Comparing Chest X-rays with Ultrasound for the Prediction of Left Atrial Size at Pretoria Academic Hospital

by

Susanna Jacoba Quinton

Presented in partial fulfilment of the requirements for the degree
Master of Science in Clinical Epidemiology

In the Faculty of Health Sciences

University of Pretoria

Pretoria

August 2006

Acknowledgements

Authorship and a Co-workers

First Author:	Dr SJ Quinton	
Co-authors:	Prof P Rheeder	(Study Supervisor)
	Prof JA Ker	(Reviewer)
	Dr A Deffur	(Reviewer)
Co-worker:	Dr P Soma	

I would like to express my gratitude and appreciation to the abovementioned doctors without whom this study would not have been possible.

Declaration

I hereby declare that this dissertation presented to the University of Pretoria for the of Masters Science in Clinical Epidemiology degree is my own work and has not been presented previously to any other tertiary institution for any degree.

Abstract

Comparing chest X-rays with ultrasound for the prediction of left atrial size at Pretoria Academic Hospital

Student: Susanna Jacoba Quinton

Promoter: Prof P Rheeder

Division: Clinical Epidemiology
Faculty of Health Sciences

Degree: MSc (Clinical Epidemiology)

Introduction:

Estimates of left atrial size in patients with suspected cardiac disease play an important role in diagnostic medicine. Left atrial size is used in predicting prognosis and events, as well as treatment decisions. Two methods are commonly used to estimate left atrial size: chest radiography and cardiac ultrasound. This study aims to determine the test characteristics of chest radiography and compare the use of radiographs to cardiac ultrasound (the gold standard test).

Methods:

Data from patients older than 18 years admitted to Pretoria Academic Hospital during 2000-2003 who had both chest X-rays and cardiac ultrasound were included in this cross-sectional, retrospective analysis. Chest X-rays were classified into three quality classes, and the sub-carinal angle (SCA) and sub-angle distance (SAD) were measured twice in all available X-rays by two observers. Intra- and interobserver variability (3 methods) as well as the predictive value of the SCA and SAD measurements were determined using logistic regression (with left atrial size determined by ultrasound as comparator). P-values < 0.05 were regarded as statistically significant for all comparisons.

Results:

Data for 159 patients were available (154 cardiac ultrasounds and 178 chest radiographs). Intraobserver variability regarding chest X-ray measurements was low with almost perfect concordance (P=0.000). Interobserver variability was higher for supine X-rays. Using logistic regression, a linear model was identified which was statistically significant only for erect X-rays. While goodness-of-fit analysis showed that the model fits the data, performance characteristics were poor, with high sensitivity and low specificity, and an area under the ROC curve of 0.62-0.63, depending on type of X-ray and measurement (SCA or SAD). Linearity in the logit of the dependent variable was assessed, and found to be present at the extremes of carinal angle measurements for the supine X-ray data and in the first three quartiles for erect X-ray data. A non-linear model determined by fractional polynomial analysis did not perform significantly better than the original linear model. Cut-off values for the SCA of 72° and 84° (erect and supine X-rays, respectively) were found to give the best compromise between sensitivity and specificity. The corresponding cut-off values for SAD were 24.1mm and 26.9mm.

Interpretation:

Assessment of either SCA or SAD to determine left atrial size is equivalent and repeatable, both within the same observer, and between two observers (less so for supine X-rays). While this measure is precise, it was found not to be very accurate. Therefore, chest X-rays are not reliable in predicting left atrial enlargement.

Abstrak

Die betroubaarheid van borskasplate in die voorspelling van linker atrium vergroting te Pretoria Akademiese Hospitaal.

Student: Susanna Jacoba Quinton

Promotor: Prof P Rheeder

Divisie: Kliniese Epidemiologie
Fakulteit van Gesondheidswetenskappe

Graad: MSc (Kliniese Epidemiologie)

Inleiding:

Die raming van die linker atrium grootte in pasiënte met vermoedelike hartsiekte speel ‘n belangrike rol in diagnostiese medisyne. Linker atrium grootte word gebruik in die voorspelling van prognose en gebeurlikhede sowel as in besluitneming rakende behandeling. Twee metodes word algemeen gebruik om linker atrium grootte te bepaal: borskasplate en hartsonar. Hierdie studie beoog om die toets karakteristieke van borskasplate te bepaal en om die bruikbaarheid van borskasplate met hartsonar (die goudstandaard) te vergelyk.

Metode:

Data van pasiënte ouer as 18 jaar wat opgeneem was te Pretoria Akademiese HospitaaL gedurende 2000 -2003 en beide borskasplate en ‘n hartsonar tydens opname ondergaan het, is ingesluit in hierdie deursnit retrospektiewe studie. Borskasplate is in drie kwaliteitsklasse geklassifiseer en die sub-carina hoek (SCH) asook die sub-hoek afstand (SHA) is tweemaal gemeet in alle beskikbare plate deur twee waarnemers. Intra- en interwaarnemer ooreenstemming (drie metodes) sowel as voorspellingswaarde van die SCH en SHA meetings is bepaal met behulp van logistiese regressie (met linker atrium grootte op hartsonar as goudstandaard). P-waardes < 0.05 is beskou as statisties betekenisvol vir alle vergelykings.

Resultate:

Data van 159 pasiënte was beskikbaar (154 hartsonars en 178 borskasplate). Intrawaarnemer ooreenstemming ten opsigte van borskasplate was uitstekend met bykans perfekte konkordansie (P=0.000). Interwaarnemer ooreenstemming was swakker ten opsigte van liggende borskasplate. Logistiese regressie het 'n lineêre model identifiseer wat statisties betekenisvol was vir staande borskasplate alleenlik. Alhoewel die model gepas was vir die data volgens die passingsanaliese, was die uitkoms karakteristieke swak, met hoë sensitiwiteit en lae spesifisiteit, en 'n sub-ROC kurwe area van 0.62 – 0.63, afhangende van die tipe borskasplaat en meting (SCH of SHA). Die afhanklike veranderlike was lineêr in sy verspreiding vir die ekstreme meetings van SCH vir liggende plate en in die eerste drie kwartiele vir staande plate. 'n Non-liniêre model soos bepaal deur fraksionele polinomiese analiese het nie beduidend beter gevaar as die oorspronklike liniêre model nie. Afsnywaardes vir die SCH van 72^{o} en 84^{o} (staande en liggende plate, respektiewelik) het die beste kompromie tussen sensitiwiteit en spesifisiteit gelewer. Die ooreenstemmende afsnywaardes vir SHA was 24.1 mm en 26.9 mm.

Interpretasie:

Meting van óf die SCH óf die SHA om die linker atrium grootte te bepaal is ekwivalent en herhaalbaar, beide vir dieselfde waarnemer sowel as tussen twee waarnemers (minder so vir liggende plate). Alhoewel hierdie meetings herhaalbaar is, is dit nie beduidend akkuraat nie. Dus, borskasplate is nie betroubaar om linker atrium vergroting te voorspel nie.

Contents

LIST OF TABLES 2

LIST OF FIGURES 3

TITLE 5

AIMS 5

LITERATURE REVIEW 6

STUDY METHODOLOGY 9

Setting 9

Subjects : 9

Study Design :..... 9

Study Outline :..... 10

Ethics :..... 10

Measurements :..... 10

Sample Size : :..... 11

Data Analysis :..... 12

RESULTS :..... 13

DISCUSSION :..... 40

CONCLUSION :..... 44

REFERENCES :..... 45

List of Tables

Table 1.1	Patient demographics	13
Table 1.2	Radiograph characteristics	13
Table 2.1	Intrarater Repeatability of Examiner1	19
Table 2.2	Intrarater Repeatability of Examiner2	22
Table 2.3	Intrarater Repeatability Between Examiner1 and Examiner2	25
Table 3.1	Linear Model of Left Atrial Size Estimators	26
Table 3.2	Logistic Regression Results	27
Table 3.3	Fractional Polynomials	33
Table 3.4	Diagnostic Characteristics of Original and Transformed Variables	34

List of Figures

Figure 1.1	Measurement of sub-carinal angle (SCA)	11
Figure 1.2	Measurement of sub-angle distance (SAD)	11
Figure 2.1	Breakdown and descriptives of study sample	14
Figure 2.1a	Breakdown of patients and echocardiograms	15
Figure 2.1b	Breakdown of female patients and their radiographs	16
Figure 2.1c	Breakdown of male patients and their radiographs	17
Figure 2.1d	Radiograph types and technical quality	18
Figure 3.1	Intrarater repeatability of Examiner1 using LIN's Correlation Coefficient	20
Figure 3.2	Intrarater repeatability of Examiner1 using LOA	20
Figure 3.3	Intrarater repeatability of Examiner1 using Bland & Altman Plot	21
Figure 3.4	Intrarater repeatability of Examiner2 using LIN's Correlation Coefficient	23
Figure 3.5	Intrarater repeatability of Examiner2 using LOA	23
Figure 3.6	Intrarater repeatability of Examiner2 using Bland & Altman plot	24
Figure 4.1	Exploring nature of relationship between Gold standard and mean SCA of Examiner1 for erect good quality X-rays	28
Figure 4.2	Exploring nature of relationship between Gold standard and mean SCA of Examiner1 for supine good quality X-rays	29
Figure 4.3	Exploring nature of relationship between Gold standard and mean SAD of Examiner1 for erect good quality X-rays	29

Figure 4.4	Exploring nature of relationship between Gold standard and mean SAD of Examiner1 for supine good quality X-rays	30
Figure 4.5	Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SCA of erect X-rays	30
Figure 4.6	Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SCA of supine X-rays	31
Figure 4.7	Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SAD of erect X-rays	31
Figure 4.8	Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SAD of supine X-rays	32
Figure 4.9	Comparison of ROC curves of original and transformed variable	34
Figure 4.10	Exploring the nature of the relationship between Gold standard and transformed mean SCA of Examiner1 for erect good quality X-rays	35
Figure 4.11	Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the transformed mean SCA of erect X-rays	35
Figure 4.12	Cut-off point for SCA on erect X-rays using *roctg2*	36
Figure 4.13	Cut-off point for SCA on supine X-rays using *roctg2*	37
Figure 4.14	Cut-off point for SAD on erect X-rays using *roctg2*	38
Figure 4.15	Cut-off point for SAD on supine X-rays using *roctg2*	*39*

TITLE

Comparing Chest X-rays with Ultrasound for Prediction of Left Atrial Size: Pretoria Academic Hospital.

AIMS

1. To determine intra- and inter-observer variability of chest X-ray measurements for left atrial size.

2. To determine the test characteristics of chest X-rays for determining left atrial size when compared with ultrasound.

LITERATURE REVIEW

Left atrial size is used in predicting prognosis and events, as well as deciding on treatment.

PROGNOSIS

As far as prognosis is concerned, several studies have been done to determine the usefulness of left atrial size.

Rossi A et al.[1] reports on the "Usefulness of left atrial size in predicting postoperative symptomatic improvements in patients with aortic stenosis". Symptoms of congestive heart failure may persist postoperatively, despite surgery proven as highly effective treatment for symptomatic relief in patients with aortic stenosis. In a group of patients with aortic stenosis characterised by a wide range of left atrial size, a correlation was found between the latter and postoperative symptomatic improvement. Left atrial size is therefore useful in predicting postoperative symptomatic improvements in patients with aortic stenosis.

An echocardiographic study, conducted by Svanegaard J et al.[2] investigated the relationship between left atrial size and atrial natriuretic peptide after acute myocardial infarction. In this study serial echocardiographic examinations of the left atrium were compared with measurements of plasma atrial natriuretic peptide. A significant correlation between the two was found at both 10-12 days post infarction, as well as at 6 months after. Based on a correlation coefficient of 0.70 it was concluded that the percentage change in the size of the left atrial can reliably predict the percentage change in atrial natriuretic peptide after an acute myocardial infarction. Morbidity and mortality can subsequently be inferred from the left atrial size.

A long term follow-up study by Modena MG et al.[3] was conducted to determine the influence that left atrial size had on prognosis of patients with dilated cardiomyopathy. In this study, echo-derived atrial dimension was found to be the major predictor of cardiac death and overall clinical outcome compared to other echo-cardiographic, clinical and hemodynamic parameters at time of entry into the study.

In “The long-term effect of successful mitral balloon valvotomy on left atrial size”, Stefadouros MA et al.[4] explored the long-term effect on left atrial size of successful mitral balloon valvotomy. This investigation leads to the conclusion that successful mitral balloon valvotomy results in significant long-term reduction in left atrial size in most patients.

LEFT ATRIAL SIZE AS A PREDICTOR OF EVENTS

A number of studies have been conducted to explore the usefulness of left atrial size as a predictor of events. According to Ozer N et al.[5] left atrial appendage area is increased in size in patients with atrial fibrillation, atrial fibrillation being a known increased risk of cardioembolic stroke. Therefore left atrial size could be useful as a predictor of cardioembolic stroke.

An increased left atrial size correlates with an increased risk of stroke in patients with sinus rhythm. This was determined in a study by Sadanandan S et al.[6] Mattioli AV et al[7] conducted a serial evaluation of left atrial dimension and function after cardioversion for atrial fibrillation. They found that left atrial size decreased after restoration of sinus rhythm in all patients and that a higher atrial Ejection Force was associated with a more marked reduction in left atrial size.

In an ethnically mixed population, Di Tullio MR et al.[8] found that left atrial size is proportionate to the risk of ischemic stroke. Left atrial size is also a major factor in initiation & maintenance of atrial fibrillation. This was concluded from a study by Sankar NM et al.[9] on left atrial reduction for chronic atrial fibrillation associated with mitral valve disease. It was also found that left atrial size is proportionate to left ventricular mass in obese patients according to Gottdiener JS et al.[10] Left ventricular mass is associated with an adverse outcome with regard to cardiac effects. Left atrial size can be used to predict the severity of mitral regurgitation as shown in a study conducted by Katz ES et al.[11]

TREATMENT

Left atrial size in atrial fibrillation predicts response to a rhythm control therapeutic strategy. In a study conducted by Brodsky MA at al.[12] to explore the possible relationship between left atrial size and the success of treatment, it was found that sinus rhythm can be maintained after conversion, at left atrial size of 45-60 mm, if Amiodarone is added. In patients with larger left atrial dimensions, atrial fibrillation will probably return despite treatment.

PREVIOUS STUDIES ON THIS SUBJECT

A literature search (1987 - 2003) yielded two articles[13,14] on determining left atrium size by measuring the tracheal SCA and comparing that to echo findings. In the first study 35 patients with enlarged left atria (> 4.5cm) and 35 paired aged-matched patients with normal atria (< 4.0cm) were selected as determined by echo. The sample consisted of 37 female and 33 male patients aged 18 - 87 years. Interbronchial angle on chest X-rays (standard and supine portable films) was measured by blinded observer using a goniometer. In 90% of cases the SCA was adequately visualised. The SCA and left atrium size measurements were plotted on a scattered diagram. Their relationship was determined using a pearson correlation. Left atrium size could be accurately predicted to be larger than 5.0 cm in diameter if the SCA was greater than 100 degrees ($r = 0.746$ with $p < 0.001$). Although not clearly stated 7 patients (10%) were probably excluded from the study due to poor quality X-rays. The sample size of 63 is therefore too small to yield statistically significant results. Standardisation of plain-film techniques was not used. No intra- or inter-observer variation was tested. The sensitivity and specificity of X-rays as diagnostic tool was not determined.

In the second study the posteroanterior chest radiographs and echocardiographs of 108 clinically stable patients were respectively reviewed. The sample consisted of 53 men and 55 women, age ranging from 27 to 85 years. 43 patients had an enlarged left atrial dimension (defined as > 40mm) on sonar findings. Correlation analysis was used to determine which angle measurement best predicted the left atrial size. A threshold angle that predicted left atrial enlargement was derived by discriminant analysis. According to this study, left atrial size correlated poorly with both interbronchial angle ($r=0.33$) and SCA ($r = 0.25$) values. An interbronchial angle of 76,4° and a SCA of 65,4° were the best discriminators between patients with normal and those with enlarged left atrial dimensions. (sensitivities: 63% and 51%, specificities: 63% and 66%, for interbronchial angle and SCA respectively.) Although a larger study, the sample size is still too small if categorical data is to be used. Standardisation of plain film techniques was used and sensitivity and specificity were determined, but intra- and inter-observer variation was not tested.

Deleted: S
Deleted: S

METHODOLOGY

A literature search on methods used to determine left atrial size was conducted. The findings are deemed to reside more appropriately in the study methodology section and are therefore discussed under that heading.

Deleted: S
Deleted: M

STUDY METHODOLOGY

This study aims to investigate the reliability of chest X-rays to predict left atrial size by comparing to measurement by ultrasound as gold standard. Continuous data will be used and analyzed as such using appropriate statistical methods rather than categorising data, thereby reducing the sample size necessary to yield statistically significant results. Furthermore the Bland and Altman method, far superior in comparing two tests, will be utilised. The study will explore test characteristics like sensitivity, specificity, intra- and inter-observer variation. The results should come to a conclusion on the usefulness of chest X-rays as diagnostic tool in this regard.

SETTING

Patients admitted to medical wards of Pretoria Academic Hospital (PAH) 2000 – 2003 who had both chest X-rays and heart sonar done during hospital stay.

SUBJECTS

A co-worker selected the sampling-frame from records at the sonar department of heart sonars done on all in-patients above the age of 18 years. With hospital numbers of patients from this sampling frame, X-rays were retrieved from the radiology department. Patients with supine and/or erect PA chest X-rays of quality such that the sub-carinal angle is adequately visible were selected.

STUDY DESIGN

Cross-sectional diagnostic study conducted retrospectively.

STUDY OUTLINE

Demographics namely "age", "gender", "race", as well as presence or absence of heart disease and left atrial size (as determined by echo) was documented for each case. A number was assigned to each case of the sample and corresponding X-ray.

An independent expert classified the X-rays into three categories according to their quality, namely good, fair and poor. The principal investigator (blinded) then measured the sub-carinal angle of anonymous X-rays numbered by the co-worker. This process was repeated to determine intra-observer variability. A second blinded co-worker also measured the numbered X-rays twice to provide a figure on intra-observer as well as inter-observer variability.

ETHICS

Both observers were blinded, all data dealt with anonymously and the principle of confidentiality adhered to. The request to conduct this study as a retrospective study on known patient data by an investigator, who works in this unit, was granted by the ethics committee.

MEASUREMENTS

1. Ultrasound:

- Two-dimensional targeted N-mode echocardiography was performed using System 3GE ultrasound machine.
- Measurements of left atrial size were taken according to American Society of Echocardiography (ASE) criteria.[15]
- To determine left atrial size, the maximal dimension was measured from the parasternal long-axis view between the leading edge of the posterior aortic wall to the leading edge of the posterior wall of the left atrial at end-systole.
- In order to have accurate measurements (left atrial size may be underestimated in the parasternal long-axis view because this chamber may enlarge longitudinally),[16] left atrial size was measured from two apical orthogonal views (four-chamber and two-chamber) as well, from the tip of the mitral valve to the posterior wall of the left atrial at end-systole, and the highest of the values was taken into consideration.

Deleted: C

- An enlarge left atrial was defined by a maximal echocardiographic dimension of greater than 40mm.

2. Chest X-rays:

- X-rays were read on a radiographic viewing-box.
- The sub-carinal angle of divergence (SCA) **α** (in degrees) of the first few centimetres of the inferior main-stem bronchi borders was measured using a protractor (*Figure 1.1*).
- The sub-angle distance (SAD) **x** (in mm), on the opposite side of the sub-carinal angle α, was measured 20mm from the SCA along the medial borders of the bronchi **y** using a tape measure *(Figure 1.2*).

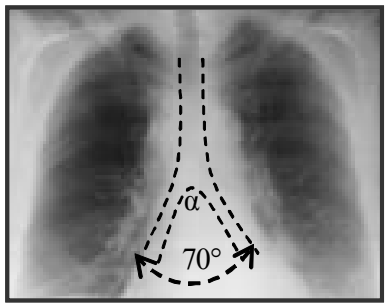


Figure 1.1 Measurement of SCA

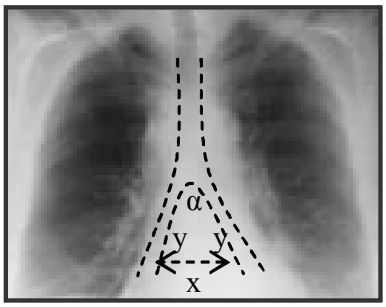


Figure 1.2 Measurement of SAD

- Observers measuring the SCA **α** using the technique described by Haskins and Goodman,[17] and distance **x** in the standardised method increases precision. In order to increase accuracy, observers were blinded.

SAMPLE SIZE

Nquery program was used to determine sample size. With α and power set at 5% and 90% respectively, δ (expected proportion of subjects with enlarged left atrial size) estimated at 70%, Ko (hypothetical perfect agreement between two methods) chosen as 90% and K1 (expected agreement between two methods) as 75%, the sample size was estimated at 106. Allowing for 10% of patients with heart sonars not having traceable chest X-rays and expecting 20% of SCA not to be clearly visible on chest X-rays, an initial sampling-frame of 150 was aimed for.

DATA ANALYSIS

1. Breakdown and descriptives of study sample portrayed as a flowchart and two tables.
2. Evaluation of intra-rater agreement between the 1st and 2nd measurement, of both the mean SCA as well as the mean SAD for both Examiner1 and Examiner2, for all types of X-rays, for only good quality X-rays, for erect X-rays, for erect X-rays excluding poor quality X-rays, for supine X-rays, for supine X-rays excluding poor quality X-rays. The inter-rater agreement of both the mean SCA as well as the mean SAD between Examiner1 and Examiner2, for all types of X-rays, for only good quality X-rays, for erect X-rays, for erect X-rays excluding poor quality X-rays, for supine X-rays, for supine X-rays excluding poor quality X-rays, using:
   a) LIN's concordance correlation coefficient.[18]
   b) Limits of agreement (Bland & Altman methodology)[19]
   c) Bland & Altman plots.
3. Logistic regression was used for erect and supine X-rays of good quality to determine the predictive value of the SCA and the SAD:
   a) Linear model.
   b) The fit of the model was assessed using Hosmer-Lemeshow Goodness-of-Fit test.[20] The diagnostic value of the angle and the hypotenuse was determined by sensitivity, specificity, positive predictive value, negative predictive value and finally by assessing the area under the ROC curve.
   c) To assess whether x was linear in the logit, three methods as proposed by Lemeshow[21] were used:
      i) Lowess Smoothing curve.
      ii) Design variables – Lincheck.
      iii) Fractional polynomials.
   d) Linear model using transformed variable.
   e) Optimal cut-points for mean SCA of examiner1 and mean SAD of Examiner1 were determined using receiver operating characteristic (ROC) curves.[22]

P values < 0.05 were regarded as statistically significant. The analysis was done using *Intercooled Stata* version 8.2.[23]

Deleted: C
Deleted: C
Deleted: C
Deleted: A
Deleted: M
Deleted: P
Deleted: T
Deleted: L
Deleted: L
Deleted: 3
Deleted: P

RESULTS

1. Study Sample

Echocardiography and radiography data were available on 159 patients. Five echocardiograms were incomplete as left atrial size was not measured, and therefore only 154 echocardiograms were included in the logistic regression analysis. As several patients had more than one chest radiograph taken, 178 chest radiographs were available for determination of intra- and inter-observer variability. The mean age of the study sample was 59 years (range 18 to 88 years).

Deleted: R

Table 1.1

Patient demographics

Patients

	Black	White	Coloured	Asian	**Total**
Male	39	34	0	1	74
Female	42	39	3	1	85
Total	81	73	3	2	**159**

Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt

Table 1.2

Formatted: Font: Bold, Italic
Formatted: Font: Bold, Italic

Radiograph characteristics

Chest radiographs

	Good	Fair	Poor	**Total**
Supine	43	1	3	47
Erect	115	10	6	131
Total	158	11	9	**178**

Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt
Formatted: Font: 12 pt

The detailed outcome of the 159 patients selected from the sonar department is portrayed in *Figure 2.1*. This figure serves as overview. Further *Figures 2.1a,b,c* and *d* are enlargements of parts of *Figure 2.1*.

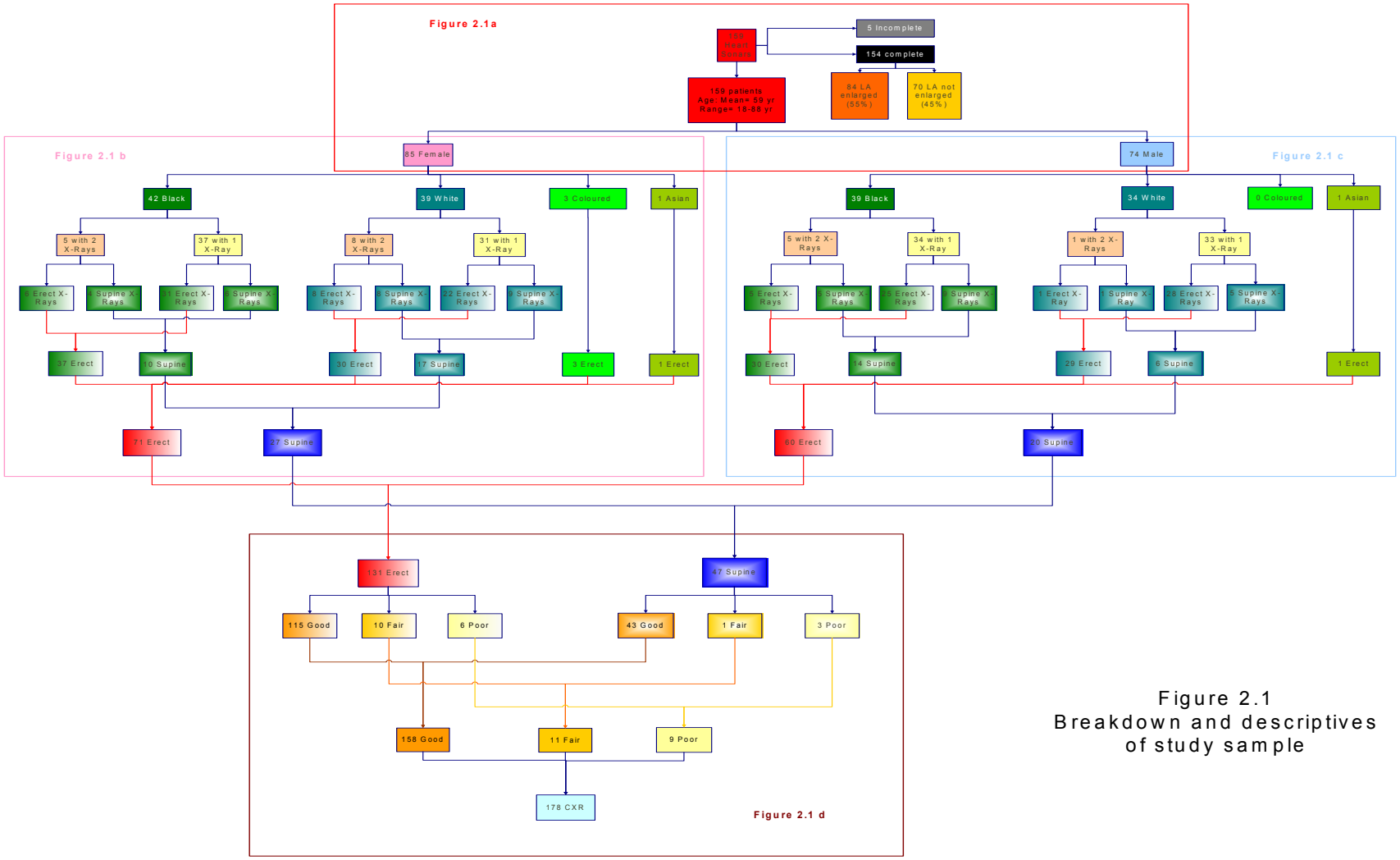


Figure 2.1
Breakdown and descriptives of study sample

Figure 2.1a
Breakdown of patients and echocardiograms

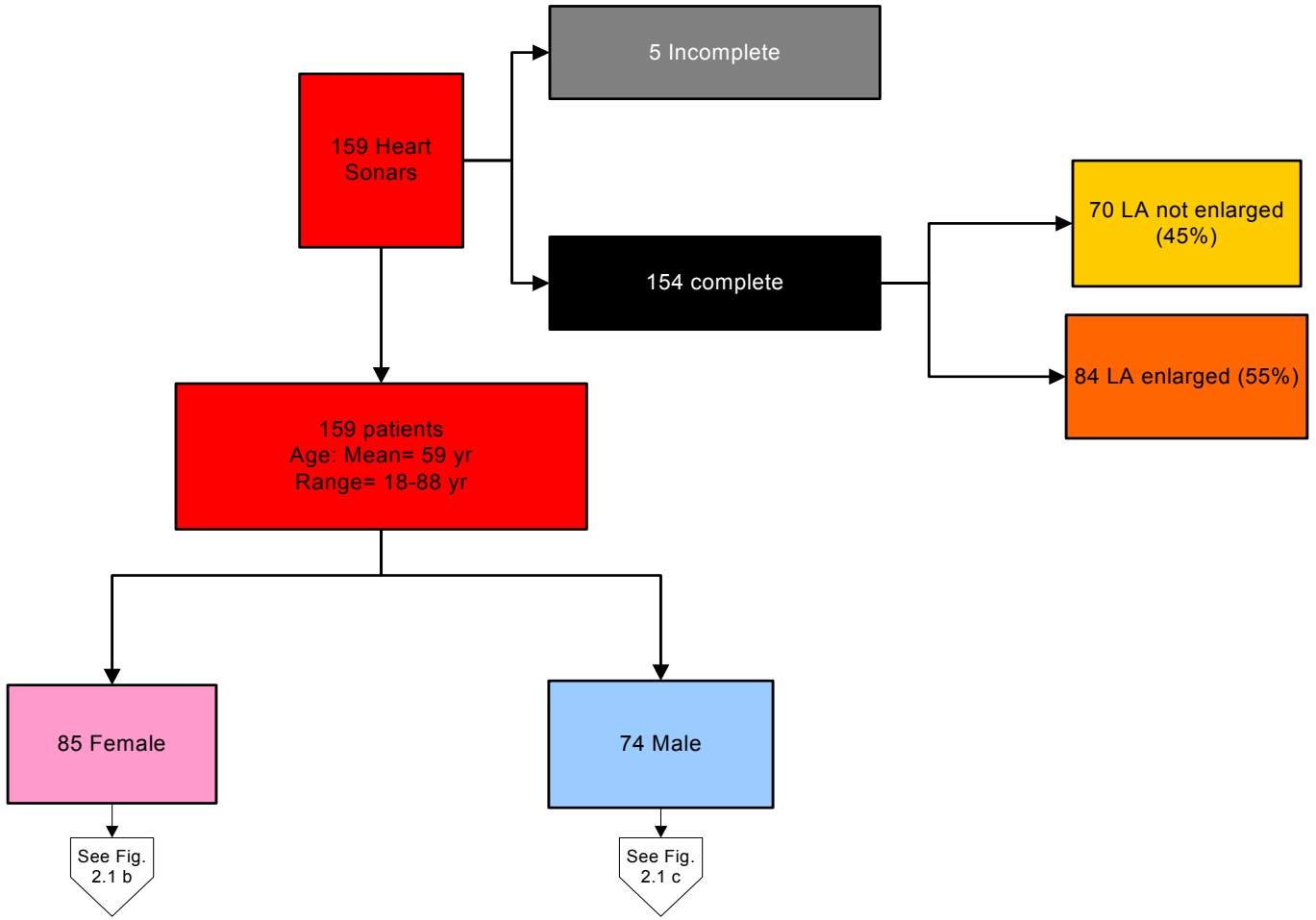


Figure 2.1b
Breakdown of female patients
and their radiographs

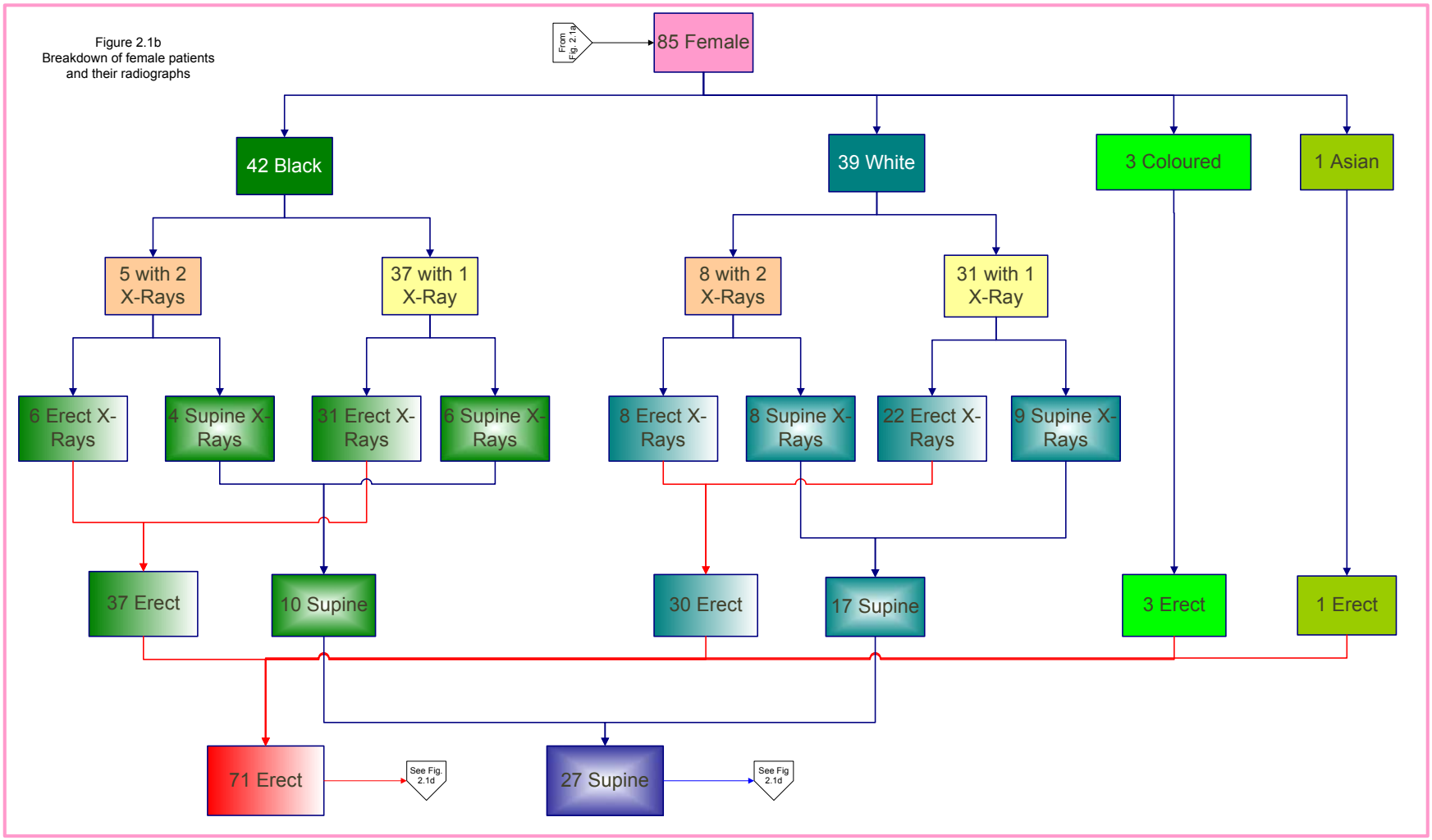


Figure 2.1c
Breakdown of male patients
and their radiographs

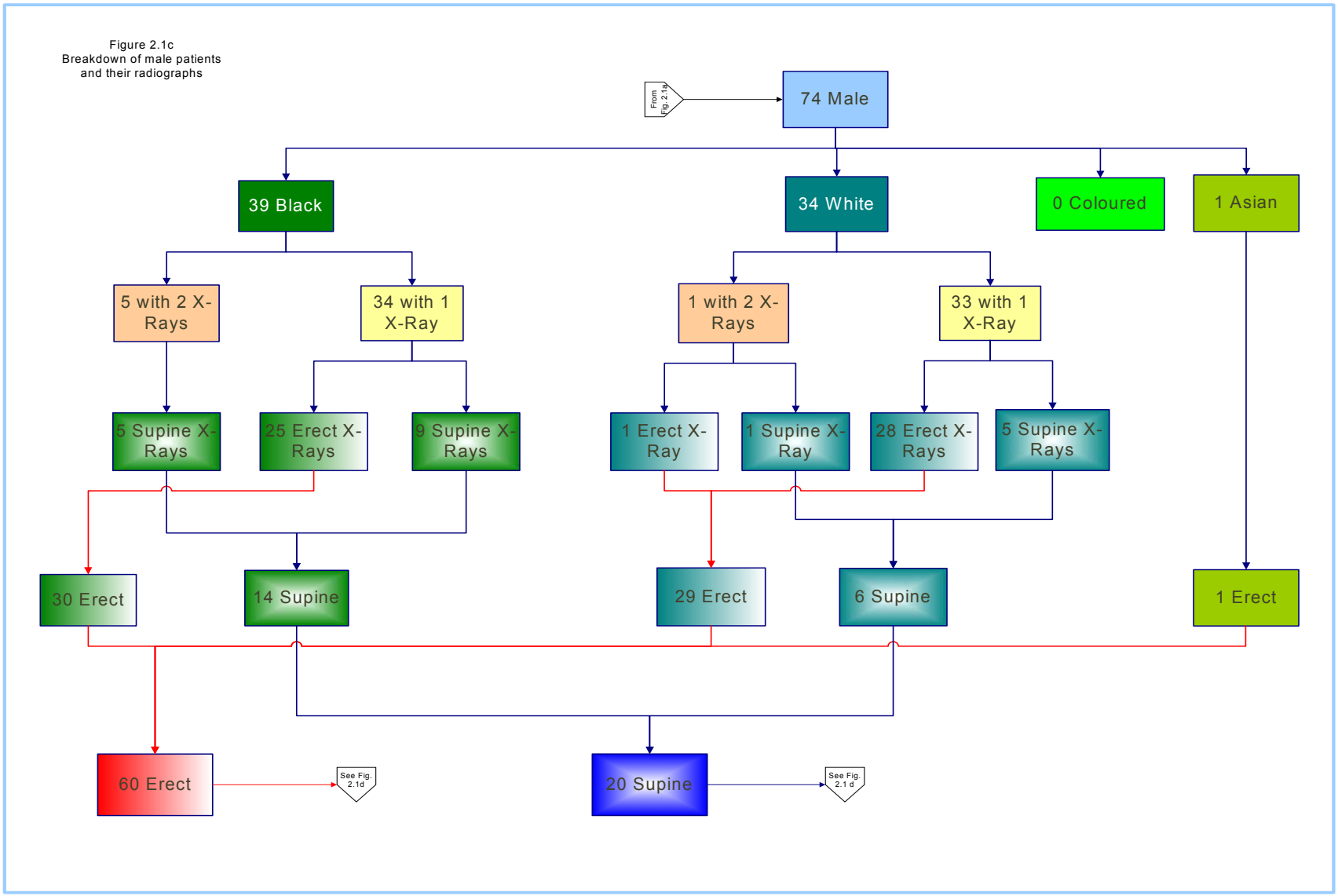
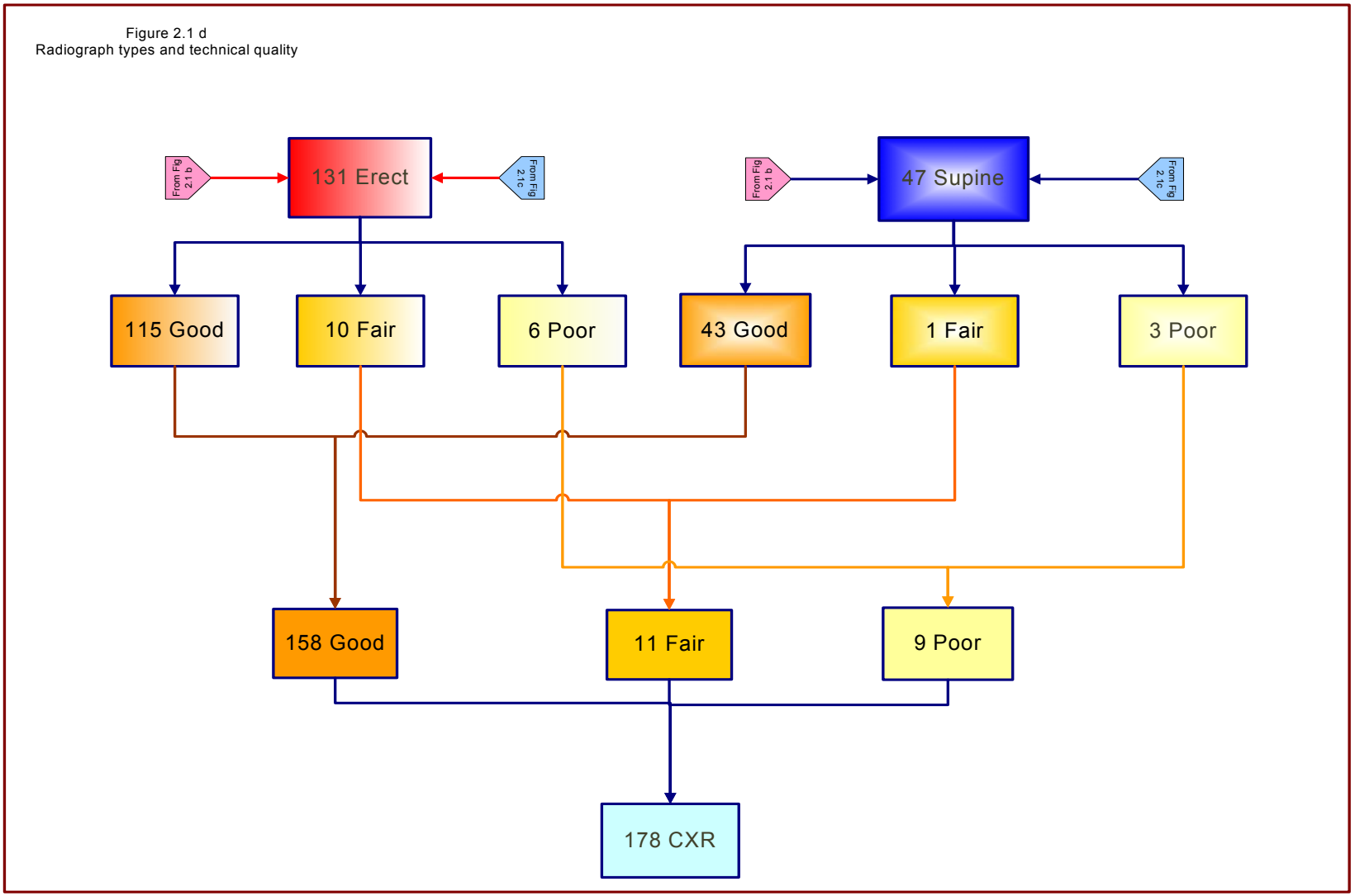


Figure 2.1 d
Radiograph types and technical quality



2. LIN's Concordance Correlation Coefficient and Bland and Altman methodology utilized to assess intra- and inter-rater agreement.

The agreement between the first and second measurements of Examiner1 varies between 0.98 and 0.99 for different combinations of X-ray types. All results are statistically significant. This is shown in *Table 2.1* and *Figures 3.1-3.3*

Deleted: Examiner 1

Table 2.1 Intrarater Repeatability of Examiner1

Intrarater Repeatability Examiner1 for SCA & SAD

X-Ray Type	Mean	Range	Obs	Corr	SE	95% CI	Diff	SD	95% LOA
For all types	SCA 75.4	26 ↔ 127	165	0.99	0.00	0.99 ↔ 0.99	0.01	2.81	-5.51 ↔ 5.52
	SAD 24.4	9 ↔ 36	165	0.98	0.00	0.97 ↔ 0.99	0.04	1.07	-2.06↔2.13
Good Quality	SCA 76.0	30 ↔ 127	154	0.99	0.00	0,99 ↔ 0.99	-0.21	2.55	-5.20↔4.78
	SAD 24.6	10 ↔ 36	154	0.98	0.00	0.98 ↔ 0.99	-0.03	1.00	-1.99↔1.93
Erect	SCA 73.8	26 ↔ 126	120	0.99	0.00	0.98 ↔ 0.99	-0.19	3.00	-6.07↔5.68
	SAD 23.9	9 ↔ 36	120	0.98	0.00	0.97 ↔ 0.99	-0.01	1.09	-2.14↔2.13
Erect & not Poor Quality	SCA 74.1	33 ↔ 126	115	0.99	0.00	0.98↔0.99	-0.26	2.99	-6.13↔5.61
	SAD 24.0	11 ↔ 36	115	0.98	0.00	0.97↔0.99	-0.04	1.06	-2.11↔2.02
Supine	SCA 79.8	28 ↔ 127	45	0.99	0.00	0.99↔1.00	0.53	2.19	-3.76↔4.83
	SAD 25.8	10 ↔ 136	45	0.98	0.01	0.97↔0.99	0.16	1.02	-1.85↔2.16
Supine & not Poor = Supine & Good	SCA 81.1	30 ↔ 127	43	0.99	0.00	0.99↔1.00	0.54	2.24	-3.86↔4.93
	SAD 26.0	10 ↔ 36	43	0.98	0.01	0.96↔0.99	0.16	1.05	-1.89↔2.21

*P=0.000 for all rows.

***Abbreviations; Obs** observations, **Corr** LIN's correlation coefficient, **SE** standard error, **CI** confidence interval, **Diff** difference, **SD** standard deviation, **LOA** limits of agreement (Bland and Altman).*

Formatted: Font: (Default) Arial, Italic
Deleted: *Table 1*
Formatted: Font: (Default) Arial, Not Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Not Bold, Italic

Intrarater Repeatability of Examiner1 for SCA of all Good Quality X-rays

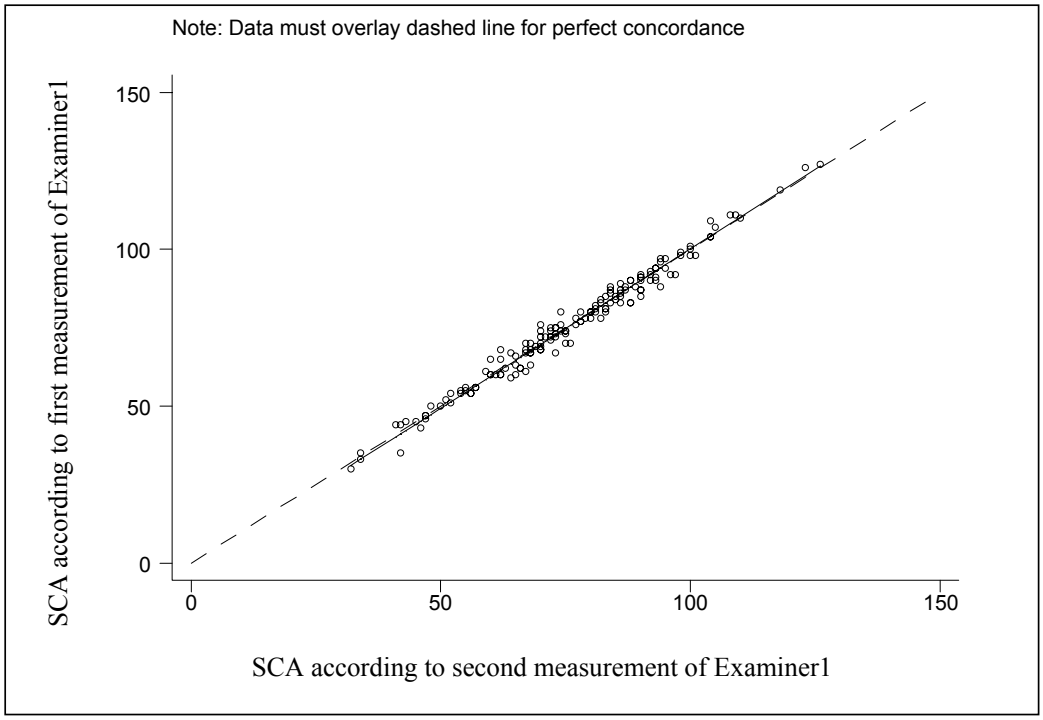
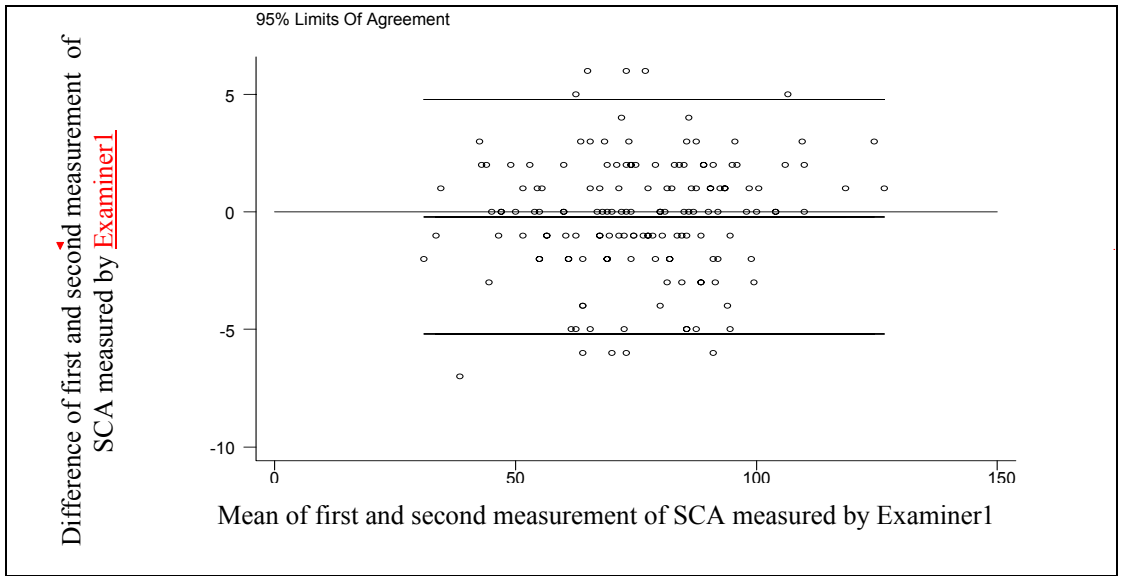


Figure 3.1 Intrarater Repeatability of Examiner1 using LIN's Correlation Coefficient



Deleted: Examiner 1

Figure 3.2 Intrarater Repeatability of Examiner1 using LOA

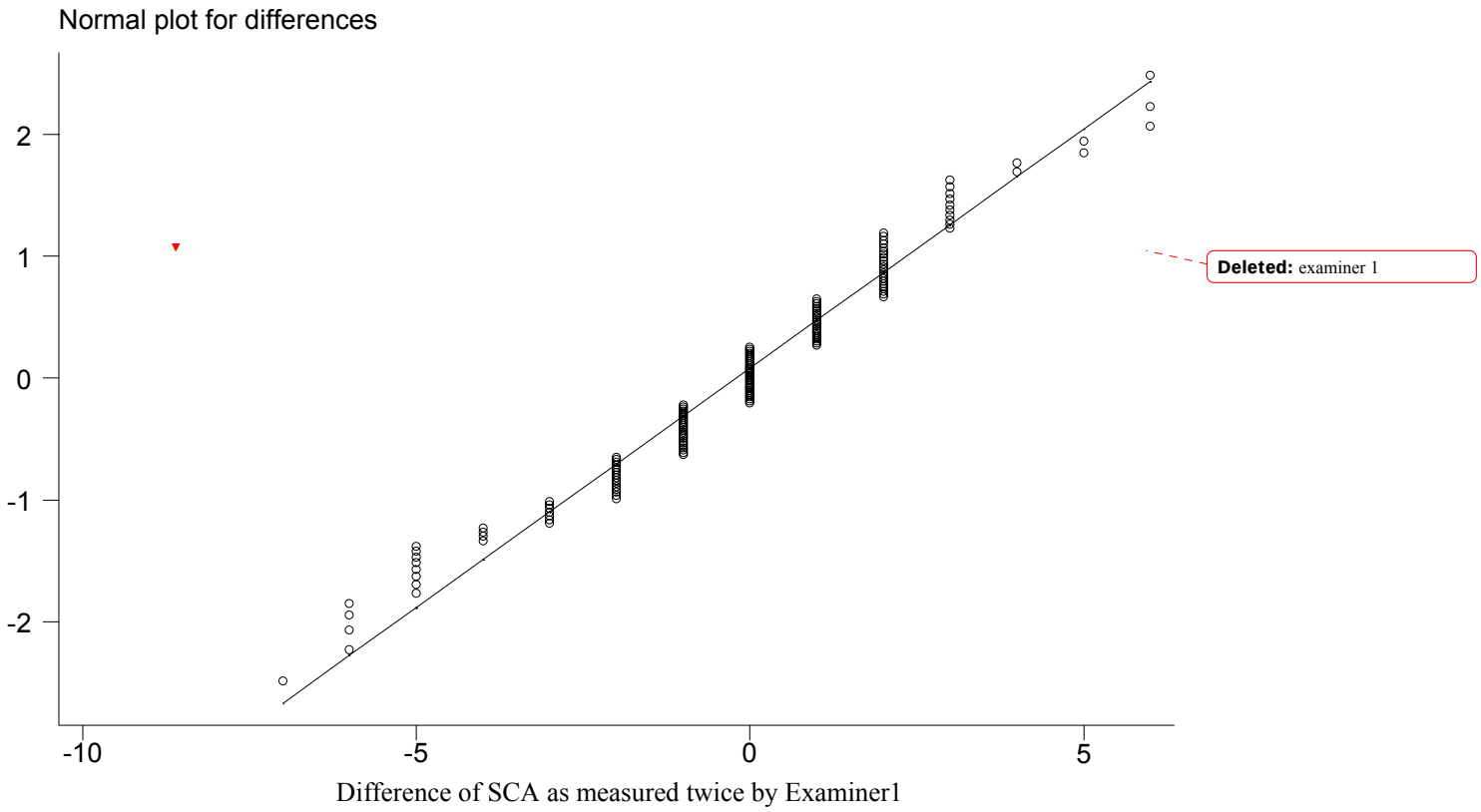


Figure 3.3 Intrarater Repeatability of Examiner1 using Bland & Altman Plot

The agreement between the first and second measurements of Examiner2 varies between 0.92 and 0.98 for different combinations of X-ray types. All results are statistically significant. This is shown in *Table 2.2* and *Figures 3.4-3.6*

Deleted: x

Table 2.2 Intrarater Repeatability of Examiner2

Intrarater Repeatability Examiner2 for SCA & SAD									
X-Ray Type	Mean	Range	Obs	Corr	SE	95% CI	Diff	SD	95% LOA
For all types	SCA 76.7	28 ↔ 126	161	0.97	0.00	0.96 ↔ 0.98	-0.51	3.99	-8.32↔ 7.30
	SAD 24.5	10 ↔ 37	161	0.94	0.01	0.92↔0.96	-0.21	1.66	-3.45↔3.04
Good quality	SCA 77.4	35 ↔ 126	151	0.97	0.00	0.97↔0.98	-0.54	3.82	-8.02↔6.94
	SAD 24.7	11 ↔ 37	151	0.94	0.01	0.92↔0.96	-0.18	1.61	-3.32↔2.97
Erect	SCA 75.8	30 ↔ 126	119	0.97	0.01	0.96↔0.98	-0.54	4.03	-8.50↔7.30
	SAD 24.3	10 ↔ 37	119	0.94	0.01	0.92↔0.96	-0.27	1.61	-3.42↔2.88
Erect & not Poor	SCA 76.1	37 ↔ 126	115	0.97	0.01	0.96↔0.98	-0.63	4.05	-8.57↔7.31
	SAD 24.4	11 ↔ 37	115	0.94	0.01	0.92↔0.96	-0.26	1.61	-3.42↔2.90
Supine	SCA 79.4	28 ↔ 11	42	0.98	0.01	0.96↔0.99	-0.26	3.90	-7.90↔7.37
	SAD 25.3	10 ↔ 35	42	0.94	0.02	0.90↔0,99	-0.02	1.79	-3.53↔3.48
Supine & not Poor = Supine & Good	SCA 80.8	35 ↔ 116	40	0.97	0.01	0.95↔0.99	-0.38	3.95	-8.12↔7.37
	SAD 25.7	12 ↔ 35	40	0.92	0.02	0.88↔0.97	-0.03	1.83	-3.62↔3.57

*P=0.000 for all rows

***Abbreviations; Obs** observations, **Corr** LIN's correlation coefficient, **SE** standard error, **CI** confidence interval, **Diff** difference, **SD** standard deviation, **LOA** limits of agreement (Bland and Altman)*

Formatted: Font: (Default) Arial, Italic
Deleted: *Table 1*
Formatted: Font: (Default) Arial, Not Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Not Bold

Intrarater Repeatability of Examiner2 for Angle of all Good Quality X-rays

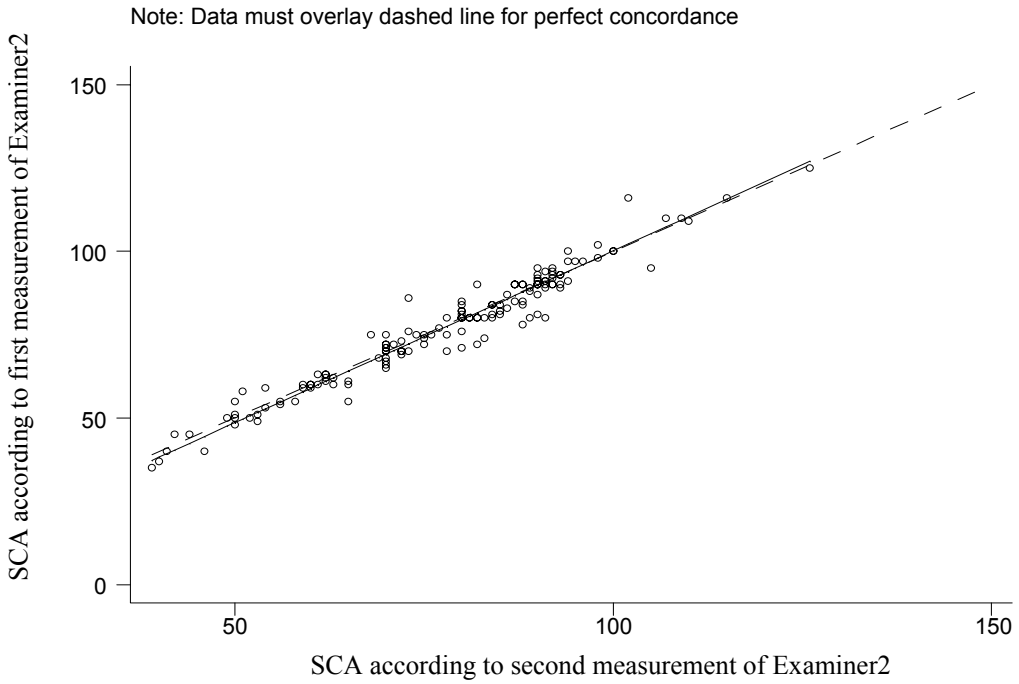
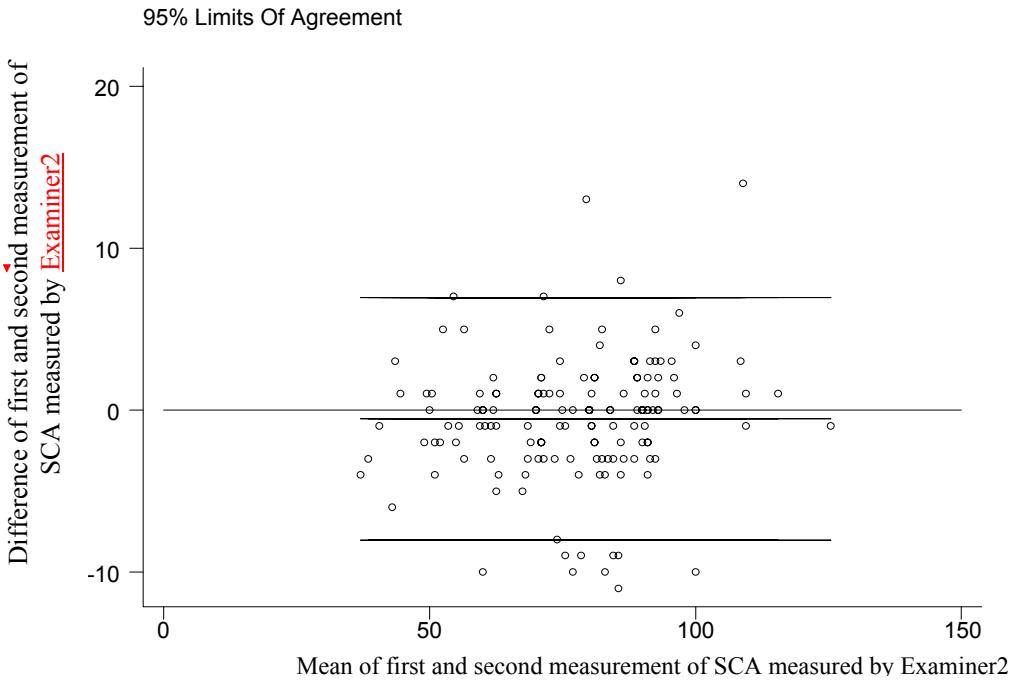


Figure 3.4 Intrarater Repeatability of Examiner2 using LIN's Correlation Coefficient



Deleted: Examiner 2

Figure 3.5 Intrarater Repeatability of Examiner2 using LOA

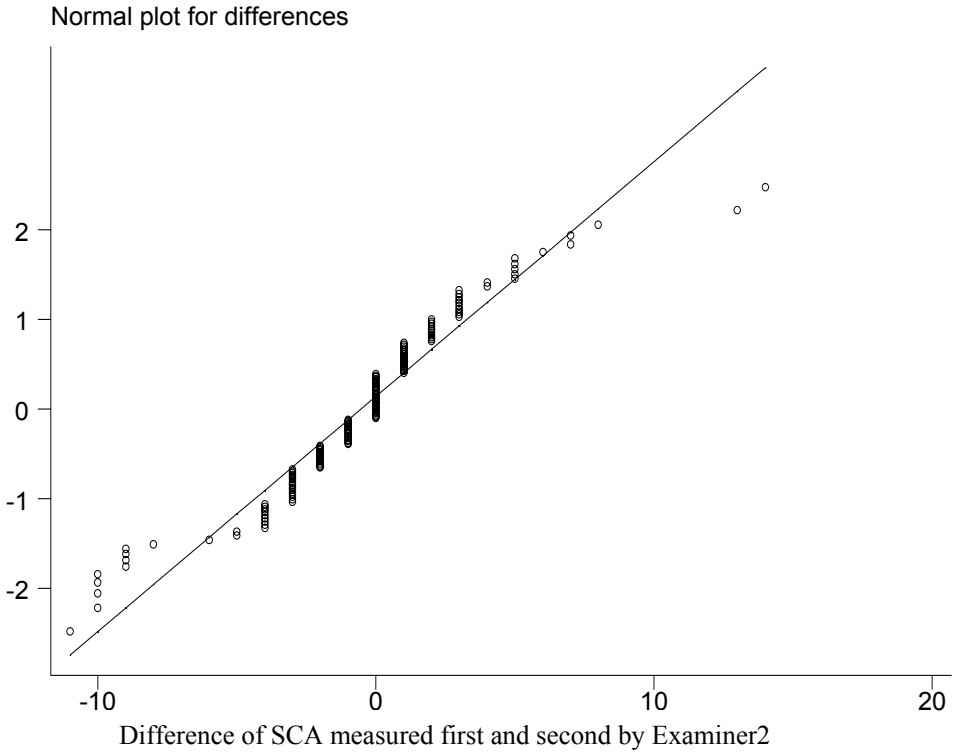


Figure 3.6 Intrarater Repeatability of Examiner2 using Bland & Altman Plot

The agreement between Examiner1 and Examiner2 as regards the measured mean SCA and mean SAD varies between 0.74 and 0.93 for different combinations of x-ray types. All results are statistically significant. This is shown in *Table 2.3*.

Table 2.3 Interrater Repeatability between Examiner1 and Examiner2

Interrater Repeatability Between Examiner1 and Examiner2: SCA & SAD

X-Ray Type	Mean	Range	Obs	Corr	SE	95% CI	Diff	SD	95% LOA
For all types	SCA 76.0	26 ↔ 126.5	160	0.91	0.01	0.89 ↔ 0.94	-1.16	7.22	-15.31 ↔ 13.00
	Hyp 24.5	9 ↔ 36	160	0.89	0.02	0.86↔0.92	-0.11	2.35	-4.71↔4.49
Good quality	SCA 76.6	31 ↔ 126.5	150	0.91	0.01	0.88↔0.93	-1.08	7.40	-15.59↔13.43
	SAD 24.6	10 ↔ 36	150	0.88	0.02	0.84↔0.91	-0.11	2.41	-4.82↔4.61
Erect	SCA 74.7	26 ↔ 125.5	118	0.93	0.01	0.90↔0.95	-1.66	6.49	-14.38↔11.07
	SAD 24.1	9 ↔ 36	118	0.92	0.01	0.90↔0.95	-0.28	1.90	-4.01↔3.44
Erect & not Poor	SCA 75.1	33.5 ↔ 125.5	114	0.92	0.01	0.90↔0.95	-1.67	6.55	-14.51↔11.17
	SAD 24.2	11 ↔ 36	114	0.92	0.02	0.89↔0.95	-0.28	1.93	-4.06↔3.49
Supine	SCA 79.6	28 ↔ 109	42	0.88	0.04	0.80↔0.95	0.25	8.90	-17.19↔17.69
	SAD 25.4	10 ↔ 36	42	0.80	0.06	0.69↔0.91	0.37	3.28	-6.06↔6.80
Supine & not Poor = Supine & Good	SCA 80.9	31 ↔ 126.5	40	0.84	0.05	0.75↔0.93	0.38	9.10	-17.46↔17.21
	SAD 25.8	10 ↔ 36	40	0.74	0.07	0.60↔0.88	0.39	3.36	-6.20↔6.98

*P=0.000 for all rows.

Abbreviations: Obs observations, **Corr** LIN's correlation coefficient, **SE** standard error, **CI** confidence interval, **Diff** difference, **SD** standard deviation, **LOA** limits of agreement (Bland and Altman)

Deleted: Table 1

Formatted: Font: (Default) Arial

Formatted: Font: (Default) Arial, Bold

Formatted: Font: (Default) Arial, Bold

Formatted: Font: (Default) Arial, Bold

Formatted: Font: (Default) Arial, Not Bold

3. Logistic Regression Results

Logistic regression was used for erect and supine X-rays of good quality to determine the predictive value of the SCA and the SAD.

3.a Linear Model

Table 3.1 Linear model of left atrial size estimators

X-Ray Type	Mean of Examiner1	Obs	Coeff	95% CI	P value
Erect	SCA	102	0.02	0.00 ↔ 0.5	0.04
	SAD	102	0.09	0.01 ↔ 0.17	0.03
Supine	SCA	37	0.03	0.01 ↔ 0.07	0.20
	SAD	37	0.10	-0.05 ↔ 0.24	0.19

Abbreviations: ***SCA*** *sub-carinal angle,* ***SAD*** *sub-angle distance,* ***Obs*** *observations*
Coeff *coefficient,* ***CI*** *confidence interval*

Formatted: Font: (Default) Arial, Italic
Formatted: Font: (Default) Arial, Not Bold, Italic
Deleted: ***Table 1***
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Not Bold, Italic

3.b Fit of the Model and Diagnostic Performance

The goodness-of-fit and diagnostic performance characteristics of the model are shown in Table 3.2.

Table 3.2 Logistic Regression Results

Logistic Regression Results				
Good quality X-rays	**Erect**		**Supine**	
Mean of Examener1	SCA	SAD	SCA	SAD
Number of Obs	102	102	37	37
Coeff (SE)	0.02 (0.01)	0.09 (0.04)	0.03 (0.02)	0.10 (0.07)
LR chi2	4.41	5.12	3.16	3.41
OR	1.02	1.09	1.03	1.10
P	0.04	0.02	0.08	0.06
95% CI	1.00 ↔ 1.05	1.01 ↔ 1.19	0.99 ↔ 1.07	0.95 ↔ 1.27
GOF	0.43	0.15	0.38	0.21
Sens	82.76 %	82.76 %	80.00 %	80.00 %
Spec	34.09 %	36.36 %	52.94 %	47.06 %
PPV	62.34 %	63.16 %	66.67 %	64.00 %
NPV	60.00 %	61.54 %	69.23% %	66.67 %
Area under curve	0.62	0.63	0.62	0.63

Abbreviations: *Obs observations, Coeff coefficient, SE standard error, LR likelihood ratio, OR odds ratio, CI confidence interval, GOF goodness-of-fit, Sens sensitivity, Spec specificity, PPV positive predictive value, NPV negative predictive value, AUC area under the curve.*

Formatted: Font: (Default) Arial, Italic
Deleted: ***Table 1***
Formatted: Font: (Default) Arial, Not Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Not Bold, Italic

3.c *Linearity of x in the logit*

Assessing whether x was Linear in the Logit, 3 methods as proposed by Lemeshow were used

i. Lowess Smoothing curve
ii. Design Variables
iii. Fractional Polynomials

i. *Lowess Smoothing curve*

Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*
mangle1 *mean value of the SCA as measured twice by Examiner1*

Figure 4.1 ***Exploring the nature of the relationship between Gold standard and mean SCA of Examiner1 for erect good quality X-rays.***

Deleted: *Examiner 1*

Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*
mangle1 *mean value of the SCA as measured twice by Examiner1*

Figure 4.2 ***Exploring nature of relationship between Gold standard and mean SCA of Examiner1 for supine good quality X-rays.***

Deleted: *Examiner 1*

Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*
mhyp1 *mean value of the SAD as measured twice by Examiner1*

Figure 4.3 ***Exploring nature of relationship between Gold standard and mean SAD of Examiner1 for erect good quality X-rays.***

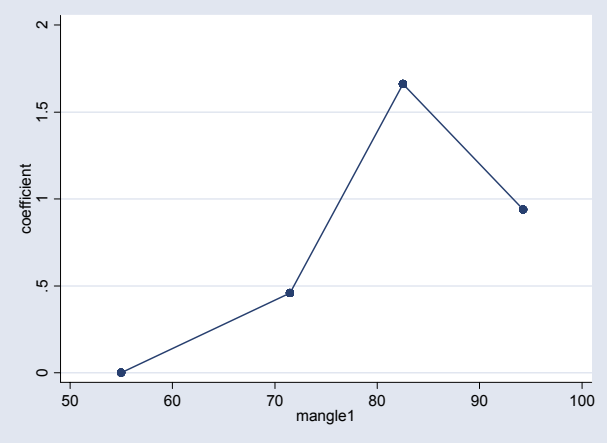
Deleted: *Examiner 1*

Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*
mhyp1 *mean value of the SAD as measured twice by Examiner1*

Figure 4.4 ***Exploring nature of relationship between Gold standard and mean SAD of Examiner1 for supine good quality X-rays.***

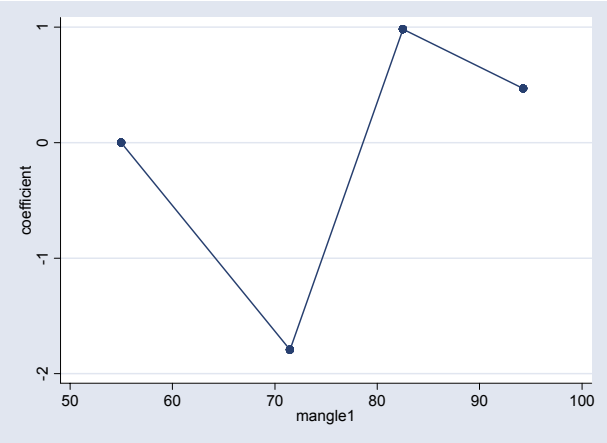
Deleted: *Examiner 1*

ii. Design Variables



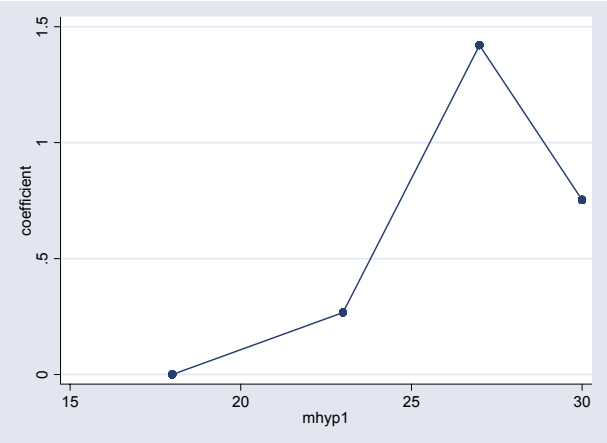
Abbreviation: ***mangle1*** *mean value of the SCA as measured twice by Examiner1*

Figure 4.5 ***Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SCA of erect X-rays.***



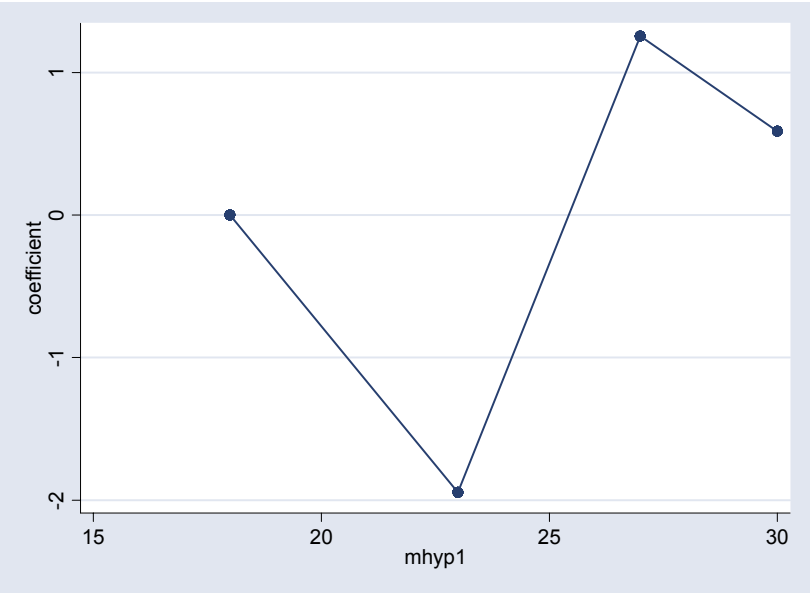
Abbreviation**:* ***mangle1 *mean value of the SCA as measured twice by Examiner1*

Figure 4.6 ***Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SCA of supine X-rays.***



Abbreviation**:* ***mhyp1 *mean value of the SAD as measured twice by Examiner1*

Figure 4.7 ***Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SAD of erect X-rays.***



***Abbreviation**:* ***mhyp1** mean value of the SAD as measured twice by Examiner1*

Figure 4.8 ***Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the mean SAD of supine X-rays.***

iii. Fractional Polynomial Model Comparisons

Table 3.3 Fractional Polynomials

Fractional Polynomial Model Comparisons:								
X-Ray type	**Mean of Examiner1**	**Obs**	**Mean**	**df**	**Deviance**	**Gain**	**P**	**Powers**
Erect & Good Quality	SCA	102	Not in model	0	139.47	--	--	
			Linear	1	134.95	0.00	0.03	1
			m=1	***2***	***130.89***	***4.06***	***0.04****	***-2***
			m=2	4	129.39	5.56	0.47	-2 3
	SAD	102	Not in model	0	139.47	--	--	
			Linear	1	134.43	0.00	0.03	1
			m=1	2	131.49	2.95	0.09	-2
			m=2	4	130.42	4.02	0.59	-2 3
Supine & Good Quality	SCA	47	Not in model	0	51.05	--	--	
			Linear	1	49.22	0.00	0.18	1
			m=1	2	49.11	0.11	0.74	2
			m=2	4	48.96	0.25	0.93	-2 3
	SAD	47	Not in model	0	51.05	--	--	
			Linear	1	49.12	0.00	0.16	1
			m=1	2	48.94	0.18	0.67	3
			m=2	4	48.83	0.29	0.95	-2 3

Deleted: Examiner 1

***Abbreviations**: **Obs** Observations, **df** degrees of freedom*

A P-value of ***0.04**** (<0.05) suggests that transformation of the variable mean SCA, as measured by Examiner1 on good quality erect X-rays, might improve the fit of the model.

3.d Linear model using transformed variable

A new variable, the inverse of the square of the mean SCA, was generated as suggested by fractional polynomials method comparison.

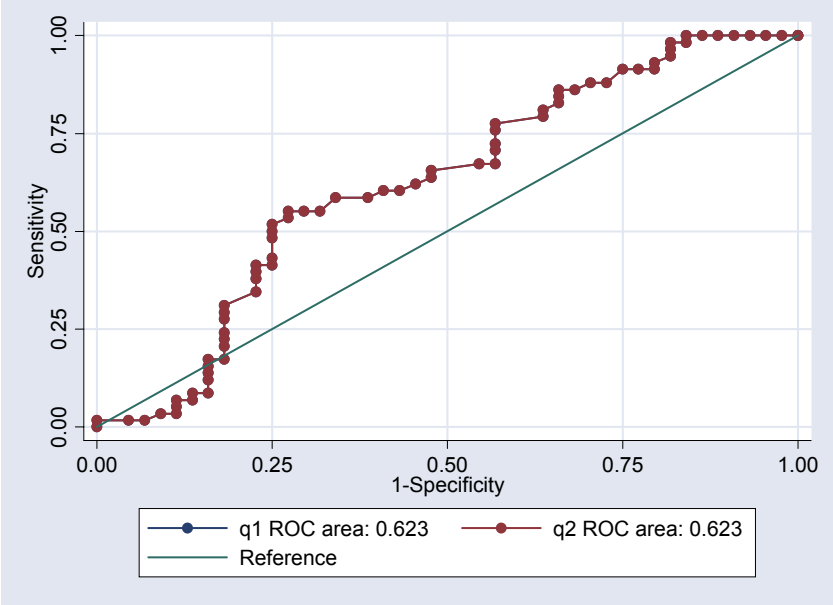
Subsequent logistic regression yielded the following results:

Table 3.4 Diagnostic characteristics of original and transformed variable

	Mean SCA	Inverse of the mean squared
GOF	0.43	0.53
Sens	82.76%	87.93%
Spes	34.09%	29.55%
PPV	62.34%	62.20%
NPV	60.00%	65.00%
AUC	0.62	0.62

Abbreviations: ***GOF*** *goodness-of-fit,* ***Sens*** *sensitivity,* ***Spec*** *specificity,* ***PPV*** *positive predictive value,* ***NPV*** *negative predictive value,* ***AUC*** *area under the curve*

The transformation does not improve the diagnostics of the test to such an extent that the transformation applied to the mean SCA is clinically or statistically useful, as evidenced by very similar model performance characteristics (see Table 3.4) and overlapping ROC curves (*Figure 4.9*).



Abbreviation: ROC *receiver operating characteristic*

Figure 4.9 Comparison of ROC curves of original and transformed variable

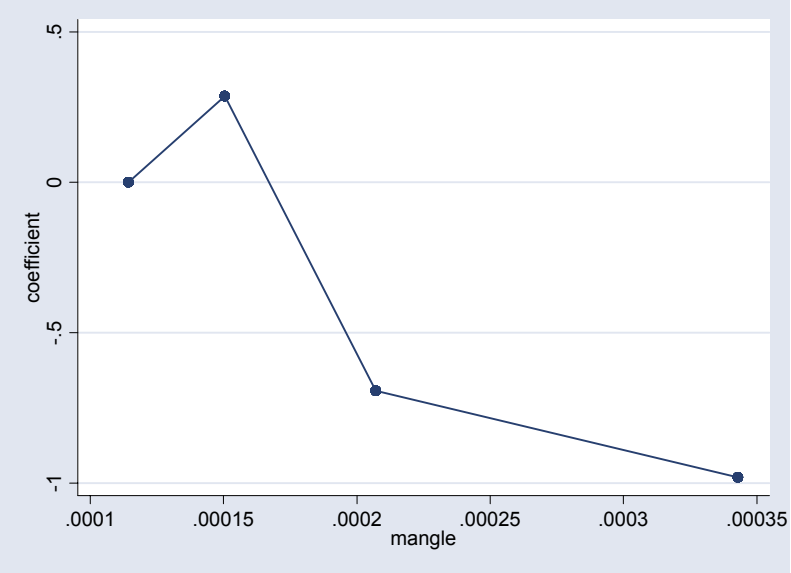
Formatted: Font: (Default) Arial, Italic
Deleted: *Table 1*
Formatted: Font: (Default) Arial, Not Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Bold, Italic
Formatted: Font: (Default) Arial, Not Bold, Italic

A Lowess Smoothing curve and Lincheck were performed to assess whether x was linear in the logit.

Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*
mangle *transformed mean value of the SCA as measured twice by Examiner1*

Figure 4.10 Exploring nature of relationship between Gold standard and transformed mean SCA of Examiner1 for erect good quality X-rays.

Deleted: *Examiner 1*



Abbreviation: ***mangle*** *transformed mean value of the SCA as measured twice by Examiner1*

Figure 4.11 Plot of estimated logistic regression coefficients vs. approximate quartile midpoints of the transformed mean SCA of erect X-rays.

3.e ROC curves (roctg2) to determine optimal cut-off points for SCA and SAD for Examiner1.

Deleted: Examiner 1

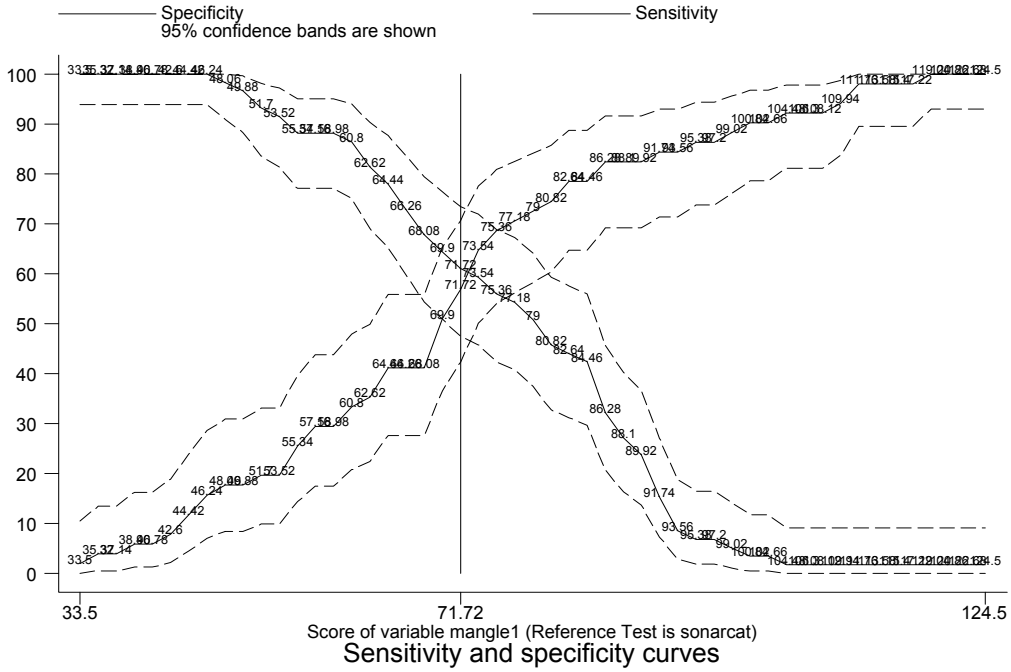
Optimal cut-off points for SCA and SAD, that would yield the best compromise between sensitivity and specificity were determined using *roctg2*.

Determining a Cut-Off Point for SCA of good quality erect X-rays.

---

sensitivity (obs) = 61.02% (47.44% - 73.45%)

specificity (obs) = 56.86% (42.25% - 70.65%)



Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*

mangle1 *mean value of the SCA as measured twice by Examiner1*

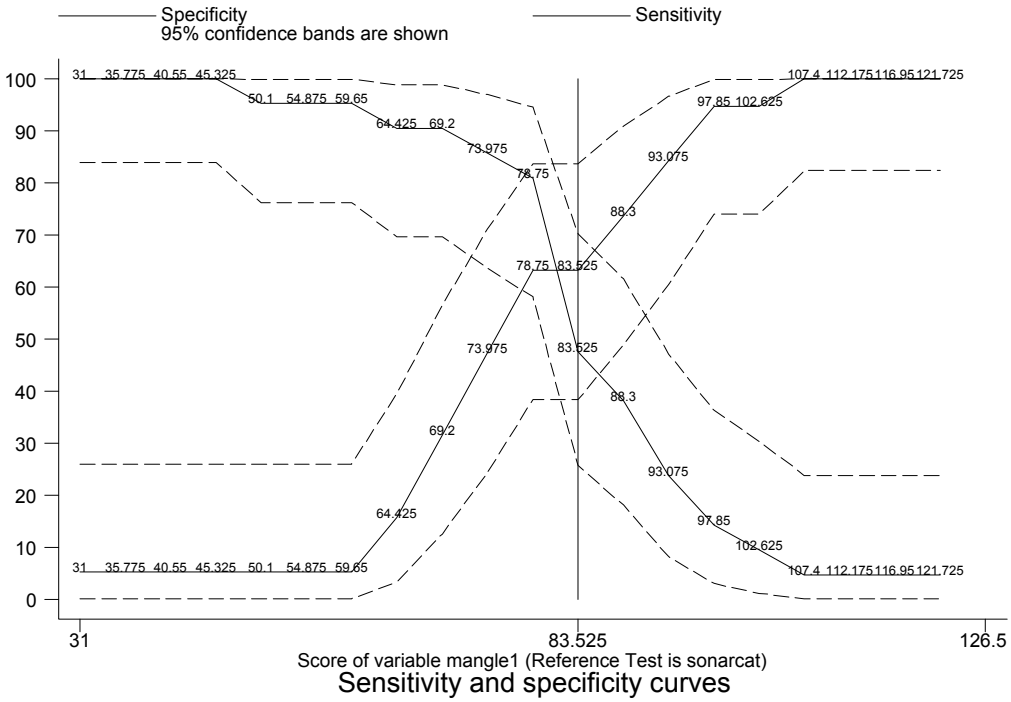
Figure 4.12 Cut-off point for SCA on erect X-rays

Determining a Cut-Off Point for SCA of good quality supine X-rays.

---

sensitivity (obs) = 47.62% (25.71% - 70.22%)

specificity (obs) = 63.16% (38.36% - 83.71%)



Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*

mangle1 *mean value of the SCA as measured twice by Examiner1*

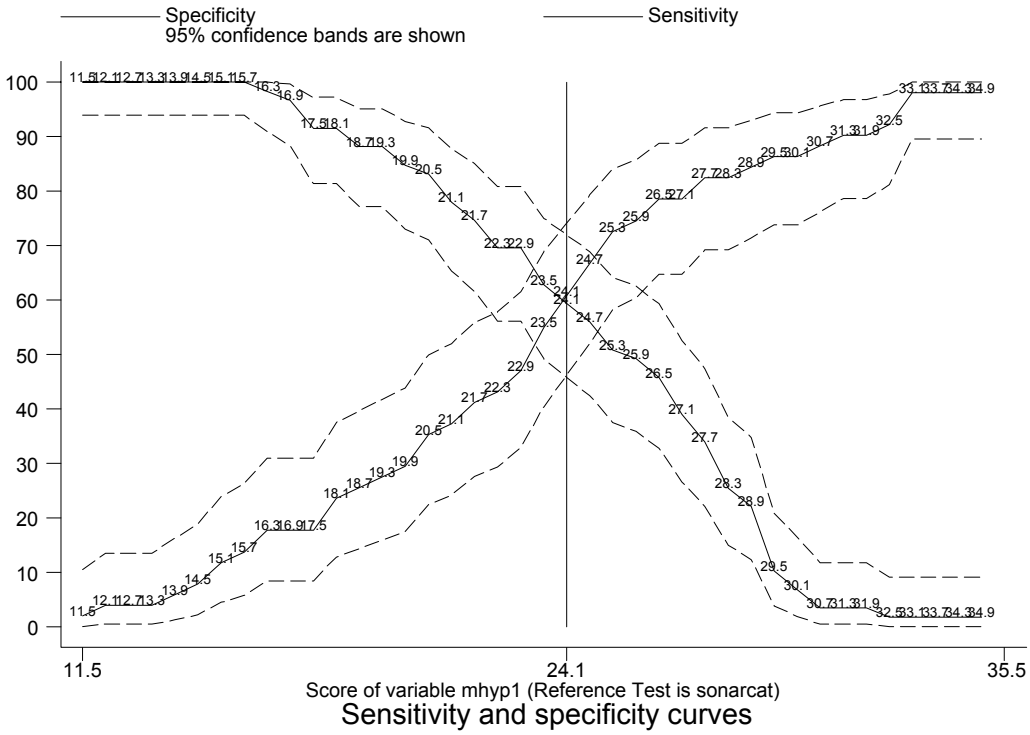
Figure 4.13 Cut-off point for SCA on supine X-rays

Determining a Cut-Off Point for SAD of good quality erect X-rays.

---

sensitivity (obs) = 59.32% (45.75% - 71.93%)

specificity (obs) = 60.78% (46.11% - 74.16%)



Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*

mhyp1 *mean value of the SAD as measured twice by Examiner1*

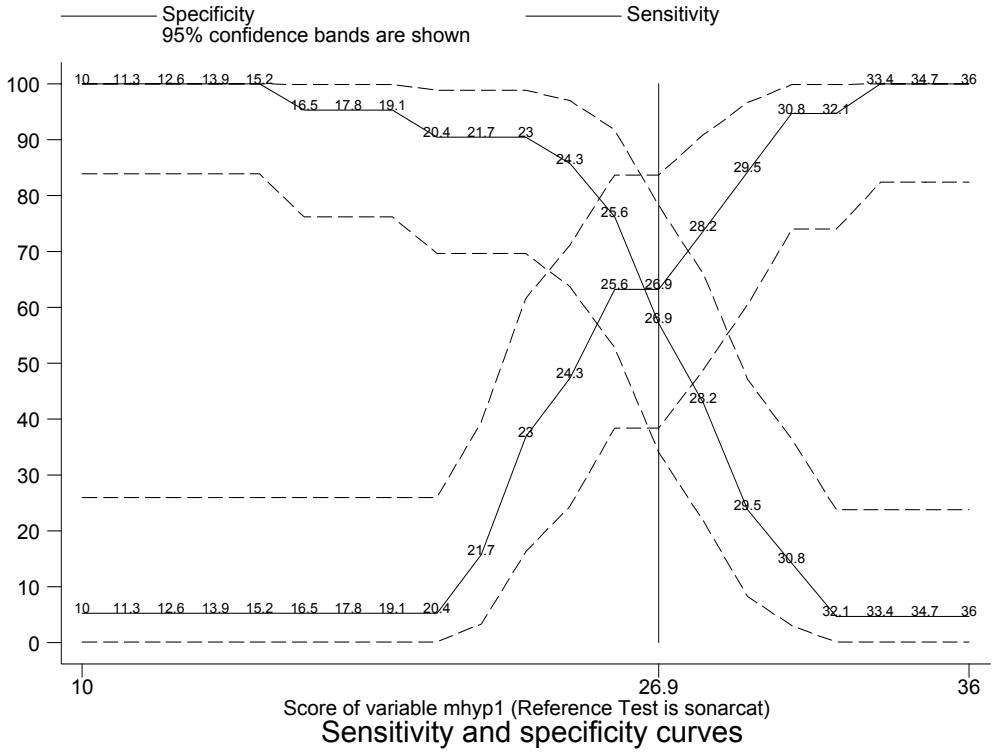
Figure 4.14 Cut-off point for SAD on erect X-rays

Determining a Cut-Off Point for SAD of good quality supine X-rays.

---

sensitivity (obs) = 57.14% (34.02% - 78.18%)

specificity (obs) = 63.16% (38.36% - 83.71%)



Abbreviations: ***sonarcat*** *sonar categories of enlarged vs non-enlarged left atria*

mhyp1 *mean value of the SAD as measured twice by Examiner1*

Figure 4.15 Cut-off point for SAD on supine X-rays

DISCUSSION

Chest X-rays are used on a daily basis as diagnostic tool in clinical settings from primary health care clinics to tertiary institutions. One common derivative (from a chest X-ray) would be an estimation of the left atrial size. From the left atrial size, inferences are then made regarding prognosis and events, as well as decisions on different available treatment choices.

Deleted: P
Deleted: H
Deleted: C
Deleted: C
Deleted: T
Deleted: I
Deleted: x

Very little is known about the test characteristics e.g. sensitivity and specificity of the chest X-ray to predict left atrial enlargement. Uncertainty exists on the reliability of chest X-rays for prediction of left atrial enlargement.

Deleted: x

Biplane 2D echo is the gold standard used for determining left atrium size.[24,25,26] This diagnostic tool is not widely available or affordable. Chest X-rays are often used to predict left atrium enlargement, being cheaper and more readily available. We base our decisions on this, ignorant of the accuracy and precision that we have traded. The question remains: “How reliable are chest X-rays to predict left atrium enlargement? ”

A literature search yielded only two articles[13,14] on this subject, as discussed in the literature review section. Both studies had rather small sample sizes and did not test inter- or intra-observer variation. Only one of the studies used standardisation of plain film techniques on determined sensitivity and specificity of X-rays as diagnostic tool. An ROC curve had not yet been established.

This study explored the intra- and interobserver variability of chest X-ray measurements. It also aimed to determine the diagnostic utility of chest X-rays compared to echocardiography in estimating left atrial size.

Data form 159 patients were sampled. Five patients had to be excluded from the final analysis due to undocumented left atrial size on echocardiography. The final number of echocardiograms used in the analysis (154) still exceeded the required sampling frame of 150 patients which was based on an estimated prevalence of enlarged left atrium of 70%. However, only good quality radiographs (102 erect and 37 supine) were used in the final analysis Furthermore, the prevalence of enlarged left atrium as determined by

echocardiogram was only 55%. This implies that the sample size was not large enough to prevent a type 2 error in this population with a lower than estimated prevalence of enlarged left atrium.

Intrarater agreement of measures of left atrial size was excellent for both examiners. This implies that the methods (using either the SCA or the SAD) are very precise. The same results were found when the two different examiners were compared. This underscores the repeatability of these methods.

The linear model obtained using logistic regression demonstrated that only erect chest X-rays were useful in predicting whether the left atrium was enlarged as supine chest X-rays did not yield statistically significant results. This may be due to the small number of supine X-rays included in the study.

Goodness-of-fit results for the above model showed that the model is a reasonable fit for both erect and supine X-rays using either variable (SCA or SAD), which implies that the model effectively describes the outcome variable. As the desired outcome for this model is the decision not to reject the null hypothesis that the model fits the data, the above conclusion is subject to a type II error and hence the power of the goodness-of-fit test is an important consideration especially when the sample size is small, as in this study.

Although the sensitivity of all four categories (erect vs. supine X-rays using SCA vs SAD) is consistently above 80%, the specificity is so poor (<53%) that the utility of this test might be limited to a rule-out strategy only. Positive and negative predictive values (as well as the area under the ROC curve) are close to 50% (+/- 60%) implying that either a positive or a negative result are hardly better than a random guess in predicting the left atrial size. This in effect means that a rule-out strategy using these variables (SCA / SAD) alone is not feasible as the model's ability to discriminate between enlarged and normal left atria is poor.

The visual representation of the relationship between the actual left atrial size category and the estimate (SCA and SAD) as portrayed in the lowess smoothing curves suggests a linear relationship only at the extremes of SCA_s (or equivalently SAD_s). Between angles of 50° (18mm) and 110° (32mm) the estimation of left atrial size is unreliable, especially for supine X-rays.

The results of the quartile analysis as plotted show a definite deviation from linearity in the fourth quartile throughout all permutations of X-ray types and variable used in analysis. In contrast, the results suggest linearity in the logit for both SCA and SAD in the first three quartiles.

Overall, fractional polynomial model comparisons showed that the best non-linear transformations are not significantly different from the linear model. Thus, the fractional polynomial analysis supports treating both variables as linear in the logit in general, with one exception: a significant P-value of 0.04 for the variable SCA suggests that the fit of the model might be improved if the variable is transformed by its inverse square. This in turn suggested that the use of the transformed variable in the logistic regression analysis might result in a superior model.

The subsequent logistic regression results of the transformed variable, when compared to the original variable, seem similar. This similarity is borne out in the near-perfect overlap of the ROC curves of the two models. The discriminating value of the diagnostic test did not improve by using a transformed variable.

The model may therefore be regarded as linear in the logit of the parameters, using either dependent variable (i.e. SCA or SAD). This means that logistic regression is the appropriate method for deriving a prediction rule for left atrial size category (enlarged/ not enlarged) using radiograph-derived left atrial size estimates.

Optimal cut-off points for the SCA and SAD of both erect and supine X-rays revealed surprising values. The commonly used cut-off value for the SCA in clinical practice thought to indicate an enlarged left atrium, is 90°. In contrast, this study found the cut-off points to be 72° for erect and 84° for supine films. In addition to this, the SAD measure seems to be interchangeable with the SCA measurement throughout the analysis. This simplifies the estimation of left atrial size as the SAD measurement does not require any equipment not usually found in a physician's pocket.

Several possible limitations arise from the retrospective nature of the study design.

Firstly, only data from patients with traceable echocardiograms and chest radiographs were included. Thus, no data is available on the number of patients excluded because only echocardiograms and no chest radiographs were available. In the tertiary setting that this study took place, such a scenario is highly implausible and if chest radiographs were not available, it would most likely be due to the odd chest X-ray gone missing. As this would be a random event rather than an occurrence in some obscure systematic manner, selection bias was not considered to pose any threat to this study.

Secondly, gold standard related measurement bias validity issues arise from the fact that echocardiographic data collected previously, were used. This was discussed with a clinical- as well as an operational expert, who both assured that the accuracy would be acceptable as measurements are recorded according to the method described in the Methodology section.

Deleted: G

Deleted: S

Thirdly, the fact that as per protocol, two weeks would be acceptable as the maximum period within which any given patient had undergone both echocardiographical and chest radiographical examinations, does raise some concerns. It would be possible for a patient's clinical picture to change somewhat during his hospital stay. It is unknown what effect that might have on his/her left atrial size as measured with echocardiogram and/or chest radiography. Due to the study design it was not possible to correct for this limitation.

Lastly, the fact that this study was conducted as an audit, limited the amount of information available on patient's e.g. clinical conditions, absence- and presence of heart disease and certain patient demographics.

These limitations could be overcome if a prospective study, with validated echocardiography and radiography taking place on the same day, is conducted.

CONCLUSION

Both SCA as well as SAD can be used interchangeably on erect chest radiographs of good quality to predict left atrial enlargement with great precision but poor accuracy. This might also be true for supine radiographs, although the sample size in this study was too small to yield a statistically significant result. Cut-off values below the traditionally used 90° were found to predict left atrial enlargement with improved diagnostic characteristics. The corresponding SAD appeared to have slightly better diagnostic discriminatory capability regarding test sensitivity, specificity and corresponding confidence intervals surrounding those measures. Given the lower than expected prevalence of enlarged left atrium, a type 2 error concealing better performance of radiographic measures to predict left atrial enlargement cannot be excluded.

The use chest radiographs in predicting left atrial enlargement, using either *SpIn* or *SnOut* strategies is not recommended due to low sensitivity *and* specificity of the determined cut-off values. The accuracy (and therefore the reliability) of radiographs to predict left atrial size may be better than suggested here, but this needs evaluation in a study with a larger sample size.

In conclusion, chest radiographs has poor diagnostic utility in predicting left atrial enlargement. Echocardiography remains the preferred method of determining left atrial size.

Deleted: Discussion¶
¶
This study aimed to determine the diagnostic utility of chest radiographs compared to echocardiography in estimating left atrial size. It also explored the intra- and interobserver variability of chest radiograph measurements. ¶
Data form 156 patients were sampled, exceeding the initial sampling frame of 150 patients (which allowed for 10% of untraceable and 20% unusable radiographs resulting in a final estimated sample size of 106 patients, given a prevalence of 70% of large left atrium on ultrasound).¶
Intraobserver agreement of measures of left atrial size was excellent for both examiners. This implies that the methods (using the carinal angle or the SAD) are very precise. Similar results were found when the two different examiners were compared (but less so for supine radiographs). This underscores the repeatability of these methods.¶
The linear model obtained using logistic regression demonstrated that only erect chest radiographs were useful in predicting whether the left atrium was enlarged as supine chest radiographs did not yield statistically significant results. This may be due to the small number of supine radiographs included in the study.¶
Goodness-of-fit results for the above model showed that the model is a reasonable fit for both erect and supine radiographs using either variable (angle or SAD), which implies that the model effectively describes the outcome variable. As the desired outcome for this model is the decision not to reject the null hypothesis that the model fits the data, the above conclusion is subject to a type II error and hence the power of the goodness-of-fit test is an important consideration especially when the sample size is small as in this study.¶
Although the sensitivity of all four categories (erect vs. supine radiographs using carinal angle vs. SAD) is consistently above 80%, the specificity is so poor (<53%) that the utility of this test may only be limited to a rule-out strategy (i.e. a normal value would rule out enlargement of the left atrium). Positive and negative predictive values (as well as the area under the ROC curve) are close to 50% (+/- 60%) implying that either a positive or a negative re... [1]

References

1. Rossi A. et al. *Usefulness of left atrial size in predicting postoperative symptomatic improvement in patients with aortic stenosis.* Am J Cardiol. 2000; 86: 567-570.
2. Svanegaard J. et al. *Correlation between serial changes in left-sided heart chambers and atrial natriuretic peptide and N-terminal pro atrial natriuretic peptide after a first myocardial infarction: An echocardiographic study.* Scand Cardiovasc J. 1999; 33: 355-361.
3. Modena MG. et al. *Left atrial size is the major predictor of cardiac death and overall clinical outcome in patients with dilated cardiomyopathy: a long term follow-up study.* Clin Cardiol. 1997; 20: 553-560.
4. Stefadouros MA. et al. *The long-term effect of successful mitral balloon valvotomy on left atrial size.* J heart valve Dis. 1999; 8: 534-550.
5. Ozer N. et al. *Left atrial appendage function in patients with cardioembolic stroke in sinus rhythm and atrial fibrillation.* J Am Soc echocardiogr. 2000; 13: 661-665.
6. Sadanandan S, Sherrid MV. C*linical and echocardiographic characteristics of left atrial spontaneous echo contrast in sinus rhythm.* J Am Coll Cardiol. 2000; 35: 1932-1938.
7. Mattioli AV et el. *serial evaluation of Left atrial dimension after cardioversion for atrial fibrillation and relation to atrial function.* Am J Cardiol. 2000; 85: 832-836.
8. Di Tullio MR. et al. *Left atrial size and the risk of ischemic stroke in an ethnically mixed population.* Stroke. 1999; 30: 2019-2024.
9. Sankar NM, Farnsworth AE. *Left atrial reduction for chronic atrial fibrillation associated with mitral valve disease.* Ann Thorac Surg. 1998; 66: 254-256.
10. Gottdiener JS. et al. *Left atrial size in hypertensive men: influence of obesity, race and age. Department of Veterans Affairs Cooperative Study Group on Antihypertensive Agents.* J Am Coll Cardiol. 1997; 29: 651-658.
11. Katz ES. et al. *Surgical left atrial appendage legation is frequently incomplete: a transesophageal echocardiographic study.* J Am Coll Cardiol. 2000; 36: 468-471.
12. Brodsky MA. et al. *Amiodarome for maintenance of sinus rhythm after conversion of atrial fibrillation in the setting of a dilated left atrial.* Am J Cardiol. 1987; 60: 572-575.
13. Taskin V. et al. *Tracheal carinal angle and left atrial size.* Arch intern Med. 1991; 151: 307-308.
14. Murray JG. et al. *Widening of the tracheal bifurcation on chest radiographs: value as a sign of left atrial enlargement.* Am J Roeutgenol. 1995; 164: 1089-1092.
15. ASE et al. *Criteria for measuring the size of the left atrial.* J Am Soc Echocardiogr. 1978; 58: 1072-1083.
16. James B. et al. *The Echo Manual.* The Mayo Foundation. 1999; 4: 37.
17. Haskin PH. et al. *Normal tracheal bifurcation angle: a reassessment.* Am J R. 1982; 139: 879-882.
18. Lin L. *A concordance correlation coefficient to evaluate reproducitility.* Biometrics 45: 255 – 268.
19. Bland JM, Altman DG. *Measurement: Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet Feb 8, 1986: 306-310.
20. Hosmer DW, Lemeshow S. *AppliedLogistic Regression* 2nd Edition*:* 11, 13, 85, 164, 183-184.

21. Hosmer DW, Lemeshow S. *AppliedLogistic Regression* 2nd Edition: 60, 61, 99-111.

22. Reichenheim ME. *Two-graph receiver operating characteristic.* The Stata Journal, 2002; 4: 351-357.

23. Stata Corp LP 2005.

24. Kircher B. et al. *Left atrial volume determination by biplane two-dimensional echocardiography: validation by cine computed tomography.* Am heart J. 1999; 121: 864-871.

25. Lester SJ. et al. *Best method in clinical practice and in research studies to determine left atrial size.* Am J Cardiol. 1999; 84: 829-832.

26. Rodevan O. et al. *Left atrial volumes assessed by three- and two-dimensional echocardiography compared to MRI estimates.* Int J Card Imaging. 1999; 15: 397-410.

Discussion

This study aimed to determine the diagnostic utility of chest radiographs compared to echocardiography in estimating left atrial size. It also explored the intra- and interobserver variability of chest radiograph measurements.

Data form 156 patients were sampled, exceeding the initial sampling frame of 150 patients (which allowed for 10% of untraceable and 20% unusable radiographs resulting in a final estimated sample size of 106 patients, given a prevalence of 70% of large left atrium on ultrasound).

Intraobserver agreement of measures of left atrial size was excellent for both examiners. This implies that the methods (using the carinal angle or the SAD) are very precise. Similar results were found when the two different examiners were compared (but less so for supine radiographs). This underscores the repeatability of these methods.

The linear model obtained using logistic regression demonstrated that only erect chest radiographs were useful in predicting whether the left atrium was enlarged as supine chest radiographs did not yield statistically significant results. This may be due to the small number of supine radiographs included in the study.

Goodness-of-fit results for the above model showed that the model is a reasonable fit for both erect and supine radiographs using either variable (angle or SAD), which implies that the model effectively describes the outcome variable. As the desired outcome for this model is the decision not to reject the null hypothesis that the model fits the data, the above conclusion is subject to a type II error and hence the power of the goodness-of-fit test is an important consideration especially when the sample size is small as in this study.

Although the sensitivity of all four categories (erect vs. supine radiographs using carinal angle vs. SAD) is consistently above 80%, the specificity is so poor (<53%) that the utility of this test may only be limited to a rule-out strategy (i.e. a normal value would rule out enlargement of the left atrium). Positive and negative predictive values (as well as the area under the ROC

curve) are close to 50% (+/- 60%) implying that either a positive or a negative result are hardly better than a random guess in predicting the left atrial size. This in effect means that a rule-out strategy using these variables (carinal angle / SAD) alone is not feasible as the model's ability to discriminate between enlarged and normal left atria is poor.

The visual representation of the relationship between the actual left atrial size and the estimate (carinal angle and SAD) as portrayed in the lowess smoothing curves suggests a linear relationship only at the extremes of carinal angles (or equivalently SAD). Between angles of 50° (18mm) and 110° (32mm) the estimation of left atrial size is unreliable, especially for supine radiographs.

The results of the quartile analysis as plotted show a definite deviation from linearity in the fourth quartile throughout all permutations of radiograph types and variable used in the analysis. In contrast, the results suggest linearity in the logit for both angle and SAD in the first three quartiles for erect radiographs.

The last measure of linearity in the logit (fractional polynomials) suggested a non-linear model using a transformed carinal angle (its inverse square). As this model’s performance did not differ significantly from the linear one, it may be concluded that no non-linear transformation can in fact improve the predictive value of the model. Therefore, the model may be regarded as linear in the logit of the parameters, using either dependent variable (i.e. subcarinal angle or SAD). This means that logistic regression is the appropriate method for deriving a prediction rule for left atrial size category (enlarged/ not enlarged) using radiograph-derived left atrial size estimates.

Optimal cut-off points for the carinal angle and SAD of both erect and supine radiographs revealed surprising values. The commonly used cut-off value for the carinal angle in clinical practice thought to indicate an enlarged left atrium is 90°. In contrast, this study found the cut-off points to be 72° for erect and 84° for supine films. In addition to this, the SAD measure seems to be interchangeable with the angle measurement throughout the analysis. Corroesponding cut-off values for SAD for erect and supine films were 24.1mm and 26.9mm, respectively. This simplifies the estimation of left atrial

size as this measurement does not require any equipment not usually found in a physician's pocket.

Conclusion

Both carinal angle as well as SAD can be used interchangeably on erect chest radiographs of good quality to predict left atrial enlargement with great precision but poor accuracy. This might also be true for supine radiographs, although the sample size in this study was too small to yield a statistically significant result. Cut-off values below the traditionally used 90° were found to predict left atrial enlargement with improved diagnostic characteristics. The corresponding SAD appeared to have slightly better diagnostic discriminatory capability regarding test sensitivity, specificity and corresponding confidence intervals surrounding those measures.

In conclusion, use of chest radiographs has poor diagnostic utility in predicting left atrial size. Echocardiography remains the preferred method of determining left atrial size.