Computer-assisted detection (CAD) methodology for early detection of response to pharmaceutical therapy in tuberculosis patients

Robert Lieberman[a], **Heston Kwong**[b], **Brent Liu**[a], and **HK Huang**[a]

[a]Image Processing and Informatics Laboratory, 1450 San Pablo St., Suite 2100, Los Angeles, CA, USA 90033

[b]Principal Medical Officer, Department of Health Hong Kong Special Administrative Government

Abstract

The chest x-ray radiological features of tuberculosis patients are well documented, and the radiological features that change in response to successful pharmaceutical therapy can be followed with longitudinal studies over time. The patients can also be classified as either responsive or resistant to pharmaceutical therapy based on clinical improvement. We have retrospectively collected time series chest x-ray images of 200 patients diagnosed with tuberculosis receiving the standard pharmaceutical treatment. Computer algorithms can be created to utilize image texture features to assess the temporal changes in the chest x-rays of the tuberculosis patients. This methodology provides a framework for a computer-assisted detection (CAD) system that may provide physicians with the ability to detect poor treatment response earlier in pharmaceutical therapy. Early detection allows physicians to respond with more timely treatment alternatives and improved outcomes. Such a system has the potential to increase treatment efficacy for millions of patients each year.

Keywords

CAD Development; Lung; Directly Observed Treatment Short Course (DOTS); Tuberculosis; Chest Radiographs; Outcomes

1. INTRODUCTION

With a prevalence of one-third of the world's population, tuberculosis (TB) is the leading cause of infectious disease in the world [1]. Eight million people annually become infected with TB, and two million people die each year from disease caused by TB. According to the World Health Organization (WHO), disease burden caused by TB disproportionately affects developing countries where many essential medical services are limited or non-existent, especially China [2].

TB is caused by Mycobacterium tuberculosis, a bacterium that spreads via inhalation. Although 90% of people infected with mycobacterium tuberculosis are asymptomatic (referred to as latent TB infection), 10% of people infected with the bacterium go on to develop tuberculosis disease in their lifetime. Although it can attack the central nervous

system, the lymphatic system, the circulatory system, the genitourinary system, the gastrointestinal system, bones, joints, and even the skin, tuberculosis primarily infects the lung tissue. Such lung infection with mycobacterium tuberculosis is referred to as pulmonary TB.

The definitive diagnosis of TB requires the positive identification of cultured mycobacterium tuberculosis in a clinical sample, commonly of patient sputum. However, when the culture method of diagnosis is not possible, a probable diagnosis can be made using a tuberculin skin test or using imaging, most commonly chest radiography [3]. Furthermore, the sensitivity of the particular strain of mycobacterium tuberculosis to the various anti-TB drugs can be determined by checking for growth of the particular bacterial strain on agar plates prepared with the anti-TB medications in question.

1.1 Radiologic Features of TB Infection

Given that TB primarily infects the lung tissue, chest radiography has become the primary modality used around the world to aid in the diagnosis and management of TB patients. The general pathology of TB infection is often described as exudative, proliferative, caseating, and often calcifying. Radiographically, the initial parenchymal focus of TB presents as granulomatous tissue or mature fibrous tissue appearing as a lung nodule. Lymph node enlargement is one of the most common presenting abnormalities on chest radiograph, most commonly unilateral and in the hilum or paratracheal region. Airspace consolidation is also common and is usually unilateral. Pleural effusion is often found unilaterally on the same side as the primary focus of TB. Other common radiographic findings include but are not limited to focal or patchy heterogeneous consolidation of the apical and posterior segments of the upper lobes and the superior segments of the lower lobes, as well as poorly defined nodules and linear opacities, cavitary lesions, and tuberculomas [4].

1.2 TB Treatment: Directly Observed Therapy — Short Course

The standard pharmaceutical treatment for tuberculosis infection involves directly observed therapy [5]. Directly observed therapy treatment for TB was created because, although TB is curable in almost all cases, most experts agree that patients' inability to adhere to sufficient uninterrupted therapy has prevented the total elimination of TB infection in the world. Directly observed therapy treatment requires the observation of the patient by a health care provider or other responsible person as the patient ingests anti-TB medications [5]. In Hong Kong, the directly observed treatment — short course (DOTS) is the preferred form of TB treatment. Patients undergo treatment with Isoniazid, Rifampicin, Pyrazinamide and either Ethambutol or Streptomycin in the initial 2 months, followed by two drugs, Isoniazid and Rifampicin, in the subsequent four months. Timely radiological evaluation, using a time series of chest x-rays, is useful in following the progression or resolution of TB disease in patients undergoing DOTS therapy.

1.3 Drug Resistant Tuberculosis

Even with the implementation of DOTS treatment for TB, drug resistance is common. Various isolates of mycobacterium tuberculosis exist that are resistant to one or more of the commonly used anti-TB drugs. Disease manifested as a result of infection with one of the

resistant TB strains is referred to as drug-resistant tuberculosis. When the particular TB strain is resistant to at least Isoniazid and Rifampicin, it is more specifically referred to as Multidrug-Resistant Tuberculosis (MDR-TB). It is estimated that China and India together represent approximately 50% of the total number of MDR-TB cases in the world [6].

Although drug-resistance can be diagnosed during the initial sputum culture, they are often inaccurate and take six to eight weeks for drug sensitivity results to be determined. By the time the culture and sensitivity results are known, the patient has most likely been taking DOTS therapy for almost two full months or almost one-third of the total duration of DOTS therapy. Furthermore, the medical supplies needed to perform culture and sensitivity tests on patients being tested for TB are not available at all medical facilities. More commonly, TB patients are not diagnosed with drug-resistant tuberculosis or MDR-TB until it is recognized that the patient's symptoms continue to worsen while undergoing DOTS therapy.

In the context of recognizing drug resistance based on patient response to DOTS therapy, chest radiography in conjunction with clinical, laboratory, and microbiological parameters play an invaluable role in the monitoring of TB patients undergoing DOTS therapy. Because many of the sophisticated medical supplies required to monitor the laboratory and microbiological parameters of TB patients are not readily available to many laboratories throughout the world [6], chest radiography can be used to monitor changes in the radiographic findings mentioned in section 1.1 of patients undergoing DOTS therapy. The radiographic changes alert a physician to the possibility of treatment failure. However, patients responding to DOTS therapy may not show favorable signs of improvement until four to six weeks following the initiation of treatment, and treatment failure may not be recognized for a significant period of time beyond this window. By the time treatment failure is recognized, the disease progression has occurred and valuable resources may have been wasted.

Therefore, timely and effective treatment of TB minimizes disease progression, decreases the contagiousness of a patient, and helps control the spread of the disease. However, the prevalence of drug-resistance and DOTS treatment failure prevent the timely and effective treatment of TB and lead to poor outcomes. The chest radiographs of TB patients undergoing DOTS treatment will be used to develop a digital image analysis algorithm that utilizes subtraction images created from the original chest radiographs. Analysis of the subtraction images will be done in order to determine which statistical parameters or image analysis correlates with a patient's response to DOTS therapy. Thresholds will be determined based on the patient subtraction images in order to assist in the early establishment of successful TB treatment response versus TB treatment resistance in future TB patients. A CAD system will be created utilizing the chest radiograph algorithm in order to provide early detection of treatment failure to physicians. With successful results, the CAD system may improve TB outcomes of millions of patients worldwide.

2. METHODS

2.1 Patient Selection

We selected 200 patients from TB clinics of the Department of Health in Hong Kong in the year 2004. The diagnosis of TB infection was determined by chest physicians specializing in TB management and confirmed with laboratory testing. Sensitivity tests were used to determine the drug sensitivity of the strain of mycobacterium tuberculosis in each patient.

2.2 TB Treatment Protocol and Image Collection

The patients underwent the six month DOTS treatment regimen for TB. Based on the sensitivity results, the appropriate anti-TB regimen was chosen. Typically, a patient took Isoniazid, Rifampicin, Pyrazinamide, and Ethambutol for two months. The patients subsequently took Isoniazid and Rifampicin for the remaining four months of treatment. During DOTS therapy, six longitudinal PA chest radiographs of each patient were taken and scanned into DICOM files. The first chest radiograph was taken at the onset of DOTS therapy (Fig. 1) and the subsequent radiographs were taken over the six month duration of the DOTS treatment. The number of days between the first chest radiograph and each subsequent radiograph was recorded.

2.3 Chest Radiograph Subtraction Image Creation

As the patient responds to DOTS therapy, the pulmonary disease improves. Consequently, the chest radiographs show improvement in the lung disease over time in those patients that respond to treatment. As the pulmonary disease continues to improve towards the end of DOTS therapy, the improvement in the lung disease becomes more apparent. Subtraction images between the chest radiographs were created to emphasize the changes occurring in the level of pulmonary disease over time.

The subtraction images were created using the six original DICOM chest radiographs mentioned in the previous section. All the DICOM chest radiographs were converted to a jpeg format. The jpeg images were then used to create the chest subtraction images. The initial chest radiograph taken at the onset of DOTS treatment was used as the reference point that all subsequent chest radiographs were compared to. The six chest radiographs were then used to create five subtraction images by subtracting the second image from the first, the third image from the first, the fourth image from the first and so forth.

The subtraction images were created in the following manner. First, the lung fields of each chest radiograph were isolated. Lung field isolation was done through several steps. Because the lung fields and background have densities and consequently pixel values lower than other tissues, the lung tissue was merged with the spine and other high density tissues to create a mask that allowed for the empty background to be removed. Once the background was removed, the next step involved isolating the lung tissue from the other tissues in the chest radiograph, such as the heart, spine, ribs, etc. Lung tissue isolation was done using the run length coding method that compares each pixel value to either eight or four neighboring pixel values. This run length coding method takes advantage of the fact that the lungs are fairly homogeneous organs in terms of tissue density. Consequently, pixels representing

lung tissue in the chest radiograph have pixel values that are almost identical to neighboring lung tissue pixel values. A random pixel, called a seed, was chosen within one of the lung fields. The run length coding method states that if this pixel and its neighboring pixels shared the same pixel value, then it was concluded that the pixel was a part of the lung fields and was added to the lung image array. If the pixel was not identical to its neighboring pixels, it was assumed to not be lung tissue and was removed from the lung field array.

Because tissues other than lung tissue may share pixel values with neighboring pixels, areas outside of the actual lung fields may be included in the lung field array as false positives. However, lung tissue generally has a very low density compared to other homogeneous tissues in the chest. Therefore, once the lung field arrays were completed, the false positive tissue was removed by zeroing out pixel values above a certain threshold. The resulting lung field array contains only pixels representing lung tissue. The process of isolating the lung tissue by creating a lung field array was performed on all of a patient's chest radiographs.

Next, we created the subtraction images. Because it was difficult to have patients stand in the exact same location with identical levels of inspiration/expiration, the lung fields in two different chest radiographs did not completely overlap. To solve the issue of overlap, the lung field arrays of the later chest radiographs were warped to overlap as much as possible with the original chest radiograph lung fields. Once the images had been warped, the pixel values from the later chest radiograph were subtracted directly from the pixel values of the original chest radiograph lung field array. The resulting image was a subtraction image with pixel values emphasizing the differences between the chest radiographs (Fig. 2). Ideally, the differences between the chest radiographs represented improvement in pulmonary disease.

Each subtraction image was then further broken down into one subtraction image with only the left lung field (Fig. 3) and one subtraction with only the right lung field. Consequently, each patient had three groups of subtraction images: left lung subtraction images, right lung subtraction images, and subtraction images containing the entire chest. In order to minimize the effects of anatomy outside of the lungs, only the left lung and right lung subtraction images were analyzed. We subsequently analyzed various aspects of the subtraction images.

2.4 Subtraction Image Analysis

Several statistical parameters were calculated for each subtraction image. The image histogram for each subtraction image was created. Furthermore, seven statistical parameters - the image mean, uniformity (average entropy), 2nd moment (variance), 3rd moment (skewness), 4th moment (relative flatness), correlation, and contrast - were calculated for each subtraction image [7]. Finally, the Fourier and power spectrum were created in order to evaluate the overall energy and the various frequencies, respectively, of each image.

Each individual lung subtraction image was then placed into one of three separate categories based on the lung's radiographic changes in response to DOTS therapy as confirmed by the physician interpreting the chest radiographs: an improved group, a worsening group, and an unchanged group. The improved group consisted of patients that showed radiographic signs of improvement in response to DOTS therapy. The worsening group consisted of patients that showed radiographic signs of worsening in response to DOTS therapy. Finally, the

unchanged group consisted of patients that did not show radiographic signs of significant improvement or decline in response to DOTS therapy. A t-test was performed to determine any difference between the statistical parameters of the three groups.

Next, the time in days between the two images used to create each subtraction image was compared to the various statistical parameters. The time in days represents the number of days between the second chest radiograph used to create the subtraction image and the original chest radiograph taken at the onset of DOTS therapy. For example, if the fourth of six chest radiographs was taken 120 days after the original chest radiograph, then the time in days was recorded as 120 days. Regression analysis was done on the subtraction images for each patient using the time in days as the independent variable and the various statistical parameters as the dependent variables.

3. RESULTS

3.1 Subtraction Image Analysis

Each of the 200 patients had six chest radiographs taken for a total of 1200 chest radiographs. All the chest radiographs were digitized and subjected to the subtraction image algorithm. When subjected to the subtraction image algorithm, 101 of the patients' chest radiographs successfully produced subtraction images for a total of 390 subtraction images. The remaining 99 patients' chest radiographs did not produce subtraction images. The right and left lungs were isolated in each of the 390 subtraction images. 390 left lung and 390 right lung subtraction images were created for a total of 780 individual lung subtraction images.

Based on physician assessment of each lung in the original chest radiographs, the individual lung subtraction images were classified into three categories: improved, worsened, or unchanged. The number of subtraction images in each group is summarized below (Table 1). It is no surprise that the total number of left lung images, 390, is identical to the total number of right lung images. It is also important to note that the number of left and right subtraction images is not identical within each category. For instance, the number of right lung subtraction images that showed improvement, 150, is not the same as the number of left lung subtraction images that showed improvement, 102. This inequality is due to the fact that when TB is treated with DOTS therapy, the pathology in each lung may respond in different ways and at different speeds, resulting in different radiologic changes. It is also important to note that the number of combined right and left subtraction images — 252, 286, and 242 respectively - in each category are roughly the same.

The mean and standard deviation were calculated for each of the seven statistical parameters mentioned above. The statistical parameters for the group of subtraction images that showed signs of radiological improvement were then compared to the statistical parameters of the other two groups. A t-test was performed in order to determine if a statistical difference exists between the group that improved and the two groups that did not improve (Tables 2 and 3). It is important to notice that the t-test shows that a significant difference exists between the image mean and the uniformity when the group that improved is compared to either the group that worsened or remained unchanged ($p<0.05$). Furthermore, the t-statistic

for the correlation of the subtraction images suggests that a significant difference between the mean correlation of the two groups likely exists (p<0.20), but may require more data to confirm. The image 2nd moment, 3rd moment, 4th moment, and contrast were not significantly different between the subtraction image groups.

3.2 Computer-Assisted Detection System

A computer-assisted detection (CAD) system was created that utilizes the subtraction image algorithm and statistical analysis results. The CAD utilizes a graphical user interface (GUI) that allows the user to load the patient's original chest radiographs using "load images" button (Figure 4).

The original chest radiographs are displayed side-by-side in order from earliest to most recent so that the physician can get a visual feel for the radiographic findings that may be changing over time as a result of DOTS therapy. Once the original chest radiographs have been loaded, the "create subtraction images" button allows the user to create subtraction images using the subtraction image algorithm. The subtraction images are displayed below the original chest radiographs in order. The subtraction images emphasize the areas of greatest difference between a more recent chest radiograph and the original baseline chest radiograph. Finally, the "analyze images" button allows the user to calculate the statistical parameters of each subtraction image. The statistical parameters are then used to determine whether the chest radiographs suggest that the patient is improving, worsening, or not changing in response to DOTS therapy. The physician may use this CAD in order to determine whether or not to change to an alternate anti-TB pharmaceutical regimen before the patient's condition worsens.

4. CONCLUSION

Currently, TB poses a serious health threat worldwide, especially in many areas of the world that have insufficient medical supplies. The treatment regimen proved to be difficult to follow and led to increased non-compliance. DOTS therapy was developed to increase the effectiveness of the anti-TB treatment regimen. However, drug resistance among various strains of mycobacterium tuberculosis and difficulty with patient compliance have decreased the effectiveness of DOTS therapy. As a result, TB continues to cause a significant amount of preventable morbidity and mortality.

200 TB patients from Hong Kong were identified. The 200 patients were treated for six months and followed with time-series chest radiographs over that time for a total of 1200 chest radiographs. 390 subtraction images were created from the 1200 original chest radiographs using a subtraction image algorithm. Each lung was isolated and seven statistical parameters were measured on each subtraction image. Image mean and image uniformity proved to be statistically different between those patients who showed radiological signs of improvement and those patients that did not.

A CAD system was successfully created. The CAD system utilizes a GUI that allows the user to load patient chest radiographs and creates subtraction images using the subtraction image algorithm. Finally, the CAD uses the subtraction images to determine whether or not

the patient is showing signs of improvement in response to DOTS therapy. Physicians can use the CAD system to respond earlier to TB treatment failure, potentially leading to improved outcomes.

REFERENCES

[1]. Cegielski JP, Chin DP, et al. The global tuberculosis situation: progress and problems in the 20th century, prospects for the 21st century. Infect. Dis. Clin. North Am. 2002; 16:1–58. [PubMed: 11917808]

[2]. Corbett EL, Watt CJ, et al. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch. Intern. Med. 2003; 163:1009–1021. [PubMed: 12742798]

[3]. Dunlap NE, Bass J, et al. Diagnostic standards and classification of tuberculosis in adults and children. This official statement of the American Thoracic Society and the Centers for Disease Control and Prevention was adopted by the ATS board of directors, July 1999. This statement was endorsed by the Council of the Infectious Disease Society of America, September 1999. Am. J. Respir. Crit. Care Med. 2000; 161:1376–1395. [PubMed: 10764337]

[4]. Jeong YJ, Lee KS. Pulmonary Tuberculosis: Up-to-Date Imaging and Management. American Journal of Roentgenology. 2008; 191:834–844. [PubMed: 18716117]

[5]. Chaulk CP, Kazandjian VA. Directly observed therapy for treatment completion of pulmonary tuberculosis. Consensus statement of the public health tuberculosis guidelines panel. J. Amer. Med. Assoc. 1998:279–943.

[6]. Sharma SK, Mohan A. Multidrug-resistant tuberculosis. Indian J. Med. Res. 2004; 120:354–376. [PubMed: 15520486]

[7]. Gonzalez, R.; Woods, R. Digital Image Processing. Second Edition. Prentice Hall; Upper Saddle River, New Jersey: 2002. p. 665-669.
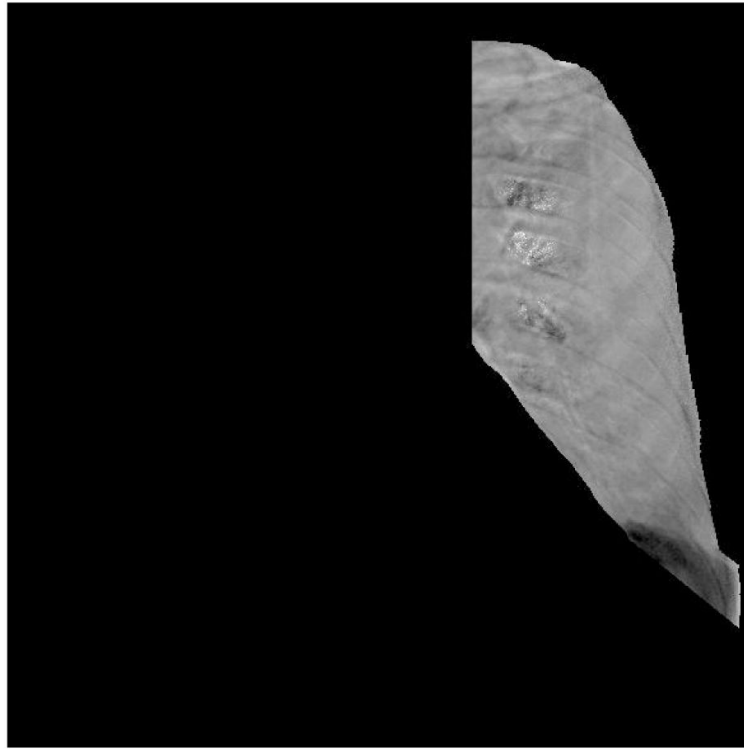
Fig. 1.
A chest radiograph of a TB patient taken at the onset of DOTS therapy.

Fig. 2.
An example of a subtraction image using the entirety of both the first and second chest radiographs of a TB patient undergoing DOTS therapy. Note that some anatomical structures can still be identified.

Fig. 3.
The left lung tissue has been isolated from figure 2 and all pixel values outside the left lung have been zeroed out.
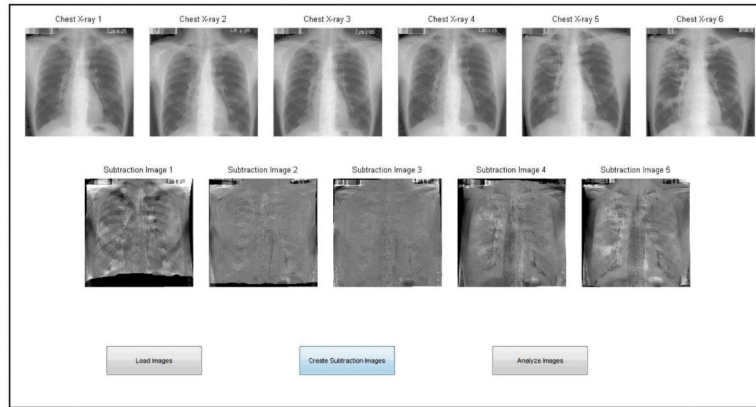
Fig. 4.
GUI utilized by the CAD system. All six chest radiographs are displayed on the top row in order of oldest to most recent. The subtraction images are displayed on the second row. The three buttons are displayed on the bottom row.

Table 1

The number of individual subtraction images in each category.

	Improved	Worsened	Unchanged	Total
Left Lung	102	123	165	390
Right Lung	150	163	77	390
Combined	252	286	242	780

Table 2

Comparison of subtraction images that improved vs. subtraction images that worsened including a t-stat for the various statistical parameters.

Parameter	Improved		Unchanged		
	Mean	Standard Deviation	Mean	Standard Deviation	t-stat
Image Mean	29.24	8.61	31.42	8.77	−2.68
Uniformity (entropy)	0.642	0.073	0.630	0.068	3.00
2nd Moment (variance)	3889.8	1456.7	4196.7	1514.4	−0.04
3rd Moment (skewness)	1.2×10^{12}	1.4×10^{13}	1.1×10^{12}	1.3×10^{13}	2.0×10^{-13}
4th Moment (flatness)	4.3×10^{17}	4.8×10^{18}	3.7×10^{17}	4.5×10^{18}	5.6×10^{-19}
Correlation	0.981	0.088	0.986	0.057	−1.45
Contrast	79.84	25.88	81.60	26.66	−0.42

Table 3

Comparison of subtraction images that improved vs. subtraction images that remained unchanged including a t-stat for the various statistical parameters.

Parameter	Improved		Unchanged		t-stat
	Mean	Standard Deviation	Mean	Standard Deviation	
Image Mean	29.24	8.61	31.05	8.81	−2.10574
Uniformity (entropy)	0.642	0.073	0.631	0.067	2.734344
2nd Moment (variance)	3889.8	1456.7	4147.7	1571.2	−0.03026
3rd Moment (skewness)	1.2×10^{12}	1.4×10^{13}	1.3×10^{12}	1.42×10^{13}	$−6.7 \times 10^{-14}$
4th Moment (flatness)	4.3×10^{17}	4.8×10^{18}	4.42×10^{17}	4.85×10^{18}	$−2 \times 10^{-19}$
Correlation	0.981	0.088	0.982	0.090	−0.07709
Contrast	79.84	25.88	80.44	27.36	−0.13619