	106	35	91	64	296	0.67	0.54	0.65	0.75	0.41	1.52	0.71
Cho *et al.*[31]	49	4	15	5	73	0.88	0.77	0.56	0.92	0.25	1.72	0.42
Zheng *et al.*[32]	5	11	4	7	27	0.33	0.56	0.39	0.31	0.64	0.91	1.14

(Continued)

Table 2. (Continued)

Comparison (author)	Patients					Prevalence of GERD according to reference test	SEN	SPE	PPV	NPV	LR+	LR–
	TP	FP	FN	TN	Total							
Lee *et al.*[33]	52	28	17	67	164	0.42	0.75	0.71	0.65	0.80	2.56	0.35
Fan *et al.*[44]	20	2	5	5	32	0.78	0.80	0.71	0.91	0.50	2.80	0.28
Des Varannes *et al.*[35]	15	6	3	5	29	0.62	0.83	0.45	0.71	0.63	1.53	0.37
Aanen *et al.*[36]	41	17	5	4	67	0.69	0.89	0.19	0.71	0.44	1.10	0.57
Dickman *et al.*[37]	12	2	4	17	35	0.46	0.75	0.89	0.86	0.81	7.13	0.28
Juul-Hansen *et al.*[38]	34	17	0	1	52	0.65	1.00	0.06	0.67	1.00	1.06	0.00
Pandak *et al.*[39]	19	7	1	11	38	0.53	0.95	0.61	0.73	0.92	2.44	0.08
Juul-Hansen *et al.*[40]	29	11	5	11	56	0.61	0.85	0.50	0.73	0.69	1.71	0.29
Fass *et al.*[41]	21	8	0	6	35	0.60	1.00	0.43	0.72	1.00	1.75	0.00
Fass *et al.*[42]	28	3	7	4	42	0.83	0.80	0.57	0.90	0.36	1.87	0.35
Bate *et al.*[43]	22	11	10	15	58	0.55	0.69	0.58	0.67	0.60	1.63	0.54
Johnsson *et al.*[44]	100	16	35	9	160	0.84	0.74	0.36	0.86	0.20	1.16	0.72
Fass *et al.*[45]	18	2	5	12	37	0.62	0.78	0.86	0.90	0.71	5.48	0.25
Carlsson *et al.*[46]	66	25	72	62	225	0.61	0.48	0.71	0.73	0.46	1.66	0.73
Galmiche *et al.*[47]	27	65	10	39	141	0.26	0.73	0.38	0.29	0.80	1.17	0.72
Hatlebakk *et al.*[48]	55	59	22	25	161	0.48	0.71	0.30	0.48	0.53	1.02	0.96
Venables *et al.*[49]	80	120	21	109	330	0.31	0.79	0.48	0.40	0.84	1.51	0.44
A *versus* G												
Ates *et al.*[50]	108	6	34	120	268	0.53	0.76	0.95	0.95	0.78	15.97	0.25
Saritas Yuksel *et al.*[51]	37	9	5	18	69	0.61	0.88	0.67	0.80	0.78	2.64	0.18
F *versus* E												
Xu *et al.*[52]	68	2	34	22	126	0.81	0.67	0.92	0.97	0.39	8.00	0.36

A: esophageal pH/pH impedance monitoring and/or endoscopy; B: baseline impedance; C: salivary pepsin; D: DIS; E: GERDQ; F: PPI test; G: mucosal impedance.
AET, acid exposure time; DIS, dilated intercellular space; FN, false negative; FP, false positive; GERD, gastroesophageal reflux disease; GERDQ, GERD questionnaire; LR, likelihood ratio; MNBI, mean nocturnal baseline impedance; NPV, negative-predictive value; PPI, proton-pump inhibitor; PPV, positive-predictive value; SAP, symptom association probability; SEN, sensitivity; SI, symptom index; SPE, specificity; TP, true positive; TN, true negative.

diagnosis was conducted. The results revealed that the baseline impedance, GERDQ and PPI test exhibited lower sensitivity and specificity when compared with esophageal pH/pH impedance monitoring or endoscopy. We also calculated the AUROC for each diagnostic test and found that the esophageal impedance and PPI test were higher than 0.70, indicating that they had relatively high

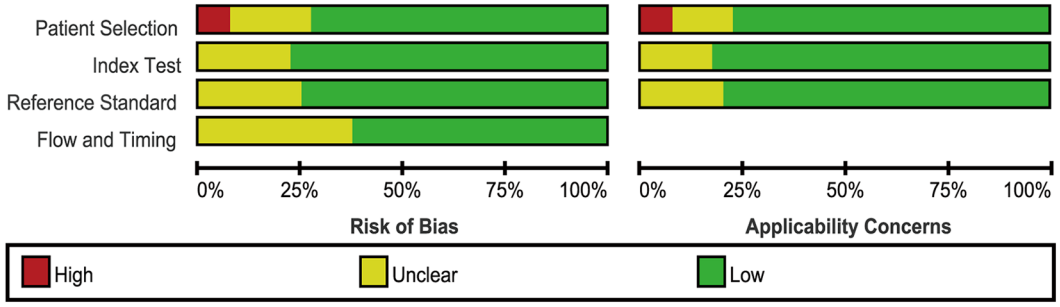

Figure 2. The evaluation of risks of bias of included studies.

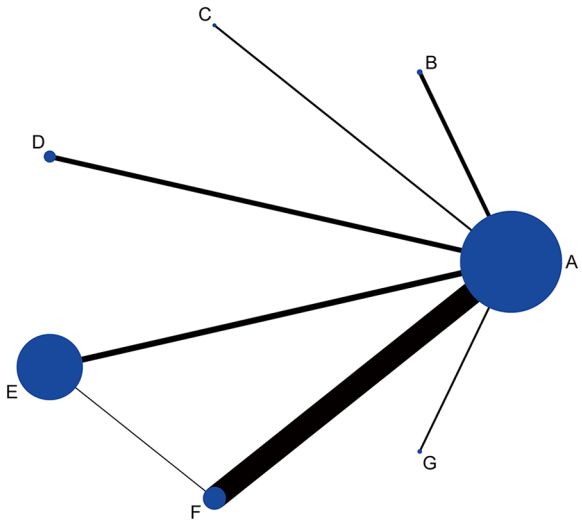

Figure 3. The evidence network structure.
A: esophageal pH/pH impedance monitoring or endoscopy; B: baseline impedance; C: salivary pepsin; D: DIS; E: GERDQ; F: PPI test; G: mucosal impedance.
DIS, dilated intercellular spaces; GERD, gastroesophageal reflux disease; GERDQ, GERD questionnaire; PPI, proton-pump inhibitor.

diagnostic value (Supplementary Table 1). The pairwise meta-analysis of DIS could not be performed successfully due to a threshold effect. The pairwise meta-analysis of salivary pepsin and mucosal impedance could not be performed either because there were only two studies included.

Evidence network of diagnostic tests for GERD

The evidence network structure included seven diagnostic tests. The highest number of evaluable patients performed the esophageal pH/pH impedance monitoring or endoscopy, and most studies compared PPI test with esophageal pH/pH impedance monitoring or endoscopy for GERD diagnosis (Figure 3).The effect of the direct comparison of different tests with esophageal pH/pH impedance monitoring or endoscopy had similar effect on the entire network meta-analysis (Supplementary Figure 2).

Main results of network meta-analysis of diagnostic tests for GERD

The NMA found no significant difference among the baseline impedance, mucosal impedance, and esophageal pH/pH impedance monitoring or endoscopy in terms of both sensitivity and specificity. It was also demonstrated that the salivary pepsin detected by Peptest had comparable specificity with esophageal pH/pH impedance monitoring or endoscopy (Figure 4).

Ranking probability of diagnostic tests for GERD

Ranking probability indicated that esophageal pH/pH impedance monitoring and/or endoscopy had the highest sensitivity, followed by baseline impedance or mucosal impedance, PPI test, salivary pepsin or DIS, and GERDQ [Figure 5(a)]. Moreover, esophageal pH/pH impedance monitoring and/or endoscopy also had the highest specificity, followed by the mucosal impedance, baseline impedance, DIS or salivary pepsin, GERDQ, and PPI test [Figure 5(b)]. The ranking probability of positive-predictive value and negative-predictive value was also provided in Figure 5(c) and 5(d).

Assessment of publication bias

The results of assessment of publication bias demonstrated symmetrical distribution, indicating no small sample effect or publication bias in this NMA (Supplementary Figure 3).

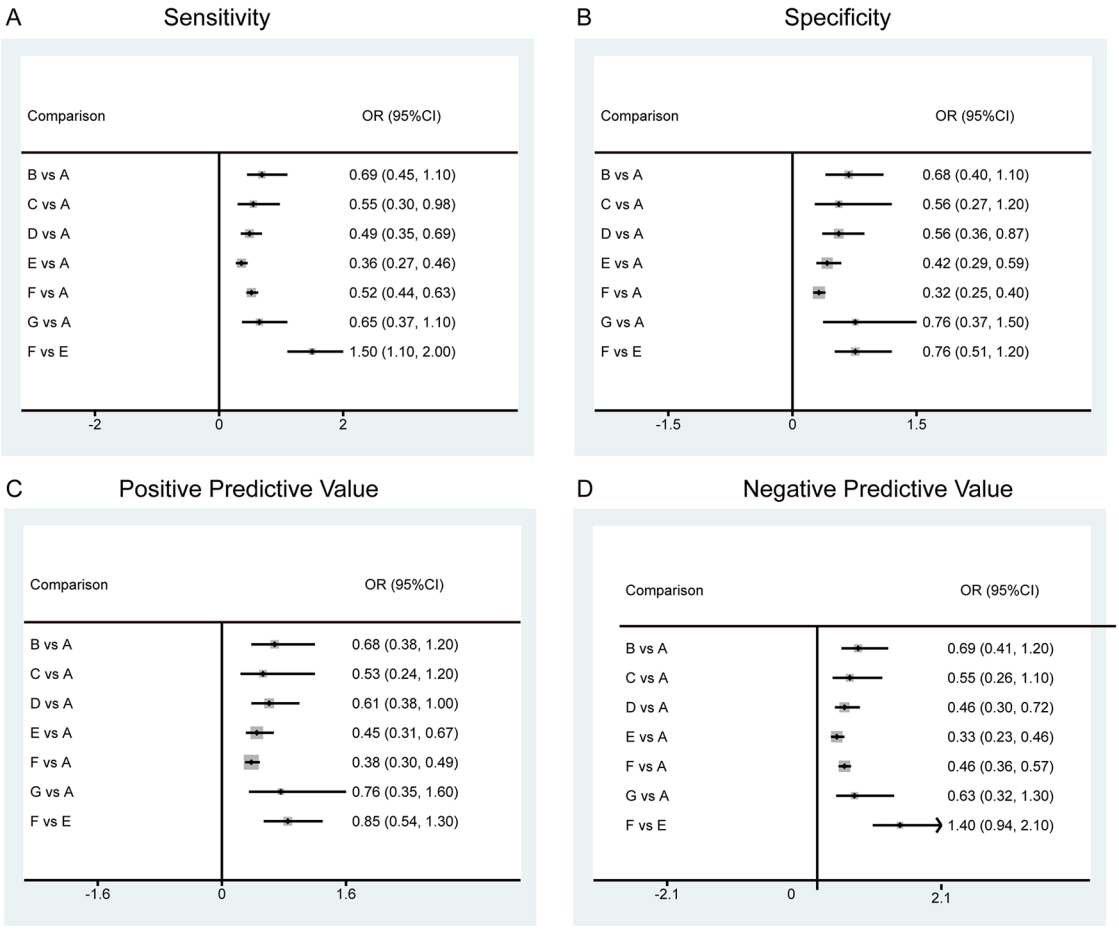

Figure 4. The forest plots based on sensitivity, specificity, positive-predictive value and negative-predictive value of different diagnostic tests for GERD.
A: esophageal pH/pH impedance monitoring or endoscopy; B: baseline impedance; C: salivary pepsin; D: DIS; E: GERDQ; F: PPI test; G: mucosal impedance.
DIS, dilated intercellular spaces; GERD, gastroesophageal reflux disease; GERDQ, GERD questionnaire; PPI, proton-pump inhibitor.

Discussion

We present the first systematic review and NMA comparing the diagnostic performance of GERDQ questionnaire, PPI test, baseline impedance, mucosal impedance, DIS, salivary pepsin and esophageal pH/pH impedance monitoring/endoscopy for GERD. The NMA and ranking probabilities reveal that the mucosal impedance and baseline impedance had comparable sensitivity and specificity with esophageal pH/pH impedance monitoring or endoscopy. GERDQ had the lowest sensitivity, and PPI test had the lowest specificity.

The results of direct pairwise comparison and NMA shows that esophageal reflux monitoring or endoscopy is superior to other diagnostic tests for GERD, which is in accordance with the recommendation in current guidelines.[59] In the current meta-analysis, the criteria for abnormal reflux varied across the included studies which may add clinical heterogeneity to our analysis to some extent. However, this heterogeneity couldn't be avoided, since the understanding toward GERD pathophysiology and the diagnostic criteria for GERD had developed and changed during the past several decades. For instance, the latest Lyon Consensus lists conclusive evidence, including Los Angeles grade C and D erosive esophagitis from upper endoscopy or AET > 6% from pH/pH impedance monitoring for the definitive diagnosis of GERD.[60] However, most previous studies diagnosed GERD based on lower AET thresholds and esophageal mucosal breaks, regardless of grades. Further studies are needed

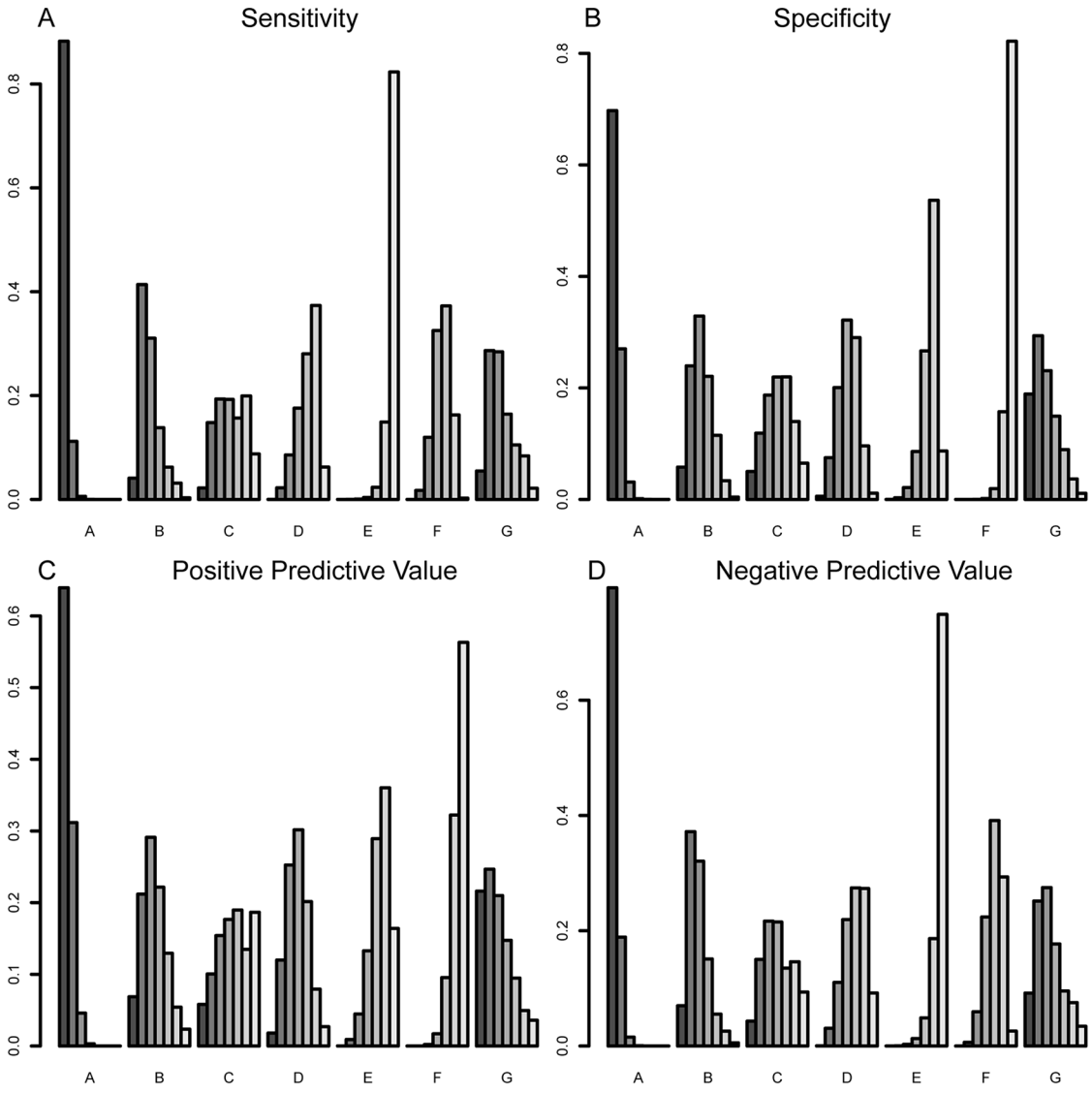

Figure 5. The ranking probability based on sensitivity, specificity, positive-predictive value and negative-predictive value of different diagnostic tests for GERD.
A: esophageal pH/pH impedance monitoring and/or endoscopy; B: baseline impedance; C: salivary pepsin; D: DIS; E: GERDQ; F: PPI test; G: mucosal impedance.
DIS, dilated intercellular spaces; GERD, gastroesophageal reflux disease; GERDQ, GERD questionnaire; PPI, proton-pump inhibitor.

to compare different grades of esophagitis and different criteria for reflux for the diagnosis and management of GERD.

We also found that the esophageal mucosal impedance as well as the baseline impedance had comparable diagnostic performance with esophageal reflux monitoring and endoscopy. Baseline impedance reflects the integrity of the esophageal mucosa,[61] with low values observed in patients having GERD, and associated with increased acid reflux, as well as DIS.[62,63] It is reported that mucosal impedance can help differentiate GERD from eosinophilic esophagitis (EoE), achalasia, and healthy controls with higher specificity (95%) when compared with reflux testing (64%).[50] Given its high diagnostic performance, the esophageal mucosal impedance, as well as the baseline impedance, may become promising diagnostic tools for GERD in the future. However, normative values for them still need to be determined. Also, mucosa impedance is currently not

commercialized and widely used despite high diagnostic utility.

Additionally, the salivary pepsin detection was initially proposed as a non-invasive method for the diagnosis of GERD. Our results demonstrated that the measurement of salivary pepsin detected by the Peptest device had comparable specificity to esophageal pH/pH impedance monitoring or endoscopy. Nevertheless, Sifrim and coworkers failed to reproduce good specificity of salivary pepsin to diagnose GERD, and they found that salivary pepsin could not differentiate GERD from functional heartburn.[64] Therefore, salivary pepsin detection cannot be recommended for clinical application at present. Besides, our results found that the measurement of DIS only had only modest diagnostic characteristics. It has been reported that esophageal mucosal changes such as DIS may help differentiate GERD from other disorders.[65,66] Even so, the measurement of DIS on electron microscopy is not ready for clinical practice yet.

There are limitations to these findings. First, esophageal pH/pH impedance monitoring and endoscopy were examined together in the current NMA, which were found to have highest sensitivity and specificity. However, upper endoscopy alone has, actually, low sensitivity for GERD, so we should interpret results with caution. Second, we only analyzed GERDQ, while there are several other questionnaires for GERD diagnosis. However, all of them have different scoring systems and putting them together may lead to high heterogeneity in the NMA. And the GERDQ is a questionnaire derived from validated questionnaires including the Reflux Disease Questionnaire, Gastrointestinal Symptom Rating Scale, and the GERD Impact Scale. It is the most commonly used questionnaire for GERD. Thus, we had to choose only one questionnaire to be examined. Also, we did not include the postreflux swallow-induced peristaltic wave (PSPW) index in the current NMA because studies evaluating the diagnostic value of the PSPW index are limited at present; future NMA studies can include it, since it is also a promising metric for GERD diagnosis. Third, the rankings and probabilities can sometimes be misleading, ignoring the basic principles of certainty evaluation in evidence during NMA. We should analyze the results of NMA with caution and take the results of pairwise comparison into account. Furthermore, there exists high heterogeneity in the study populations, including various baseline characteristics, different procedure of the testing, and different outcome measure, etc., which may make the results incomparable, to an extent. Finally, the results of pairwise comparison and NMA were associated with wide confidence intervals and some included studies were of moderate-to-low quality.

In conclusion, the current systematic review and NMA shows that esophageal mucosal impedance and baseline impedance has a high diagnostic performance similar to esophageal reflux monitoring or endoscopy. The future direction of GERD management should be developing and improving techniques with high diagnostic performance; not only for a precision phenotype definition but also for a tailored treatment strategy.

Acknowledgment

Mengyu Zhang: data acquisition and analysis, manuscript drafting; John E Pandolfino: critical revision of the manuscript; Xuyu Zhou: statistical analysis; Niandi Tan: data acquisition and analysis; Yuwen Li: data acquisition and analysi;, Minhu Chen and Yinglian Xiao: study design, data analysis, study supervision, and finalizing the manuscript.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was supported by grants from the National Natural Science Foundation of China (81770544).

Conflict of interest statement

The authors declare that there is no conflict of interest.

ORCID iDs

Mengyu Zhang https://orcid.org/0000-0002-2838-1103

Minhu Chen https://orcid.org/0000-0001-9925-135X

Supplemental material

Supplemental material for this article is available online.

References

1. Vakil N, Van Zanten SV, Kahrilas P, *et al.* The Montreal definition and classification

of gastroesophageal reflux disease: a global evidence-based consensus. *Am J Gastroenterol* 2006; 101:1900–1920; quiz 1943.

2. El-Serag HB, Sweet S, Winchester CC, *et al*. Update on the epidemiology of gastro-oesophageal reflux disease: a systematic review. *Gut* 2014; 63: 871–880.
3. Peery AF, Dellon ES, Lund J, *et al*. Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology* 2012; 143: 1179–1187.e3.
4. Katz PO, Gerson LB and Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. *Am J Gastroenterol* 2013; 108: 308–328; quiz 29.
5. Kahrilas PJ, Shaheen NJ and Vaezi MF. American Gastroenterological Association Institute technical review on the management of gastroesophageal reflux disease. *Gastroenterology* 2008; 135: 1392–1413, 1413.e1-5.
6. Richter JE and Rubenstein JH. Presentation and epidemiology of gastroesophageal reflux disease. *Gastroenterology* 2018; 154: 267–276.
7. Bytzer P, Jones R, Vakil N, *et al*. Limited ability of the proton-pump inhibitor test to identify patients with gastroesophageal reflux disease. *Clin Gastroenterol Hepatol* 2012; 10: 1360–1366.
8. Roman S, Gyawali CP, Savarino E, *et al*. Ambulatory reflux monitoring for diagnosis of gastro-esophageal reflux disease: update of the Porto consensus and recommendations from an international consensus group. *Neurogastroenterol Motil* 2017; 29: 1–15.
9. Vaezi MF and Sifrim D. Assessing old and new diagnostic tests for gastroesophageal reflux disease. *Gastroenterology* 2018; 154: 289–301.
10. Deville WL, Bezemer PD and Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000; 53: 65–69.
11. Ioannidis JP, Chew P and Lau J. Standardized retrieval of side effects data for meta-analysis of safety outcomes. A feasibility study in acute sinusitis. *J Clin Epidemiol* 2002; 55: 619–626.
12. Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.
13. Frazzoni M, de Bortoli N, Frazzoni L, *et al*. The added diagnostic value of postreflux swallow-induced peristaltic wave index and nocturnal baseline impedance in refractory reflux disease studied with on-therapy impedance-pH monitoring. *Neurogastroenterol Motil* 2017; 29.
14. Ravi K, Geno DM, Vela MF, *et al*. Baseline impedance measured during high-resolution esophageal impedance manometry reliably discriminates GERD patients. *Neurogastroenterol Motil* 2017 May; 29(5). DOI: 10.1111/nmo.12974. Epub 2016 Oct 24.
15. Frazzoni M, Savarino E, De Bortoli N, *et al*. Analyses of the post-reflux swallow-induced peristaltic wave index and nocturnal baseline impedance parameters increase the diagnostic yield of impedance-pH monitoring of patients with reflux disease. *Clin Gastroenterol Hepatol* 2016; 14: 40–46.
16. Kandulski A, Weigt J, Caro C, *et al*. Esophageal intraluminal baseline impedance differentiates gastroesophageal reflux disease from functional heartburn. *Clin Gastroenterol Hepatol* 2015; 13: 1075–1081.
17. Hayat JO, Gabieta-Somnez S, Yazaki E, *et al*. Pepsin in saliva for the diagnosis of gastro-oesophageal reflux disease. *Gut* 2015; 64: 373–380.
18. Saritas Yuksel E, Hong SK, *et al*. Rapid salivary pepsin test: blinded assessment of test performance in gastroesophageal reflux disease. *Laryngoscope* 2012; 122: 1312–1316.
19. Cui R, Zhang H, Zhou L, *et al*. Diagnostic value of dilated intercellular space and histopathologic scores in gastroesophageal reflux disease. *Dis Esophagus* 2015; 28: 530–537.
20. Zhou LY, Wang Y, Lu JJ, *et al*. Accuracy of diagnosing gastroesophageal reflux disease by GerdQ, esophageal impedance monitoring and histology. *J Dig Dis* 2014; 15: 230–238.
21. Cui R, Zhou L, Lin S, *et al*. The feasibility of light microscopic measurements of intercellular spaces in squamous epithelium in the lower-esophagus of GERD patients. *Dis Esophagus* 2011; 24: 1–5.
22. Mastracci L, Spaggiari P, Grillo F, *et al*. Microscopic esophagitis in gastro-esophageal reflux disease: individual lesions, biopsy sampling, and clinical correlations. *Virchows Archiv* 2009; 454: 31–39.
23. Vela MF, Craft BM, Sharma N, *et al*. Refractory heartburn:comparison of intercellular space diameter in documented GERD vs. functional heartburn. *Am J Gastroenterol* 2011; 106: 844–850.
24. Zavala-Gonzales MA, Azamar-Jacome AA, *et al*. Validation and diagnostic usefulness of

gastroesophageal reflux disease questionnaire in a primary care level in Mexico. *J Neurogastroenterol Motil* 2014; 20: 475–482.

25. Jonasson C, Wernersson B, Hoff DA, *et al.* Validation of the GerdQ questionnaire for the diagnosis of gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 2013; 37: 564–572.
26. Lacy BE, Chehade R and Crowell MD. A prospective study to compare a symptom-based reflux disease questionnaire to 48-h wireless pH monitoring for the identification of gastroesophageal reflux (revised 2-26-11). *Am J Gastroenterol* 2011; 106: 1604–1611.
27. Jones R, Junghard O, Dent J, *et al.* Development of the GerdQ, a tool for the diagnosis and management of gastro-oesophageal reflux disease in primary care. *Aliment Pharmacol Ther* 2009; 30: 1030–1038.
28. Bai Y, Du Y, Zou D, *et al.* Gastroesophageal Reflux Disease Questionnaire (GerdQ) in real-world practice: a national multicenter survey on 8065 patients. *J Gastroenterol Hepatol* 2013; 28: 626–631.
29. Lee SH, Jang BI, Jeon SW, *et al.* A multicenter, randomized, comparative study to determine the appropriate dose of lansoprazole for use in the diagnostic test for gastroesophageal reflux disease. *Gut Liver* 2011; 5: 302–307.
30. Dent J, Vakil N, Jones R, Bytzer P, *et al.* Accuracy of the diagnosis of GORD by questionnaire, physicians and a trial of proton pump inhibitor treatment: the Diamond study. *Gut* 2010; 59: 714–721.
31. Cho YK, Choi MG, Lim CH, *et al.* Diagnostic value of the PPI test for detection of GERD in Korean patients and factors associated with PPI responsiveness. *Scand J Gastroenterol* 2010; 45: 533–539.
32. Zheng J, Du ZM, Chen MH, *et al.* [Diagnosis of gastroesophageal reflux disease-related noncardiac chest pain]. *Zhonghua yi xue za zhi* 2008; 88: 1390–1393.
33. Lee YC, Lin JT, Wang HP, *et al.* Influence of cytochrome P450 2C19 genetic polymorphism and dosage of rabeprazole on accuracy of proton-pump inhibitor testing in Chinese patients with gastroesophageal reflux disease. *J Gastroenterol Hepatol* 2007; 22: 1286–1292.
34. Fan YH, Lu B, Zhan LX, *et al.* [Oesophageal acid exposure test in non-erosive gastroesophageal reflux disease and the diagnostic value of rabeprazole]. *Zhonghua nei ke za zhi* 2007; 46: 475–477.
35. Des Varannes SB, Sacher-Huvelin S, Vavasseur F, *et al.* Rabeprazole test for the diagnosis of gastro-oesophageal reflux disease: results of a study in a primary care setting. *World J Gastroenterol* 2006; 12: 2569–2573.
36. Aanen MC, Weusten BL, Numans ME, *et al.* Diagnostic value of the proton pump inhibitor test for gastro-oesophageal reflux disease in primary care. *Aliment Pharmacol Ther* 2006; 24: 1377–1384.
37. Dickman R, Emmons S, Cui H, *et al.* The effect of a therapeutic trial of high-dose rabeprazole on symptom response of patients with non-cardiac chest pain: a randomized, double-blind, placebo-controlled, crossover trial. *Aliment Pharmacol Ther* 2005; 22: 547–555.
38. Juul-Hansen P and Rydning A. Endoscopy-negative reflux disease: what is the value of a proton-pump inhibitor test in everyday clinical practice? *Scand J Gastroenterol* 2003; 38: 1200–1203.
39. Pandak WM, Arezo S, Everett S, *et al.* Short course of omeprazole: a better first diagnostic approach to noncardiac chest pain than endoscopy, manometry, or 24-hour esophageal pH monitoring. *J Clin Gastroenterol* 2002; 35: 307–314.
40. Juul-Hansen P, Rydning A, Jacobsen CD, *et al.* High-dose proton-pump inhibitors as a diagnostic test of gastro-oesophageal reflux disease in endoscopic-negative patients. *Scand J Gastroenterol* 2001; 36: 806–810.
41. Fass R, Ofman JJ, Sampliner RE, *et al.* The omeprazole test is as sensitive as 24-h oesophageal pH monitoring in diagnosing gastro-oesophageal reflux disease in symptomatic patients with erosive oesophagitis. *Aliment Pharmacol Ther* 2000; 14: 389–396.
42. Fass R, Ofman JJ, Gralnek IM, *et al.* Clinical and economic assessment of the omeprazole test in patients with symptoms suggestive of gastroesophageal reflux disease. *Arch Intern Med* 1999; 159: 2161–2168.
43. Bate CM, Riley SA, Chapman RW, *et al.* Evaluation of omeprazole as a cost-effective diagnostic test for gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 1999; 13: 59–66.
44. Johnsson F, Weywadt L, Solhaug JH, *et al.* One-week omeprazole treatment in the diagnosis of gastro-oesophageal reflux disease. *Scand J Gastroenterol* 1998; 33: 15–20.
45. Fass R, Fennerty MB, Ofman JJ, *et al.* The clinical and economic value of a short course of omeprazole in patients with noncardiac chest pain. *Gastroenterology* 1998; 115: 42–49.

46. Carlsson R, Dent J, Watts R, *et al*. Gastro-oesophageal reflux disease in primary care: an international study of different treatment strategies with omeprazole. International GORD Study Group. *Eur J Gastroenterol Hepatol* 1998; 10: 119–124.

47. Galmiche JP, Barthelemy P and Hamelin B. Treating the symptoms of gastro-oesophageal reflux disease: a double-blind comparison of omeprazole and cisapride. *Aliment Pharmacol Ther* 1997; 11: 765–773.

48. Hatlebakk JG, Hyggen A, Madsen PH, *et al*. Heartburn treatment in primary care: randomised, double blind study for 8 weeks. *BMJ (Clinical research ed)* 1999; 319: 550–553.

49. Venables TL, Newland RD, Patel AC, *et al*. Omeprazole 10 milligrams once daily, omeprazole 20 milligrams once daily, or ranitidine 150 milligrams twice daily, evaluated as initial therapy for the relief of symptoms of gastro-oesophageal reflux disease in general practice. *Scand J Gastroenterol* 1997; 32: 965–973.

50. Ates F, Yuksel ES, Higginbotham T, *et al*. Mucosal impedance discriminates GERD from non-GERD conditions. *Gastroenterology* 2015; 148: 334–343.

51. Saritas Yuksel E, Higginbotham T, Slaughter JC, *et al*. Use of direct, endoscopic-guided measurements of mucosal impedance in diagnosis of gastroesophageal reflux disease. *Clin Gastroenterol* 2012; 10: 1110–1116.

52. Xu X, Chen Q, Liang S, *et al*. Comparison of gastroesophageal reflux disease questionnaire and multichannel intraluminal impedance pH monitoring in identifying patients with chronic cough responsive to antireflux therapy. *Chest* 2014; 145: 1264–1270.

53. Chen LX, Li YL, Ning GZ, *et al*. Comparative efficacy and tolerability of three treatments in old people with osteoporotic vertebral compression fracture: a network meta-analysis and systematic review. *PloS One* 2015; 10: e0123153.

54. Tu YK, Needleman I, Chambrone L, *et al*. A Bayesian network meta-analysis on comparisons of enamel matrix derivatives, guided tissue regeneration and their combination therapies. *J Clin Periodontol* 2012; 39: 303–314.

55. Zhu GQ, Shi KQ, Huang S, *et al*. Systematic review with network meta-analysis: the comparative effectiveness and safety of interventions in patients with overt hepatic encephalopathy. *Aliment Pharmacol Ther* 2015; 41: 624–635.

56. Chaimani A, Higgins JP, Mavridis D, *et al*. Graphical tools for network meta-analysis in STATA. *PloS One*. 2013; 8: e76654.

57. Salanti G, Ades AE and Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; 64: 163–171.

58. Mavridis D, Giannatsi M, Cipriani A, *et al*. A primer on network meta-analysis with emphasis on mental health. *Evid Based Ment Health* 2015; 18: 40–46.

59. Aziz Q, Fass R, Gyawali CP, *et al*. Functional esophageal disorders. *Gastroenterology* 2016; 150: 1368–1379.

60. Gyawali CP, Kahrilas PJ, Savarino E, *et al*. Modern diagnosis of GERD: the Lyon Consensus. *Gut* 2018; 67: 1351–1362.

61. Farre R, Blondeau K, Clement D, *et al*. Evaluation of oesophageal mucosa integrity by the intraluminal impedance technique. *Gut* 2011; 60: 885–892.

62. Kessing BF, Bredenoord AJ, Weijenborg PW, *et al*. Esophageal acid exposure decreases intraluminal baseline impedance levels. *Am J Gastroenterol* 2011; 106: 2093–2097.

63. Zhong C, Duan L, Wang K, *et al*. Esophageal intraluminal baseline impedance is associated with severity of acid reflux and epithelial structural abnormalities in patients with gastroesophageal reflux disease. *J Gastroenterol* 2013; 48: 601–610.

64. Woodland P, Singendonk MMJ, Ooi J, *et al*. Measurement of salivary pepsin to detect gastroesophageal reflux disease is not ready for clinical application. *Clin Gastroenterol Hepatol* 2019 Feb; 17(3): 563–565. doi: 10.1016/j.cgh.2018.05.016. Epub 2018 May 18

65. Savarino E, Zentilin P, Mastracci L, *et al*. Microscopic esophagitis distinguishes patients with non-erosive reflux disease from those with functional heartburn. *J Gastroenterol* 2013; 48: 473–482.

66. Van Malenstein H, Farre R and Sifrim D. Esophageal dilated intercellular spaces (DIS) and nonerosive reflux disease. *Am J Gastroenterol* 2008; 103: 1021–1028.