Comparison of Chest X-ray findings in ambulatory and hospitalised children with pulmonary TB

Dr Gregory Baker

0303131H

MMed (RadD)

A research report submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Medicine in Diagnostic Radiology

Johannesburg, 2017

Declaration

I, Dr Gregory Baker, declare that this research report is my own work. It is being submitted for the degree of MMed (RadD) at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

DR GREGORY BAKER

On this 19th day of December 2017

For my Delia and Eva

Publications and presentations

This work has never been published.

It was presented:

Congress: The 53rd Annual Meeting & 39th Post Graduate Course of the European Society of Paediatric Radiology (ESPR)

Title: Comparison of Chest X-ray findings in ambulatory and hospitalised children with suspected pulmonary TB

Authors: G Baker, S Andronikou, S Lucas, H Zar

Presenter: Prof S Andronikou

Dates: 1-3/6/17

Location:Davos, Switzerland

Abstract

INTRODUCTION:

Pulmonary TB is common in South Africa, with many children affected. Diagnosis can be challenging and chest x-ray remains fundamental for diagnosis. Interpretation is difficult and shown to have wide inter-reader variability. No study however has compared CXR findings and inter-reader agreement between ambulatory and hospitalised patients.

AIM:

This study compares the frequency of CXR changes, as well as inter-reader agreement in ambulatory compared to hospitalised children with suspected TB.

METHOD:

A pre-existing database containing CXR data of children worked up for PTB from 2008-2013 was used. Retrospective analysis of 69 ambulatory and 112 hospitalised patients, aged 0-12 years from Nolungile clinic and Red Cross Children's Hospital respectively was done. Each sample contained 50% proven TB and 50% negative controls. Two paediatric radiologists and one paediatrician served as blinded, independent readers for the database using standardised tick sheets.

RESULTS:

Finding frequency; Fleiss/Free marginal Kappa (ambulatory and hospitalised respectively):

Overall TB: 27.5% and 35.7%; -0.07 and 0.12 /-0.06 and 0.12

Parenchymal change: 34.8% and 67.9%; 0.24 and 0.49/0.25 and 0.60

Lymphadenopathy: 24.6% and 33.9%; 0.01 and 0.13/0.01 and 0.13

Pleural effusion: 7.3% and 18.8%; 0.27 and 0.61/0.84 and 0.80

Airway compression: 11.6 and 17.9%; 0.26 and 0.20/0.73 and 0.52

CONCLUSIONS:

Our study demonstrated no significant difference in lymphadenopathy, but an increase in parenchymal change in the hospitalised group. We otherwise showed comparable results to literature regarding finding frequency, but poor inter-observer agreement. If the least expert reader were removed, results were comparable with available literature. This

highlights the need for development and study of explicit CXR criteria for lymphadenopathy to improve the value of CXR for paediatric TB in all settings.

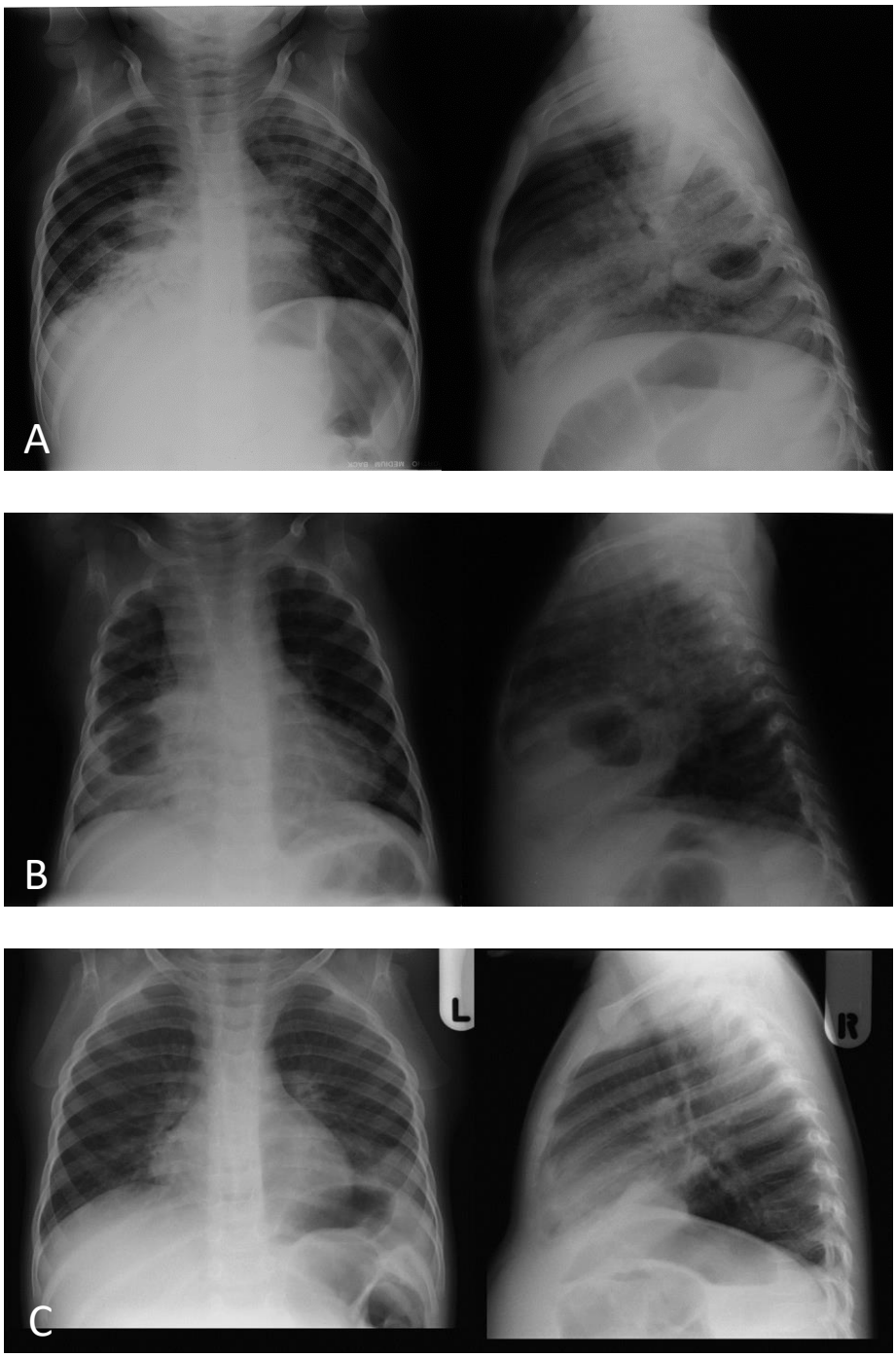
Acknowledgements

Professors Andronikou & Zar, and Dr Lucas.

Thank you for your help and wisdom on a long and occasionally difficult journey.

Table of Contents

List of Figures



A: Hospital patient with lab proven TB. Good inter-reader agreement for lymphadenopathy, parenchymal disease and overall diagnosis of PTB.
B: Hospital patient with lab proven TB. Poor inter-reader agreement for lymphadenopathy and overall diagnosis of PTB.
C: Clinic patient without TB disease. Poor agreement on the presence of lymphadenopathy and the overall diagnosis.

List of Tables

1. Introduction

Pulmonary Tuberculosis (PTB) is very common in South Africa, with a large proportion of affected individuals being children. Diagnosis can be challenging and chest X-ray (CXR) is a fundamental part of the diagnostic algorithm. Interpretation of CXR of children with suspected PTB is difficult and findings may depend on factors such as the severity of disease and age of the child. Further, radiological signs for diagnosis of PTB using CXR have been shown to be subject to wide inter- and intra-observer agreement for the diagnosis of PTB in children. In addition the sensitivity and specificity have been poor against a strong gold standard(1).

However, the patient selection, or non-selection for these studies may affect the outcome. CXR may be more useful for severely ill patients requiring hospitalisation or presenting with large airways disease as a result of tuberculous lymphadenopathy, while it may be less useful for patients with early PTB or suspected TB infection where it is often used as a screening test in patients with a TB contact.

The aim of this study was to compare CXR changes in children with ambulatory disease compared to hospitalised patients with confirmed pulmonary TB.

1.1. Epidemiology of TB

Despite the various preventative and treatment measures in place to combat TB, it still represents a significant source of morbidity and mortality globally, with lower than target declines in incidence of 1.5% reported in 2015 from 2014(2). While the situation is much worse in low or middle income countries (LMICs) such as South Africa (2), the disease has also shown increasing incidence in some high income countries such as Denmark over the last 20 years, mainly due to immigration (3). A similar situation has been identified in the USA with immigrant and adopted children from certain high burden countries showing higher risk of TB disease (4).

The majority of TB cases worldwide in 2015 were in LMICs, with six countries accounting

for 60% of new cases - India, Indonesia, China, Nigeria, Pakistan and South Africa. Global progress in combatting TB depends on advances made in these countries(2).

The WHO 2016 report estimates South Africa's TB incidence to be greater than 10000 new cases per year, one of the highest in the world (2).

Children are one of the groups at highest risk of TB disease. The contribution they make to worldwide TB burden is however generally poorly documented, in particular in high prevalence countries, due to poor confirmation and reporting of cases(5, 6). According to the World Health Organization's report in 2016, of the 10.4 million new cases of TB worldwide, approximately 1.4 million deaths resulted (97000 in South Africa) and that 6.3% of new cases of TB occurred in children under 15 years of age (2). In a study done by Marais et al in 2006 in South Africa, children aged 13 contributed 13.7% of the total local TB burden. Children are therefore an important target population group in the prevention and treatment of TB. In a meta-analysis of 22 high burden LMICs (reported to harbour 80% of the global burden), the paediatric burden has been estimated to be between 4-21% of each country's total burden, significantly higher than the number of notifications(6)

1.2. Pathology/Pathogenesis of TB

Primary tuberculosis disease begins with the inhalation of *M. tuberculosis* (TB) bacilli in an infectious droplet. This then proceeds to the terminal airways and alveoli and a localized pneumonic process occurs, known as the Ghon focus (7). Lymphatic drainage to ipsilateral lymph nodes then causes nodal involvement and lymphadenopathy, which in combination with the Ghon focus, as well as in some cases associated pleural thickening, is referred to as the Ghon complex (7).

The Ghon complex can then progress to full blown primary pulmonary tuberculosis (PTB), as is often seen in older children. However, in most cases infection remains subclinical and becomes dormant, as the bacilli are phagocytosed by local macrophages (7). They are, however not killed by this process (7). The Ghon focus therefore remains a site of possible dissemination and is of importance in cases of so-called 'reactivation PTB' (7).

Progression of TB can occur via one or more of several routes, including local progression from the Ghon focus, via lymphatics and via haematogenous spread (7). Local progression involves penetration through anatomical planes and later caseation and cavitation, once the material is discharged into a bronchus (7). This is most commonly seen in adults (7). Lymphatic spread is far more common in children under 10 years, where lymphobronchial patterns of spread are often seen (7, 8). Children over 10 years of age tend to have adult type disease. Haematogenous spread occurs to some degree in all cases, with 2 ends of a spectrum identified (7). The first occurs early on with bacilli lodging in capillaries and going on to cause disease locally or to further disseminate downstream (7). The second involves erosion into a vessel by more advanced disease and subsequent dissemination (7).

PTB is therefore represented by a spectrum of nodal, airway and airspace disease in children depending on the degree of each method of spread (7).

1.3. Diagnosis of PTB

The diagnosis of PTB has been based upon a combination of clinical features, positive Mantoux test, radiologic investigation and microbiology (9, 10). Diagnosis may be difficult in children due to non-specific, variable symptoms and difficulties in making a microbiologic diagnosis (11, 12). HIV infection further affects disease pattern, complicating diagnosis (12, 13). Clinical features include cough, fever, loss of weight and rarely haemoptysis (9). On history there may be contact with a known TB patient or household contact (9).

Mantoux tuberculin skin testing remains important, with one study showing that each 1mm increase in diameter over the positive threshold represented a 4% increased likelihood of a positive sputum culture (14). This has not been confirmed in other studies. Mantoux testing can however frequently give false negatives, cannot distinguish between exposure versus disease, and BCG vaccine may cause false positives as well (2, 10, 15)

The development of tests with increased accuracy in the diagnosis of TB in children are vital (11). Currently new tests such as Interferon gamma release assays do not allow for the differentiation between active disease and infection (14). Rapid PCR tests such as Xpert have become an important part of the diagnostic process in children suspected of TB, due to their speed (median 1 vs. 16 days for culture) and higher sensitivity (64%) than smear microscopy (10.7%), as well as their ability to detect rifampicin resistance (12). However even in optimal settings only 50% of children with clinically diagnosed PTB have positive microbiological confirmation using Xpert and culture.

The World Health organization estimated that in 2014, despite many people gaining access to effective treatment, up to 37% of new cases were missed (16). In 2016, they reported that case fatality rates for TB were over 20% in many African countries, compared with as low as 5% in the best performing countries, indicating continued inequalities in TB diagnosis and access to treatment(2). Marais et al found that an alarming number of severe cases of paediatric TB were being missed on presentation at local clinics in Cape Town, South Africa (17). Furthermore, of the missed cases then diagnosed at referral centers, most were then hospitalized at TB hospitals due to the severity of their disease (17). The fact that so many severely infected children were missed serves to underline the great difficulty in the diagnosis of paediatric TB and as such, great reliance has fallen onto the role of chest imaging in diagnosis, in particular CXR (13, 18, 19).

1.4. Chest X-ray in Paediatric PTB

CXR is an important part of the diagnosis of PTB in children for the reasons discussed above (13, 18, 19). There are however difficulties arising from the differing anatomy such as the thymus, and physiology of children (13). Differing radiological techniques are also used compared to adults, such as lower dose exposures (13), which may further complicate interpretation, especially for less experienced readers. Variation in the quality of the films taken may also contribute (13).

Further sources of difficulty with the use of CXR for TB diagnosis include effects of HIV/AIDS, nonspecific radiographic signs, and inability to assess drug resistance. High cost was also noted to be a concern and this is particularly so in the setting of resource-limited, LMICs such as South Africa where a very high disease burden is found (18).

Chest x-ray features of pulmonary TB in children overlap significantly with those of non-tuberculous (bacterial, fungal, viral) pneumonia, with few distinguishing features apart from the presence of lymphadenopathy (18). The primary feature used in the radiographic diagnosis of PTB is therefore the presence or absence of mediastinal lymphadenopathy (18-20). A key factor complicating paediatric TB x-ray diagnosis is the absence of standardized, measurable terminology for the spectrum of disease (7). In particular, there are no measurable chest x-ray criteria for stating the presence or absence of lymphadenopathy (19).

1.5 Frequency of findings in paediatric CXR interpretation for TB

Several studies have reported on the frequency of various CXR findings in paediatric PTB, table 1. Only one South African study by Lamont et al that looked specifically at finding frequency in 154 laboratory confirmed TB positive children's CXRs, was found, confirming lymphadenopathy as the single most common finding in TB, found in 63% of cases, and 75% of cases in which the patient was younger than 1 year (21). This was also found to be closely associated with localised, patchy consolidation and reticular interstitial parenchymal disease, with 91% of patients demonstrating lymphadenopathy, also showing these lung changes (21). Miliary pattern was found in (29%), 64% of which were less than 1 year of age.

Table 1.1: Frequency of findings on CXR for paediatric patients with TB

Study (year)	Sample size	Lymphadenopathy	Parenchyma	Pleural effusion
Lamont et al (1986)	154	63%	Overall – 81% Linear interstitial -29% Consolidation – 27% Miliary – 18%	12%
Khatami, et al (2008)	30	90%	Airspace – 43% Reticulonodular and miliary – 40%	6.7%
Du Toit, Swingler, et al (2002)	100	Present 18% Absent 62% Equivocal 20%		
Swingler, Du Toit, et al (2005)	100	47% (overall, incl. confident and equivocal findings)		

A second smaller, Iranian study on 30 patients demonstrated a higher prevalence of all findings especially lymphadenopathy at 90% (22).
All publications report the frequency of findings on CXR of known, TB culture positive children, which may have led to bias and skewed data compared to real world practice, where the imaged patients are unlikely to be confirmed TB.

In their studies on diagnostic accuracy and inter-observer agreement, Du Toit and Swingler provide data on the frequency of the finding of lymphadenopathy in their studies as well, with the study done in 2002 showing lymphadenopathy in 18% of cases, with equivocal findings in 20%. Their study in 2005 combined positive and equivocal findings for lymphadenopathy, which were reported in 47% of their study population (18, 19).

1.6. Inter-observer agreement in Paediatric CXR interpretation for pulmonary TB

There are relatively few studies that examine reliability through determining the level of inter-observer agreement for paediatric chest X-ray diagnosis of TB and/or mediastinal lymphadenopathy, table 1.2.

Table 1.2: Inter-observer agreement for CXR findings in paediatric patients with TB

Study (year)	Sample size	Findings assessed	No of readers	Inter-observer agreement (average weighted Kappa)	Intra-observer agreement
Du Toit, Swingler et al (2002)	100	Lymphadenopathy	4 (paediatric pulmonologists)	0.33	0.55
Swingler, Du Toit et al (2005)	100	Lymphadenopathy	6 (3 paediatricians and 3 primary care physicians)	0.36	
Johnson, Kline (2010)	8 (including 12 'diversion' radiographs)	Consolidation – any cause	6 (2 paediatric radiologists, 2 senior EM physicians, and 2 junior paediatric EM doctors)	0.51 (Paediatric radiologists) 0.55 (Senior EM physicians) 0.37 (Junior paediatric EM doctors)	0.87 (Paediatric radiologists) 0.68 (Senior EM physicians) 0.62 (Junior paediatric EM doctors)
Xavier-Souza et al (2013)	803	Parenchymal change - any cause	2	0.683	

Both studies by Du Toit and Swingler in 2002 and 2005 demonstrated a low to moderate (kappas of 0.33 and 0.36 respectively), by their own categorization, level of inter-observer agreement in the detection of lymphadenopathy, amongst experienced observers with an interest in TB as well as clinic based primary care doctors. There was also a lack of consistency for individual observers noted i.e. a wide intra-observer variability noted between cases, and this was felt to be due to the lack of formal standardized criteria for identifying lymphadenopathy on chest x-ray (18). No statistically significant improvement in agreement was identified with the use of a lateral view in addition to the standard AP, with an estimated 121 lateral views needed to improve detection by one additional case, and 74 needed to remove one false positive (19).

Due to the fact that no reference values were used, validity of the observers' findings were not able to be assessed, and the relatively poor level of agreement was felt to be due to limitation of chest X-ray as a modality in the detection of paediatric TB/lymphadenopathy (19). Weaknesses of the studies include low threshold criteria for TB diagnosis, varied expertise in the interpretation of CXR for TB and the lack of confirmation of TB by culture.

Despite this relatively low level of agreement, actual levels of agreement in 'real world' practice may be even lower, (18) especially as the people reporting on CXR findings in studies were experienced in diagnosing TB on chest X-ray (18). Therefore, diagnosis of paediatric PTB based on the finding of mediastinal lymphadenopathy on chest X-ray, in the absence of other supporting clinical features should be made with caution.

Better agreement was noted for the studies dealing with pneumonic changes, likely due to the simpler, more clear-cut nature of the findings involved. However agreement levels were still only moderate, based on the categorization implemented by Du Toit and Swingler (18, 19).

1.7. This project in context

The abovementioned studies on the frequency of CXR findings in paediatric TB include only hospitalized children that were microbiologically confirmed TB (21). Both abovementioned studies on inter-observer agreement, and the study assessing the frequency of CXR findings in TB took place at Red Cross Children's Hospital in the Western Cape province, South Africa, on children admitted with PTB. As this is a referral hospital, children would be likely to have severe or more advanced disease. The vast majority of PTB cases are, seen at primary clinic level and this therefore, is where most chest x-rays will be performed and viewed. As discussed above, in this 'real world' setting, finding frequency and levels of inter-observer agreement are likely to be lower than reported (23). No studies have been published comparing frequency of CXR findings and inter-observer agreement for the diagnosis of PTB on CXR between a clinic and hospital setting. This study aims to compare the CXR findings in children presenting with differing severities of microbiologically confirmed PTB i.e. hospitalized vs. ambulatory disease. The study investigates the frequency of CXR findings by severity and category of disease as well as inter-observer agreement.

1.8. Aim

This study aims to compare the frequency of radiological findings as well as the inter-observer agreement in the diagnosis of paediatric pulmonary TB in ambulatory versus hospitalised disease.

The primary endpoint of the study is the comparison of the frequency of overall CXR changes in ambulatory compared to hospitalised children with PTB.

1.8.1. Study Objectives

-To compare the frequency of findings on CXR for a hospital as well as an outpatient population in children with confirmed PTB

-To investigate the inter-observer agreement for findings on CXR for a hospital as well as an outpatient population in children with confirmed PTB

2. Materials and Methods

2.1. Research paradigm

The study is a secondary analysis of data of paediatric chest x-rays (CXRs) collected as part of a prospective study on improved TB diagnosis in children(10, 12). CXRs were obtained at the time of enrolment in the parent study and read by 2 paediatric radiologists as well as a paediatrician using a standardised reporting form, in two distinct populations: (a) hospital patients and (b) clinic patients.

2.2. Sample

The study sample was drawn from previously performed CXRs in paediatric patients enrolled in a TB diagnostic study, performed between February 2008 and January 2013 in the Western Cape in 2 separate locations: (a) a hospital setting (Red Cross Children's Hospital) in Cape Town, Western Cape and (b) a clinic setting (Nolungile clinic in Khayalitsha, Western Cape). Both settings serve populations with similar demographics and socioeconomic backgrounds, with Nolungile falling within Red Cross' drainage area. For our study, all microbiologically proven cases of TB from the clinic setting were included, with a randomly selected, equal number of non-TB (patients diagnosed with non-tuberculous lower respiratory disease) cases as controls for further analyses. A larger sample of proven TB cases was randomly selected from the hospitalised group for improved statistical power, with an equal number of non-TB/ other respiratory disease controls.

2.2.1. Inclusion criteria

All radiographs read and captured in the existing database of children with suspected TB were considered for inclusion in the study.

Inclusion criteria for this study:

- Microbiologically confirmed TB
- Chest X-ray of diagnostic technical quality

Inclusion criteria for the original study were:

- Age younger than 15 years
- Cough and one of the following:
    - Household TB contact
    - Failure to thrive or loss of weight within the previous 3 months
    - A positive PPD

2.2.2. Exclusion criteria

Radiographs reported as inadequate for diagnostic purposes by 2 out of 3 readers, as well as those that have incompletely captured data were excluded.

Exclusion criteria for the main study were:

- TB treatment/prophylaxis within the preceding 72hrs
- Inability to be followed up
- No/Unobtainable consent
- Sputa (induced or nasopharyngeal) unobtainable

2.3. Materials and Methods

CXRs were taken of children with suspected PTB in either a clinic (Nolungile) or hospital (Red Cross Children's Hospital) setting using various X-ray machines with a variety of operators. The CXR's were read by 2 experienced paediatric radiologists and 1 paediatrician according to a standardised reporting form. The findings of the readers were entered onto a tick sheet (Appendix B), and these findings were entered onto an Excel spread sheet.

The diagnoses of the patients of whom the films were taken were later confirmed TB positive or negative either microbiologically or with Xpert. The readers were blinded to any microbiological results, if available, at the time of reading.

The features used in evaluating the CXRs for TB included:

- Parenchymal changes [air-space and nodular]
- Mediastinal lymphadenopathy
- Pleural effusion

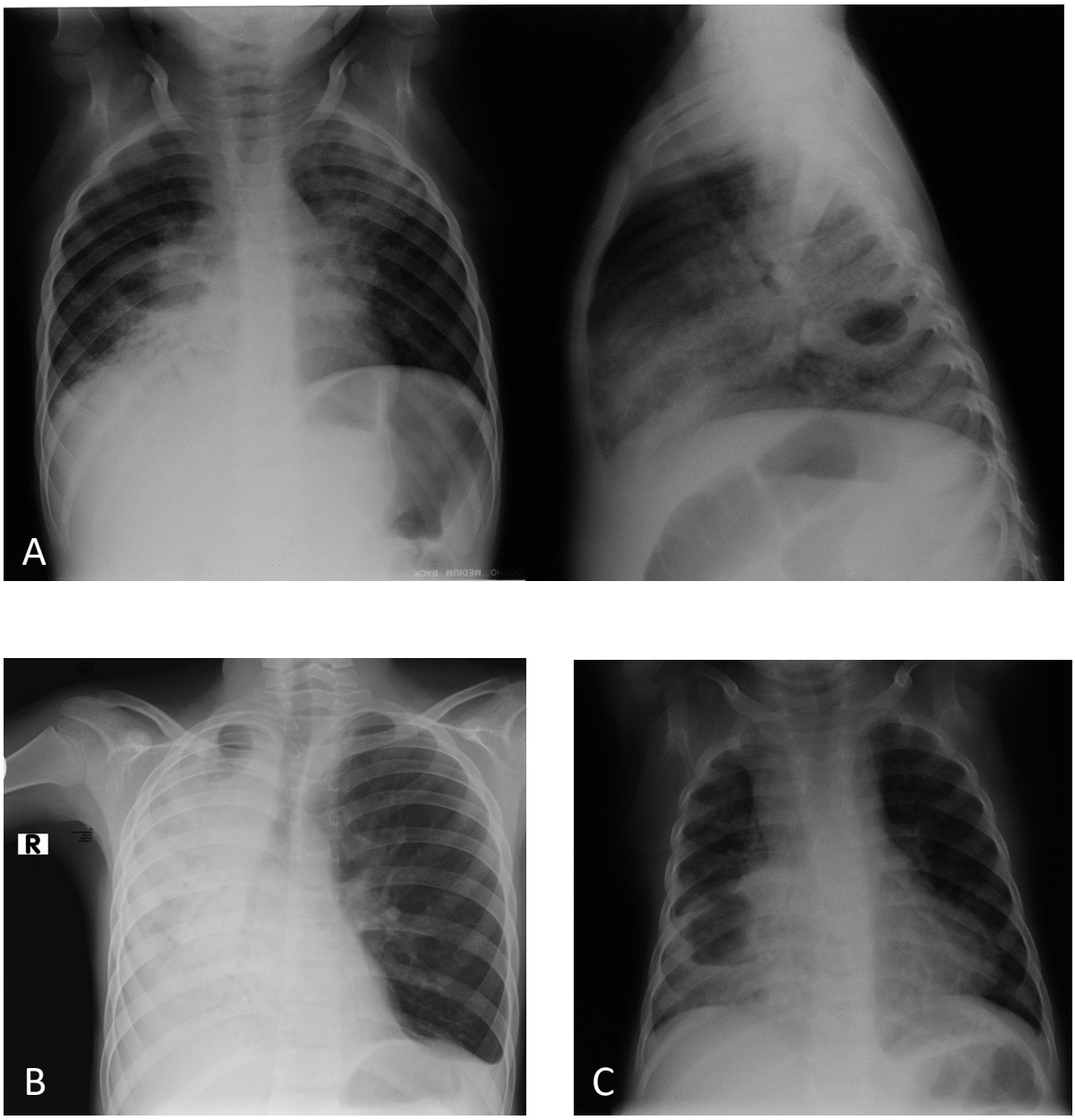
- Airway compression

Examples of chest X-rays with these features are shown in Figure 2.1

The subcategories of the above groups are included in Appendix B, however in our study we looked at each group of findings as a whole. Each reader then decided whether the CXR was consistent with TB, equivocal/~~equivocal~~ inconclusive or normal.

Criteria used by the readers for the CXR diagnosis of TB were:

- Definite radiological TB - defined as the presence of mediastinal lymphadenopathy or miliary pattern, regardless of the presence of other findings.
- Equivocal/Inconclusive – Any positive finding (parenchymal changes or effusion) without the presence of the above 2 findings
- Normal – normal CXR



A: Parenchymal change (Cavitary pneumonia) and lymphadenopathy
B: Pneumonia and pleural effusion
C: Lymphadenopathy, airway compression and cavitary pneumonia

Figure 2.1. Examples of database chest X-rays, with features assessed at reading

2.4. Data collection

The existing image interpretation data was collected using a pre-developed tick sheet detailing each feature evaluated during reading of the films, as well as the diagnostic adequacy/quality of the films (Appendix B) and was captured into an Excel spread sheet, which was used for this study.

2.5. Reliability and validity

Reliability of X-ray interpretation was ensured by using the same 3 readers for the evaluation of all radiographs. Readers were blinded to any clinical and laboratory results. The use of an appropriately sized sample also addressed reliability.

Validity of the results was ensured by the use of microbiological confirmation as well as the newer Xpert TB test. CXR is the first line imaging procedure in the diagnosis of TB (18, 24, 25), and has been found to be a valid test in the screening and diagnosis of TB (26)

2.6. Bias

The literature review has been limited to studies done in English, regardless of date up to and including May 2017 on Pubmed, using the search terms 'children', OR 'paediatric', 'tuberculosis' OR 'PTB' AND 'chest x-ray', which may cause relevant information to be missed. Few studies in languages other than English were, however, found on the topics searched. All relevant South African studies found were in English and were included.

Bias may have been introduced to the study due to the fact that TB is endemic in the study population (Western Cape), as well as the fact that the X-rays are known to be of patients with suspected TB. That knowledge may have potentially affected the threshold of the readers to 'call' a feature that might otherwise be regarded as absent. This is partially negated by the fact that this is a true clinical scenario faced by the readers daily in clinical practice, as well as by the fact that the readers are experienced and thus will have developed their own diagnostic thresholds that should be relatively unaffected as compared with a less experienced reader.

2.7. Data analysis and statistics

Based on the prevalence of radiological findings suggestive of TB of 82% in the clinic population and 63% in the hospital populations in 2 previous studies using the entire dataset (10, 12) and to detect a difference of 22%, in the frequency of the overall CXR diagnosis of PTB the minimum sample size necessary for statistical significance was determined to be 52 in each of the hospitalised and ambulant groups.

The frequency of overall TB diagnosis as well as each of the above feature groups were determined for each population and compared.

The radiological findings and their frequency are summarised by frequency and percentage tabulation, and illustrated by means of bar charts. Continuous variables were summarised by the mean, standard deviation, median and interquartile range, and their distribution illustrated by means of histograms.

For the CXR outcome, 'Equivocal~~//Inconclusive~~Inconclusive' was grouped with 'not TB'.

The comparison of gender, HIV status and CXR characteristics between the two patient types was carried out using Fisher's exact test. The relationship between age and patient type and TB status was assessed by the Wilcoxon rank sum test, since the data did not meet the assumptions of the independent samples t-test. The relationship between gender and TB status was compared using Fisher's exact test.

Sensitivity and specificity were compared between the two patient types using the clinical/laboratory categories as the gold standard, with the z-test for proportions.

Inter-observer agreement was calculated for the final diagnosis and for each radiological feature group of disease presence by Fleiss' kappa for multiple readers and a binary outcome, for each of the population groups (27):

$$\hat{\mathrm{K}}_f = 1 - \frac{\sum_{i=1}^{k} y_i(n_i - y_i)}{kn(n-1)\hat{\pi}(1-\hat{\pi})}$$

Where

k = number of subjects

y_i = number of positive ratings for subject i

n_i =number of readers for subject I

$\hat{\pi} = \sum_{i=1}^{k} \frac{y_i}{nk}$

Inter-reader reliability between pairs of readers was determined by Fleiss' chance-corrected kappa. McNemar's test and the free-marginal equivalent, Byrt's PABAK, were used to determine inter-reader bias.

The free-marginal multi-rater kappa was also calculated to address two limitations of Fleiss' kappa, namely its dependence on the prevalence of the feature in question, and its assumption that the distribution of the cases amongst the possible categories is fixed.

$$\kappa_{free} = \frac{[\frac{1}{Nn(n-1)}(\sum_{i=1}^{N}\sum_{j=1}^{k} n_{ij^2} - Nn)] - [\frac{1}{k}]}{1 - [\frac{1}{k}]}$$

k - values of >0.81 represent almost perfect, 0.61-0.80 substantial, 0.41-0.60 moderate, 0.21-0.40 fair, 0.01- 0.20 slight and <0.0 less than chance inter-observer agreement.

The statistical inter-observer agreement was also compared for the two groups.

Data analysis was carried out using SAS v9.4 for Windows. The 5% significance level was used i.e. p-values <0.05 indicate significant results.

3. Results

There were 82 (41 TB and 41 non-TB controls) ambulatory and 120 (60 TB and 60 non-TB controls) hospitalised cases identified in the database. Due to exclusions due to inadequate CXR technical quality and reader error (mislabelling, loss of forms) during CXR reading, the final included sample was 69 ambulatory and 112 hospitalised cases.

The number of patients with laboratory confirmed TB in these groups was 31 (44.9%) for the ambulatory group with 38 non-TB controls and 56 (50.0%) confirmed TB and 56 non-TB controls for the hospitalised group.

3.1.1. Age, HIV status and Demographics

3.1.1.1. Age

There was no significant difference in the median age of the hospitalised vs ambulatory patients (Wilcoxon rank sum test; p=0.20) with the median age of 37months (IQR 19-62m;). The age distributions of overall and for each location are presented in figures 3.1, 3.2 and 3.3. There was also no significant association between the different age stratifications and the radiological diagnosis of TB overall or for either location. This is presented in table 3.1.

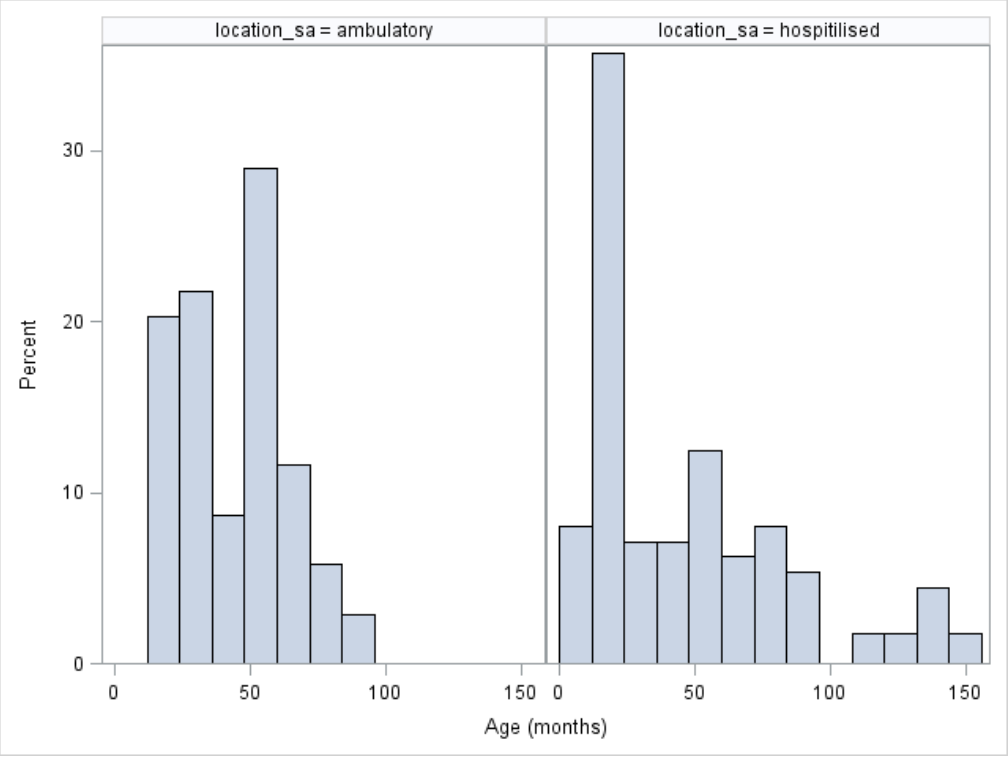


Figure 3.1. Overall Age Distribution

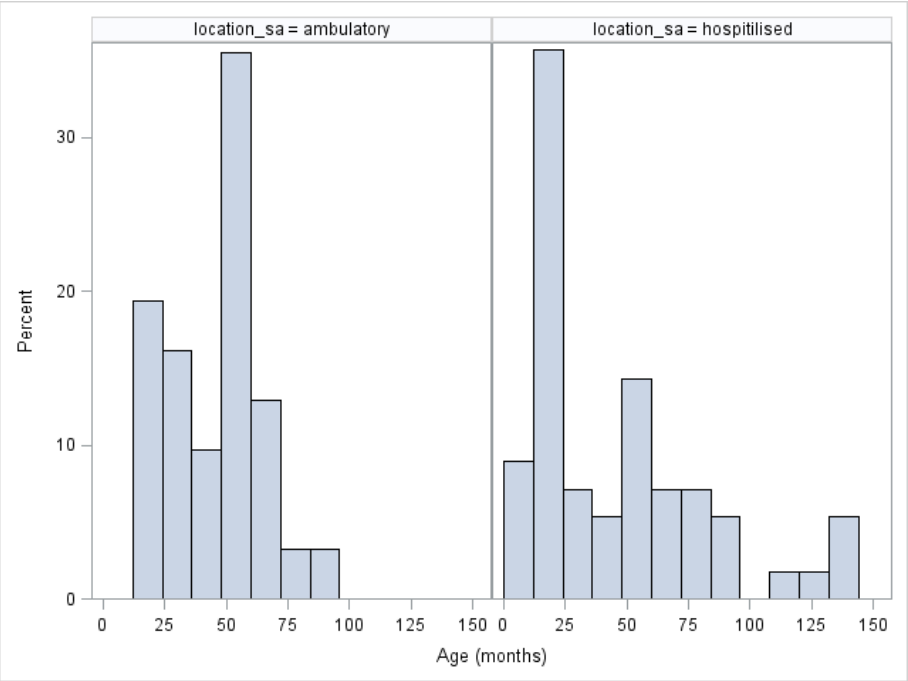


Figure 3.2. Age Distribution in TB patients

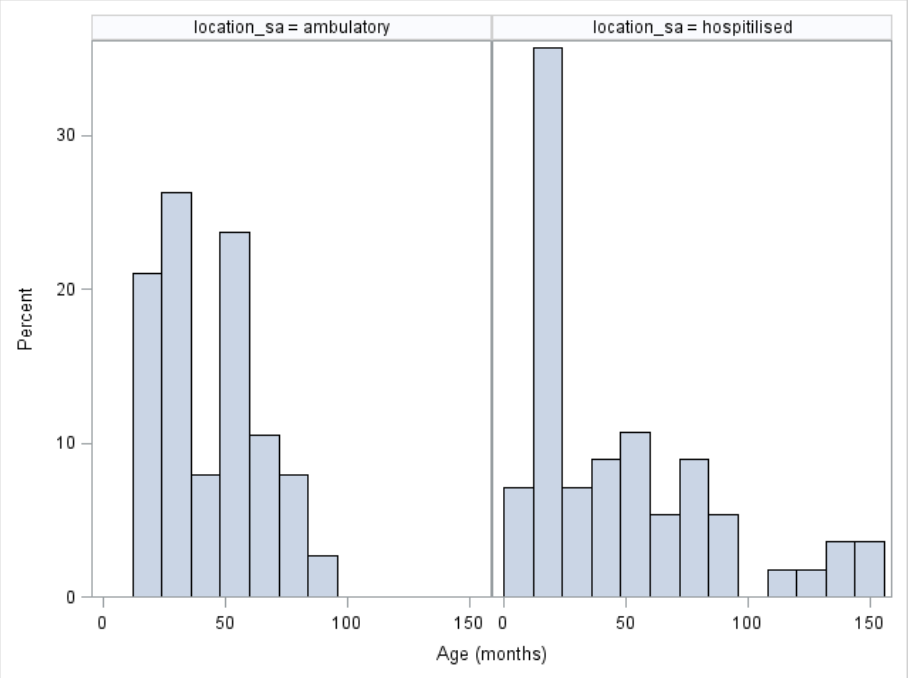


Figure 3.3. Age Distribution TB negative patients

Table 3.1. Comparison of Age and Lab Proven TB status Overall and by Location

Age (months)										p-value for between-group test
Location	**N Obs**	**N**	**Mean**	**Std Dev**	**Median**	**Interquartile range**		**Minimum**	**Maximum**	
Overall	**181**	181	45.5	32.2	36.8	19.2	62.2	2.8	155.5	
Ambulatory	**69**	69	43.7	19.3	47.0	27.4	56.2	15.2	88.5	0.20
Hospitalized	**112**	112	46.6	38.0	33.0	16.6	66.8	2.8	155.5	
TB cases	**87**	87	45.2	31.4	37.6	18.9	61.4	2.8	143.4	
Ambulatory	**31**	31	44.9	18.8	50.8	27.7	56.5	15.2	86.5	0.30
Hospitalized	**56**	56	45.5	36.8	33.0	16.3	64.6	2.8	143.4	
non-TB cases	**94**	94	45.7	33.0	36.7	20.1	63.3	3.1	155.5	
Ambulatory	**38**	38	42.7	20.0	39.6	24.8	56.2	15.2	88.5	0.46
Hospitalized	**56**	56	47.7	39.5	34.2	16.8	71.9	3.1	155.5	

3.1.1.2. Gender

Overall, there was no significant difference in the gender composition of the two locations (Fisher's exact test; p=0.54); the study group comprised 97 (53.6%) male and 84 (46.4%) female patients, presented in Table 3.2. There was no significant difference in proven TB status between genders, either overall or for ambulatory or hospitalized patients, table 3.3.

3.1.1.3. HIV status

There was no significant difference in the HIV status of patients by location (Fisher's exact test; p=0.30); overall the study group comprised 15.5% HIV-infected patients. These results are summarised in Table 3.2.

Table 3.2. Gender, HIV and Lab Proven TB status

Variable	Category	Overall		Overall					TB cases (n=87)							non-TB cases(n=94)						
				Location						Location							Location					
		Overall		Ambulatory		Hospitalised		p-value	Overall		Ambulatory		Hospitalised		p-value	Overall		Ambulatory		Hospitalised		p-value
N		**181**		**69**		**112**			**87**		**31**		**56**			**94**		**38**		**56**		
		n	**%**	**n**	**%**	**n**	**%**		**n**	**%**	**n**	**%**	**n**	**%**		**n**	**%**	**n**	**%**	**n**	**%**	
Gender	Female	84	46.4	30	43.5	54	48.2	0.54	38	43.7	13	41.9	25	44.6	0.83	46	48.9	17	44.7	29	51.8	0.53
	Male	97	53.6	39	56.5	58	51.8		49	56.3	18	58.1	31	55.4		48	51.1	21	55.3	27	48.2	
HIV status	Infected	28	15.5	8	11.6	20	17.9	0.30	14	16.1	4	12.9	10	17.9	0.76	14	14.9	4	10.5	10	17.9	0.39
	Un-infected	153	84.5	61	88.4	92	82.1		73	83.9	27	87.1	46	82.1		80	85.1	34	89.5	46	82.1	
Xpert/Culture TB status	TB disease	87	48.1	31	44.9	56	50.0	0.54	87	100.0	31	100.0	56	100.0								
	Non-TB	94	51.9	38	55.1	56	50.0		0	0.0						94	100.0	38	100.0	56	100.0	

Table 3.3. Comparison of Gender and Lab Proven TB status Overall and by Location

Gender	TB status				
	TB		Non-TB		p-value for between-group test
	n	**%**	**n**	**%**	
	Overall				
Female	38	43.7	46	50.0	0.55
Male	49	56.3	46	50.0	
	Ambulatory				
Female	13	41.9	17	44.7	>0.99
Male	18	58.1	21	55.3	
	Hospitalized				
Female	25	44.6	29	51.8	0.57
Male	31	55.4	27	48.2	

3.1.2. Frequency of findings and bias

3.1.2.1 Comparison of Frequency of Findings

The frequency of CXR findings, overall and by location (separately for TB cases and non-TB controls) are presented in table 3.4 and summarised in figures 3.4, 3.5 and 3.6. The percentage of cases with parenchymal change, and pleural effusion, was significantly higher in the hospitalised than the ambulatory group. [TB cases 42(75%) vs 16(51.6%), p=0.032; Non-TB controls 34(60.7%) vs 8(21.1%), p=0.0002] There were no significant differences between locations for the other features. Airway compression was present in 28 (15.5%) patients overall, with 24 of those patients confirmed TB versus only 4 in the non-TB controls, as seen in, Table 3.4. There was no significant difference between locations.

Table 3.4. Comparison of Finding Frequency Overall and by Location

Variable	Category	Overall		Overall					TB cases (n=87)							non-TB control (n=94)						
				Location						Location								Location				
				Ambulatory		Hospitalised		p-value	Overall	Ambulatory		Hospitalised		p-value		Overall	Ambulatory		Hospitalised		p-value	
n		**181**		**69**		**112**			**87**	**31**		**56**				**94**	**38**		**56**			
		n	**%**	**n**	**%**	**n**	**%**		**n**	**%**	**n**	**%**	**n**	**%**		**n**	**%**	**n**	**%**	**n**	**%**	
CXR	Consistent with TB	59	32.6	19	27.5	40	35.7	0.33	42	48.3	15	48.4	27	48.2	>0.99	17	18.1	4	10.5	13	23.2	0.17
	Parenchymal change	100	55.2	24	34.8	76	67.9	<0.0001	58	66.7	16	51.6	42	75.0	0.034	42	44.7	8	21.1	34	60.7	0.00002
	Airway compression	28	15.5	8	11.6	20	17.9	0.30	24	27.6	8	25.8	16	28.6	>0.99	4	4.3	0	0.0	4	7.1	0.14
	Nodes	55	30.4	17	24.6	38	33.9	0.24	39	44.8	13	41.9	26	46.4	0.86	16	17.0	4	10.5	12	21.4	0.26
	Pleural effusion	26	14.4	5	7.2	21	18.8	0.028	22	25.3	5	16.1	17	30.4	0.20	4	4.3	0	0.0	4	7.1	0.14

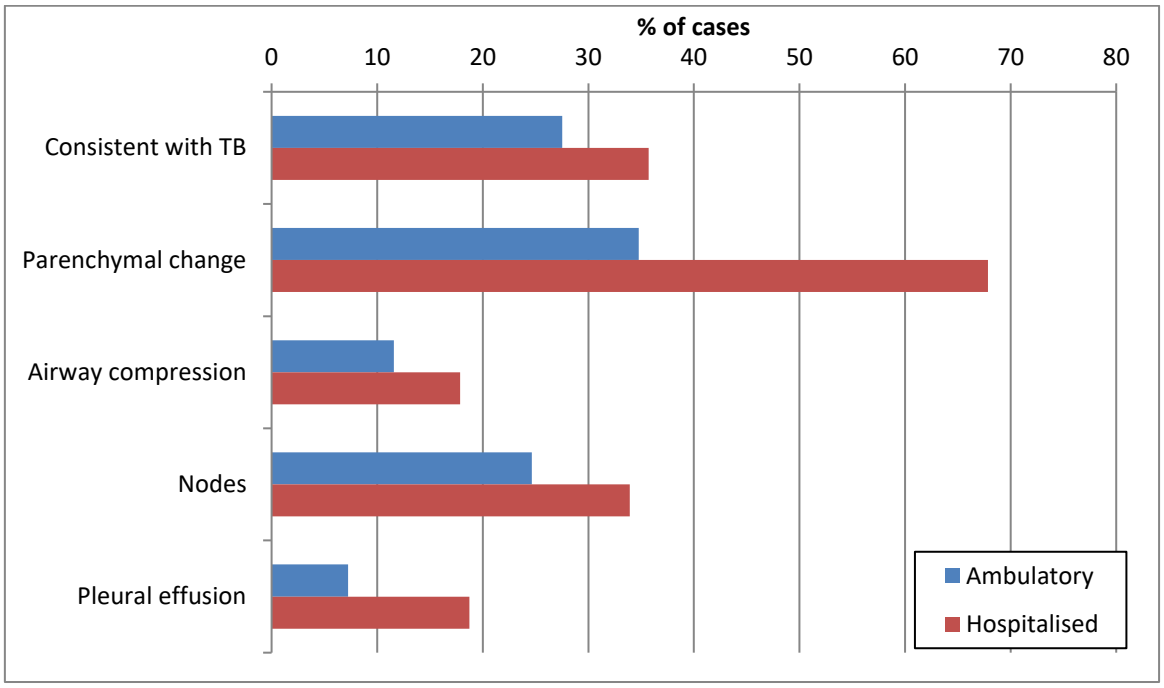


Figure 3.4. Frequency CXR Findings Overall

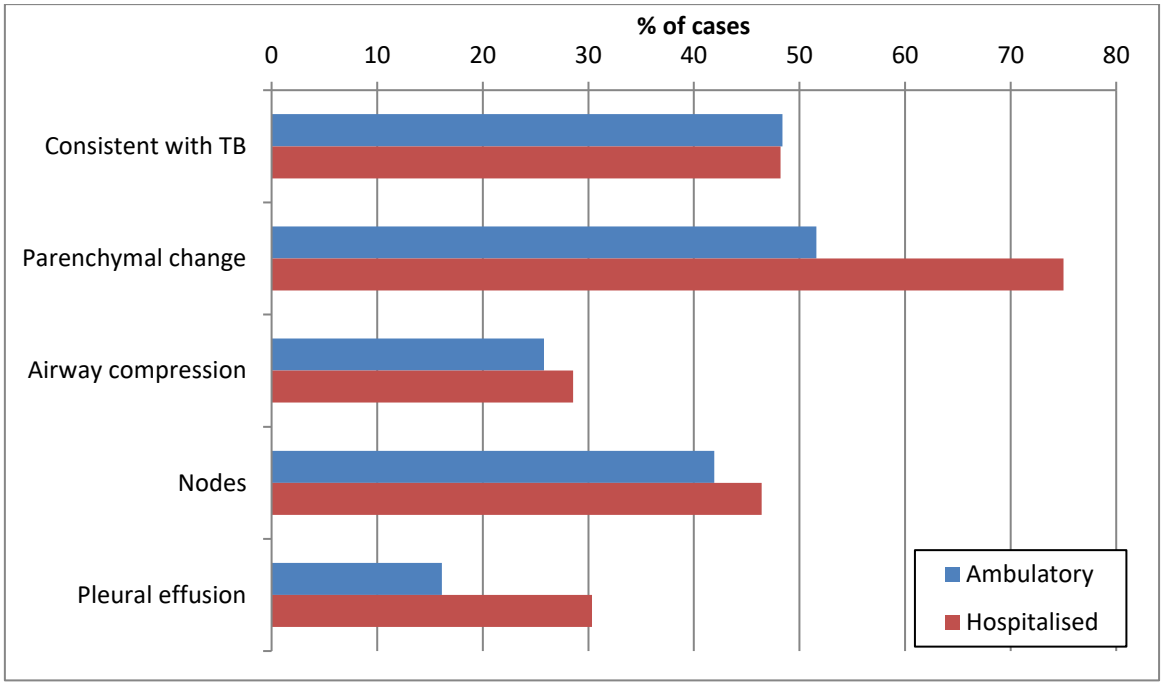


Figure 3.5. Frequency of CXR findings in confirmed TB cases

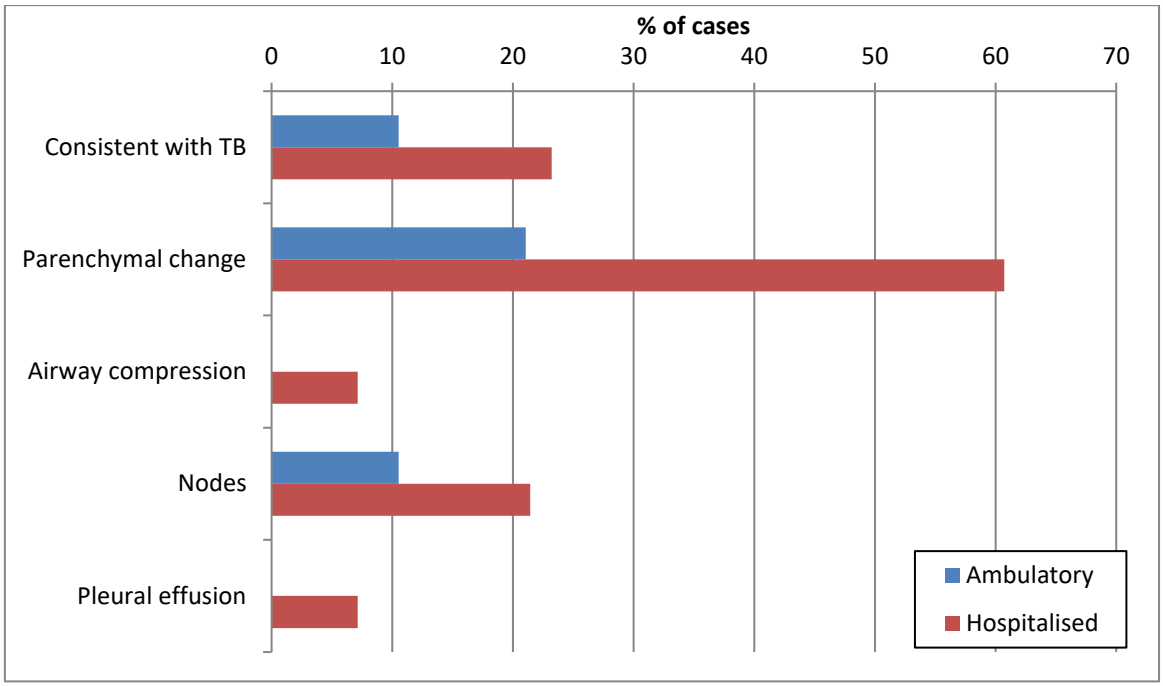


Figure 3.6. Frequency of CXR findings in non-TB cases

3.1.2.2. CXR diagnosis of TB

The frequency of X-ray diagnosis of TB according to each reader and for ambulatory and hospitalised groups is summarised below in table 3.5. Reader C stands out as having identified a higher proportion of the patients as having a CXR with TB, compared to the other readers (as well as to the true, microbiologically confirmed value). Conversely, the other two readers have identified a lower proportion of patients as having a CXR with TB, compared to the true value.

Table 3.5. Frequency of CXR diagnosis of TB

Reader	% patients with TB (number)	
	ambulatory (n=69)	hospitalised (n=112)
Culture/Xpert proven TB	**44.9 (31)**	**50.0 (56)**
A	27.5 (19)	35.7(40)
B	18.8(13)	30.4(34)
C	89.9(62)	74.1(83)

Inter-observer agreement and bias for CXR diagnosis of TB is summarised in table 3.5 below. The raw agreement between all three readers was 20.3% and 33.9% for the Ambulatory and Hospitalised groups respectively.

The Fleiss' kappa chance-corrected measure of agreement was -0.07 (Ambulatory) and 0.12 (Hospitalised), which corresponds to 'less than chance' and 'slight' agreement respectively. The respective free-marginal kappas were -0.06 (Ambulatory) and 0.12 (Hospitalised), again corresponding to 'less than chance' and 'slight' agreement.

Inter-reader bias was tested using Cochran's Q statistic. This was significant ($p<0.0001$) for both locations, indicating significant bias between the readers in determining the presence / absence of TB from the CXR.

For each pair of readers, the cross-tabulation of ratings, the chance-corrected kappa and the test for inter-reader bias (McNemar's test for two readers) are summarised in table 3.6.

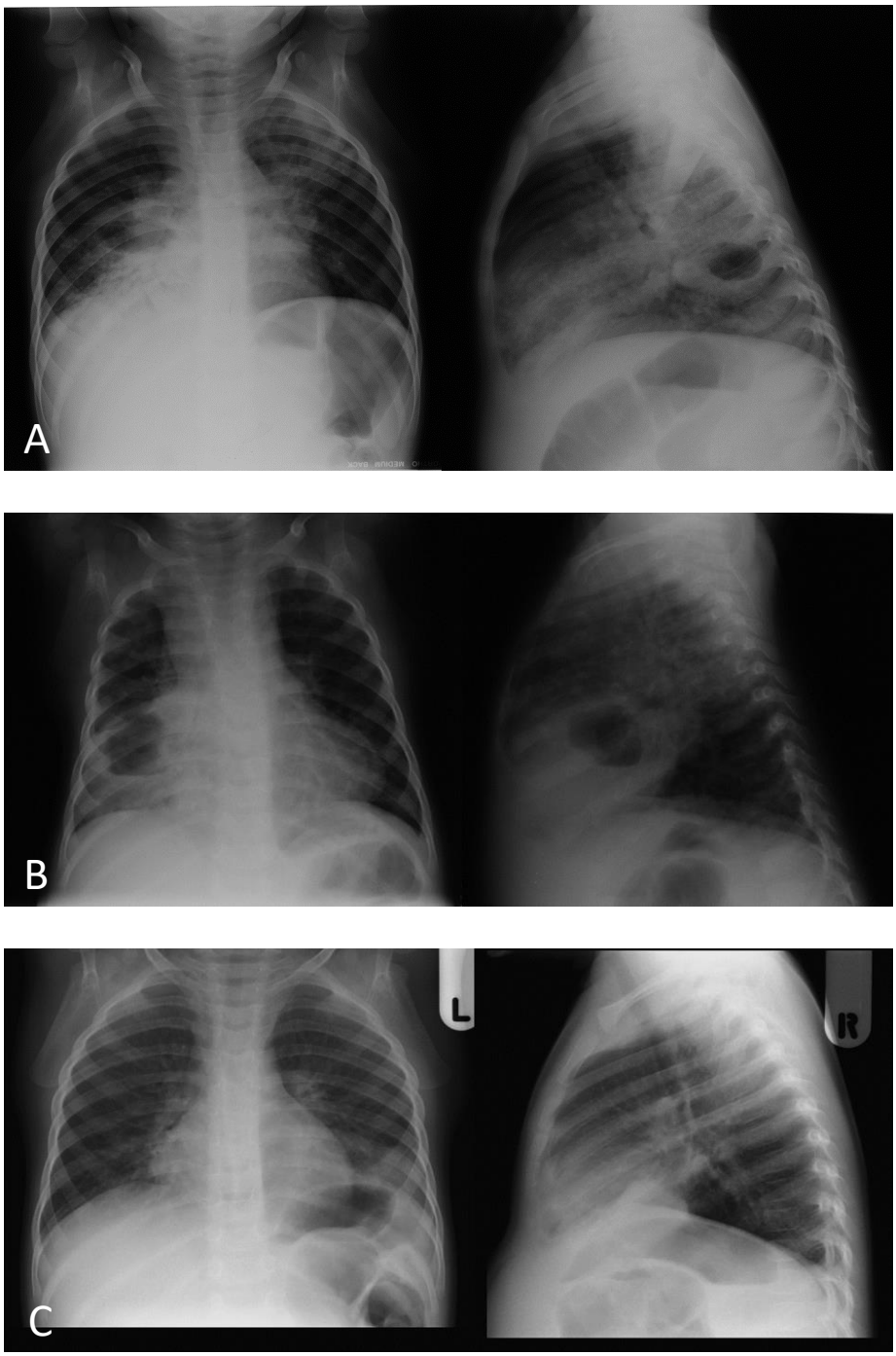
Of note, is that reader C is rating significantly differently to the other readers for both locations. If C were to be removed, the agreement for the other 2 readers (both experienced paediatric radiologists) was k=0.53 and 0.33 for the hospitalised and ambulatory groups respectively, which compares much more favourably with the literature.

The ratings therefore cannot be combined to give a majority decision rating, as this would be statistically equivalent to randomly generating the majority decision for each case.

If reader C were to be removed, a majority decision cannot be calculated, as there would only be 2 readers left.

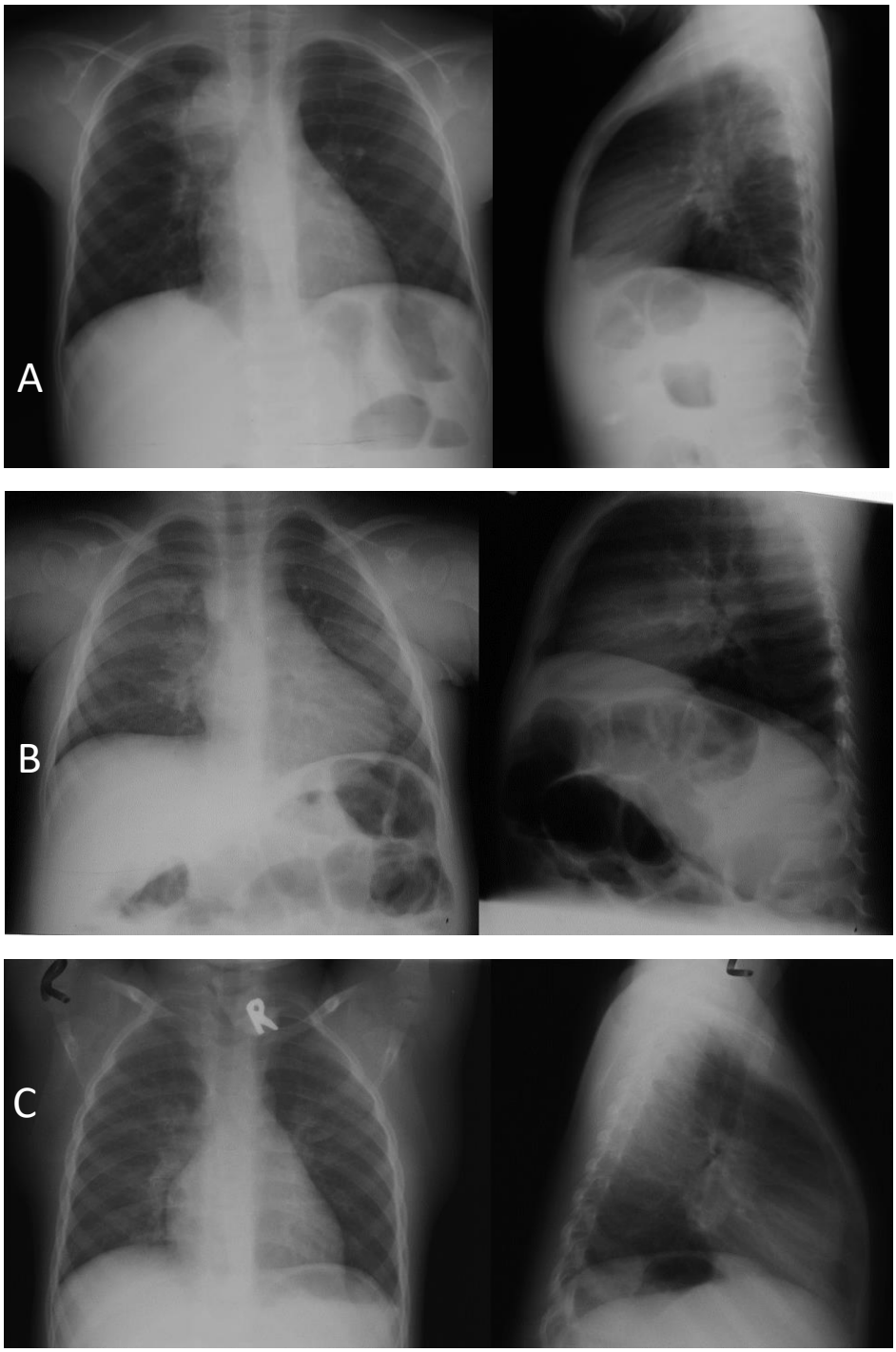
It was therefore concluded that a majority decision cannot be created out of the data available and it was decided to use the ratings of the single most experienced reader (A) for the finding frequency analyses, as well as sensitivity and specificity.

Examples of chest X-rays of the Ambulatory and Hospitalised patients with agreement for each are given in Figure 3.7 and Figure 3.8.



A: Hospital patient with lab proven TB. Good inter-reader agreement for lymphadenopathy, parenchymal disease and overall diagnosis of PTB.
B: Hospital patient with lab proven TB. Poor inter-reader agreement for lymphadenopathy and overall diagnosis of PTB.
C: Clinic patient without TB disease. Poor agreement on the presence of lymphadenopathy and the overall diagnosis.

Figure 3.7. Examples of chest X-rays of the Hospitalised group



A: Clinic patient with lab proven TB. Good inter-reader agreement for lymphadenopathy and overall diagnosis of PTB.
B: Clinic patient with lab proven TB. Poor inter-reader agreement for lymphadenopathy and overall diagnosis of PTB.
C: Clinic patient without TB disease. Poor agreement on the presence of lymphadenopathy and the overall diagnosis.

Figure 3.8. Examples of CXR's of the Ambulatory group

Table 3.6. Inter-observer agreement and bias for CXR diagnosis of TB

Location	Raters	Fleiss' kappa	Interpretation of kappa	Free-marginal kappa	interpretation of kappa	p-value
Ambulatory	Overall	-0.07	less than chance agreement	-0.06	less than chance agreement	<0.0001
	A vs B	0.44	moderate agreement	0.59	moderate agreement	0.11
	A vs C	0.04	slight agreement	-0.30	**less than chance agreement**	<0.0001
	B vs C	0.01	slight agreement	-0.48	**less than chance agreement**	<0.0001
Hospitalised	Overall	0.12	slight agreement	0.12	slight agreement	<0.0001
	A vs B	0.32	fair agreement	0.39	fair agreement	0.30
	A vs C	0.17	slight agreement	0.05	slight agreement	<0.0001
	B vs C	0.08	slight agreement	-0.09	**less than chance agreement**	<0.0001

3.1.2.3. Parenchymal change

The frequency of parenchymal change as recorded individually by the three readers is summarised in table 3.7 below. Reader C again stands out as having identified a higher proportion of the patients as having parenchymal change in both groups.

Table 3.7. Frequency of Parenchymal change

Reader	% patients with Parenchymal change(number)	
	ambulatory (n=69)	hospitalised (n=112)
A	34.8 (24)	67.9 (76)
B	37.7 (26)	69.6 (78)
C	68.1 (47)	81.3 (91)

Agreement:

The raw agreement between all three readers was 43.5% and 69.6% for the Ambulatory and Hospitalised groups respectively. This is summarised in table 3.8 below.

The Fleiss' kappa chance-corrected measure of agreement was 0.24 (Ambulatory) and 0.49 (Hospitalised), which corresponds to 'fair' and 'moderate' agreement respectively. The respective free-marginal kappas were 0.25 (Ambulatory) and 0.60 (Hospitalised), which again corresponded to 'fair' and 'moderate' agreement.

Table 3.8. Inter-observer agreement and bias for Parenchymal change on CXR

Location	Raters	Fleiss' kappa	interpretation of kappa	Free-marginal kappa	interpretation of kappa	p-value
Ambulatory	Overall	0.24	fair agreement	0.25	fair agreement	<0.0001
	A vs B	0.56	moderate agreement	0.59	moderate agreement	0.59
	A vs C	0.24	fair agreement	0.16	**slight agreement**	<0.0001
	B vs C	0.07	slight agreement	-0.01	**less than chance agreement**	0.0004
Hospitalised	Overall	0.49	moderate agreement	0.60	moderate agreement	0.0029
	A vs B	0.71	substantial agreement	0.75	substantial agreement	0.59
	A vs C	0.33	fair agreement	0.48	**moderate agreement**	0.0053
	B vs C	0.41	moderate agreement	0.55	moderate agreement	0.0093

Inter-reader bias:

The inter-reader bias was significant for both groups ($p<0.0001$ and 0.0029 for Ambulatory and Hospitalised, respectively), indicating significant bias between the readers in determining the presence / absence of parenchymal change.

For both locations, reader C is rating significantly differently to the other readers.
The agreement between the remaining two readers is k=0.56 and 0.71 for the Ambulatory and Hospitalised groups respectively.

3.1.2.4. Lymphadenopathy

The frequency of lymphadenopathy on CXR as recorded by the individual readers is summarised in table 3.9 below. Reader C again stands out as having identified a higher proportion of the patients as having lymphadenopathy in both groups, compared to the other readers.

Table 3.9. Frequency of Lymphadenopathy on CXR

Reader	% patients with Lymphadenopathy(number)	
	ambulatory (n=69)	hospitalised (n=112)
A	24.6(17)	33.9(38)
B	24.6 (17)	40.2 (45)
C	89.9 (62)	75.9 (85)

Agreement:

The raw agreement between all three readers was 26.1% and 34.8% for the Ambulatory and Hospitalised groups respectively. This is summarised in table 3.10 below.

The Fleiss' kappa chance-corrected measure of agreement was 0.01 (Ambulatory) and 0.13 (Hospitalised), which corresponds to 'slight' agreement in both cases. The corresponding free-marginal kappas were identical.

Table 3.10. Inter-observer agreement and bias for Lymphadenopathy on CXR

Location	Raters	Fleiss' kappa	interpretation of kappa	Free-marginal kappa	interpretation of kappa	p-value
Ambulatory	Overall	0.01	slight agreement	0.01	slight agreement	<0.0001
	A vs B	0.53	moderate agreement	0.65	**substantial agreement**	>0.99
	A vs C	0.07	slight agreement	-0.30	**less than chance agreement**	<0.0001
	B vs C	0.07	slight agreement	-0.30	**less than chance agreement**	<0.0001
Hospitalised	Overall	0.13	slight agreement	0.13	slight agreement	<0.0001
	A vs B	0.33	fair agreement	0.38	fair agreement	0.24
	A vs C	0.19	slight agreement	0.05	slight agreement	<0.0001
	B vs C	0.06	slight agreement	-0.04	**less than chance agreement**	<0.0001

Inter-reader bias:

The inter-reader bias was significant for both groups ($p<0.0001$ for both groups), indicating significant bias between the readers in determining the presence / absence of lymphadenopathy.

For both locations reader C is rating significantly different.

If reader C were to be excluded, the agreement between the remaining two readers is $k=0.53$ and 0.33 for Ambulatory and Hospitalised groups respectively.

3.1.2.5. Pleural Effusion

The frequency of pleural effusion by each individual reader is summarised in table 3.11 below. Reader C does not rate significantly differently from the other readers for this feature.

Table 3.11. Frequency of Pleural Effusion on CXR

Reader	% patients with Pleural effusion (number)	
	ambulatory (n=69)	hospitalised (n=112)
A	7.3 (5)	18.8 (21)
B	1.5 (1)	12.5 (14)
C	7.3 (5)	14.3 (16)

Agreement:

The raw agreement between all three readers was 88.4% and 84.8% for the Ambulatory and Hospitalised groups respectively. This is summarised in table 3.12. below.

The kappa chance-corrected measure of agreement was 0.27 (Ambulatory) and 0.61 (Hospitalised), which corresponds to 'fair' and 'substantial' agreement respectively. The respective free-marginal kappa measures of agreement were 0.84 (Ambulatory) and 0.80 (Hospitalised), which corresponded to 'almost perfect' agreement in both cases.

Table 3.12. Inter-observer agreement and bias for Pleural Effusion on CXR

Location	Raters	Fleiss' kappa	interpretation of kappa	Free-marginal kappa	interpretation of kappa	p-value
Ambulatory	Overall	0.27	fair agreement	0.84	**almost perfect agreement**	0.10
	A vs B	0.32	fair agreement	0.88	**almost perfect agreement**	0.046
	A vs C	0.35	fair agreement	0.83	**almost perfect agreement**	>0.99
	B vs C	0.32	fair agreement	0.88	**almost perfect agreement**	0.046
Hospitalised	Overall	0.61	substantial agreement	0.80	**almost perfect agreement**	0.10
	A vs B	0.63	substantial agreement	0.80	**almost perfect agreement**	0.035
	A vs C	0.52	moderate agreement	0.73	**substantial agreement**	0.20
	B vs C	0.69	substantial agreement	0.86	**almost perfect agreement**	0.48

Inter-reader bias:

The inter-reader bias was not significant for the either group (p=0.10 for both groups).

Although there was no significant inter-reader bias at overall level for the Ambulatory group, the pairwise comparisons suggest that the ratings from B could be different to those of the other two readers and although there was also no significant inter-reader bias at overall level for the Hospitalised group, the pairwise comparisons suggest that the ratings from A and B were significantly biased.

3.1.2.6. Airway Compression

The frequency of airway compression on CXR as recorded by the individual readers is summarised in table 3.13 below. In this case Reader C has identified a lower proportion of patients as having airway compression compared with the other readers.

Table 3.13. Frequency of Airway Compression on CXR

Rater	% patients with Airway compression	
	ambulatory (n=69)	hospitalised (n=112)
A	11.6	17.9
B	15.9	25.9
C	2.9	10.7

Agreement:

The raw agreement was 79.7% and 64.3% for the Ambulatory and Hospitalised groups respectively, summarised in Table 3.14 below.

The Fleiss' kappa chance-corrected measure of agreement was 0.26 (Ambulatory) and 0.20 (Hospitalised), which corresponds to 'fair' agreement in both cases.
The corresponding free-marginal kappa measures of agreement were 0.73 (Ambulatory) and 0.52 (Hospitalised), which corresponded to 'substantial' and 'moderate' agreement respectively. Here the difference between the two sets of kappa coefficients is evident as the Fleiss' (fixed-marginal) kappa is affected by the relatively low prevalence of airway compression.

Inter-reader bias:
The inter-reader bias was significant for both groups (p=0.011 and 0.0044 for Ambulatory and Hospitalised groups respectively), indicating significant bias between the raters in determining the presence / absence of airway compression.

For the Ambulatory group, it is clear that reader C is reading differently compared to the other two readers.

Although the free-marginal kappas appear better than the Fleiss'/fixed-marginal kappas (due to low-prevalence), the inter-reader bias remains, and complicates attempts to draw conclusions from the kappa values.

In the Hospitalised group, although there was significant inter-reader bias, there is no single reader at fault – the ratings are simply disparate.

Table 3.14. Inter-observer agreement and bias for Airway Compression on CXR

Location	Raters	Fleiss' kappa	interpretation of kappa	Free-marginal kappa	Interpretation of kappa	p-value for Cochrane's Q / McNemar's test
Ambulatory	Overall	0.26	fair agreement	0.73	**substantial agreement**	0.011
	A vs B	0.33	fair agreement	0.68	**substantial agreement**	0.37
	A vs C	0.16	slight agreement	0.77	**substantial agreement**	0.034
	B vs C	0.27	fair agreement	0.74	**substantial agreement**	0.0027
Hospitalised	Overall	0.20	fair agreement	0.52	**moderate agreement**	0.0044
	A vs B	0.35	fair agreement	0.55	**moderate agreement**	0.072
	A vs C	0.13	slight agreement	0.57	**moderate agreement**	0.10
	B vs C	0.11	slight agreement	0.45	**moderate agreement**	0.0023

3.1.3. Sensitivity and Specificity

The sensitivity and specificity of CXR for the detection of TB for each location, are shown in table 3.15. The values in parentheses denote the 95% confidence interval for the estimate (the confidence intervals are very wide due to the small sample size).

Table 3.15. Sensitivity and Specificity of CXR (Reader A) for the diagnosis of TB

	ambulatory (n=69)	**hospitalised (n=112)**
Sensitivity (%)	48.4 (30.2-66.9)	48.2 (34.7-62.0)
Specificity (%)	89.5 (75.2-97.1)	76.8 (63.6-87.0)

Sensitivity was lower than specificity for both locations.

There was no significant difference in the sensitivity or specificity between the two locations (p=0.99 and 0.12 respectively).

4. Discussion

4.1. Results in context

Our study was consistent with the literature in demonstrating lymphadenopathy as the most common extra parenchymal manifestation of pulmonary TB in children, but did not show a significant difference its frequency between the ambulatory and hospitalised groups. There was however a significantly higher frequency of less specific findings (parenchymal change and pleural effusions) in the hospitalised group. Overall agreement between readers was poor, however with the least experienced reader removed, agreement between the remaining two readers compared more favourably with literature.

4.1.1 TB

Pulmonary tuberculosis remains a major public health problem in the Cape Town area, with the drainage areas of Red Cross Children's Hospital and Nolungile clinic being amongst the worst affected (21).

Due to a number of factors including a poorly developed immune response, most cases of pulmonary tuberculosis cases seen in infants and young children are due to primary infection (22). Our study findings are consistent with this, with hilar and mediastinal lymphadenopathy being the most common specific finding for TB.

The high rate of parenchymal lung changes in our study and in particular in the hospitalised group, most of which are severe and non-specific is also consistent with literature which states that PTB in children is a potentially more serious disease with a higher risk of severe complications than in adults (22). Treatment should therefore begin on suspicion, rather than on culture confirmation as therapy is rapidly effective, and potentially lifesaving (21, 22).

Diagnosis of PTB in children is difficult, imprecise and is confirmed in no more than 40% of cases due to difficulty in sputum production and sampling, with a large number of the children in our database never receiving microbiological confirmation (and thus were not eligible for inclusion) and requiring empiric treatment on clinical and CXR grounds.

Further investigations such as CXR are therefore very important.

4.1.2 CXR

In the paediatric age group the chest radiograph (CXR) is of major importance in early diagnosis due to the aforementioned difficulties with bacteriological confirmation (9, 18, 21, 22)

Computerized tomography (CT) with contrast injection is regarded as more accurate than CXR in detecting lymphadenopathy and is the current gold standard for detection, with significantly better sensitivity and specificity than we found in our study (9, 18, 22).

In a clinic setting however CT is not easily available and attempts to utilize it routinely would only further delay diagnosis due to long referral processes and waiting times.

Even in settings where CT is readily available, attempts should be made to define and improve the accuracy of chest radiography, given its faster turn-around time, lower radiation dose, lower cost and the lack of need for contrast injection compared with CT (9, 18, 22).

4.1.3 Overall TB and Feature frequency and agreement

The frequency of overall TB diagnosis by CXR could not be determined by majority decision for the reasons outlined in section 3.1.2.2. The overall frequency of TB diagnosis on CXR by the most experienced reader, Reader A, was 27.5% and 35.7%, compared with the microbiologically proven TB positive values of 44.9% and 50% for ambulatory and hospitalised patients respectively. As expected, these results correlated best with the patients' actual, XPert/culture proven TB status of all 3 readers.

Our study shows significantly higher frequency of parenchymal change and pleural effusion in the hospitalised patients for both confirmed TB and non-TB patients with other respiratory disease, compared with the ambulatory group. This most likely relates to increased severity of disease in the TB positive cases and the presence of other organisms in the non-TB patients causing their symptoms and thus admission.

The lack of a significant difference in the frequency of lymphadenopathy between the hospitalised ambulatory groups may relate to lymphadenopathy being difficult to detect on CXR until large; further large nodes are not always associated with airway compression and erosion or clinically severe disease, as evidenced by the low prevalence of airway compression relative to nodes in our study. In the non-TB patients, lymphadenopathy is not expected to be a major/common feature of disease and thus a significant difference between the hospitalised and ambulatory groups is not expected.

Despite the lack of statistical significance, of interest is the trend toward better agreement on the presence of lymphadenopathy and the overall diagnosis of TB in the ambulatory group, particularly between readers A and B. This is contrary to our initial hypothesis and may indicate that the more severe parenchymal change in the hospitalised group could obscure more specific findings of TB.

Airway compression was present in a higher number of TB versus non-TB patients, which is expected given its relationship with the presence of nodes. Although not reaching statistical significance, a higher number of hospitalised TB patients also demonstrated airway compression relative to ambulatory TB patients.

The poor and/or statistically insignificant agreement between readers along with the degree of bias found, does however call into question the reliability of the findings by some of the readers, and opens up new avenues for study on attempts at maximizing the sensitivity and specificity for the detection of TB on this modality through the development of clear criteria regarding the presence or absence features of TB, as well as comparing CXR with CT, which is the gold standard (18, 20).

4.1.4. Comparison of Finding Frequency (most experienced reader – Reader A) with the Literature

The available literature concerns only hospitalised patients with confirmed TB and is summarised in Table 4.1.
Our findings regarding the frequency of lymphadenopathy in both ambulatory and hospitalised patients with confirmed TB are comparable with those found by Swingler et al in their 2005 study in the same region. The lower overall frequency of lymphadenopathy in our study can be attributed to the presence of TB negative controls.

Table 4.1. Comparison of Frequency of findings on CXR with literature

Study (year)	Sample size	Lymphadenopathy		Parenchymal change		Pleural effusion	
		Amb	Hosp	Amb	Hosp	Amb	Hosp
Our study (overall)	181	24.6%	33.9%	34.8%	67.9%	7.3%	18.8%
Our study (confirmed TB)	87	41.9%	46.4%	51.6%	75%	16.1%	30.4%
Lamont et al (1986)	154	63%		81%		12%	
Kim et al (2006)	25	96%		72%			
Khatami et al (2008)	30	90%		83.3%		6.7%	
Du Toit, Swingler, et al (2002)	100	Present 18% Absent 62% Equivocal 20%					
Swingler, Du Toit, et al (2005)	100	47% (overall, incl. confident and equivocal findings)					

Key: Amb = ambulatory; Hosp = Hospitalised

The remaining studies demonstrate a higher percentage of lymphadenopathy in the case of Khatami et al, Kim et al and Lamont et al, while Du Toit et al found a lower percentage. These differences may be due to differing levels of reader experience, differing reader criteria for the detection of lymphadenopathy and demographic/regional differences. In the case of Khatami et al and Kim et al, differences may be due to the very small sample size in these studies. Incidences reported in Nigeria were 17% and 59% in Korean children (21)

Lymphadenopathy was however still the most common extra-parenchymal manifestation of TB in all the reviewed studies, a result which is consistent with ours and important in the clinical setting due to this finding being the most specific for the diagnosis of TB on CXR (18, 19, 21, 22).

Our ambulatory group demonstrated 51.6% of patients with parenchymal change on CXR, which is likely due to the earlier/less severe disease expected in most clinic presentations. In our hospitalised group however, our study found parenchymal involvement in 75% of TB patients, similar to the 81% and 83% found by Lamont et al and Khatami et al respectively, but lower than the 96% found by Kim et al. Again, the small sample size used by Kim et al may be the reason for this. The differential in parenchymal change between ambulatory and hospitalised patients warrants further study, with exploration of the varying subtypes of parenchymal change found, which was outside of the scope of this study.

Pleural effusion is not a common feature of primary pulmonary tuberculosis in young children, and is rare in infants (15). This is confirmed by the results shown by Lamont et al and Khatami et al. Our study however showed a higher percentage of pleural effusion, the reason for which is unclear. Further study is however needed in our South African setting, to determine whether pleural effusion may have a larger than previously known role in distinguishing between paediatric pneumonia and PTB.

4.1.5. Comparison of Inter-observer agreement for CXR findings with literature

Comparison of inter-observer agreement with the available literature is summarised in table 4.2.

Du Toit et al concluded that lymphadenopathy could be detected on chest radiography with fair inter-observer agreement (18). Our study however demonstrated very poor inter-observer agreement for this feature. Factors that may have influenced this include the fact that our study, unlike those of Swingler et al and Du Toit et al, included both laboratory confirmed TB and non-TB patients. As such our study may better represent actual real-world results for a TB endemic area. A further influence on our study is the fact that one of our readers, C (a clinician), read consistently and significantly differently from our other two, more expert/experienced readers. If C were removed, the agreement for the other 2 readers (both experienced paediatric radiologists) compares more favourably with the literature.

For parenchymal change, our level of agreement overall compares favourably with those of Johnson et al, but are lower than that reported by Xavier-Souza et al. If reader C is again removed, the agreement for the other 2 readers is k=0.56 and 0.71 for ambulatory and hospitalised patients respectively. This compares more favourably with Xavier-Souza et al.

Table 4.2: Comparison of Inter-observer agreement for CXR findings with literature

Study (year)	Sample size	Findings assessed	No of readers	Inter-observer agreement (average weighted Kappa)	**Inter-observer agreement for our study overall/Reader A and B only (for the equivalent feature)**	
					Amb	Hosp
Du Toit, Swingler et al (2002)	100	Lymphadenopathy	4 (paediatric pulmonologists)	0.33	0.01/0.53	0.13/0.33
Swingler, Du Toit et al (2005)	100	Lymphadenopathy	6 (3 paediatricians and 3 primary care physicians)	0.36		
Johnson, Kline (2010)	8	Consolidation – any cause	6 (2 paediatric radiologists, 2 senior EM physicians, and 2 junior paediatric EM doctors)	0.51 (Paediatric radiologists) 0.55 (Senior EM physicians) 0.37 (Junior paediatric EM doctors)	0.24/ 0.56	0.49/0.71
Xavier-Souza et al (2013)	803	Parenchymal change – any cause	2	0.683		

Key: Amb = ambulatory; Hosp = Hospitalised

For pleural effusion, our readers show Fleiss‘ kappa values of 0.27 for the ambulatory and 0.61 for hospitalised patients. These represent fair and substantial agreement respectively, but in this instance the statistical measures are flawed due to the low prevalence of pleural effusion. The free-marginal kappa values of 0.84 (Ambulatory) and 0.80 (Hospitalised), which corresponded to ‘almost perfect’ agreement in both cases, demonstrated more accurately the true level of agreement for this feature in our study. Pleural effusion is however an infrequent and non-specific feature of primary TB and, on its own, of little help in the CXR diagnosis of TB. Furthermore, the literature with which we are comparing results use Fleiss’ kappa, making this a more appropriate comparative measure.

The issue of reader C rating so differently from the other 2 more experienced readers highlights a concern for real world use of CXR for the diagnosis of TB in that the expertise and experience of the readers so vastly affects the interpretation. In a clinic setting doctors of widely varying experience interpret CXR’s for TB. Their experience and expertise in CXR interpretation will however invariably be less than that of qualified paediatric radiologists and far poorer levels of agreement can be expected. This may be less of an issue in a setting such a Red Cross Children’s’ Hospital, where studies are routinely reported by paediatric radiologists, however a similar, if less marked, difference in expertise can be expected in lower level, less specialised hospital settings where only a general radiologist may be available.

Sensitivity and Specificity

Our study showed no significant difference in sensitivity or specificity for the diagnosis of TB on CXR between the Ambulatory and Hospitalised groups. This goes against our hypothesis that the hospitalised patients would demonstrate more severe changes and therefore be more reliably diagnosed on X-ray. It does however correlate well with our findings of no significant difference in the frequency of the most specific CXR feature, lymphadenopathy, between the 2 groups. The hospitalised group did indeed show increased frequency of parenchymal change and pleural effusion, but unfortunately these features have considerable overlap with other chest pathologies and are thus much less

specific for TB. If only 2 readers (A and B) are considered, there is however better reliability within the hospitalised group, suggesting that for more expert readers at least, CXR signs may be more easily agreed upon in these sicker patients.

Sensitivity was lower than specificity for both locations. This is unfortunate in the context of our study, as, even with the most experienced reader, whose results were used for the generation of sensitivity and specificity, approximately half of the cases that presented would go undiagnosed on the basis of CXR alone.

A normal CXR therefore cannot exclude the presence of PTB in children. When present however, the radiographic findings, although not commonly specific for PTB, reflect the pathological changes taking place and greatly assist in diagnosis and treatment decisions (21).

4.2. Current applications

Despite the fact that in the diagnosis of paediatric pulmonary TB, CXR is commonly the single most useful investigation, our study as well as the available literature demonstrates that the absence of any abnormality can be misleading (21).

Our study therefore supports the current clinical practice of using CXR as an adjunctive investigation along with clinical and laboratory evaluation. If the CXR is normal however and there is the suspicion of TB, the clinician should not cease the search for the disease, but should rather turn to more advanced investigations such as CT or MRI for clarification. Newer laboratory investigations such as Xpert have been shown to improve detection from the poor sputum and gastric lavage yields in children and should also be used as far as possible, as currently no single investigation demonstrates acceptable sensitivity and specificity and adequate diagnostic accuracy can currently only be achieved using a combination of the above.

4.3. Limitations of the current study

Limitations of our study include a relatively small sample size due to the relative scarcity of microbiologically confirmed TB patients available within the database, as well as the need to have all CXR's read by all 3 readers.

Bias was shown to be a significant problem in our study (except for pleural effusion) and likely relates to the readers knowing that the CXR's were being assessed specifically for TB, as well as the fact that the study was being done in a region and population with endemic TB (12, 21). This is difficult to address and eliminate. A similar bias can also be expected in practicing clinicians in the region. The effect of this degree of bias may be an area for further study.

The database utilised also included multiple subgroups for each of the CXR features groups (parenchymal, lymphadenopathy, effusion), allowing for more in depth analysis of the frequency and agreement each feature group. This was however deemed beyond the scope of the current study.

A further limitation is the inclusion of CXRs read as inconclusive as normal for the purpose of analysis. This is due to the fact that in a practical, clinical setting an inconclusive CXR is unhelpful in the diagnosis or exclusion of TB. Further study and more in-depth analysis of inconclusive CXRs along with the abovementioned feature subgroups may help further clarify features most helpful in TB diagnosis, and assist in minimising inconclusive studies going further.

4.4. Future applications

Further development and study of explicit criteria for the detection of the various feature groups, especially that of lymphadenopathy may improve the sensitivity and specificity and thus the clinical value, of CXR in the diagnosis of TB in children.
Comparison with CT as the gold standard would also help develop and validate such criteria.

5. Conclusion

Our study demonstrated similar findings to prior publications with lymphadenopathy the most common extra-parenchymal finding on CXR in paediatric TB. Of interest, however is the lack of a significant difference in the frequency of lymphadenopathy in TB when comparing ambulatory and hospitalised patients. Parenchymal changes, while showing the expected increase in frequency in the hospitalised group, are unfortunately relatively non-specific for TB as shown by the concomitant increase in parenchymal change in the control group. Combined with other features however there was a slight, but statistically insignificant, higher percentage of patients in the hospitalised group diagnosed with TB on CXR. Exploration and comparison of the subtypes of parenchymal change seen in TB, compared with those found in non tuberculous infection, is outside of the scope of this study, but may be of value for future study.

Pleural effusions were relatively common in our study compared with previous literature, the reason for which is uncertain, but may relate to co-infection with other pathogens or indicate a larger than known role of pleural effusion in the differentiation between paediatric pneumonia and TB, and warrants further study.

Our study, like previous literature showed poor inter-observer agreement, particularly when considering our least 'expert' reader, and this raises concerns about the use of CXR in a clinic setting where the interpreting doctors are likely to be relatively junior, with less paediatric experience than any of our readers. However, when considering only two readers the inter-observer agreement trended toward being higher for the ambulatory group for lymphadenopathy and overall CXR diagnosis of PTB. This is contrary to our initial hypothesis and may indicate that the more severe parenchymal changes in the hospitalised group could obscure more specific findings of TB.

This highlights the need for the development and study of explicit criteria, in particular for the presence of lymphadenopathy, and also for those parenchymal changes more specific for TB, in order to improve the diagnostic value of CXR for paediatric TB in all settings.

Appendix A: Ethics Clearance Certificate

R14/49 Dr Gregory Baker

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M150755

NAME: (Principal Investigator)	Dr Gregory Baker
DEPARTMENT:	Radiology Red Cross Hospital and Nolungile Clinic, Cape Town Charlotte Maxeke Johannesburg Academic Hospital
PROJECT TITLE:	Comparison of Chest X-Ray of Clinic-Based and Hospital-Based Children with regard to Findings and Inter-Observer Agreement
DATE CONSIDERED:	31/07/2015
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Dr Su Lucas, Prof Savvas Andronikou and Prof Heather Zar
APPROVED BY:	Professor P Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	28/08/2015

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature

Date 5/9/2015

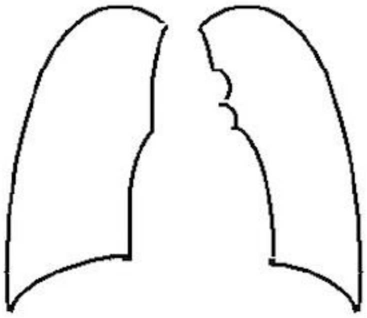
PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix B: Standardised reporting sheet

CXR CRF Baseline vs1
26-aug-2010

CXR results form Baseline (enrolment)

10.	Chest x-ray date		
11.	Is the chest x-ray Normal (go to end) Abnormal (go to x)	N	A
12.	Indicate the distribution of abnormalities		



Coder/Reporter:			Name:								
PARENCHYMA			Lung area involved (tick applicable boxes)								
Alveolar (consolidaton)	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Ghon focus	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Expansile	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Cavity (number, size, thickness)	Y	N	RUL	RML	RLL	LUL	LING	LLL			
Nodular infiltr (>2mm)	Y	N	Peri-hilar			Peripheral				Right	Left
Miliary infiltr (up to 2mm)	Y	N	Peri-hilar			Peripheral				Right	Left
Volume loss	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Hyperinflation	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Calcification (lung)	Y	N	RUL	RML	RLL	LUL	LING	LLL		Segm	Lobar
Peri-hilar streakiness	Y	N	Right	Left							

AIRWAY COMPRESSION AND OR TRACHEAL DISPLACEMENT				
Right paratraceal	y	n	maybe	Not visible
Right main bronchus	y	n	maybe	Not visible
Right bronchus intermedius	y	n	maybe	Not visible
Left main bronchus	y	n	maybe	Not visible

NODES											
Peri –hilar (P-H)	Y	N	R	L	SC	U					
Paratracheal (PT)	Y	N	R	L	U						
Calcification (nodes)	Y	N	P-H	PT							
PLEURA											
Effusion	Y	N	R	L	B-L	Sm	Lge		Loculated	Y	N

HEART							
Enlargement	Y	N	U	Pericardial Effusion	Y	N	U

TECHNICAL QUALITY				
AP view	Acceptable	Poor but readable	Not acceptable	Not readable
Lateral view	Acceptable	Poor but readable	Not acceptable	Not readable

REPORTER				
	Report completed by/ investigator code:			
	Has digital image been recorded	**Y**	**N**	**DK**
	Is the CXR consistent with TB?	**Y**	**N**	**Inconclusive**
Comments				

6. References

(1) Andronikou S, Van der Merwe DJ, Goussard P*, et al.* 2012. Usefulness of lateral radiographs for detecting tuberculous lymphadenopathy in children – confirmation using sagittal CT reconstruction with multiplanar cross-referencing. *South African Journal of Radiology*, 16(3).

(2) WHO. 2016. Global Tuberculosis Report. Available: http://www.who.int/tb/publications/global_report/en/ [Accessed on 23/05/2017].

(3) Rahman N, Pedersen KK, Rosenfeldt V*, et al.* 2012. Challenges in diagnosing tuberculosis in children. *Danish medical journal*, 59(7):A4463.

(4) George SA, Ko CA, Kirchner HL*, et al.* 2011. The role of chest radiographs and tuberculin skin tests in tuberculosis screening of internationally adopted children. *The Pediatric infectious disease journal*, 30(5):387-91.

(5) Marais BJ, Schaaf HS. 2014. Tuberculosis in children. *Cold Spring Harbor perspectives in medicine*, 4(9):a017855.

(6) Dodd PJ, Gardiner E, Coghlan R*, et al.* 2014. Burden of childhood tuberculosis in 22 high-burden countries: a mathematical modelling study. *The Lancet Global Health*, 2(8):e453-e59.

(7) Marais BJ, Gie RP, Schaaf HS*, et al.* 2004. A proposed radiological classification of childhood intra-thoracic tuberculosis. *Pediatric radiology*, 34(11):886-94.

(8) Lucas S, Andronikou S, Goussard P*, et al.* 2012. CT features of lymphobronchial tuberculosis in children, including complications and associated abnormalities. *Pediatric radiology*, 42(8):923-31.

(9) Boloursaz MR, Khalilzadeh S, Baghaie N*, et al.* 2010. Radiologic manifestation of pulmonary tuberculosis in children admitted in pediatric ward-Massih Daneshvari Hospital: a 5-year retrospective study. *Acta medica Iranica*, 48(4):244-9.

(10) Nicol MP, Workman L, Isaacs W*, et al.* 2011. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in Cape Town, South Africa: a descriptive study. *The Lancet Infectious Diseases*, 11(11):819-24.

(11) Graham SM, Ahmed T, Amanullah F*, et al.* 2012. Evaluation of Tuberculosis Diagnostics in Children: 1. Proposed Clinical Case Definitions for Classification of Intrathoracic Tuberculosis Disease. Consensus From an Expert Panel. *JID*, 205 (Suppl 2):10.

(12) Zar HJ, Workman L, Isaacs W*, et al.* 2013. Rapid diagnosis of pulmonary tuberculosis in African children in a primary care setting by use of Xpert MTB/RIF on respiratory specimens: a prospective study. *The Lancet Global Health*, 1(2):e97-e104.

(13) Andronikou S, Vanhoenacker FM, De Backer AI. 2009. Advances in imaging chest tuberculosis: blurring of differences between children and adults. *Clinics in chest medicine*, 30(4):717-44, viii.

(14) Luabeya KK, Mulenga H, Moyo S*, et al.* 2012. Diagnostic features associated with culture of Mycobacterium tuberculosis among young children in a vaccine trial setting. *The Pediatric infectious disease journal*, 31(1):42-6.

(15) Kim WS, Choi JI, Cheon JE*, et al.* 2006. Pulmonary tuberculosis in infants: radiographic and CT findings. *AJR American journal of roentgenology*, 187(4):1024-33.

(16) WHO. 2015. Global Tuberculosis Report. Available: http://www.who.int/tb/publications/global_report/gtbr2015_executive_summary.pdf [Accessed on 23/05/2017].
(17) Marais BJ, Hesseling AC, Gie RP*, et al.* 2006. The burden of childhood tuberculosis and the accuracy of community-based surveillance data. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 10(3):259-63.
(18) Du Toit G, Swingler G, Iloni K. 2002. Observer variation in detecting lymphadenopathy on chest radiography. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 6(9):814-7.
(19) Swingler GH, du Toit G, Andronikou S*, et al.* 2005. Diagnostic accuracy of chest radiography in detecting mediastinal lymphadenopathy in suspected pulmonary tuberculosis. *Archives of disease in childhood*, 90(11):1153-6.
(20) Andronikou S, Joseph E, Lucas S*, et al.* 2004. CT scanning for the detection of tuberculous mediastinal and hilar lymphadenopathy in children. *Pediatric radiology*, 34(3):232-6.
(21) Lamont AC, Cremin BJ, Pelteret RM. 1986. Radiological patterns of pulmonary tuberculosis in the paediatric age group. *Pediatric radiology*, 16(1):2-7.
(22) Khatami A, Sabouri S, Ghoroubi J*, et al.* 2008. Radiological Findings of Pulmonary Tuberculosis in Infants and Young Children. *Iran J Radiol*, 5(4):4.
(23) WHO. 2013. Global Tuberculosis Report. Available: http://www.who.int/tb/publications/global_report/en/ [Accessed on 31/7/2014].
(24) Kwong JS, Carignan S, Kang EY*, et al.* 1996. Miliary tuberculosis. Diagnostic accuracy of chest radiography. *Chest*, 110(2):339-42.
(25) Marais BJ, Gie RP, Schaaf HS*, et al.* 2006. Childhood pulmonary tuberculosis: old wisdom and new challenges. *Am J Respir Crit Care Med*, 173(10):1078-90.
(26) Mor Z, Leventhal A, Weiler-Ravell D*, et al.* 2012. Chest radiography validity in screening pulmonary tuberculosis in immigrants from a high-burden country. *Respiratory care*, 57(7):1137-44.
(27) Shoukri MM, ed. Measures of Interobserver Agreement and Reliability. Second ed: CRC Press, 2010.