COPD JOURNAL OF CHRONIC OBSTRUCTIVE PULMONARY DISEASE

informa healthcare

COPD, 9:563–570, 2012
ISSN: 1541-2555 print / 1541-2563 online

DOI: 10.3109/15412555.2012.692000

ORIGINAL RESEARCH

Diagnostic Value of Computed Tomography in Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-analysis

Jian-sheng Li,[1,2] Hai-long Zhang,[1,3] Yun-ping Bai,[1] Yan-fang Wang,[1] Hai-feng Wang,[2] Ming-hang Wang,[2] Su-yun Li,[2] and Xue-qing Yu[2]

1 Institute of Geriatrics, Henan University of Traditional Chinese Medicine, Zhengzhou 450008, China

2 Department of Respiratory Diseases, the First Affiliated Hospital of Henan University of Traditional Chinese Medicine, Zhengzhou 450000, China

3 Dongzhimen Hospital Affiliated to Beijing University of Chinese Medicine, Beijing 100700, China

Abstract

Introduction: **Computed tomography (CT) has been approved for diagnosing chronic obstructive pulmonary disease (COPD). The diagnostic accuracy, however, has never been examined in a systematic review. Therefore, we conducted a meta-analysis to evaluate the accuracy of CT in diagnosing COPD.** ***Methods:*** **Articles reporting diagnostic accuracy of CT for COPD were searched from seven electronic databases and hand searching. Two reviewers independently extracted data and assessed methodological quality. Sensitivity (SEN), specificity (SPE), positive and negative likelihood ratios (LR+ and LR–, respectively), and diagnostic odds ratios (DOR) were pooled using a bivariate model. The diagnostic performance of overall test also was assessed using the visual power of the ROC plot to present the bivariate model. Potential between-study heterogeneity was explored using subgroup analyses.** ***Results:*** **Data were extracted from 8 studies that met the inclusion criteria. All summary measures were grossly heterogeneous and therefore would not be appropriately summarized. These studies were further stratified by type of imaging technique and test index. The summary estimates of CT for COPD were as follows: SEN, 0.83(95% CI, 0.73-0.89); SPE, 0.87(95% CI, 0.70-0.95); LR+, 6.2(95% CI, 2.5-15.5); LR–, 0.20(95% CI, 0.12-0.34); and DOR, 31(95% CI, 8-116). The five summary estimates of CT on the lung density were 0.80 (95% CI, 0.74-0.84), 0.77(95% CI, 0.58-0.89), 3.5(95% CI, 1.8-6.9), 0.26(95% CI, 0.20-0.34) and 13(95% CI, 6-32), respectively. Conclusions: The current meta-analyses suggest that quantitative measures of CT may be useful to diagnose COPD. Developed CT technology may improve the accuracy of diagnosis. Further studies assessed diagnostic performance of CT are needed.**

Abbreviations

COPD	chronic obstructive pulmonary disease
PFT	pulmonary function tests
CT	computed tomography
LDCT	low-dose computed tomography
HRCT	high-resolution computed tomography
MSCT	multi-slice computed tomography
QUADAS	quality assessment for studies of diagnostic accuracy
SEN	sensitivity
SPE	specificity
DOR	diagnostic odds ratio
SROC	summary receiver operating characteristic
LR+	positive likelihood ratio
LR–	negative likelihood ratio
AUC	area under the curve
TP	true positive

Keywords: Chronic Obstructive Pulmonary Disease, Diagnosis, Computed tomography, Meta-analysis.

Correspondence to: Jian-sheng Li, The Geriatric Department of Henan University of Traditional Chinese Medicine, No. 1 Jinshui Road, Jinshui District , Zhengzhou 450008, China, Tel: +86 371 65676568. Fax: 86-10-65944307. email: li_js8@163.com

FP	false positive
FN	false negative
TN	true negative
ALD ex	the full expiration average lung density
EI ex	emphysema index in expiration
PI ex	pixel index in maximum expiratory
BF	blood flow
BV	blood volume;
LAA%	the percentage of the low attenuation area
LD	lung density.

Background

Chronic obstructive pulmonary disease (COPD) is an important cause of mortality throughout the world (1). It is the fourth-leading cause of death worldwide (2), and the prevalence of it has been gradually increasing in recent years and is expected to further increase in the future, becoming the third leading cause of mortality in 2020 (3).

To decrease the economic and social burden of COPD, an accurate diagnosis and timely treatment in the early stages of COPD is very important. The diagnosis of COPD is conventionally based upon spirometry (1), pulmonary function tests (PFT) is a well-established method for the diagnosis and assessment of clinical stage of COPD. However, COPD is an insidious disease, with many years between the development of pulmonary function abnormalities with an irreversible airflow limitation and the onset of serious respiratory symptoms, such as severe breathlessness. As much as 30% of the lung may be destroyed by emphysema before either symptoms or abnormalities become evident on pulmonary function tests (4). So, during the early stage of the disease, conventional spirometry may reveal no abnormality as the earliest changes in COPD affect the alveolar walls and small airways (5). The pulmonary function test is limited as it is only a global measure of all the changes occurring in COPD.

During the last few decades, with the advent of high-resolution computed tomography (HRCT), low-dose computed tomography (LDCT), and the development of Multi-slice computed tomography (MSCT) scanning techniques as a new diagnostic modality, computed tomography (CT) has become a very popular technique for the noninvasive assessment of airway disease in COPD (6-9), and has been established as a sensitive diagnostic modality for the detection of early symptomatic and asymptomatic COPD (10).

Because the role of the CT for diagnosing COPD has not been well established, we undertook this systematic review and meta-analysis to evaluate the accuracy of CT in diagnosing COPD.

Methods

Study identification and selection

We searched the PUBMED (1966.1~2011.10), EMBASE(1974.1~2011.10), CNKI(1979.1~2011.10),VIP (1989.1~2011.10), CBM(1978.1~2011.10), WANFANG (1983.1~2011.10) and The Cochrane Library (2011; Issue 4) with the following search terms: "Chronic obstructive pulmonary disease", "COPD", "Computed Tomography", "CT", "diagnosis", etc; and we also hand-searched the references of relevant studies without date limitation. The searches were limited in English and Chinese publications on human subjects.

To be included, the study had to meet the following criteria: (1) The type of research was a diagnostic test that assessed the diagnostic accuracy of CT, HRCT, LDCT or MDST for COPD; (2) sensitivity and specificity were reported or a 2 × 2 contingency table could be (re-)constructed; (3) Diagnostic method for evaluation of test was CT imaging diagnosis, and reference standard was PFT; and (4) the publication was a full report. When different studies from the same institution used the same patients because one author published several reports, the article with the most details or the most recent article was chosen. Two investigators (W. Y. F. and L. S.Y.) independently selected the studies and disagreements were resolved by consensus with a third reviewer (W. H.F.).

Data extraction and quality assessment

Two reviewers (W. M. H. and Y. X. Q.) independently extracted the following information: author, year of publication, sample size, mean age, imaging technique, text index, and outcome data. We performed the quality assessment of included studies using an updated Quality assessment tool "QUADAS-2" (11), an improved, redesigned tool that is based on both experience using the original tool and new evidence about sources of bias and applicability of primary diagnostic accuracy studies. Risk of bias is judged as "low," "high," or "unclear." If the answers of all questions for a domain are"yes," then risk of bias can be judged low. If any signaling question is answered "no," potential for bias exists. The "unclear" category should be used only when insufficient data are reported to permit a judgment. Concerns about applicability are rated as "low," "high," or"unclear" (11). Any disagreements were resolved by discussion with a third author (B. Y. P.) to reach a consensus.

Statistical analysis and data synthesis

Most of study in the meta-analysis contributed a pair of numbers: sensitivity and specificity, the number of true positives, false positives, false negatives and true negatives were calculated according to each sensitivity and specificity by Review Manager (version 5.1). By using a bivariate regression approach, we computed the overall sensitivity (SEN), specificity (SPE) with 95% CI as the main outcome measures. We also used the visual power of the ROC plot to present the resules of the bivariate model. At the same time, we calculated positive likelihood ratio (LR+) and negative likelihood ratio (LR–) respectively (12, 13). These measures were pooled using the random effects models (14-16).

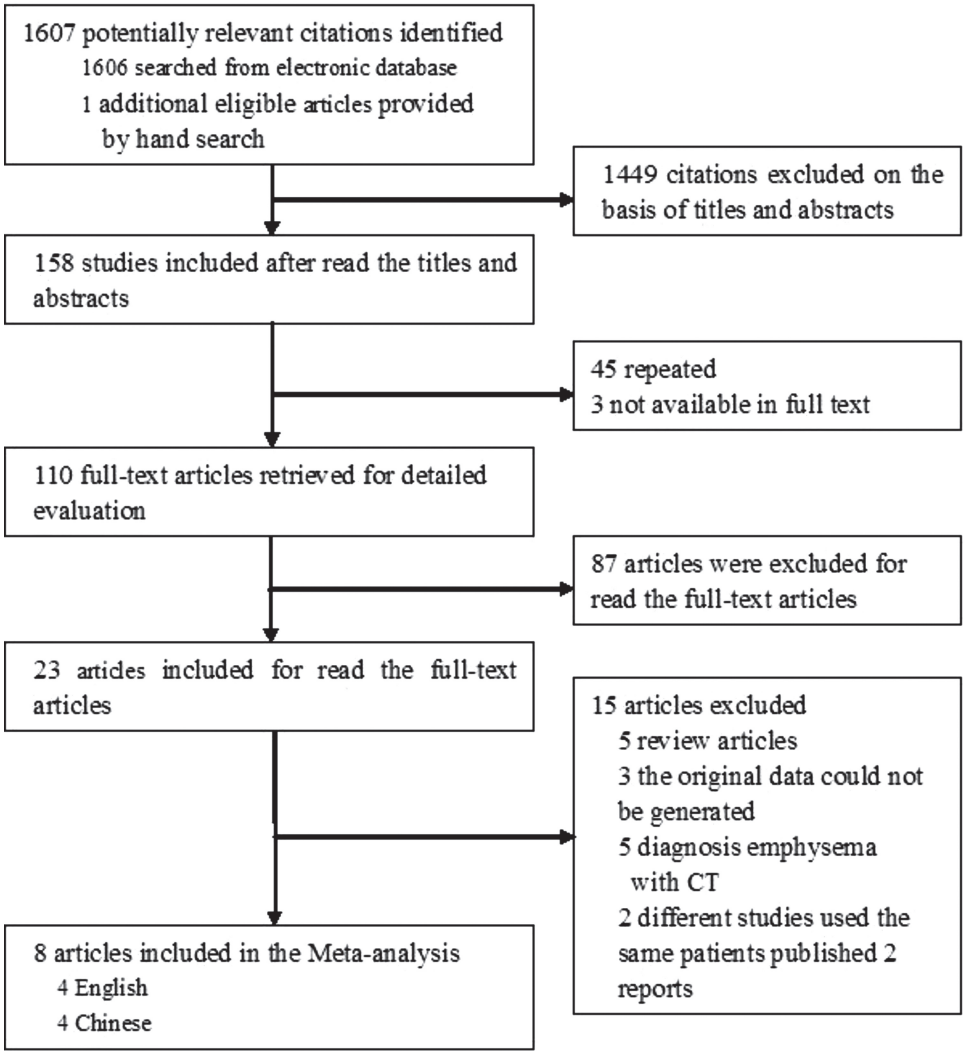

Figure 1. A flow chart shows the results of the literature search and selection for this systematic review.

Heterogeneity in meta-analyses refers to the degree of variability in results across studies (17). The primary causes of heterogeneity in test accuracy studies are threshold effect and sources other than threshold effect (18). In this study, we calculated the spearman correlation coefficient between the logit of sensitivity and logit of 1-specificity assessment for threshold effect, a strong positive correlation would suggest threshold effect; We tested for sources other than threshold effect amongst various studies in way of Chi-square tests, which are automatically implemented during analysis to evaluate if the differences across the studies are greater than expected by chance alone (19). A p-value less than 0.1 will suggest presence of heterogeneity beyond what could be expected by chance alone. If heterogeneity due to threshold effect were present, the accuracy data can be pooled by fitting a ROC curve and summarising that curve by means of the Area Under the Curve (AUC) (20). If there is heterogeneity due to sources other than threshold effect, potential between-study heterogeneity was explored by subgroup analyses (21, 22), then pooling should only be attempted within homogeneous subsets.

Heterogeneity in test accuracy between studies is likely to arise due to differences in patient characteristics, test methods, study design and other factors (23). This study were further stratified by type of imaging technique and test index, and tried to merge in Subgroup of homogeneous. The analyses were performed using Meta-Disc, Version 1.4 (19), STATA version 10 (Stata corporation, Texas), and SAS 9.2.

Results

Characteristics of included studies and quality assessment

The computer search yielded 1607 citations: 597 from PubMed, 466 from EMBASE, 2 from the Cochrane Library, 146 from CBM, 259 from CNKI, 116 from WANFANG and 20 from VIP, 1 additional eligible studies provided by hand search, of which 8 articles (24–31) ultimately were included in this review (Figure 1). Study characteristics are presented in Tables 1 and 2. Quality assessment of all included studies based on the updated QUADAS-2 is shown in Table 3. Overall, certain studies

Table 1. Characteristics of included studies

Study	Yearı	Cases (male/female)		Mean age (years)		Imaging technique	Test index	TP	FP	FN	TN	SEN	SPE
		COPD	Normal	COPD	Normal								
Li Kai et al. (24)	2008	40(23/17)	26(16/10)	63.5	61.5	16 MSCT	ALD ex	–	–	–	–	73.9%	93.7%
Chen Shu-jing. (25)	2009	23(23/0)	8(8/0)	67.8	51.8	LDCT	EI ex	–	–	–	–	91.3%	100%
Long Li-ling et al. (26)	2008	40(23/17)	26(16/10)	63.5	61.5	16 MSCT	PI ex	–	–	–	–	96.6%	100%
Miao Fei. (27)	2010	37(26/11)	32(23/9)	64.2	53.4	16 MSCT	BF	–	–	–	–	81.0%	75.0%
Kurashima K et al. (28)	2005	(462/54)		69.0		HRCT	LD	228	102	52	134	-	-
Tsushima K et al. (29)	2010	48(40/8)	2199(1319/880)	61.1	53.5	LDCT	LAA%	–	–	–	–	81.3%	87.5%
Marsh S et al. (30)	2007	22(19/3)	185(88/97)	64.7	54.4	Unclear	LD	–	–	–	–	83.3%	62.8%
Mets OM et al. (31)	2011	1140		63.5		LDCT	Diagnostic Model	274	85	163	618	–	–

TP = number of true positives; FP = number of false positives; FN = number of false negatives; TN = number of true negatives. MSCT=multislice computed tomography; ALD ex=the full expiration average lung density; LDCT=low-dose computed tomography; HRCT=high resolution computed tomography; EI ex=emphysema index in expiration; PI ex=pixel index in maximum expiratory; BF=blood flow; BV=blood volume; LAA%=the percentage of the low attenuation area; LD=lung density; Diagnostic Model=CT emphysema, CT air trapping, BMI, pack-years, and smoking status.

had some methodologic limitations. For 6 studies, the investigators explained that operators interpreted CT results with the results of PFT.

Diagnostic accuracy of CT

In this study, the spearman correlation coefficient between the logit of sensitivity and logit of 1-specificity was –0.238, the *P* value was 0.570, so there are no threshold effect in the study. The bivariate model directly provides summary estimates of (logit) sensitivity and specificity with corresponding 95% CI for the all of included studies (see Table 4) (Forest plots not shown). The SEN, SPE, PLR, NLR and DOR with associated 95% CI were 0.83 (0.73–0.89), 0.87(0.70-0.95), 6.2(2.5–15.5), 0.20(0.12–0.34), and 31(8-116), respectively. We used the visual power of the ROC plot to present the results of the bivariate model. Because the bivariate approach estimates the strength and the shape of the correlation between sensitivity and specificity, we can draw a 95% confidence ellipse and a 95% ellipse and a 95% prediction ellipse. These ellipse clearly show the area under the SROC (AUC) was 0.90 (0.87-0.93) (see Figure 2), indicating that the CT has a high discrimination ability for COPD. Table 4 presents the measures were grossly heterogeneous ($P < 0.1$). The heterogeneity in these studies was explored by subgroup analyses. Subgroup was divided into several groups by two factors (imaging technique and test index), the two factors seems strongly associated with the observed heterogeneity.

Subgroup analysis of CT index with or without the lung density was divided into two groups (Done, and Not done). Four of the studies evaluated with the lung density index (24, 28-30), the others evaluated different of index, such as, emphysema index in expiration (EIex) (25), pixel index in maximum expiratory (PIex) (26), blood flow (BF) (27) and Diagnostic Model (31). A more description of the test index can be found the Appendix. Sensitivity and specificity with corresponding 95% CI of the two groups divided by text index can be estimated using the bivariate model (see Table 5). We also test the difference of the sensitivity or specificity between the two groups using the bivariate model. The results show that the SEN, SPE, LR+, LR– and DOR with associated 95% CI of the lung density index were 0.80(0.74–0.84), 0.77(0.58–0.89), 3.5(1.8–6.9), 0.26 (0.20–0.34) and 13 (6–32), respectively.The SEN, SPE, LR+,LR- and DOR with associated 95% CI of others were 0.87(0.64–0.96), 0.95(0.66–0.99), 17.5(1.8–171.5), 0.14(0.04–0.45) and 127(5–2982). There was no statistically significant difference in the mean value of sensitivity between two groups ($P > 0.05$).

Because the number of the studies is less than 4 for each types of imaging technique: three of the studies evaluated the MSCT test (24, 26, 27), three the LDCT test (25, 29, 31), one the HRCT test (28), and one not clear (30), the analysis of the bivariate model can't be used. Table 6 presents the results of the summary ROC approach. The SEN, SPE, LR+, LR– and DOR with associated 95% CI of LDCT were 0.66(0.61–0.70), 0.88(0.86–0.89), 5.89(4.57–7.59), 0.25(0.12–0.53) and

Table 2. The number of true positives, false positives, false negatives and true negatives were calculated according to each sensitivity and specificity by Review Manager 5.1

Study	COPD	Normal	TP	FP	FN	TN
Li Kai et al. (24)	40	26	30	2	10	24
Chen Shu-jing. (25)	23	8	21	0	2	8
Long Li-ling et al. (26)	40	26	39	0	1	26
Miao Fei. (27)	37	32	30	8	7	24
Tsushima K et al. (29)	48	2199	39	275	9	1924
Marsh S et al. (30)	22	185	18	69	4	116

Table 3. QUADAS-2 results for studies performed with CT

Studies	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference Standard
Li Kai et al. (24)	↑	↑↑	↑	↑	↑	↑	↑
Chen Shu-jing. (25)	↑	↑↑	↑	↑	↑	↑	↑
Long Li-ling et al. (26)	↑	↑↑	↑	↑	↑	↑	↑
Miao Fei. (27)	↑	↑↑	↑	↑	↑	↑	↑
Kurashima K et al. (28)	↑	↑↑	↑	↑	↑	↑	↑
Tsushima K et al. (29)	↑	↑↑	?	↑	↑	↑	↑
Marsh S et al. (30)	↑	?	?	↑	↑	↑	↑
Mets OM et al. (31)	↑	↑	↑↑	↑	↑	↑	↑

↑= Low risk, ↑↑ = high risk, ? = unclear risk.

21.27(8.47–53.41), respectively. The summary measures of specificity and LR+ were very high and homogeneous. All other measures were highly heterogeneous. A meta-analysis was also not conducted because too much heterogeneity among these studies with MSCT ($P < 0.1$) and too few studies were identified with HRCT test (Table 6).

Publication Bias

There was no publication bias. The Egger test was not statistically significant ($P = 0.142$), with an symmetrical funnel plot (Figure 3).

Discussion

Early diagnosis and treatment for COPD are clearly desirable because of both the clinical and socioeconomic implications of the disease (32). However, early diagnosis sometimes can be difficult because the presentation of COPD usually is insidious, and many patients are undiagnosed until the disease is far advanced (32, 33). CT is also a diagnostic useful method for patients with COPD because it can easily quantitatively assess airway and parenchymal pathology (34). We performed this systematic review and meta-analysis to summarize the evidence on accuracy of CT for the diagnosis of COPD.

Exploring heterogeneity is a critical issue to understand the possible factors that influence accuracy estimates, and to evaluate the appropriateness of statistical pooling of accuracy estimates from various studies (19). An exploration of the reason for heterogeneity, rather than the computation of summary measures, is an important goal of meta-analyses (35, 36). In this study, the spearman correlation coefficient between the logit of sensitivity and logit of 1-specificity was –0.238, P value was 0.570, so not exist threshold effect.

The summary Rceiver Operating Characteristic (SROC) approach adopted the diagnostic odds ratio to compare the accuracy of diagnostic tests, and neglected different thresholds to define positive and negative test results, covered up the true diagnostic performance of the testing. The bivariate model with random effects method to estimate the sensitivity and specificity estimates and their respective 95% CI, it can also produce summary estimates of sensitivity and specificity, acknowledging any possible (negative) correlation between these two measures. This provides a very important information for the heterogeneity in the result of the study.

In this study, we analyzed the accuracy of the CT diagnosing COPD with the bivariate model. Our review shows that CT had high sensitivity and specificity. The more accurate the test, the closer of the curve approached

Table 4. Summary estimates for sensitivity, specificity, and diagnostic odds ratio for all studies from the bivariate model

Measure of all test accuracy	Pooled summary measure (95% CI)	Test for heterogeneity	
		Chi-squared value	p- value
Sensitivity	0.83(0.73, 0.89)		
Specificity	0.87(0.70, 0.95)		
Positive Likelihood Ratio(LR+)	6.2(2.5, 15.5)	111.970	<0.001
Negative likelihood Ratio(LR–)	0.20(0.12, 0.34)		
Diagnostic odds ratio(DOR)	31(8,116)		

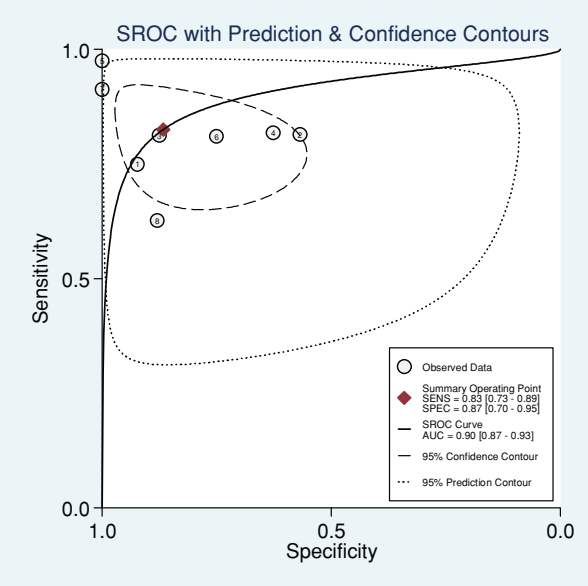

Figure 2. Bivariate summary estimates of sensitivity and specificity for all of the studies and the corresponding 95% confidence ellipse around these mean values.

the top left-hand corner of the graph. The value of the AUC closer to 1 for any test indicates that the test is more excellent. As seen in Figure 2, the SROC curve showed the value of AUC is 0.90 for all of the studies, indicating that CT scan has a better diagnostic capability. But after

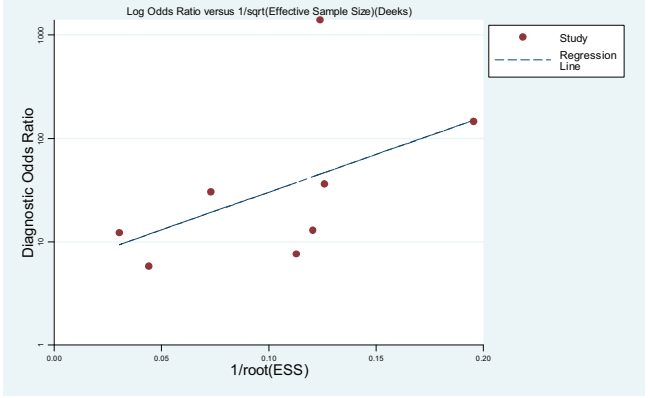

Figure 3. Funnel plot for evaluation of publication bias in all studies.

Chi-square test, *P* values of the measures were < 0.1, show that the heterogeneity caused by sources other than threshold effect. These reasons include chance as well as variations in study population (e.g., severity of disease and co-morbidities), index test differences in technology, assays, operator etc.), reference standard, and the way a study was designed and conducted (37). The subgroup analyses identified that account for some of the observed heterogeneity in our results.

One probable cause of the heterogeneity is that different imaging techniques were used in different studies.

Table 5. Summary estimates for sensitivity, specificity, and diagnostic odds ratio for different text index from the bivariate model

Imaging modality	sensitivity (95% CI)	specificity (95% CI)	LR+ (95% CI)	LR– (95% CI)	DOR (95% CI)
LD (N = 4)	0.80(0.74–0.84)	0.77(0.58–0.89)	3.5(1.8–6.9)	0.26(0.20–0.34)	13(6–32)
Others (N = 4)	0.87(0.64–0.96)	0.95(0.66–0.99)	17.5(1.8–171.5)	0.14(0.04–0.45)	127(5–2982)
P-value LD vs. Others	0.8653	0.2355	–	–	0.3100

Comparison between two test indexes.
Abbreviations: LD, Lung Density; Others refers to the 4 studies with the test index (emphysema index in expiration, pixel index in maximum expiratory, blood flow and Diagnostic Model).

Table 6. Summary estimates for accuracy measures for types of imaging technique from the sROC approach

Measure of test accuracy	Pooled summary measure (95% CI)	Test for heterogeneity	
		Chi-squared value	p- value
16 MSCT(N = 3)			
Sensitivity	0.85(0.77, 0.91)	10.23	0.006
Specificity	0.88(0.79, 0.94)	11.23	0.004
Positive Likelihood Ratio(LR+)	8.42(1.64, 43.23)	8.43	0.015
Negative likelihood Ratio(LR–)	0.18(0.07, 0.44)	6.96	0.031
Diagnostic odds ratio(DOR)	52.06(6.71, 403.95)	7.71	0.021
LDCT(N=3)			
Sensitivity	0.66(0.61, 0.70)	15.75	<0.001
Specificity	0.88(0.86, 0.89)	2.20	0.333
Positive Likelihood Ratio(LR+)	5.89(4.57, 7.59)	4.46	0.107
Negative likelihood Ratio(LR–)	0.25(0.12, 0.53)	10.28	0.006
Diagnostic odds ratio(DOR)	21.27(8.47, 53.41)	7.18	0.028

The SROC analysis showed that LDCT in diagnosing COPD have high specificity, low and widely varying sensitivities. These test properties suggest a potential role for LDCT in confirming the diagnosis of COPD, but unhelpful in ruling out it. The radiation dose from LDCT scan is limited (38, 39), LDCT is commonly performed to screen for lung cancer in high-risk subjects (40), so an additional benefit may be got in early COPD. The accuracy of MSCT was heterogeneous across studies, and thus meaningful summary measures of accuracy could not be determined. Another cause of the heterogeneity is that these studies with different type of test index. The results of the bivariate model showed that CT had a high sensitivity (80%) in diagnosing COPD with the lung density index. The lung density has significant correlations with PFT and it has been proven to be a good predictor of lung function (41, 42). These studies showed that the lung density has a potential role in diagnosing COPD.

Likelihood ratios are metrics that take into account the interaction between the SEN and the SPE in their calculation, and LR+>10 and LR−<0.1 are considered convincing evidence to rule in or rule out diagnoses, respectively, in most circumstances (13). In this study, the result of pooled LR+ is 6.2 < 10, suggested that the results of CT imaging were positive, there is the possibility of suffering from COPD; the result of pooled LR- is 0.20 > 0.1, indicated that the results of CT imaging were negative, the possible of suffering from COPD can not excluded. With advances in CT technology, new methods will undoubtedly be developed. CT with some quantitative CT indices may play an important role in diagnosing COPD at the early stage.

There are some limitations in this study. First, most of the studies in our review did not using blind method. This is known as review bias, and may lead to inflated measures of diagnostic accuracy. Second, different of detecting instrument, operation procedure, the quality control and operating rules in included studies could have influenced the accuracy of the results. Third, although we adopted widely of search strategy and aimed to retrieve additional data from investigators, and failed to find some missing and unpublished data, it is unavoidable that potential publication bias.

Conclusions

In conclusion, our results suggest that quantitative measures in CT scans may be useful to identify suspected subjects with COPD, although there was heterogeneity among these studies. Because the early stages of COPD are substantially under diagnosed, early detection of airflow limitation with chest CT and early intervention can improve outcomes for patients with COPD.

Declaration of Interest

Li Jian-sheng, Zhang Hai-long, Bai Yun-ping, Wang Yan-fang, Wang Hai-feng, Wang Ming-hang, Li Su-yun and Yu Xue-qing have no conflicts of interests. The authors alone are responsible for the content and writing of the paper.

Appendix: Definition of Terms Used in This Review.

1. Lung density: The potential of lung density measurements to diagnose emphysema in vivo was first noted by comparison of frequency distribution curves of density measurements between those with and without subsequent pathological diagnoses of emphysema. Density threshold techniques use a computer program to identify the percentage of total lung area occupied by areas of low attenuation (the radiological equivalent of tissue loss) below a predetermined value. The original study using this technique used a threshold of -910 Hounsfield Units (HU) whilst subsequent workers used -950(HU), both measurements showing a close correlation with pathological features of emphysema.
2. Blood flow (BF) refer to the rate of blood flow within tissues and organs, is one of the perfusion parameters, mainly related to tissue blood volume, draining veins, lymphatic drainage and tissue oxygen consumption factors.
3. Emphysema index (EI) was defined as the ratio of the emphysema volume (EV) in a range of threshold after 3D reconstruction to the total lung volume (TLV) at the definite width and level.
4. Pixel index (PI) was defined as the percentage of pixel in CT with an attenuation below -900 HU.
5. The Diagnostic Model included 5 factors independently associated with obstructive pulmonary disease: CT emphysema, CT air trapping, body mass index, pack-years, and smoking status. Computed tomographic emphysema was defined as the percentage of voxels in inspiratory CT with an attenuation below -950 HU. The HU value at the 15th percentile of the attenuation distribution curve as a measure of CT emphysema was calculated for secondary analysis. Computed tomographic air trapping was defined as the expiratory: inspiratory ratio of mean lung density.

References

1. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet 2007 Sep; 370(9589):765–73.
2. Calverley PM, Walker P. Chronic obstructive pulmonary disease. Lancet 2003 Sep; 362 (9389):1053–1061.
3. Slusarcick AL, McCaig LF. National Hospital Ambulatory Medical Care Survey: 1998 outpatient department summary. Adv Data 2000 Jul; (317):1–23.
4. Pratt PC, K ilburn KH . A modern concept of the emphysemas based on correlations of structure and function. Hum Pathol 1970 Sep; 1(3):443–63.
5. Macklem PT, Hogg JC, Thurlbeck WM. The flow resistance of central and peripheral airways in human lungs, Form and

Function in the Human Lung. Edited by G Cumming, LB Hunt, Baltimore Williams and Wilkins Company; 1968: 76–78.
6. de Jong PA, Muller NL, Pare PD, et al. Computed tomographic imaging of the airways: relationship to structure and function. Eur Respir J 2005 Jul; 26(1):140–52.
7. Hasegawa M, Nasuhara Y, Onodera Y, et al. Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2006 Jun; 173 (12):1309–15.
8. Nakano Y, Muro S, Sakai H, et al. Computed tomographic measurements of airway dimensions and emphysema in smokers. Correlation with lung function. Am J Respir Crit Care Med 2000 Sep; 162(3 Pt 1):1102–8.
9. Niimi A, Matsumoto H, Amitani R, et al. Airway wall thickness in asthma assessed by computed tomography. Relation to clinical indices. Am J Respir Crit Care Med 2000 Oct; 162 (4 Pt 1):1518–23.
10. Thurlbeck WM, Müller NL. Emphysema: definition, imaging, and quantification. AJR Am J Roentgenol 1994 Nov; 163(5):1017–25.
11. 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 Oct; 155(8):529–36.
12. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005 Oct; 58(10):982–90.
13. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994 Mar; 271(9):703–7.
14. Devillé WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. BMC Med Res Methodol 2002 Jul; 2:9.
15. Cochrane Methods Group on Systematic Review of Screening and Diagnostic Tests: Recommended Methods. Cochrane Collaboration 1996. Available from: http://www.cochrane.org/ Cochrane / sadtdoc1.htm.
16. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. Br Med J 2001 Jul; 323(7305):157–62.
17. Pai M, Flores LL, Hubbard A, et al. Nucleic acid amplification tests in the diagnosis of tuberculous pleuritis: a systematic review and meta-analysis. BMC Infect Dis 2004 Feb; 4:6.
18. Zhang T.-S., Zhong W-Z. Meta-DiSc Software in Meta-Analysis of Diagnostic Test. J Evidence-Based Med 2008 Apr; 8(2):97–100.
19. Zamora J, Abraira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006 Jul; 6:31.
20. Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. Stat Med 2002 May; 21(9):1237–56.
21. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med 2002 Jun; 21(11):1525–37.
22. Petitti DB. Approaches to heterogeneity in meta-analysis. Stat Med. 2001 Dec; 20 (23):3625–33.
23. Macaskill P, Gatsonis C, Deeks JJ, et al. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from: http://srdta.cochrane.org/.
24. Li K, Long Li-L, Huang Z-K. The preliminary study of multislice CT for lung density measurement in forecasting diagnosis of COPD. Journal of Clinical Radiology. 2008; 27(9):1226–1230.
25. Chen S-J. The economic burden of Chronic Obstructive Pulmonary Disease (COPD) in Shanghai and the feasibility of 3D reconstruction using Low-dose CT scan in diagnosis and classification of COPD. Shanghai: Zhongshan Hospital Fudan University, 2009 Apr: 36–58.
26. Long L-L, Li K, Huang Z-K. Preliminary Clinical Study of Pixel Index (PI) in Diagnosing COPD. J Clin Radiol 2008; 27(11):1479–1483.
27. Miao F. The Correlative Investigation of Function Test with First-pass Perfusion CT Indexes in Patients With Chronic Obstructive Pulmonary Disease. Ningxia: Ningxia Medical University, 2010 Apr:1–39.
28. Kurashima K, Takayanagi N, Sato N, et al. High resolution CT and bronchial reversibility test for diagnosing COPD. Respirology 2005 Jun; 10(3):316–22.
29. Tsushima K, Sone S, Fujimoto K, et al. Identification of occult parechymal disease such as emphysema or airway disease using screening computed tomography. COPD 2010 Apr; 7(2):117–25.
30. Marsh S, Aldington S, Williams MV, et al. Utility of lung density measurements in the diagnosis of emphysema. Respir Med 2007 Jul; 101(7):1512–20.
31. Mets OM, Buckens CF, Zanen P, et al. Identification of chronic obstructive pulmonary disease in lung cancer screening computed tomographic scans. JAMA 2011 Oct; 306(16):1775–81.
32. Bellamy D, Smith J. Role of primary care in early diagnosis and effective management of COPD. Int J Clin Pract 2007 Aug; 61(8):1380–9.
33. Man SF, McAlister FA, Anthonisen NR, et al. Contemporary management of chronic obstructive pulmonary disease: clinical applications. JAMA 2003 Nov; 290 (17):2313–6.
34. Akane Haruna, Toru Oga, Shigeo Muro, et al. Relationship between peripheral airway function and patient-reported outcomes in COPD: a cross-sectional study. BMC Pulm Med. 2010 Mar; 10:10.
35. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med 2002 Jun; 21(11):1525–37.
36. Petitti DB. Approaches to heterogeneity in meta-analysis. Stat Med 2001 Dec; 20(23):3625–33.
37. Dinnes J, Deeks J, Kirby J, et al. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. Health Technol Assess. 2005 Mar; 9(12):1–113.
38. Picozzi G, Paci E, Lopez Peqna A, et al. Screening of lung cancer with low dose spiral CT: results of a three year pilot study and design of the randomized controlled trial "Italung – CT". Radiol Med. 2005 Jan–Feb; 109(1–2):17–26.
39. Brenner DJ, Doll R, Goodhead DT, et al. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. Proc Natl Acad Sci USA 2003 Nov; 100(24): 13761–6.
40. Veronesi G, Bellomi M, Mulshine JL, et al. Lung cancer screening with low-dose computed tomography: a non-invasive diagnostic protocol for baseline lung nodules. Lung Cancer 2008 Sep; 61(3):340–9.
41. Akira M, Toyokawa K, Inoue Y, et al. Quantitative CT in chronic obstructive pulmonary disease:inspiratory and expiratory assessment. AJR Am J Roentgenol 2009 Jan;192(1):267–72.
42. Lee YK, Oh YM, Lee JH, et al. Quantitative assessment of emphysema, airtrapping, and airway thickening on computed tomography. Lung 2008 May-Jun; 186(3):157–65.