

# Inter-Observer Variability in the Detection and Interpretation of Chest X-Ray Anomalies in Adults in an Endemic Tuberculosis Area

Boniface Moifo<sup>1,2\*</sup>, Eric Walter Pefura-Yone<sup>1,3,4</sup>, Georges Nguéfack-Tsague<sup>1,5</sup>,  
Marie Laure Gharingam<sup>1</sup>, Jean Roger Moulion Tapouh<sup>2</sup>, André-Pascal Kengne<sup>6</sup>,  
Samuel Nko'o Amvene<sup>1,2</sup>

<sup>1</sup>Faculty of Medicine and Biomedical Sciences (FMBS), The University of Yaoundé I, Yaoundé, Cameroon

<sup>2</sup>Department of Radiology and Radiation Oncology, FMBS, Yaoundé, Cameroon

<sup>3</sup>Department of Internal Medicine and Specialties', FMBS, Yaoundé, Cameroon

<sup>4</sup>Service of Pneumology A, Jamot Hospital of Yaoundé, Yaoundé, Cameroon

<sup>5</sup>Department of Public Health, FMBS, Yaoundé, Cameroon

<sup>6</sup>South African Medical Research Council & University of Cape Town, Cape Town, South Africa

Email: \*[bmoifo@yahoo.fr](mailto:bmoifo@yahoo.fr), [pefura2002@yahoo.fr](mailto:pefura2002@yahoo.fr), [nguefacksague@yahoo.fr](mailto:nguefacksague@yahoo.fr), [mclor01@yahoo.fr](mailto:mclor01@yahoo.fr),  
[tapouh@yahoo.fr](mailto:tapouh@yahoo.fr), [apkengne@yahoo.com](mailto:apkengne@yahoo.com), [nkoo\\_as@yahoo.com](mailto:nkoo_as@yahoo.com)

Received 25 June 2015; accepted 9 September 2015; published 11 September 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

**Purpose:** To assess the inter-observer agreement in reading adults chest radiographs (CXR) and determine the effectiveness of observers in radiographic diagnosis of pulmonary tuberculosis (PTB) in a tuberculosis endemic area. **Methods:** A quasi-observational study was conducted in the Pneumology Department of Yaounde Jamot Hospital (Cameroon) from January to March 2014. This included six observers (two chest physicians, two radiologists, two end-training residents in medical imaging) and 47 frontal CXRs (4 of diffuse interstitial lung disease, 6 normal, 7 of lung cancers, 7 of bacterial pneumonia, 23 of PTB). The sample size was calculated on the basis of an expected 0.47 Kappa with a spread of 0.13 ( $\alpha = 5\%$ , CI = 95%) for six observers and five diagnostic items. The analysis of concordance was focused on the detection of nodules, cavitary lesions, pleural effusion, adenomegaly and diagnosis of PTB and lung cancer. These intervals of kappa coefficient were considered: discordance (<0.0), poor agreement (0.0 - 0.20), fair (0.21 - 0.40), moderate (0.41 - 0.60), good (0.61 - 0.80), excellent (>0.81). **Results:** The average score for the detection of caverns was the highest (58.3%) followed by that of the correct diagnosis of tuberculosis (49.3%). Pneumologists had the highest proportions of correct diagnosis of tuberculosis (69.6% and 73.9%) and better inter-observer agreement ( $k = 0.71$ ) for PTB diagnosis. Observers were

---

\*Corresponding author.

more in agreement for the detection of nodules (0.32 - 0.74), adenomegalies (0.43 - 0.69), and for the diagnosis of cancer (0.22 - 1) than for the diagnosis of tuberculosis (0.19 - 0.71). Disagreements were more frequent for the detection of pleural effusions (-0.08 - 0.73). Conclusion: The inter-observer agreement varies with the type of lesions and diagnosis. Pneumologists were most effective for the diagnosis of pulmonary tuberculosis. Observers were more in agreement for the detection of nodules and the diagnosis of cancer than for the diagnosis of pulmonary tuberculosis.

## Keywords

Inter-Observer Variability, Concordance, Pulmonary Tuberculosis, Nodules, Caverns, Lung Cancer, Chest Radiography, *Kappa*

## 1. Introduction

Chest X-ray (CXR) is the most prescribed radiography in developing countries. It plays a major role in management of many thoracic diseases [1]. It is widely used in the preoperative workup and screening [2] [3], and it provides evidence for prescription of many chest CT-scans [1] [4].

The complexity of CXR image is source of great variability in the detection and interpretation of pulmonary anomalies between readers [1] [5]-[7]. Several studies have been conducted to assess the variability in the diagnosis of pulmonary tuberculosis (PTB) [8]-[10], of pneumoconiosis [2] [3], of lung cancers [4] [11], of pneumonia [8] [12] [13] and of lung nodules [7]. CXR is an important tool in the diagnosis of pulmonary tuberculosis [10] [14] [15]. In developing countries, where CXR is usually the only available or accessible chest imaging test and where PTB is endemic [16] [17], inter-observer variability in the interpretation of CXR has not been studied. The purpose of this study was to assess the concordance in reading adult CXRs between radiologists, pneumologists and senior residents in medical imaging and to determine their effectiveness in suggesting diagnosis of pulmonary tuberculosis among other chest diseases.

## 2. Materials and Methods

It was a cross-sectional quasi-observational study carried out in Yaounde (capital city of Cameroon) from January to March 2014, including six observers and 47 CXRs.

### 2.1. Selection of Observers

Six observers, all working in university-affiliated hospitals, selected by convenience agreed to participate in the study: two pneumologists of six and 14 years of experience, two radiologists of two and six year's experience, and two senior residents in medical imaging. Readers were categorized as "radiologist", "pneumologists" and "residents". Pneumologists were from the department of pneumology B of the Yaounde Jamot Hospital (YJH). Arbitrarily, observer designated by "1" was the one who had the best score for the detection and diagnosis in its category. The inter-observer agreement was calculated between readers 1 and 2 for each category, and between reader 1 of a giving category and reader 1 of the other category.

### 2.2. Selection of Chest Radiographs

Technically adequate frontal chest radiographs of patients 25 years old and above were selected in the department of pneumology A of the YJH, the highest referral and treatment center for respiratory diseases in Yaounde and its neighborhoods. All abnormal CXRs had a definitive diagnosis of concerned disease by suitable means (e.g. PTB confirmed by positive sputum smear).

Radiographs of patients with pulmonary tuberculosis (PTB) were chosen among the first 50 PTB patients hospitalized in the service during the year 2013. Normal CXRS or including diseases other than tuberculosis were consecutively selected from the files of patients treated in outpatient pneumology. Informatics treatments were performed to cancel name of patients and date of examination on all the CXR images.

### 2.3. Reading of Radiographs

The consensus interpretation was obtained by reviewing all selected CXRs by a group consisting of one radiologist and one pneumologist (nine years experience each) and one senior resident in radiology. Members of this committee did not participate as readers to the study. Consensus interpretation determined for each case basic radiographic lesions and radiological diagnosis. The following were selected for analysis of concordance: two types of parenchymal lesions (16 cases of nodules and 12 cases of cavitary lesions), 6 CXRs with pleural effusion and 6 others with hilar or mediastinal adenomegaly, 23 cases with the diagnosis of PTB and 7 of lung cancer.

A total of 47 radiographs were selected for this study: 4 of diffuse infiltrative lung disease, 6 normal, 7 of lung cancer, 7 of bacterial pneumonia and 23 of pulmonary tuberculosis. Reading of CXRs by the six observers was performed on the same computer and using a report form with part for description of detected lesions and part for radiological diagnosis. Reading time was not limited, and each observer chose its convenient time to read the same 47 CXRs.

### 2.4. Data Collections and Analysis

The sample size [18]-[22] was calculated using the “KappaSize” R Version 2.13.0 statistical software. Based on expected Kappa of  $0.47 \pm 0.13$  [8], and a type I error of 0.05, the minimum sample size was 46 radiographs for six observers and five diagnostic possibilities. The data were entered and analyzed using SPSS 17 software. These Kappa (K) intervals and thresholds [23] were used to measure inter-observer agreement: discordance ( $<0.0$ ), low (0.0 - 0.20), poor (0.21 - 0.40), moderate (0.41 - 0.60), good (0.61 - 0.80), excellent ( $>0.81$ ). For each observer, the score of detection of a given lesion was the number of correctly detected lesions over the total number of that lesion detected during consensus reading.

This study was approved by the Ethics Committee of the Faculty of Medicine and Biomedical Sciences and the administrative authorities of the Yaounde Jamot Hospital.

## 3. Results

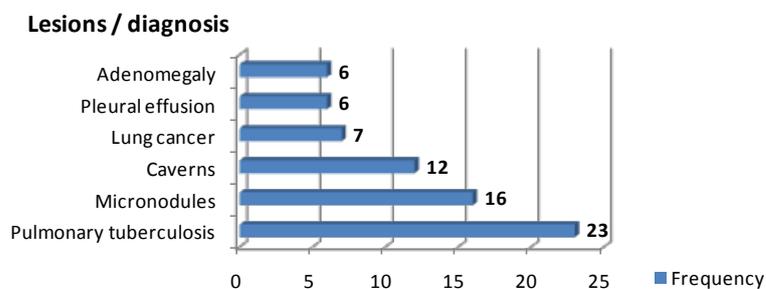
The most common lesion was pulmonary nodules (16/47). **Figure 1** shows anomalies and diagnosis on which the Kappa coefficients of agreement were calculated.

The performances in detection and diagnosis of CXRs anomalies for each observer are shown in **Table 1**.

The average score of correct results was 42.3% with variable proportions between different observers, and for the same reader from one lesion or diagnosis to another. The radiologist 1 had the highest average score of correct results (53.5%) with excellent detection of pleural effusions. The average score for the detection of caverns was the highest (58.3%). Pneumologists had the best proportions of correct diagnosis of tuberculosis (69.6% and 73.9%).

Kappa coefficients of inter-observer agreement in the detection of elementary lesions between different observers are shown in **Table 2**.

The highest Kappa coefficient was found in the agreement between the residents for the detection of pleural effusions ( $k = 0.73$ ) and between radiologist and pneumologists in the detection of nodules ( $k = 0.74$ ). Observers were more in agreement for the detection of nodules and adenomegalies than for the detection of caverns and pleural effusions, with frequent disagreement in the detection of pleural effusions. The inter-observer agreement



**Figure 1.** Frequency of lesions and diagnosis considered for study of inter-observer variability.

**Table 1.** Proportion of correct detected lesion and correct diagnosis for each observer.

Observers	ADP (n = 6)	Pleural eff. (n = 6)	Cavern* (n = 12)	Micronodule (n = 16)	Lung cancer (n = 7)	PTB (n = 23)	Mean score
	%	%	%	%	%	%	%
Resident 1	16.7	50.0	58.3	12.5	57.1	39.1	38.9
Resident 2	16.7	33.3	58.3	6.3	57.1	34.8	34.4
Radiologist 1	50.0	100	50.0	43.7	42.8	34.8	53.5
Radiologist 2	16.7	16.7	66.7	62.5	14.3	43.5	36.7
Pneumologist 1	33.3	50.0	41.7	56.2	57.1	73.9	52.0
Pneumologist 2	16.7	0	75.0	37	28.6	69.6	37.9
Mean score	25.0	41.7	58.3	36.4	42.8	49.3	42.3

ADP = hilar or mediastinal adenomegaly; Pleural eff. = pleural effusion; caverns = cavitary lesions; PTB = pulmonary tuberculosis.

**Table 2.** Kappa coefficient (CI 95%) of inter-observer agreement in the detection of elementary lesions.

Observers	Adenomegaly	Cavern	Pleural effusion	Micronodule
Resident 1/resident 2	0.69 (0.36 - 1.02)	0.26 (-0.08 - 0.60)	0.73 (0.37 - 1.09)	0.69 (0.47 - 0.91)
Radiologist 1/radiologist 2	0.55 (0.20 - 0.89)	0.25 (-0.09 - 0.59)	-0.08 (-0.14 - -0.02)	0.32 (0.01 - 0.64)
Pneumologist 1/pneumologist 2	0.61 (0.30 - 0.91)	0.45 (0.10 - 0.79)	0.34 (-0.05 - 0.73)	0.53 (0.25 - 0.81)
Resident 1/radiologist 1	0.43 (0.05 - 0.80)	0.09 (-0.22 - 0.40)	-0.08 (-0.14 - -0.02)	0.64 (0.40 - 0.87)
Resident 1/pneumologist 1	0.69 (0.36 - 1.01)	0.50 (0.14 - 0.85)	0.23 (-0.22 - 0.68)	0.47 (0.14 - 0.68)
Radiologist 1/pneumologist 1	0.55 (0.20 - 0.89)	0.25 (-0.09 - 0.59)	-0.07 (-0.12 - -0.01)	0.74 (0.41 - 0.98)

ranged from poor to good ( $k = 0.32$  to  $0.74$ ) for the detection of nodules, and moderate to good ( $k = 0.43$  to  $0.69$ ) for the detection of adenomegalies.

Kappa coefficients of inter-observer agreement in the diagnosis of pulmonary tuberculosis and bronchopulmonary cancer among readers are shown in **Table 3**.

The agreement for the diagnosis of tuberculosis was higher among pneumologists ( $k = 0.71$ ). Observers were more consistent for the diagnosis of cancer than for that of tuberculosis. The inter-observer agreement was excellent ( $k = 1$ ) between resident and pneumologist for the diagnosis of lung cancer and good between pneumologists ( $k = 0.71$ ) for the diagnosis of tuberculosis.

#### 4. Discussion

This study shows that the agreement between observers varies with the type of lesion and diagnosis. Observers were more in agreement for the detection of nodules and adenomegalies (ADP). Disagreement was most frequent regarding the detection of pleural effusions. Observers more agreed for the diagnosis of cancer and for that of tuberculosis.

Cascade *et al.* [6] in a study of competence of chest and nonchest radiologists in interpreting chest radiographs founded no difference in clinically important missed diagnoses among chest radiologists, but a statistically significantly higher rate of seemingly obvious misdiagnoses for nonchest specialty radiologists. When evaluating the reliability and validity of chest radiographs in the diagnosis of tuberculosis by 25 physicians of varying qualifications, Kumar *et al.* [15] in Nepal founded the overall sensitivity and specificity of CXR of 78% and 51% respectively, and a poor agreement between the best physician and the best radiologist. They concluded of an unsatisfactory sensitivity and specificity of chest x-rays in the diagnosis of pulmonary tuberculosis.

In establishing the performance of chest X-ray (CXR) in all suspects of tuberculosis (TB), a study by van Cleeff *et al.* [14] showed 89% agreement ( $K = 0.75$ ) for the combined scores "TB" or "no-TB".

**Table 3.** Kappa coefficient (CI 95%) of inter-observer in the diagnosis of tuberculosis and lung cancer.

Observers	Pulmonary tuberculosis	Lung cancer
Resident 1/resident 2	0.60 (0.25 - 0.94)	<b>0.72</b> (0.23 - 1.20)
Radiologist 1/radiologist 2	0.57 (0.24 - 0.89)	<b>0.66</b> (0.13 - 1.18)
Pneumologist 1/pneumologist 2	<b>0.71</b> (0.40 - 1.01)	0.22 (-0.197 - 0.63)
Resident 1/radiologist 1	0.25 (-0.15 - 0.65)	<b>0.72</b> (0.23 - 1.2)
Resident 1/pneumologist 1	0.34 (0.08 - 0.59)	<b>1</b>
Radiologist 1/pneumologist 1	0.19 (-0.08 - 0.46)	<b>0.72</b> (0.23 - 1.20)

For the detection of ADP, inter-observer agreement was good ( $K = 0.61$ ) between pneumologists and moderate ( $K = 0.55$ ) between radiologists. In a similar study conducted in 2010 in London, Abubakar *et al.* [24] reported a poor agreement between both pneumologists and radiologists. The frequent association of ADP with pulmonary tuberculosis and the high prevalence of tuberculosis in our environment could explain this difference.

The average score for the detection of caverns was the highest (58.3%). This can be justified by the fact that in an endemic TB area, our observers are used to see the caverns that are very common in pulmonary tuberculosis in the tropics [16] [17] [25] [26]. However, the agreement in the detection of caverns ranged from poor to moderate ( $K = 0.25$  to  $0.50$ ). It could be that the observers were in disagreement on some caverns. In addition, the detection of these depends on their location, size, and content, and wall thickness. Balabanova *et al.* [10] had found the similar inter-observer agreement between pneumologists and radiologists in Russia.

Disagreements were common in the detection of pleural effusions in contrast to the results from Shinsaku *et al.* [5] in Japan and Abubakar *et al.* [24] in London who found a moderate to excellent level of agreement. The predominance of low abundance pleural effusions in our sample ( $n = 5/6$ ) and their association with other abnormalities could explain these discordances.

Radiologists had a poor agreement in the detection of nodules ( $K = 0.32$ ). Dawson *et al.* [27] founded similar results ( $K = 0.32$ ) between radiologists in South Africa. Anna Ralph *et al.* [28] observed low concordance ( $K = 0.12$ ) between radiologists in Australia. In our study, observers were generally more consistent for the detection of nodules than for the detection of caverns even if individual performances were better in the detection of caverns. Both lesions are very common in pulmonary tuberculosis [25] [26].

Pneumologists had good agreement ( $K = 0.71$ ) in the diagnosis of tuberculosis, while radiologists had moderate agreement ( $K = 0.57$ ). This difference could be explained by the experience of pneumologists (6 and 14 compared to 2 and 6 years for radiologists) but also by their specific activities allowing them more experience in PTB patients. Indeed, in our setting, almost all TB patients converge to pneumologists who would on average seen much more chest radiographs of tuberculosis than radiologists. Dawson *et al.* reported a good agreement between radiologists in South Africa in 2010 [27]. Abubakar *et al.* [24] observed a moderate agreement for both radiologists and pneumologists. Balabanova *et al.* [10] had reported a poor agreement between pneumologists and moderate between radiologists. However, in our study, the pneumologists because of their specific activities were of risk of “*hindsight bias*”, especially since there were 23/47 cases of PTB. The two pneumologists were from the department B while the cases were selected in the department A to avoid “*memory bias*”.

In general, our observers were more consistent for the diagnosis of lung cancer than for that of tuberculosis. This could be justified by evocative radiological aspect of selected cases of cancer, and the highly variable radiological presentation of pulmonary tuberculosis.

In this study, the total number of cases was relatively low even if the distribution of anomalies and diseases, as well as observers recalled the normal daily exercise. This is to our knowledge, one of the few studies of inter-reader agreement in the interpretation of CXRs in Africa. It backs in the saddle the difficulties of interpretation of current imaging test regardless of experience and qualification of the observer, and it encourages more interdisciplinary collaboration.

The limitations of this study are the lack of use of a standardized interpretation grid and variable experience of readers (years of practice).

## 5. Conclusion

The inter-observer agreement varies with the type of lesion and diagnosis. Pneumologists were most effective for the diagnosis of tuberculosis. Observers were more in accord for the detection of nodules and the diagnosis of cancer than for the detection of pleural effusions and diagnosis of tuberculosis. The use of a standardized interpretation scheme is recommended to improve detection and reading concordance between different observers.

## Declaration of Interest

The authors declare to have no competing interest in relation to this article.

## References

- [1] Remy, J., Remy-Jardin, M., Bonnel, F., Masson, P. and Mastora, I. (2001) L'interprétation d'une radiographie thoracique revue et corrigée par la tomodensitométrie. *Journal de Radiologie*, **82**, 1067-1679.
- [2] Gonsu Fotsin, J., Blackett Ngu, K., Ndjitoyap Ndam, E.C., Youmbissi, J.T., Simo Moyo, J. and Malonga, E. (1990) La radiographie thoracique pré-opératoire à l'Hôpital Central de Yaoundé: Etude analytique de 1969 clichés. *Médecine d'Afrique Noire*, **37**, 310-313.
- [3] Moifo, B., Tambe, J., Pefura Yone, E.W., Zeh, O.F., Gonsu Kamga, J.E. and Gonsu Fotsin, J. (2012) Assessing the Role of Routine Chest Radiography in Asymptomatic Students During Registration at a University in an Endemic Area of Tuberculosis. *Annals of Tropical Medicine and Public Health*, **5**, 419-422. <http://dx.doi.org/10.4103/1755-6783.105122>
- [4] Neossi Guena, M., Moifo, B., Pefura Yone, E.W., Mankaa Wankie, M., Rémy-Jardin, M., Rémy, J., *et al.* (2011) Influence des protocoles d'examen sur la qualité et les performances diagnostiques d'une TDM thoracique: Expérience du service d'imagerie thoracique du CHRU de Lille (France). *Journal Africain d'Imagerie Médicale*, **6**, 277-289.
- [5] Sakurada, S., *et al.* (2012) Inter-Rater Agreement in the Anormal Chest X-Ray Findings for Tuberculosis between Two Asian Countries. *BMC Infectious Disease*, **12**, 1-8. <http://dx.doi.org/10.1186/1471-2334-12-31>
- [6] Cascade, P.N., Kazerooni, E.A., Gross, B.H., Quint, L.E., Silver, T.M., Bowerman, R.A., Pernicano, P.G. and Gebremariam, A. (2001) Evaluation of Competence in the Interpretation of Chest Radiographs. *Academic Radiology*, **8**, 315-321. [http://dx.doi.org/10.1016/S1076-6332\(03\)80500-7](http://dx.doi.org/10.1016/S1076-6332(03)80500-7)
- [7] Armato III, S.G., McNitt-Gray, M.F., Reeves, A.P., Meyer, C.R., McLennan, G., *et al.* (2007) The Lung Image Database Consortium (LIDC): An Evaluation of Radiologist Variability in the Identification of Lung Nodules on CT Scans. *Academic Radiology*, **14**, 1409-1421. <http://dx.doi.org/10.1016/j.acra.2007.07.008>
- [8] Den Boon, S., Bateman, E.D., Enarson, A.D., Borgdorff, N.W., Verver, S., Lombard, C.J., *et al.* (2005) Development and Evaluation of a New Chest Radiograph Reading and Recording for Epidemiological Surveys of Tuberculosis and Lung Disease. *International Journal of Tuberculosis and Lung Disease*, **9**, 1088-1096.
- [9] Graham, S., Das Gupta, K., Hidvegi, R.J., Hanson, R., Kosiuk, J., Al Zahrani, K., *et al.* (2002) Chest Radiograph Abnormalities Associated with Tuberculosis: Reproducibility and Yield of Active Cases. *International Journal of Tuberculosis and Lung Disease*, **6**, 137-142.
- [10] Balabanova, Y., Cocker, R., Fedorin, I., Zakharova, S., Plavinskij, S., Krukov, N., *et al.* (2005) Variability in Interpretation of Chest Radiographs among Russian Clinicians and Implications for Screening Programmes: Observational Study. *British Medical Journal*, **331**, 1-4. <http://dx.doi.org/10.1136/bmj.331.7513.379>
- [11] Lorentz Quekel, G.B.A., Alphons Kessels, G.H., Reginald, G. and Engelshoven, M.J.V. (2001) Detection of Lung Cancer on the Chest Radiograph: A Study on Observer Performance. *European Journal of Radiology*, **39**, 111-116. [http://dx.doi.org/10.1016/S0720-048X\(01\)00301-1](http://dx.doi.org/10.1016/S0720-048X(01)00301-1)
- [12] Hopstaken, R.M., Witbraad, T., Van Engelshoven, J.M.A. and Dinant, G. (2004) Inter-Observer Variation in the Interpretation of Chest Radiographs for Pneumonia in Community-Acquired Lower Respiratory Tract Infections. *Clinical Radiology*, **58**, 478-481.
- [13] Cherian, T., Mulholland, E.K., Carlin, J.B., Ostensen, H., Amin, R., de Campo, M., *et al.* (2005) Standardized Interpretation of Paediatric Chest Radiographs for the Diagnosis of Pneumonia in Epidemiological Studies. *Bulletin of the World Health Organization*, **191**, 353-359.
- [14] van Cleeff, M.R., Kivihya-Ndugga, L.E., Meme, H., Odhiambo, J.A. and Klatser, P.R. (2005) The Role and Performance of Chest X-Ray for the Diagnosis of Tuberculosis: A Cost-Effectiveness Analysis in Nairobi, Kenya. *BMC Infectious Diseases*, **5**, 111. <http://dx.doi.org/10.1186/1471-2334-5-111>
- [15] Kumar, N., Bhargava, S.K., Agrawal, C.S., George, K., Karki, P. and Baral, D. (2005) Chest Radiographs and Their Reliability in the Diagnosis of Tuberculosis. *Journal of Nepal Medical Association*, **44**, 138-142.

- [16] OMS (2013) Global Tuberculosis Report 2012. [http://apps.who.int/iris/bitstream/10665/75938/1/9789241564502\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/75938/1/9789241564502_eng.pdf)
- [17] Boulahbal, F. and Chaulet, P. (2004) La tuberculose en Afrique: Epidémiologie et mesures de lutte. *Médecine Tropicale*, **64**, 224-228.
- [18] Blum, A., Feldmann, L., Brester, F., Jouanny, P., Briançon, S. and Régert, D. (1995) Intérêt du calcul du coefficient Kappa dans l'évaluation d'une méthode d'imagerie. *Journal of Radiology*, **7**, 441-443.
- [19] Fermanian, J. (1984) Mesure de l'accord entre deux juges. Cas qualitatif. *Revue d'Epidémiologie et de Santé Publique*, **32**, 140-147.
- [20] Anthony Viera, J. and Garrett, J.M. (2005) Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*, **37**, 360-363.
- [21] Sim, J. and Wright, C.C. (2005) The Kappa Statistic in Reliability Studies: Interpretation, and Sample Size Requirements. *Physical Therapy*, **85**, 257-268.
- [22] Whitley, E. and Ball, J. (2002) Sample Size Calculations. *Critical Care*, **6**, 335-341. <http://dx.doi.org/10.1186/cc1521>
- [23] Landis, J.R. and Koch, G. (1977) The Measurement of Observer Agreement for Categorical Data. *Biometrics*, **33**, 159-174. <http://dx.doi.org/10.2307/2529310>
- [24] Abubakar, I., Story, A., Lipman, M., Bothamley, G., Van Hest, R., Andrews, N., et al. (2010) Diagnosis Accuracy of Digital Chest Radiograph for Pulmonary Tuberculosis in a UK Urban Population. *European Respiratory Journal*, **35**, 689-692. <http://dx.doi.org/10.1183/09031936.00136609>
- [25] Zteou, B., Ghadouani, F., Tizniti, S., Mahla, H., Amara, B., Elbiaze, M. and Benjelloun, M.C. (2007) La tuberculose thoracique dans sa forme typique et atypique. SFR 2007. <http://pe.sfrnet.org/Data/ModuleConsultationPoster/pdf/2007/1/8156b94b-1ee7-4896-9966-d3b8021dc2da.pdf>
- [26] McAdams, H.P., Erasmus, J. and Winter, J.A. (1995) Radiologic Manifestations of Pulmonary Tuberculosis. *Radiologic Clinics of North America*, **33**, 655-678.
- [27] Dawson, R., Masuka, P., Edwards, D.J., Bateman, E.D., Bekker, L.G., Wood, R., et al. (2010) Chest Radiograph Reading and Recording System: Evaluation for Tuberculosis Screening in Patients with Advanced HIV. *International Journal of Tuberculosis and Lung Disease*, **14**, 52-58.
- [28] Ralph, A.P., Ardian, M., Wiguna, A., Maguire, G.P., Becker, N.G., Drogumuller, G., et al. (2010) A Simple, Valid, Numerical Score for Grading Chest X-Ray Severity in Adult Smear-Positive Pulmonary Tuberculosis. *Thorax*, **65**, 863-869. <http://dx.doi.org/10.1136/thx.2010.136242>