Epidemiology and association rules analysis for pulmonary tuberculosis cases with extrapulmonary tuberculosis from age and gender perspective: a large-scale multi-center observational study in China

Wanli Kang

Beijing Chest Hospital <https://orcid.org/0000-0002-3329-7096>

Jian Du

Beijing Chest hospital

Yunqing Chang

Beijing Chest Hospital

Hongyan Chen

Shenyang chest Hospital

Jianxiong Liu

Guang Zhou Chest Hospital

Jinshan Ma

Chest Hospital of Xinjiang

Mingwu Li

The Third People's Hospital of Kunming

Jingmin Qin

Shandong provincial Chest Hospital

Wei Shu

Beijing Chest Hospital

Peilan Zong

Jiangxi chest (third people) Hospital

Xiaofeng Yan

Chongqing Public Health Medical Center

Yi Zhang

Chang Chun Infectious Diseases Hospital

Yongkang Dong

Taiyuan Fourth People's Hospital

Zhiyi Yang

Fuzhou Pulmonary Hospital of Fujian

Zaoxian Mei

Tianjin Haihe Hospital

Qunyi Deng

Third People's Hospital of Shenzhen

Pu Wang

The First Affiliated Hospital of Chongqing Medical

Wenge Han

Weifang NO.2 People's Hospital

Meiying Wu

The Fifth People's Hospital of Suzhou

Ling Chen

Affiliated Hospital of Zunyi Medical College

Xinguo Zhao

The Fifth People's Hospital of Wuxi

Lei Tan

TB Hospital of Siping City

Fujian Li

Baoding hospital for Infectious disease

Chao Zheng

The First Affiliated Of XiaMen University

Hongwei Liu

Shenyang chest Hospital

Xinjie Li
Guang Zhou Chest Hospital

Ertai A
Chest Hospital of Xinjiang

Yingrong Du
The Third People's Hospital of Kunming

Fenglin Liu
Shandong provincial Chest Hospital

Song Yang
Chongqing Public Health Medical Center

Wenyu Cui
Chang Chun Infectious Diseases Hospital

Quanhong Wang
Taiyuan Fourth People's Hospital

Xiaohong Chen
Fuzhou Pulmonary Hospital of Fujian

Junfeng Han
Tianjin Haihe Hospital

Qingyao Xie
Third People's Hospital of Shenzhen

Yanmei Feng
The First Affiliated Hospital of Chongqing Medical

Wenyu Liu
Weifang NO.2 People's Hospital

Peijun Tang
The Fifth People's Hospital of Suzhou

Jianyong Zhang
Affiliated Hospital of Zunyi Medical College

Jian Zheng
The Fifth People's Hospital of Wuxi

Dawei Chen
Baoding hospital for Infectious disease

Xiangyang Yao
The First Affiliated Of XiaMen University

Tong Ren
Shenyang Chest Hospital

Yang Li
Guang Zhou Chest Hospital

Yuanyuan Li
Chest Hospital of Xinjiang

Lei Wu
The Third People's Hospital of Kunming

Qiang Song
Shandong provincial Chest Hospital

Mei Yang
Chongqing Public Health Medical Center

Jian Zhang
Chang Chun Infectious Diseases Hospital

Yuanyuan Liu
Tianjin Haihe Hospital

Shuliang Guo
The First Affiliated Hospital of Chongqing Medica

Kun Yan
Weifang NO.2 People's Hospital

Xinghua Shen
The Fifth People's Hospital of Suzhou

Dan Lei
Affiliated Hospital of Zunyi Medical College

Yanli Zhang

Baoding Hospital for Infectious Disease

Liang Li

Beijing Chest Hospital

Shenjie Tang (✉ tangsj1106@hotmail.com)

Beijing Chest Hospital <https://orcid.org/0000-0002-9131-4165>

Research Article

Keywords: pulmonary tuberculosis (PTB), extrapulmonary tuberculosis (EPTB), epidemiology, association

Posted Date: January 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-61565/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background Tuberculosis (TB), a multi-systemic disease with protean presentation, remains a major global health problem. Although concurrent pulmonary tuberculosis (PTB) and extrapulmonary tuberculosis (EPTB) cases are commonly observed clinically, knowledge regarding concurrent PTB-EPTB is limited. Here a large-scale multi-center observational study conducted in China aimed to study epidemiology of concurrent PTB-EPTB cases by diagnostically defining TB types then implementing association rules analysis.

Methods This study was conducted at 21 hospitals in 15 provinces in China and included all inpatients with confirmed TB diagnoses admitted from Jan 2011 to Dec 2017. Association rules analysis was conducted for cases with concurrent PTB and various types of EPTB using the Apriori algorithm.

Results Evaluation of 438,979 TB inpatients indicated PTB was the most commonly diagnosed (82.05%) followed by tuberculous pleurisy (23.62%). Concurrent PTB-EPTB was found in 129,422 cases (29.48%) of which tuberculous pleurisy was the most common concurrent EPTB type observed. Fully adjusted multivariable logistic regression models demonstrated that odds ratios of concurrent PTB-EPTB cases varied by gender and age group. For PTB cases with concurrent EPTB, the strongest association was found between PTB and concurrent bronchial tuberculosis (lift=1.09). For EPTB cases with concurrent PTB, the strongest association was found between pharyngeal/laryngeal tuberculosis and concurrent PTB (lift=1.11). Confidence and lift values of concurrent PTB-EPTB cases varied with gender and age.

Conclusions Numerous concurrent PTB-EPTB case types were observed, with confidence and lift values varying with gender and age. Clinicians should screen for concurrent PTB-EPTB in order to improve treatment outcomes.

Background

Tuberculosis (TB) is an infectious disease caused by the bacillus *Mycobacterium tuberculosis*. Despite eradication efforts, TB remains a major global health problem and cause of serious illness for millions of people each year. According to the World Health Organization (WHO), the estimated global incidence of TB was approximately 10.0 million cases in 2019 [1]. TB typically affects the lungs (pulmonary TB, PTB), but can also affect other sites (extrapulmonary TB, EPTB) that include pleura, lymph nodes, abdomen, genitourinary tract, skin, joints and bones, meninges, etc. [2-4].

In recent years, considerable effort has been expended to gain a deeper understanding of TB [5-7], a multi-systemic disease with a protean presentation. In fact, in clinical practice PTB and EPTB have been both detected in the same patient [8,9]. It was reported that about 10%-50% of EPTB patients have concomitant pulmonary involvement [8]. The effective treatment regimens for concurrent PTB-EPTB cases were shown to be more difficult to administer and different from effective regimens used for treating single PTB or EPTB cases [10]. As information pertaining to concurrent PTB-EPTB is scarce, research efforts are needed to diagnostically define TB case types and explore the epidemiology and association rules between PTB and different types of concurrent EPTB. In this work we addressed this gap in knowledge by conducting a large-scale multi-center observational study. The resulting acquired knowledge will be used to alert clinicians to the common occurrence of concurrent PTB-EPTB disease and improve treatment regimens used for such cases.

Methods

Study subjects

The study was performed at 21 Hospitals in 15 provinces in China. All inpatients with confirmed TB diagnoses from 2011 to Dec 2017 were included in the study. TB was mainly categorized by lesion site. Diagnosis of TB was made based on WHO guidelines [11] and the Clinical Diagnosis Standard of TB issued by the Chinese Medical Association [12]. In general, TB was diagnosed using both traditional and modern methods based on clinical symptoms and physical signs together with results obtained using bacteriological methods (including sputum smear microscopy, bacterial culture and molecular diagnostic methods), tuberculin skin testing (TST) or purified protein derivative (PPD) testing, X-ray findings, T-SPOT.TB and/or Gene Xpert MTB/RIF assay results, outcomes of treatment with a course of anti-tuberculosis chemotherapy, etc.

Data management and statistical analysis

Measures taken to guarantee data quality included both the implementation of a standardized study protocol and standardized training of research staff. Moreover, trained health workers collected medical information using a standardized questionnaire. Meanwhile, we obtained clinical characteristics of TB-afflicted inpatients (e.g., age, gender, site of disease, etc.) from medical records. Descriptive statistical analysis included frequencies and proportions with 95% confidence intervals (CIs) for categorical variables. Multivariable logistic regression analysis was used to examine the associations of gender and age group and determinations of odds ratios were conducted for concurrent PTB-EPTB cases. $P < 0.05$ was the threshold for statistical significance.

Association rules analysis, a technique used to discover relationships hidden in large databases, was used in this work. This technique, which was developed for the computer sciences, has been used in a variety of other fields [13-15]. The Apriori algorithm makes it possible to apply a set of association rules to data mining. The principle of Apriori is based on two steps: the first step searches for item sets that exceed the minimum support threshold value, while the second step generates association rules then filters them by selecting "confidence" item sets (based on a threshold) from item sets found in the first step [16,17]. For the association rule of A concurrent with B, support, confidence and lift were defined as: $\text{Support} = P(A)$, $\text{Confidence} = P(B|A)$, $\text{Lift} = P(A \cap B) / [P(A) * P(B)]$ whereby A

is antecedent and B is consequent. Lift was used to evaluate the magnitude of association rules whereby lift>1 indicates a positive association rule. Association rules for concurrent PTB and diverse types of EPTB were analyzed using the Apriori algorithm by setting the minimum support degree and the minimum confidence degree.

All data were collected in MS Office Excel (Microsoft, Redmond, WA, USA) datasheets and all analyses were conducted using SPSS software for Windows version 13 (Chicago, USA) and SPSS Modeler version 14.1 (IBM Corp, Armonk, NY, USA).

Results

TB patient characteristics

A total of 438,979 TB inpatients were included from Jan 2011 to Dec 2017 at 21 hospitals from 15 provinces in China, most of which were specialized tuberculosis hospitals. The ratio of male to female patients was 1.83. In these patients, 83 tuberculous lesion types were detected at 604,114 sites, with each patient harboring, on average, 1.38 TB lesion types. The most common types of TB were PTB (82.05%, 95%CI: 81.94%-82.16%), followed by tuberculous pleurisy (23.62%, 95%CI: 23.49%- 23.74%) and bronchial tuberculosis (7.01%, 95%CI: 6.94%-7.09%). Types of TB found in proportions of $\geq 0.1\%$ of total cases are shown in Table 1.

Patients with concurrent PTB-EPTB

A total of 129,422 (29.48%, 95%CI: 29.35%-29.62%) concurrent PTB-EPTB inpatients were found within the entire group of 438,979 TB inpatients. An age-gender percent-bar-graph of concurrent PTB-EPTB is shown in Figure S1 in the Supplementary Appendix. In <15, 15-24, 25-34, 35-44, 45-55 years age group, percents of concurrent PTB-EPTB female cases were greater than corresponding percents of male cases. In fully adjusted multivariable logistic regression models, patients who were female (OR=1.119, 95%CI: 1.104- 1.134), <15 years (OR=1.602, 95%CI: 1.521-1.686), 15-24 years (OR=1.831, 95%CI: 1.792-1.871), 25-34 years (OR=1.700, 95%CI: 1.664- 1.738), 35-44 years (OR=1.358, 1.327-1.391) and 45-54 years (OR=1.084, 95%CI: 1.059-1.109) were more likely to have concurrent PTB-EPTB, while those who were aged 55-64 years (OR=0.971, 95%CI: 0.949-0.994) were less likely to have concurrent PTB-EPTB (Table 2).

The most common concurrent PTB-EPTB types

According to association rules analysis of concurrent PTB-EPTB cases, the 20 most common concurrent PTB-EPTB case types are listed in Table S1 in the Supplementary Appendix, sorted by case number. Case numbers of concurrent PTB and tuberculous pleurisy (15.35%, 95%CI: 15.25%-15.46%) and concurrent PTB and bronchial tuberculosis (6.28%, 95%CI: 6.20%-6.35%) were greater than numbers of other types of concurrent PTB-EPTB cases.

Association rules analysis of concurrent PTB-EPTB cases

In order to find most of the possible association rules for Antecedent=PTB, the minimum confidence degree was set to 1.00%. After executing the association model, six association rules were obtained. The association rules are shown in Table 3 and were sorted by confidence. The first rule row (ID=1) in Table 3 was interpreted as demonstrating that for a total of 360,187 PTB cases (Instances), PTB accounted for 82.05% of all TB cases (Support), while PTB with concurrent tuberculous pleurisy accounted for 18.71% of PTB cases (Confidence). The confidence value for concurrent bronchial tuberculosis in PTB cases was next highest (7.65%), followed by tuberculous meningitis (2.72%). The strongest association rule for PTB with concurrent EPTB was found for PTB with concurrent bronchial tuberculosis (lift=1.09), whereby the lift value of 1.09 indicated that PTB was positively associated with bronchial tuberculosis.

In order to find most of the possible association rules for Consequent=PTB, the minimum support degree was set to 0.1% and the minimum confidence degree was set to 40%. After executing the association model, 22 association rules were obtained, including five rules with confidence values above 70%. The association rules shown in Table 4 are sorted by confidence. The first rule row (ID=1) in Table 4 was interpreted as showing that 2,382 cases (Instances) of pharyngeal/laryngeal tuberculosis accounted for 0.54% of all TB cases (Support), while pharyngeal/ laryngeal tuberculosis with concurrent PTB accounted for 91.23% of pharyngeal/ laryngeal tuberculosis cases (Confidence). The confidence value of concurrent PTB in bronchial tuberculosis cases was next highest (89.51%), followed by that of tuberculosis of mediastinal lymph nodes (77.57%). The strongest association rule for EPTB with concurrent PTB was found for pharyngeal/laryngeal tuberculosis (lift=1.11), indicating that pharyngeal/laryngeal tuberculosis was positively associated with PTB.

Association rules of concurrent PTB-EPTB types with gender

Most types of TB can be found in both males and females, with obvious exceptions of ovarian tuberculosis, oviduct tuberculosis, etc. We found association rules in males and females by setting the minimum support degree and the minimum confidence degree (Tables S2&3 in the Supplementary Appendix). In males, tuberculous empyema with concurrent PTB was the strongest association rule (lift=1.20), followed by costal tuberculosis with concurrent PTB (lift=1.16). In females, bronchial tuberculosis with concurrent PTB was the strongest association rule (lift=1.64), followed by supraclavicular lymph node tuberculosis with concurrent PTB (lift=1.56).

Association rules of concurrent PTB-EPTB types with age

We found association rules in all age groups by setting the minimum support degree and the minimum confidence degree (Tables S4~S10 in the Supplementary Appendix). In patients <15 years of age, tuberculous meningitis with concurrent PTB was the strongest association rule (lift=3.89), followed by tuberculosis of axillary lymph nodes with concurrent PTB (lift=3.78) and cervical vertebra tuberculosis with concurrent PTB (lift=3.61). In patients aged 15-24 years, splenic tuberculosis with concurrent PTB was the strongest association rule (lift=2.23), followed by tuberculous myelitis with concurrent PTB (lift=2.18). In patients aged 25-34, oviduct tuberculosis with concurrent PTB was the strongest association rule (lift=3.09), followed by endometrial tuberculosis with concurrent PTB (lift=2.17). In patients aged 35-44 years, the strongest association rule was again oviduct tuberculosis with concurrent PTB (lift=1.71), with endometrial tuberculosis with concurrent PTB (lift=1.69) ranked next. In patients aged 45-54 years, the strongest association rule was vocal cord tuberculosis with concurrent PTB (lift=1.61) that was followed by wrist joint tuberculosis with concurrent PTB (lift=1.57). In TB patients aged 55-64 years, the strongest association rule was ankle joint tuberculosis with concurrent PTB (lift=1.55), followed by adrenal tuberculosis with concurrent PTB (lift=1.51). In TB patients aged ≥65 years, the strongest association rule was tuberculous pericarditis with concurrent PTB (lift=1.42), with hilar lymph nodes with concurrent PTB (lift=1.30) ranked next.

Discussion

TB is transmitted when people afflicted with PTB expel *M. Tuberculosis* bacteria into the air; these airborne bacilli are subsequently breathed in by another person then settle in the lungs and grow. From there they can be disseminated throughout the lymphatic or hematogenous systems and subsequently infect single or multiple extrapulmonary sites, such as the pleura, lymph nodes, meninges, bones and joints, etc. Although PTB is the most common presentation of TB, EPTB contributes considerably to morbidity, lifelong sequelae, and mortality [18]. The mechanisms for EPTB dissemination are complicated [19]. Notably, cases of PTB concurrent with EPTB are common in clinical practice, but data regarding such cases are limited. Given the variety of clinical presentations and the nonspecific systemic symptoms of TB, a more profound understanding of the site distribution of TB is needed. In this study, we delineated the diagnostic types of TB observed in clinical practice and explored the association rules of concurrent PTB-EPTB cases in order to alert clinicians to the frequency of concurrent PTB-EPTB cases within a large TB sample population.

Tuberculous pleurisy is one of the most common forms of EPTB [20]. Here we found that tuberculous pleurisy (23.62%), thought to primarily result from a hypersensitivity reaction to tuberculous protein [21], was the second most common type of TB observed. Previous studies had also noted concurrent PTB-EPTB cases [8,9], as Boonsamgskul et al. [8] demonstrated that 12.2% of TB cases had concurrent PTB-EPTB (120/986). In this study, concurrent PTB-EPTB was found in about 30% in TB patients. We also found that the strongest association rule of PTB concurrent with EPTB was found for PTB concurrence with bronchial tuberculosis (lift=1.09); indeed, here 7.65% of PTB cases had concurrent bronchial tuberculosis. This association may stem from the fact that bronchi are adjacent to the lungs and this close proximity predisposes PTB patients to bronchial tuberculosis. However, this outcome is not inevitable, as the proportion of EPTB cases with concurrent PTB varied depending on whether patients were primarily viewed as EPTB or PTB patients. Meanwhile, pharyngeal/laryngeal tuberculosis was an infrequent manifestation of EPTB and was usually seen as a complication of PTB [22]. The strongest association rule of EPTB concurrent with PTB was pharyngeal/laryngeal tuberculosis concurrent with PTB (lift=1.11). In this study, 91.23% of pharyngeal/laryngeal tuberculosis patients had concurrent PTB, reflecting the fact that expulsion of *M. tuberculosis* by PTB patients would likely affect the larynx. Consistent with this supposition, most pharyngeal/laryngeal tuberculosis patients had concurrent PTB. These findings suggest that once PTB was diagnosed, patients should be evaluated for bronchial tuberculosis first. Conversely, once a diagnosis of pharyngeal/laryngeal tuberculosis or bronchial tuberculosis is confirmed, attention should be paid to evaluating the patient for PTB. In essence, the diagnosis of TB must be carefully undertaken to avoid misdiagnosis.

Although most types of tuberculous lesions can be found both in males and females, females (OR=1.119, 95%CI: 1.104-1.134) were more likely to have concurrent PTB-EPTB than males in this study. Jung et al. [23] also found female gender to be an independent predictor of concomitant endobronchial TB in patients with active PTB (OR=4.35, 95%CI: 1.78-10.63). Meanwhile, the magnitude of association rules of concurrent PTB-EPTB also varied with gender. For males in this study, tuberculous empyema with concurrent PTB was the strongest association rule (lift=1.20), with proportions of tuberculous empyema with concurrent PTB in males observed that exceeded 70%. In females, bronchial tuberculosis with concurrent PTB was the strongest association rule (lift=1.64), with a proportion of bronchial tuberculosis with concurrent PTB in females observed that was greater than 55%, although the difference of association rules was not clear, warranting additional investigation.

TB affects all age groups, but the overall best estimate for 2019 revealed that about 90% of cases were adults (aged ≥15 years), while TB prevalence has been shown to be strongly associated with age [1]. Here we found that the strongest association rule in children and adolescents (<15 years) was PTB with concurrent tuberculous meningitis (lift=3.89), a result that may reflect particular physiological characteristics and immunological mechanisms unique to children and adolescents. We also found that the strongest association rule in TB patients aged 15-24 years was splenic tuberculosis with concurrent PTB (lift=2.23). Splenic tuberculosis is mainly caused by hematogenous dissemination of a small number of bacteria that directly spread to the spleen via lymphatic pathways and adjacent organs [24]. Meanwhile, our results showed that oviduct tuberculosis with concurrent PTB was the strongest association rule in groups of patients aged 25-34 years (lift=3.09) and 35-44 years (lift=1.71). Indeed, oviduct tuberculosis is an important chronic pelvic disease and an etiological factor underlying infertility. Transmission of infection usually occurs via hematogenous systems, while direct spread from abdominal organs and

the peritoneum are also possible [25]. Other EPTB types including isolated vocal cord tuberculosis and ankle tuberculosis were rarely reported; however, vocal cord tuberculosis with concurrent PTB was the strongest association rule in patients aged 45-54 years (lift=1.61), while ankle joint tuberculosis with concurrent PTB was the strongest association rule in patients aged 55-64 years (lift=1.55). Yet another form of EPTB, tuberculous pericarditis, refers to a M. tuberculosis infection of the membrane that covers the heart (pericardium). In endemic TB areas, tuberculous pericarditis has been found in 1% to 2% of people afflicted with pulmonary TB [26]. We found that tuberculous pericarditis with concurrent PTB was the strongest association rule in ≥65-year-old patients (lift=1.42).

This study had several strengths, including its large-scale multi-center representative sample, its detailed analysis of diagnostic types of TB and its novel determinations of confidence/lift values of concurrent PTB-EPTB cases. However, the study also had several limitations. First, it may have been subject to Berkson's bias, since the study population included hospitalized TB patients of which concurrent PTB-EPTB cases had a high likelihood of being hospitalized causing the subsequent overestimation of their proportions within the general population. Therefore, data collected from whole population-based studies are still needed to clarify the associations. Secondly, most hospitals in our study were TB-specialized hospitals; thus, our findings may not represent the general TB patient population or apply to settings elsewhere in the country. Thirdly, some TB-specialized hospitals in China do not admit pediatric TB patients and therefore our results may underestimate proportions of pediatric TB cases. Lastly, the analysis did not consider low-frequency disease complications and comorbidities in China (e.g., HIV) and other factors such as new or retreatment, income, smoking, etc.

Conclusions

In conclusion, the present study revealed several types of concurrent PTB-EPTB cases and analyzed the association rules between PTB and EPTB for the first time in a large sample population. Concurrent PTB and tuberculous pleurisy was the most common concurrent PTB-EPTB case type observed. The strongest association rule in PTB with concurrent EPTB was PTB with concurrent bronchial tuberculosis. The strongest association rule in EPTB with concurrent PTB was pharyngeal/laryngeal tuberculosis with concurrent PTB. Confidence and lift values of concurrent PTB-EPTB cases varied with gender and age. Clinicians should be aware that concurrent PTB-EPTB cases are common and that these patients require administration of customized treatment regimens in order to achieve the best outcomes.

Declarations

Ethical Approval and Consent to participate: Given that the medical information of inpatients was recorded anonymously by case history, which would not bring any risk to the participants, the Ethics Committee of Beijing Chest Hospital, Capital Medical University approved this study with a waiver of informed consent from the patients.

Consent for publication: Not applicable.

Availability of supporting data: Data are not publicly available. However, de-identified data can be obtained on a reasonable request to correspondence.

Competing interests: The authors declare that they have no competing interests.

Funding: Supported by Key Project of Chinese National Programs (No.2015ZX10003001), Tongzhou District development support plan for high-level talent (No.YHLD2019035) and Tongzhou District Science and Technology Committee (No. KJ2017CX054).

Contributors: TSJ and LL designed the investigation. TSJ, LL, KWL, DJ, CYQ, CHY, LJX, MJS, LMW, QJM, SW, ZPL, YXF, ZY, DYK, ZY, MZX, DQY, WPHW, WMY, CL, ZXG, TL, LFJ, ZC, LHW, LXJ, AET, DYR, LFL, YS, CWY, WQH, CHX, HJF, XQY, FYM, LWY, TP and ZYL participated in the implementation of the investigation. TSJ, LL, KWL, DJ, CYQ, CHY, LJX, MJS, LMW, QJM, SW, ZPL, YXF, ZY, DYK, ZY, MZX, DQY, WPHW, WMY, CL, ZXG, TL, LFJ, ZC, LHW, LXJ, AET, DYR, LFL, YS, CWY, WQH, CHX, HJF, XQY, FYM, LWY, TP, ZJY, ZJ, CDW, YXY, RT, LY, LYY, WL, SQ, YM, ZJ, LYY, GSL, YK, SXH, LD and ZYL participated in data collection and data management. TSJ, LL, KWL, DJ, CYQ, CHY, LJX, MJS, LMW, QJM participated in data analysis and wrote the initial draft of the report and all authors revised and approved the final report.

Acknowledgments: We acknowledge the outstanding contributions from Innovation Alliance on Tuberculosis Diagnosis and Treatment (Beijing) and the technicians and nursing staff at 21 Hospitals in China. We also thank Douglas Lowrie of Shanghai Public Health Clinical Center for his invaluable help editing the manuscript.

Supplementary information: Supplementary information accompanies this paper.

Authors' information:

Wanli Kang, Jian Du, Yunqing Chang, Wei Shu, Liang Li, Shenjie Tang: Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China;

Hongyan Chen, Hongwei Liu, Tong Ren: Shenyang chest Hospital, No. 11 Beihai street, Dadong district, Shenyang 110044, China;

Jianxiong Liu, Xinjie Li, Yang Li: Guang Zhou Chest Hospital, No. 62, Heng Zhi Gang Road, Yuexiu District, Guangzhou, Guangdong 510095, China;

Jinshan Ma, Ertai A, Yuanyuan Li: Chest Hospital of Xinjiang, No. 106, Yan 'An Road, Tianshan District, Urumqi, Xinjiang 830049, China;

Mingwu Li, Yingrong Du, Lei Wu: The Third People's Hospital of Kunming, No. 319 Wu Jing Road, Kunming City, Yunnan Province 650041, China;

JingminQin,Fenglin Liu, Qiang Song: Shandong provincial Chest Hospital, No.12,Lieshishandong Road, Licheng District, Jinan, Shandong250000, China;

PeilanZong: Jiangxi chest (third people) Hospital, No.346 Dieshan Road, Donghu District, Nanchang City, Jiangxi Province330006, China;

Xiaofeng Yan,Song Yang,Mei Yang: Chongqing Public Health Medical Center, No.109,Baoyu Road, Geleshan Town, Shapingba District, Chongqing400036, China;

Yi Zhang,Wenyu Cui,Jian Zhang: Chang Chun Infectious Diseases Hospital, No. 2699, Sandao Section, Changji South Line, Erdao District, Changchun City, Jilin Province130123, China;

Yunqing Chang,Yongkang Dong, Quanhong Wang:Taiyuan Fourth People's Hospital, Number 231,Xikuang Street,WanBailin District, Taiyuan City, Shanxi Province 030024, China;

Zhiyi Yang,XiaohongChen:Fuzhou Pulmonary Hospital of Fujian, No.2, lakeside, Cangshan District, Fuzhou350008,China;

Zaoxian Mei,Junfeng Han,Yuanyuan Liu: Tianjin Haihe Hospital, Number 890, Shuanggangzhenjingu Road, Jinnan district, Tianjin City 300350, China;

Qunyi Deng,QingyaoXie:Third People's Hospital of Shenzhen, 29 Bulan Road, District Longgang, Shenzhen518112, China;

Pu Wang,Yanmei Feng,ShuliangGuo:The First Affiliated Hospital of Chongqing Medical University, No. 1 Youyi Road, Yuzhong District, Chongqing400016, China;

Wenge Han,Wenyu Liu,Kun Yan: Weifang NO.2 People's Hospital, NO.7th Yuanxiao Street Kuiwen District261041, China;

Meiying Wu,Peijun Tang,XinghuaShen:The Fifth People's Hospital of Suzhou, NO.10 Guangqian Road, Suzhou City Jiangsu Province215000,China;

Ling Chen,Jianyong Zhang,Dan Lei:Affiliated Hospital of Zunyi Medical College, NO.149 Delian Road, Zunyi,Guizhou563000, China;

Xinguo Zhao,Jian Zheng:The Fifth People's Hospital of Wuxi, No.1215, GuangRui Road, Wuxi214001, China;

Lei Tan: TB Hospital of Siping City, No. 10 Dongshan Road, Tiedong District, Siping City, Jilin Province136001, China;

Fujian Li,Dawei Chen,YanliZhang:Baoding hospital for Infectious disease, No. 608 Dongfeng East Road, Lianchi District, Baoding City, Hebei Province071000, China;

Chao Zheng,Xiangyang Yao: The First Affiliated Of XiaMen University, ZhenhaiRoud, Siming District, Xiamen City, Fujian Province, 361003 China.

References

1. World Health Organization. Global tuberculosis report 2020.Geneva: World Health Organization, 2020.
2. Xu JJ,PeerS, Papsin BC, Kitai I, PropstEJ.Tuberculous lymphadenitis of the head and neck in Canadian children: Experience from a low-burden region. *Int J PediatrOtorhinolaryngol*.2016; 91:11-14.
3. Bourgi K, Fiske C, Sterling TR. Tuberculosis Meningitis. *Curr Infect Dis Rep*. 2017;19:39.
4. Held MFG,HoppeS,LaubscherM, et al. Epidemiology of Musculoskeletal Tuberculosis in an Area with High Disease Prevalence. *Asian Spine J*.2017; 11:405-411.
5. du Preez K, du Plessis L,O'Connell N, Hesselning AC. Burden, spectrum and outcomes of children with tuberculosis diagnosed at a district-level hospital in South Africa. *Int J Tuberc Lung Dis*.2018;22:1037-1043.
6. Quattrocchi A, Barchitta M, Nobile CGA,etal.Determinants of patient and health system delay among Italian and foreign-born patients with pulmonary tuberculosis: a multicentre cross-sectional study.*BMJ Open*.2018;8:e019673.
7. Glaziou P, Floyd K, RaviglioneMC.Global Epidemiology of Tuberculosis. *SeminRespirCrit Care Med*.2018;39:271-285.
8. Boonsarngsuk V, Mangkang K, Santanirand P. Prevalence and risk factors of drug-resistant extrapulmonarytuberculosis.*ClinRespir J*.2018;12:2101-2109.
9. Sotgiu G, Falzon D, Hollo V,et al. Determinants of site of tuberculosis disease: An analysis of European surveillance data from 2003 to 2014.*PLoS One*. 2017;12: e0186499.
10. National Health Commission of the People's Republic of China Technical Specifications for Tuberculosis Prevention and Control in China, 2020 Edition,<http://tb.chinacdc.cn/ggl/202004/P020200414515703939844.pdf>.
11. WHO:Tuberculosis(TB). Available from:<https://www.who.int/tb/areas-of-work/laboratory/en/>
12. Chinese Medical Association. Clinical diagnosis standard of TB for clinical technology operation (TB volumes).People's Medical Publishing House; 2005.
13. Al-Turaiki I, Badr G, MathkourH.TrieAMD: a scalable and efficient Apriori motif discovery approach.*Int J Data Min Bioinform*.2015;13:13-30.
14. Tang JY, Chuang LY, Hsi E, Lin YD, Yang CH, Chang HW.Identifying the association rules between clinicopathologic factors and higher survival performance in operation-centric oral cancer patients using the Apriorialgorithm.*Biomed Res Int*.2013; 2013:359634.
15. Pattanaprateep O, McEvoy M, Attia J, Thakkinian A. Evaluation of rational nonsteroidal anti-inflammatory drugs and gastro-protective agents use; association rule data mining using outpatient prescription patterns.*BMC Med Inform DecisMak*.2017;17:96.

16. Ibrahim H, Saad A, Abdo A, SharafEldin A. Mining association patterns of drug-interactions using post marketing FDA's spontaneous reporting data. *Journal of Biomedical Informatics*.2016;60:294–308.
17. Available from: https://www.ibm.com/support/knowledgecenter/en/SS3RA7_15.0.0/com.ibm.spss.modeler.help/web_output_plottab.htm.
18. Holden IK, Lillebaek T, Andersen PH, Bjerrum S, Wejse C, Johansen IS. Extrapulmonary Tuberculosis in Denmark From 2009 to 2014; Characteristics and Predictors for Treatment Outcome. *Open Forum Infect Dis*. 2019;6:ofz388.
19. Qian X, Nguyen DT, Lyu J, Albers AE, Bi X, Graviss EA. Risk factors for extrapulmonary dissemination of tuberculosis and associated mortality during treatment for extrapulmonary tuberculosis. *Emerg Microbes Infect*.2018: 7:102.
20. Vorster MJ, Allwood BW, Diacon AH, Koegelenberg CF. Tuberculous pleural effusions: advances and controversies. *J Thorac Dis*.2015;7:981-991.
21. Amer S, Hefnawy AE, Wahab NA, Okasha H, Baz A. Evaluation of different laboratory methods for rapid diagnosis of tuberculous pleurisy. *Int J Mycobacteriol*.2016;5:437-445.
22. Cengiz A, Göksel S, Başal Y, TaşGülen Ş, Döğler F, Yürekli Y. Laryngeal Tuberculosis Mimicking Laryngeal Carcinoma on 18F-FDG PET/CT Imaging. *Mol Imaging Radionucl Ther*.2018;27: 81-83.
23. Jung SS, Park HS, Kim JO, Kim SY. Incidence and clinical predictors of endobronchial tuberculosis in patients with pulmonary tuberculosis. *Respirology*. 2015;20:488-495.
24. Gupta A. Splenic tuberculosis: a comprehensive review of literature. *Pol Przegl Chir*. 2018;90:49-51.
25. Sharma JB, Sharma E, Sharma S, Dharmendra S. Female genital tuberculosis: Revisited. *Indian J Med Res*. 2018;148(Suppl):S71-S83.
26. Gambhir S, Ravina M, Rangan K, et al. Imaging in extrapulmonary tuberculosis. *Int J Infect Dis*. 2017;56:237-247.

Tables

Table 1. The proportion of TB diagnosis (proportion $\geq 0.1\%$)

TB types	Frequency	Proportion(95%CI) (%)
pulmonary tuberculosis	360187	82.0511(81.9372-82.1645)
tuberculous pleurisy	103680	23.6184(23.4929-23.7444)
bronchial tuberculosis	30779	7.0115(6.9361-7.0874)
tuberculous meningitis	15711	3.5790(3.5242-3.6344)
tuberculous lymphadenitis of the neck	15282	3.4813(3.4272-3.5359)
tuberculous peritonitis	10059	2.2915(2.2474-2.3362)
tuberculous empyema	7341	1.6723(1.6346-1.7107)
lumbar vertebra tuberculosis	7190	1.6379(1.6006-1.6759)
tuberculous pericarditis	5842	1.3308(1.2971-1.3652)
thoracic vertebra tuberculosis	5317	1.2112(1.1791-1.2440)
tuberculous polyserositis	4870	1.1094(1.0786-1.1408)
intestinal Tuberculosis	4711	1.0732(1.0429-1.1041)
chest wall tuberculosis	4639	1.0568(1.0267-1.0875)
tuberculosis of mediastinal lymph nodes	3482	0.7932(0.7672-0.8199)
renal tuberculosis	2793	0.6362(0.6129-0.6602)
pharyngeal and laryngeal tuberculosis	2382	0.5426(0.5221-0.5648)
pelvic tuberculosis	1835	0.4180(0.3991-0.4376)
tuberculosis of axillary lymph nodes	1521	0.3465(0.3293-0.3643)
knee joint tuberculosis	1179	0.2686(0.2535-0.2843)
tuberculosis of hilar lymph nodes	1117	0.2545(0.2398-0.2698)
hip joint tuberculosis	989	0.2253(0.2115-0.2398)
pleural tuberculoma	917	0.2089(0.1956-0.2229)
cutaneous tuberculosis	871	0.1984(0.1855-0.2120)
epididymal tuberculosis	794	0.1809(0.1685-0.1939)
tuberculosis of abdominal lymph nodes	651	0.1483(0.1371-0.1601)
sacroiliac joint tuberculosis	574	0.1308(0.1203-0.1419)
tuberculous abscess of psoas major	528	0.1203(0.1102-0.1310)
cervical vertebra tuberculosis	513	0.1169(0.1070-0.1274)
testicular tuberculosis	495	0.1128(0.1031-0.1231)

Note: proportion=frequency*100%/438979

Table 2 The associations between PTB concurrent with EPTB and gender, age group

Characteristics	No. of PTB concurrent with EPTB (%)	aOR (95%CI)	P
gender			
female	49209(31.7)	1.119(1.104 -1.134)	<0.001
male	80213(28.3)	Reference	
Age group(years)			
<15	2417(34.2)	1.602(1.521-1.686)	<0.001
15-24	27094(37.2)	1.831(1.792-1.871)	<0.001
25-34	26806(35.6)	1.700(1.664-1.738)	<0.001
35-44	17193(30.5)	1.358(1.327 -1.391)	<0.001
45-54	18154(25.8)	1.084(1.059 -1.109)	<0.001
55-64	16835(23.7)	0.971(0.949 -0.994)	0.014
≥65	20923(24.3)	Reference	

Table 3. The association rules of PTB with concurrent EPTB where Antecedent = PTB, Minconfidence = 1.00%						
Consequent	Antecedent	ID	Instances	Support (%)	Confidence (%)	Lift
tuberculous pleurisy	pulmonary tuberculosis	1	360187	82.05	18.71	0.79
bronchial tuberculosis	pulmonary tuberculosis	2	360187	82.05	7.65	1.09*
tuberculous meningitis	pulmonary tuberculosis	3	360187	82.05	2.72	0.76
tuberculous lymphadenitis of neck	pulmonary tuberculosis	4	360187	82.05	1.93	0.56
tuberculous peritonitis	pulmonary tuberculosis	5	360187	82.05	1.59	0.69
tuberculous empyema	pulmonary tuberculosis	6	360187	82.05	1.05	0.63

Table 4. The association rules of EPTB with concurrent PTB where Consequent=PTB, Minsupport=0.1%, Minconfidence = 40%						
Consequent	Antecedent	ID	Instances	Support (%)	Confidence (%)	Lift
PTB	pharyngeal/laryngeal tuberculosis	1	2382	0.54	91.23	1.11*
PTB	bronchial tuberculosis	2	30779	7.01	89.51	1.09*
PTB	tuberculosis of mediastinal lymph nodes	3	3482	0.79	77.57	0.95
PTB	tuberculosis of hilar lymph nodes	4	1117	0.25	74.93	0.91
PTB	intestinal tuberculosis	5	4711	1.07	71.75	0.87
PTB	tuberculosis of abdominal lymph nodes	6	651	0.15	66.36	0.81
PTB	tuberculous pleurisy	7	103680	23.62	65.01	0.79
PTB	tuberculous meningitis	8	15711	3.58	62.43	0.76
PTB	tuberculous polyserositis	9	4870	1.11	59.22	0.72
PTB	tuberculous pericarditis	10	5842	1.33	58.18	0.71
PTB	tuberculous peritonitis	11	10059	2.29	56.99	0.69
PTB	testicular tuberculosis	12	495	0.11	54.75	0.67
PTB	tuberculous empyema	13	7341	1.67	51.53	0.63
PTB	sacroiliac joint tuberculosis	14	574	0.13	50.87	0.62
PTB	chest wall tuberculosis	15	4639	1.06	48.35	0.59
PTB	thoracic vertebra tuberculosis	16	5317	1.21	46.00	0.56
PTB	tuberculous lymphadenitis of the neck	17	15282	3.48	45.58	0.56
PTB	pleural tuberculoma	18	917	0.21	45.37	0.55
PTB	epididymal tuberculosis	19	794	0.18	45.21	0.55
PTB	cutaneous tuberculosis	20	871	0.20	45.01	0.55
PTB	renal tuberculosis	21	2793	0.64	43.04	0.52
PTB	pelvic tuberculosis	22	1835	0.42	42.78	0.52

Notes: PTB:pulmonary tuberculosis.

The first column represents the consequents (the "then" part of the rule), while the next column represents the antecedents (the "if" part of the rule).

ID displays the sequence of the association rules.

Instances display the cases of TB.

*lift>1

Figures

Figure 1

Geographical distribution of TB inpatient enrolled in this study (*) Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFile.docx](#)
- [GraphicalAbstract.jpg](#)