The diagnostic accuracy of CT based “Banner cloud sign” for dural ossification in patients with thoracic ossification of the ligamentum flavum: a prospective, blinded, diagnostic accuracy study protocol

Guanghui Chen
Peking University Third Hospital
Baoliang Zhang
Peking University Third Hospital
Liyuan Tao
Peking University Third Hospital
Zhongqiang Chen
Peking University Third Hospital
Chuiguo Sun (sunchuiguo@163.com)
Peking University Third Hospital

Study protocol

Keywords:

Posted Date: May 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-25406/v1

Abstract

Background

Ossification of the ligamentum flavum (OLF) is the most common cause of thoracic spinal stenosis, and responds poorly to conservative treatment, making surgery the only effective method. The existence of dural ossification (DO) makes surgery challenging and increases the risk of intra-/post-operative complications. To date, several methods have been proposed to identify DO, but either the diagnostic accuracy is low or the feasibility is poor. Therefore, the aim of this study is to propose a new imaging sign (Banner cloud sign, BCs), evaluate its diagnostic accuracy in the diagnosis of DO, and provide reliable evidence-based data for its application in clinical practice.

Methods

This will be a prospective, blinded, diagnostic accuracy study to assess and compare the accuracy of BCs in the diagnosis of DO with other radiologic signs (Tram track sign (TTs) and Comma sign (Cs)). A total of 120 patients who were diagnosed with OLF and underwent decompression in Peking University Third Hospital from January 2018 to June 2019 will be enrolled. Patients' medical records and imaging data will be retrieved from the hospital database server. An observational group consisting of six spine surgeons of different seniority and two epidemiological researchers will read patients' images to determine whether there are typical imaging signs and thus the presence of DO. Surgical records will be reviewed to confirm the presence of DO, and the results will serve as the reference standard for estimating accuracy. The primary outcome of the study is to determine the accuracy of BCs for DO diagnosis, and the secondary outcome is to compare the sensitivity, specificity, area under the ROC curve and inter-observer reliability of each imaging sign. The time taken and level of confidence in DO diagnosis of each observer will also be compared.

Discussion

This study represents the first large investigation of the diagnostic value of "BCs", "TTs" and "Cs" in the diagnosis of DO, and will provide convincing evidence about their clinical application.

Trial registration:

Registered on 29 February 2020. Trial number is ChiCTR2000030380

Introduction

Ossification of the ligamentum flavum (OLF) is the most common cause of thoracic spinal stenosis and especially reported in Eastern Asian countries with a prevalence rate of 3.8%-63.9%[1, 2]. It has progressive natural course and responds poorly to conservative treatment, making surgery the only effective option. Previous studies showed that decompressive surgery on the thoracic spine was of great challenge and the existence of dural ossification (DO) makes surgery more difficult and demanding[3–5]. Complications secondary to DO, such as neurological deterioration and cerebrospinal fluid (CSF) leakage, were frequently reported.

In our previous study, we found that DO was the leading cause of CSF leakage in OLF patients, with an incidence of 78.8% [6]. Complications related to CSF leakage, including pseudocyse, central nervous system infection, and wound dehiscence were not uncommon and might even lead to catastrophic events if these complications were not handled properly[6, 7]. Miyakoshi et al [8] reported that the incidence of dural tears and CSF leaks was higher in patients with DO and dural adhesions, which represented a deleterious factor for preoperative and short-term postoperative neurological status. Muthukumar et al [9] analyzed the prognostic implications of DO in OLF patients and concluded that DO was not only a detrimental factor for surgical complications but also for prognosis. Additionally, a systematic review by Nebiyu et al [5] found that dural tears were associated with longer hospital stay and considered as an independent driver of cost. Given the clinical and economic ramifications in regard to dural tears, there is of great interest for surgeons to accurately identify the presence of DO before operation and prepare adequately for the treatment of dural tear.

The incidence of DO in OLF patients was varied, ranging from 11–62%. Aizawa et al [10] reported that 9 of 72 OLF patients had dural tears and 8 of those had an DO. Miyakoshi et al [8] reported a higher incidence of dural adhesion in 34 patients (62%) and found that it was proportionately increased with the severity of OLF. In our single-institution, we have reported that the incidence of DO was ranged from 25.2–39% [11, 12], which was comparable to previous studies. These results indicated that DO was not a rare complication of OLF, and the diagnosis of DO should be considered in the management of OLF patients.

Currently, several methods have been used to diagnose DO, and each method has its own disadvantage in terms of diagnostic accuracy and feasibility. In 2009, Muthukumar et al [9] first described the radiologic characteristics of DO and proposed two types of radiologic signs: tram track sign (TTs) and comma sign (Cs). However, due to the relatively small number of patients in the study, the lack of understanding of these specific imaging signs makes it impossible to obtain an accurate diagnosis of DO. Sun et al [7] reported that the diagnostic specificity of TTs was only 59%, and the combination of other methods might be helpful to achieve the effective diagnosis. To improve the diagnostic accuracy of DO, Li et al [11] described a “bridge sign” (Bs), defined and excluded four kinds of false TTs, and then used TTs, Cs and Bs for combined diagnosis, with sensitivity and specificity of 94.23% and 94.21% respectively. Zhou et al [12] analyzed the risk factors of DO and proposed an imaging grading system to predicate DO, with a sensitivity of 76.0% and specificity of 91.0%. The results of these studies indicated that combined use of three or more imaging signs or parameters could significantly improve the diagnostic accuracy. In spite of this, no valuable imaging signs have been proposed and validated. Recently, Prasad et al [13] reported that MRI-T2 ring sign was highly correlated with intraoperative DO (90% sensitivity, 100% specificity) and was of benefit for surgeons in preoperative diagnosis of DO. However, this is a newly proposed MRI-based imaging sign, involving a limited number of patients, and its accuracy and feasibility need to be further studied.

As one of the largest research centers of thoracic OLF in China, we treated more than one hundred OLF patients every year. In clinical practice, we have found a new typical imaging sign, which is considered to be a specific sign of DO, namely "Banner cloud sign" (BCs). This sign is named because the morphology of DO on the sagittal CT reconstruction is similar to the natural landscape of the Banner Clouds on mountain peaks [14] (Fig. 1). Its effectiveness and accuracy in the diagnosis of DO have been preliminarily verified in our daily clinical practice. However, statistics are lacking to confirm this. Therefore, we designed this prospective, blinded and diagnostic accuracy study to evaluate the diagnostic accuracy of BCs, and compare it with that of TTs and Cs, and to explore the critical role of BCs in the diagnosis of DO.

Materials And Methods

Study design

This is a prospective, outcome assessor blinded, diagnostic accuracy study. This study will evaluate and compare the accuracy of BCs, TTs and Cs in the diagnosis of DO in a consecutive series of OLF patients in our single center. Patients' medical records and imaging data will be retrieved from the hospital database server. An observation group comprising of six spine surgeons of different seniority and two epidemiological researchers will read all patients' images to determine whether there are typical imaging signs and thus the presence of DO. After imaging evaluation, surgical records will be reviewed to confirm the presence of DO, and the results will serve as the reference standard for estimating accuracy. The study workflow are present in (fig.2).

Eligibility criteria

Inclusion criteria

After reviewing medical records and image data, patients who meet the following criteria are eligible.

- Patients with thoracic OLF and undergoing posterior decompression surgery
- The operation period is from January 2018 to June 2019
- Medical records and operation notes are complete, which can be used to determine whether there is DO
- Being willing to give informed consent

Exclusion criteria

The exclusion criteria are as follows:

- OLF patients with thoracic trauma, infection, tumor or deformity
- OLF combined with Diffuse idiopathic skeletal hyperostosis (DISH), Scheuermann's disease, Ankylosing spondylitis (AS), Skeletal fluorosis or severe rheumatism
- Unwilling to sign informed consent

Study population

Patients who were diagnosed with thoracic OLF and have received surgical treatment during the period of January 2018 and June 2019 in Peking University Third Hospital will be recruited. Two surgeons will identify eligible patients based on medical records and image data. Eligible patients will be informed about the study and invited to participate.

Interventions.

Two doctors, who are not observers, are responsible for image data collection, and then statisticians randomly number all cases through the Excel sheets. Before reading images, all observers are required to receive unified training. The principal investigator (PI) will elaborate the typical image features of "BCs", "TTs" and "Cs" in the form of PPT, so that all observers are familiar with the three typical signs mentioned above. Three times of training with an interval of one week will be performed. An observation group including six spine surgeons of different seniority (2 senior titles, 2 intermediate titles, 2 residents) and 2 epidemic researchers with no experience in spine surgery will read the images according to the imaging features of "BCs", "TTs" and "Cs" to determine whether there is DO.

Outcome measurements

Primary outcome measurements

Since the main purpose of this study is to determine the accuracy of BCs in the diagnosis of DO, the primary outcome measurements are to calculate the sensitivity, specificity, positive and negative predicative value of “BCs” in the diagnosis of DO, as well as the area under the ROC curve.

Secondary outcome measurements

To further evaluate its diagnostic value, we will calculate and compare it with "TTs" and "Cs" on the sensitivity, specificity, positive predictive value, negative predictive value, Youden's index, likelihood ratio and the area under the ROC cure. As the inter-observer reliability is a critical index to test inter-observer agreement, we will compare the intra-class correlation coefficient (ICC) and Kappa values of each imaging sign. Since an ideal diagnostic method should be universal and can be mastered by clinicians with different experiences, the time taken and level of confidence of using “BCs”, “TTs” and “Cs” for DO diagnosis will also be compared.

Blinding

For the purpose of quality control, strict blind strategy is necessary. In the reading process, all observers will be blinded to the image reports and any other medical records that can be used to identify the existence of DO.

Data collection and management

We will use electronic data capture system (EDC) for data collection and management. Each of the observers will be assigned a separate ID number as the unique identity to log into the EDC. After logging into the system, the observers will randomly receive the patients' image number. With this number, the observers can access the hospital picture and archiving system (PACS) to retrieve the patients' image data and read them. The reading results will be directly uploaded to the data management system. A brief description of data collection form is shown in (Fig 3). Two statistical professionals are responsible for supervising, collating and managing the collected data.

Reference standard

The surgical records will be used as the reference standard to determine the presence of DO.

Data analysis

Data analysis will primarily focus on the diagnostic accuracy of "BCs" in DO diagnosis. Sensitivity, specificity, positive and negative predicative value, positive and negative likelihood ratios will be calculated with corresponding 95% confidence intervals, by comparing the results of BCs, as read by 8 observers, with the reference diagnosis recorded in the surgical records.

To compare the diagnostic accuracy of different imaging signs, we will calculate the sensitivity, specificity, positive and negative predicative value, Youden index, likelihood ratio and area under the ROC curve of each sign in detecting DO.

To evaluate and compare inter-observer agreement, we will calculate Interclass correlation coefficients (ICCs) and Kappa values of each imaging sign in the diagnosis of DO.

To assess how easy it is to master different imaging signs, we will calculate the time it takes each observer to read the images. The level of confidence of each observer in identifying each imaging sign will be recorded using the five-point Likert scales (none, mild, moderate, very, extreme). Differences between each observer as well as each imaging sign will be tested for statistical significance.

Sample size calculation

We will calculate the sample size based on the area under the ROC curve. According to our preliminary experimental results, the AUC of BCs was 0.85. We assume that the AUC of "TTs" and "CS" as the control group is 0.65. In our previous studies, we have reported that the incidence of DO was ranged between 25.2% and 39% [6, 7]. Therefore, a conservative estimate incidence of 25% will be guaranteed, and the ratio of sample sizes in positive/negative groups is 1:3. With the type I error $\alpha=0.05$ and type II error $\beta=0.1$, all the above relevant date are input into PASS 14.0 software to calculate the sample size, and a study group of 96 patients (24 patients with DO and 72 without DO）is required. We assume that in a consecutive series of patients, 20 percent will not meet the inclusion criteria, so a total of 120 patients will be recruit.

Statistical methods

The research data will be managed by epidata 3.1 software. After the data are exported, SPSS 25.0 software is used for statistical analysis. The quantitative data in accordance with the normal distribution are statistically described by the mean ±standard deviation, while those that do not conform to the normal distribution are described by the median (25%, 75%), and the counting data are described by the number of cases (%). The comparison of quantitative data will be carried out by independent sample t-test, and the comparison of counting data will be analyzed by chi-square test or Fisher’s exact test. Two-sided *P* value of less than 0.05 is defined as statistically significant with two-sided 90% confidence intervals (CIs).

Quality control

For quality control, all observers are required to receive unified training prior to the clinical trial. The principal investigator (PI) will elaborate the typical image features of "BCs", "TTs" and "Cs" in the form of PPT, and make sure that all observers are familiar with the three typical signs mentioned above. Training will be conducted three times at an interval of one week. To minimize the interference between different imaging signs in the process of image reading, which may produce psychological cue effect and affect the judgment of the result, it is stipulated that observers can only judge one kind of image sign in each reading process. To guarantee the quality of the whole trial, rigorous monitoring will be performed by three trained quality supervisors. During the trial, supervisors will check the data entry of observers once a week to ensure that the data are true and valid.

Ethics

This study will be conducted in compliance with the principles of the Declaration of Helsinki for Clinical Research. This trial protocol was reviewed and approved by the Research Ethical Committee of Peking University Third Hospital. Informed consent will be obtained from all individual participants included in the study, and the protocol has been registered on Clinical Trials (Trial number is ChiCTR2000030380).

Discussion

Ossification of the ligamentum flavum (OLF) is the major cause of thoracic spinal stenosis and has been frequently reported in East Asian countries. Nowadays, with the accumulation of clinical experience and the improvement of diagnostic tools, this entity is being increasingly recognized in Caucasians. In view of the progressive natural course, OLF responds so poorly to conservative treatment that surgical intervention becomes the only effective method when the spinal cord is severely compressed and neurological impairment occurs. However, in spite of the advances in surgical techniques and instruments, surgery-related complication rate is still high, especially in the case of DO fused with OLF. The involvement of dura membrane in ossification makes surgery more difficult and significantly increases the risks of spinal cord damage and complications. Therefore, there is an urgent need for surgeons to make an accurate diagnosis of DO preoperatively and to make full preparations for the management of intraoperative dural laceration.

To date, there are limited methods to identify DO, and specific imaging signs based on CT or MRI evidence are the most commonly used. However, due to the lack of sample size and limited understanding of the characteristics of DO, the diagnostic accuracy of these existing image signs needs to be further improved. In addition, because of different experience of doctors and their mastery of typical image features, there is a large intra-group difference in the process of reading images, which may affect the accuracy and consistency of diagnosis results. Theoretically, an ideal diagnostic imaging sign is not only highly specific for DO, but also easier to master for surgeons of different seniority, so as to improve the diagnostic accuracy.

To the best of our knowledge, this is the first prospective study with a large number of participants to evaluate the diagnostic accuracy of this novel imaging sign and compare it with the diagnostic value of TTs and Cs. Our primary objective is to assess the diagnostic accuracy of BCs for DO by comparing observers' reading results with the reference standard, and to preliminarily understand its value in DO diagnosis. To further confirm its diagnostic value, we will compare BCs with TTs and Cs on a series of diagnostic indexes such as sensitivity, specificity, positive and negative predicative value, positive and negative likelihood ratios and area under the ROC curve. If these results are superior to TTs and Cs, then we can conclude that BCs has a higher diagnostic value and can replace previously reported diagnostic methods.

The optimal diagnostic method should not only have high diagnostic accuracy, but also be universal, which is easy for the observers with different experiences to grasp. Therefore, our secondary objective is to compare the inter-observation reliability of different methods by using ICC and Kappa values. Inter-observer reliability is determined by comparing the initial responses of all the eight observers. ICC and κ values will be interpreted as follows: 0.00 to 0.20 indicated slight agreement; 0.21 to 0.40 indicated fair agreement; 0.41 to 0.60 indicated moderate agreement; 0.61 to 0.80 indicated substantial agreement; and 0.81 to 1.00 indicated almost perfect agreement [15,16]. Additionally, to further evaluate the universality of each imaging sign to clinicians with different seniority, we will compare the time required to read images between inter-observers and between groups, as well as the level of confidence in diagnostic accuracy. If the results of these two indexes showed no statistics difference in different observers, then we can conclude that this imaging sign applies to all clinicians regardless of their experience. In contrast, if there are statistical difference in either of the time or the level confidence between different observers or imaging signs, it indicates that these imaging signs have different clinical application value, and the possible differences should be interpreted and analyzed according to the specific situation.

We hope that this study can provide adequate assessment of the diagnostic value of each imaging sign in the diagnosis of DO with the maximum number of patients. Quality control is crucial to the overall study. To perform a reliable study, we will carry out unified training to make sure all observers involved in this study fully understand characteristics of each imaging sign before the start of the study. The fact that all observers are blinded to the identity of the patients and the strict rule that allows the observer to read only one imaging sign at a time further ensure the quality of the study. Hopefully, this prospective diagnostic accuracy study will deepen clinicians' knowledge of the value of each imaging sign in diagnosing DO and provide reliable evidence-based data for their application in clinical practice. The novel imaging sign BCs will significantly improve the diagnostic accuracy of DO preoperatively, and will aid clinicians to make full preparations for the management of DO.

Abbreviations

OLF: Ossification of the ligamentum flavum, DO: dural ossification, TTs: Tram track sign, Cs: Comma sign, CSF: Cerebrospinal fluid leakage

Declarations

Ethics approval and consent to participate

The study design, procedure and informed consent procedure were approved by the Peking University Third Hospital. Consent to participate will be obtained from the participants.

Consent for publication

Not applicable

Competing interests

Not applicable

Disclosure of Funding

None

Authors' contributions

CGH and ZBL are co-first authors of this manuscript, contributing equally to the design, conduct of the trials, and drafting of the manuscript. CZQ and SCG have made substantial contributions to conception and design, and supervised the study. TLY participated in experimental design and statistical analysis. All authors read and approved the finial manuscript.

Acknowledgments:

Not applicable

References

1. JJ G, KD L. J K, et al. Prevalence, distribution, and morphology of ossification of the ligamentum flavum: a population study of one thousand seven hundred thirty-six magnetic resonance imaging scans. Spine. 2010;35:51–6.
2. HS NL, HL Y. W, et al. Epidemiological survey of ossification of the ligamentum flavum in thoracic spine: CT imaging observation of 993 cases. European spine journal: official publication of the European Spine Society the European Spinal Deformity Society the European Section of the Cervical Spine Research Society. 2013;22:857–62.
3. XQ C, HL Y, GL W, et al. Surgery for thoracic myelopathy caused by ossification of the ligamentum flavum. Journal of clinical neuroscience: official journal of the Neurosurgical Society of Australasia. 2009;16:1316–20.
4. DA P, SW K. SM L, et al. Symptomatic myelopathy caused by ossification of the yellow ligament. Korean Journal of Spine. 2012;9:348–51.
5. NS O, ZB C, AK H, et al. Outcomes and Complications Following Laminectomy Alone for Thoracic Myelopathy due to Ossified Ligamentum Flavum: A Systematic Review and Meta-Analysis. *Spine* 2018;43:E842-E8.
6. Sun X, Sun C, Liu X, et al. The Frequency and Treatment of Dural Tears and Cerebrospinal Fluid Leakage in 266 Patients With Thoracic Myelopathy Caused by Ossification of the Ligamentum Flavum. Spine. 2012;37:E702-E7.
7. Sun X-Z, Chen Z-Q, Qi Q, et al. Diagnosis and treatment of ossification of the ligamentum flavum associated with dural ossification Clinical article. J Neurosurg-Spine. 2011;15:386–92.
8. N M, Y S, T S, et al. Factors related to long-term outcome after decompressive surgery for ossification of the ligamentum flavum of the thoracic spine. *Journal of neurosurgery* 2003;99:251–6.
9. N M. Dural ossification in ossification of the ligamentum flavum: a preliminary report. *Spine* 2009;34:2654–61.
10. T A, T S. H S, et al. Thoracic myelopathy caused by ossification of the ligamentum flavum: clinical features and surgical results in the Japanese population. J Neurosurg Spine. 2006;5:514–9.
11. G BL. Q, S G, et al. Dural ossification associated with ossification of ligamentum flavum in the thoracic spine: a retrospective analysis. BMJ open. 2016;6:e013887.
12. Zhou SY, Yuan B, Chen X-s, et al. Imaging grading system for the diagnosis of dural ossification based on 102 segments of TOLF CT bone-window data. *Sci Rep*;7:2983.
13. GL P. Thoracic spine ossified ligamentum flavum: single-surgeon experience of fifteen cases and a new MRI finding for preoperative diagnosis of dural ossification. *British journal of neurosurgery* 2019:1–9.

14. Schween JH, Kuettner J, Reinert D, et al. Definition of "banner clouds" based on time lapse movies. Atmos Chem Phys. 2007;7:2047–55.
15. L Y, B L, Y Y, et al. The Relationship Between Dural Ossification and Spinal Stenosis in Thoracic Ossification of the Ligamentum Flavum. *The Journal of bone and joint surgery. American volume* 2019;101:606–12.
16. C X, M Y, Z S, et al. An Independent Inter-observer Reliability and Intra-observer Reproducibility Evaluation of SINS Scoring and Kostuik Classification Systems for Spinal Tumor. *World neurosurgery* 2020.

Figures

Figure 1

Definition of Banner Cloud sign. (A) Natural landscape of the Banner Cloud on Mount peak. (B) A typical imaging sign of dural ossification on sagittal CT reconstruction. (C) micro-CT scanning illustrates the morphology of dural ossification. (Yellow arrow indicates the ossified dural mater).

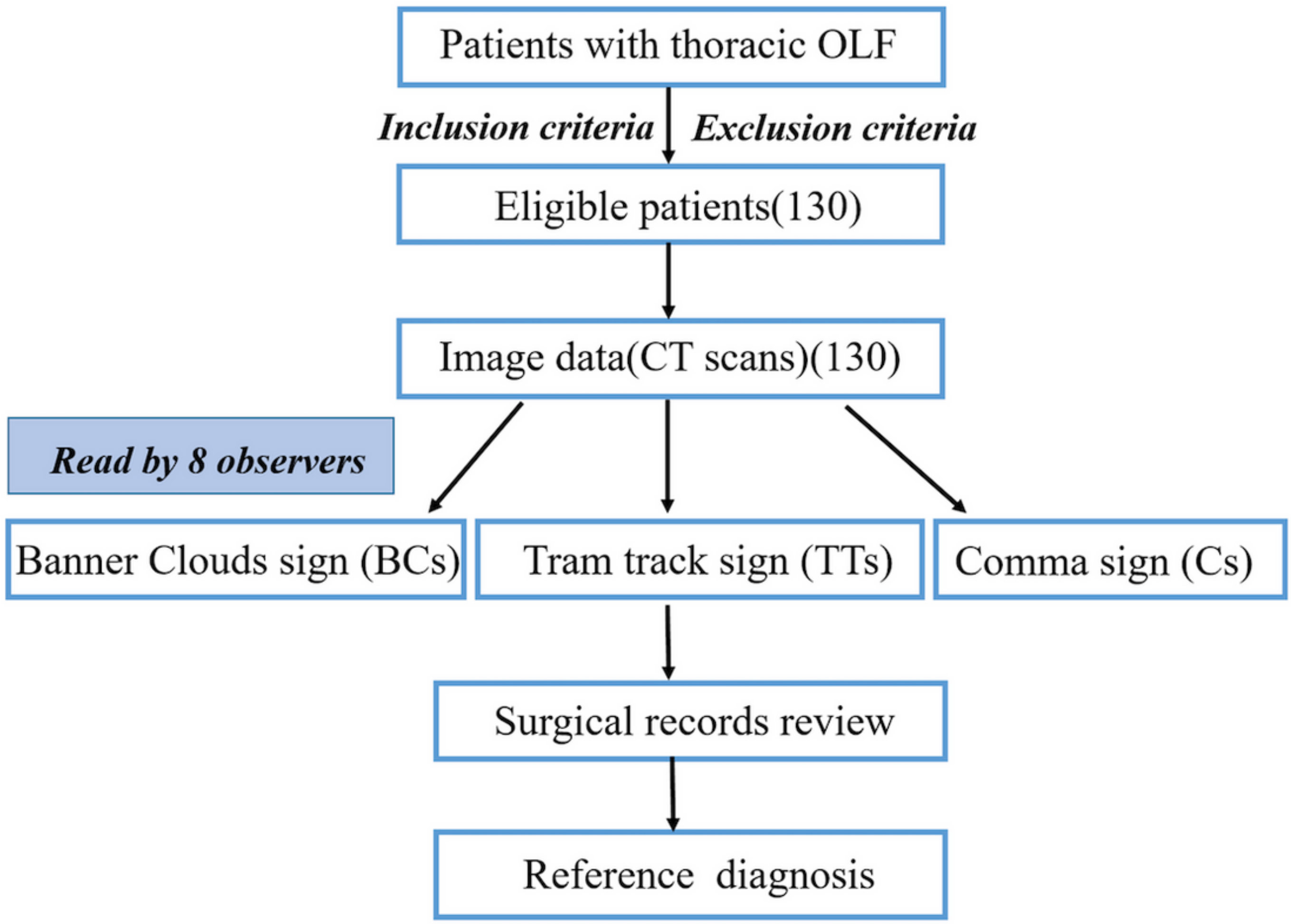

Figure 2

A study workflow

Observer ID: XXX											
Random No.		1	2	3	4	5	…	…	…	129	130
Image ID		**2418955**	xxxx	xxxx	xxxx	xxxx	…	…	…	xxxx	xxxx
Tram track sign	Yes										
	No										
	Time										
How certain are you about the diagnosis ?	1 Uncertain 2 Slightly certain 3 Moderately certain 4 Very certain 5 Extremely certain										
re-Random No.		1	2	3	4	5	…	…	…	129	130
Image ID		xxxx	xxxx	xxxx	xxxx	xxxx	…	**2418955**	…	xxxx	xxxx
Comma sign	Yes										
	No										
	Time										
How certain are you about the diagnosis?	1 Uncertain 2 Slightly certain 3 Moderately certain 4 Very certain 5 Extremely certain										
re-Random No.		1	2	3	4	5	…	…	…	129	130
Image ID		xxxx	xxxx	xxxx	xxxx	xxxx	…	…	…	**2418955**	xxxx
Banner clouds sign	Yes										
	No										
	Time										
How certain are you about the diagnosis?	1 Uncertain 2 Slightly certain 3 Moderately certain 4 Very certain 5 Extremely certain										

Figure 3

A brief description of data collection form