

**THE QUALITY OF DIGITAL SCREENING &
DIAGNOSTIC MAMMOGRAPHIC IMAGES USING
STANDARD PGMI CRITERIA IN
HOSPITAL UNIVERSITI SAINS MALAYSIA**

By

DR SHARIPAH INTAN SHAFINA BT SYED ABAS

**Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Medicine
(Radiology)**



UNIVERSITI SAINS MALAYSIA

2015

In the name of Allah,
the Most Gracious and the
Most Merciful

Acknowledgements

Firstly, I wish to express my sincere gratitude to my principal supervisor, **Associate Professor Dr Mohd Ezane Abd Aziz** for encouraging me to undertake this research title and instructing me in the art of scientific investigation. I thank him for his guidance, tolerance and constructive criticism. He has been instrumental in helping me achieve my biggest learning curve and I cannot thank him enough.

I am deeply indebted to Director of Hospital USM **Dato' Dr Zaidun Kamari** and **Dr Nik Munirah Nik Mahdi** as Head of Department of Radiology HUSM for allowing me access to mammogram images. My appreciation to **Dr Fairos Mutalib**, Radiologist HRPZII and Dr Mohd Ariff Abas as Head of Radiology Department HRPZII for supporting this study.

My sincere thanks go to the mammographers of Radiology Departments of Hospital Universiti Sains Malaysia - **Munirah, Salwa, Afidah, Hidayah**, and HRPZII - **Siti Sarah** and **Salwany** for volunteering in this study. I would like to specially thank you **Nurhazwani Hamid** for her constructive statistical understanding and comments during data analysis of this thesis and my special thanks to **Encik Zakaria**, IT officer HUSM for helping me with technical aspects in relation to mammogram images in PACS HUSM.

To my husband, **Dr Mohd Rahimi Noh**, my parents, **Syed Abas Syed Ibrahim & Zabariah Mat Piah**, mother-in-law **Siti Aishah Othman** and all my children, **Mariam Sofiah, Maisarah Anisah, Solehah and Salahuddin** who have offered me support, understanding, patience and love. My gratitude to all my lecturers and friends especially **Dr Siti Mayzura** and **Dr Siti Noorul Arisah**, for giving me an opportunity to work together and a wonderful friendship. I am honoured by the presence of all of you my life and I send you all my deepest appreciation.

JAZAKALLAH KHAIR.....

Table of Contents

Contents	Page
Acknowledgements.....	ii
List of Tables.....	vi
List of Charts.....	vi
List of Figures.....	vii
Abbreviations.....	viii
Abstrak Bahasa Malaysia.....	ix - x
English Abstract.....	xi - xii
1.0 INTRODUCTION.....	1
2.0 LITERATURE REVIEW.....	2 - 24
2.1 What is PGMI	4 - 5
2.1.1 Classification of Images (CC view)	5 - 7
2.1.2 Classification of Images (MLO view)	11 - 12
2.2 Quality standards of mammogram	17
2.3 Studies on PGMI: A perspective.....	18 - 24

3.0	AIMS AND OBJECTIVES	
3.1	Aim.....	26
3.2	Objective	
3.2.1	General objective.....	27
3.2.2	Specific objective.....	27
3.3	Hypothesis.....	28
3.4	Research questions.....	28
4.0	METHODOLOGY	
4.1	Study design.....	30
4.2	Population sampling method	
4.2.1	Reference population.....	30
4.2.2	Study population.....	30
4.2.3	Inclusion criteria.....	30
4.2.4	Exclusion criteria.....	31
4.3	Sampling method and sample size.....	32

4.4	Subject selection.....	33
4.5	Materials and methods.....	34 - 35
4.6	Image evaluation.....	36 - 37
4.7	Statistical analysis.....	37
5.0	RESULTS	
5.1	Sample characteristics	
5.1.1	Age.....	39
5.1.2	Race.....	39
5.1.3	Year of mammogram done.....	40
5.1.4	Purpose of mammogram- diagnostic/screening.....	40
5.1.5	Reason for mammogram.....	40
5.2	Specific objective 1.....	44
5.3	Specific objective 2.....	48
5.4	Specific objective 3.....	51
6.0	DISCUSSION.....	54 - 65
7.0	SUMMARY AND CONCLUSION.....	67
8.0	SUGGESTIONS.....	69
9.0	LIMITATIONS OF THE STUDY.....	71
	REFERENCES.....	73 - 75
	APPENDIX 1: PGMI criteria for MLO/CCI view.....	77 - 78
	APPENDIX 2: Classification of PGMI images.....	79 - 82
	APPENDIX 3: Mammographic views/technique.....	83 - 86
	APPENDIX 4: Sample of data collection sheet.....	87

APPENDIX 5: Sample of coding selection.....	88
APPENDIX 6: Sample of coding selection.....	89
APPENDIX 7: Results of PGMI rating for every set.....	90 – 93
APPENDIX 9: Study protocol (Flow Chart).....	95
APPENDIX 10: Gantt Chart.....	96

List of Tables	Page
Table 1: Sample characteristics.....	39
Table 2: Proportion of mammograms classified as <i>perfect, good, moderate, inadequate</i> based on standard PGMI criteria.....	45
Table 3: PGMI score (R1 – R7).....	47
Table 4: : Inter-observer reliability of PGMI - comparisons between ratings from six radiographers' assessing mammograms.....	48
Table 4.1: Sub-division of results as according to coding system of reasons of image assessment.....	50
Table 5: Inter-observer reliability of PGMI - comparisons between ratings among six radiographers' & a Radiologist assessing mammograms.....	51

List of Charts

Chart 1: Demography: Age.....	41
Chart 2: Reasons for mammogram.....	42

Chart 3: Demography: Racial distribution.....	43
Chart 4: Proportion of mammogram according to PGMI rating.....	46
Chart 5 & 6 : Percentages of mammogram by expert observer as according to PGMI rating.....	

List of Figures	Page
Figure 1 : PGMI criteria for cranio-caudal view.....	8
Figure 2 : Mammogram image in left CC view (<i>Perfect</i>).....	9
Figure 3 : Mammogram image in Right CC view (<i>Moderate</i>)	10
Figure 4: PGMI criteria for Mediolateral Oblique view	13
Figure 5: Mammogram image in right MLO view (<i>Perfect</i>).....	14
Figure 6: Mammogram image in left MLO view (<i>Moderate</i>).....	15
Figure 7: Mammogram image in Right MLO view (<i>Inadequate</i>).....	16

Abbreviations

ACR	American College of Radiology
BIRADS	Breast Imaging-Reporting and Data System
BSE	Breast self examination
CC	Craniocaudal
IMF	Inframammary fold
MLO	Mediolateral oblique
PACS	Picture archiving and communication system
PGMI	Perfect, Good, Moderate, Inadequate

Symbols

K	Kappa
---	-------

Abstrak

Tajuk: Kajian berkaitan kualiti imej digital saringan dan diagnostik mammogram menggunakan kriteria standard PGMI di Hospital Universiti Sains Malaysia.

Tujuan dan objektif:

Di HUSM, audit mammogram menggunakan PGMI tidak pernah dilakukan bagi penarafan kualiti imej. Penilaian PGMI berdasarkan kriteria tertentu yang menentukan imej mammogram tersebut sebagai *sempurna*, *baik*, *sederhana* dan *tidak mencukupi*. Menurut tatacara Kolej Radiologi Malaysia (The College of Radiology Malaysia) mengesyorkan > 97% daripada imej berada dalam kategori yang *sempurna*, *baik* atau *sederhana* dengan keseluruhan 75% dalam kumpulan *sempurna & baik*; > 3% dalam kumpulan *sempurna*; dan < 3% daripada imej dikalsifikasikan sebagai *tidak mencukupi*.

Metodologi:

Sebanyak 107 imej digital diagnostik dan saringan mammogram yang dibuat di HUSM telah diberikan kepada 2 kumpulan penilai yang terdiri dari 4 orang juru x-ray HUSM (R1, R2, R3, R4) dan 2 orang juru x-ray HRPZII (R5, R6). Set ujian telah dipilih secara sistematik dan mempunyai pelbagai nilai dari setiap kategori PGMI dan secara bebas dihakimi oleh enam orang juru x-ray tersebut. Setiap penilai (juru x-ray) mempunyai pengalaman menggunakan PGMI. Setiap imej dinilai secara individu

mengikut sistem PGMI. Sebarang pemarkahan kurang daripada sempurna (*perfect*) perlu disertakan dengan alasan mengapa ianya bersifat demikian. Semua set akan dibandingkan sesama mereka dan juga dengan seorang pemerhati pakar (R7).

Keputusan:

PGMI standard: secara keseluruhan penilaian di HUSM bagi *sempurna*, *baik* dan *sederhana* (PGM) penilaian adalah dikira sebagai 98%.

Peratusan Ini boleh diterima dan seperti tahap yang ditentukan oleh Kolej Radiologi (> 97%). Bagi kategori *sempurna & baik*, ianya lebih rendah berbanding standard sedia ada iaitu 52% (QA yang standard adalah 75%). Bagi imej *sempurna*, peratusannya ialah 15% (QA Standard > 3%) dan *tidak mencukupi* adalah 2% (QAP standard adalah < 3%).

Kebolehpercayaan di kalangan pemerhati PGMI sesama juru x-ray (*interobservers' reliability among radiographers*) adalah rendah iaitu $K = 0.18$. Kebanyakan imej yang menunjukkan *poor interobservers' reliability* adalah imej yang mempunyai rating PGMI yang baik dan sederhana.

Kebolehpercayaan di kalangan pemerhati PGMI dan penilai pakar (*interobservers' reliability between radiographers' and Radiologist*): *poor interobservers' reliability* adalah juga rendah; $\kappa = 0.20$. Kebanyakan imej yang menunjukkan *poor interobservers' reliability* adalah imej yang mempunyai rating baik, sederhana dan tidak mencukupi.

Kesimpulan:

Kepercayaan terhadap penilaian PGMI mesti dipertingkatkan kerana kajian semasa menunjukkan kepelbagaian dalam pemahaman serta tafsiran yang bersifat subjektif. Penambahbaikan atau pengubahsuaian perlu dilakukan agar PGMI kekal efisien.

Abstract

Title: The quality of digital screening and diagnostic mammographic images using PGMI standard criteria in Hospital Universiti Sains Malaysia.

Introduction and Objectives:

In HUSM, the audit of mammogram using PGMI was never performed for rating image quality. PGMI rating comprises of criterias determining mammograms as *perfect*, *good*, *moderate* and *inadequate*. The College of Radiology Malaysia guidelines recommends >97% of images to be in *perfect*, *good* or *moderate* categories with overall 75% in the *perfect* & *good* groups; >3% in the *perfect* group; and <3% of images to be classified *inadequate*.

Methodology:

Digital mammograms from 107 consecutively screened and diagnostic mammograms were sourced in two centers; namely HUSM and HRPZII. Test sets were enriched with mammograms from each PGMI category and independently scored by six radiographers, each with ≥ 3 years' experience, using PGMI. Each image was

individually scored P, G, M, or I. Reasons for scoring less than *perfect* were documented and each mammogram assigned an overall PGMI score. Test sets were compared with an expert observer.

Results:

PGMI standard: The overall rating in HUSM for *perfect*, *good* and *moderate* (PGM) rating were calculated as 98%. This was acceptable to the standard set by Malaysian College of Radiology (>97%). While in the *perfect* & *good* groups (PG); it was lower than standard which was 52% (QA standard is 75%). As for *perfect*, its proportion was 15% (QA standard is >3%) and for *inadequate* was 2% (QAP standard is < 3%) .

PGMI inter-observer reliability among radiographers' : Overall poor agreement with $\kappa = 0.18$. Most images with poor agreement were related to good and moderate image rating whereas PGMI inter-observer reliability radiographers' vs. Radiologist: Overall poor agreement with $\kappa = 0.20$ and most images with poor agreement were related to good, moderate and inadequate images.

Conclusion:

Reliability of PGMI must be improved as current study showed its variability and subjective interpretation. Efforts must be made for improvements or modifications of PGMI in order to reduce its subjectivity and maintains its efficiency.

Abstract

THE QUALITY OF DIGITAL SCREENING AND DIAGNOSTIC MAMMOGRAPHIC IMAGES USING PGMI STANDARD CRITERIA IN HOSPITAL UNIVERSITI SAINS MALAYSIA

Dr Sharipah Intan Shafina bt Syed Abas

MMed Radiology

Department of Radiology

School of Medical Sciences, Universiti Sains Malaysia

Health Campus, 16150 Kelantan, Malaysia

Introduction: The mammography has its own unique quality assurance program in order to produce a constantly high quality images. It emphasizes on quality control (QC) such as direct equipment assessment (mammogram machines) and quality audits i.e. film reject analysis and PGMI which acts as an indirect assessment tool for personnel involved; mainly radiographers. PGMI is a part of standard quality audit for mammogram and widely practiced in many countries.

Objectives: In HUSM, the audit of mammogram using PGMI was never performed for rating image quality. PGMI rating comprised of criteria determining mammograms as *perfect*, *good*, *moderate* and *inadequate*. The College of Radiology Malaysia guidelines recommends >97% of images to be in *perfect*, *good* or *moderate* categories with overall 75% in the *perfect* & *good* groups; >3% in the *perfect* group; and <3% of images to be classified *inadequate*.

Methodology: Digital mammograms from 107 consecutively screened and diagnostic mammograms were sourced in two centers; namely HUSM and HRPZII. Test sets were enriched with mammograms from each PGMI category and independently scored by six radiographers, each with ≥ 3 years' experience, using PGMI. Each image was individually scored P, G, M, or I. Reasons for scoring less than *perfect* were documented and each mammogram assigned an overall PGMI score. Test sets were compared with an expert observer.

Results: PGMI standard: The overall rating in HUSM for *perfect*, *good* and *moderate* (PGM) rating were calculated as 98%. This was acceptable to the standard set by Malaysian College of Radiology ($>97\%$). While in the *perfect* & *good* groups (PG); it was lower than standard which was 52% (QA standard is 75%). As for *perfect*, its proportion was 15% (QA standard is $>3\%$) and for *inadequate* was 2% (QAP standard is $< 3\%$). PGMI inter-observer reliability among radiographers' shows overall poor agreement with $\kappa = 0.18$. Most images with poor agreement were related to good and moderate image rating whereas PGMI inter-observer reliability radiographers' vs. Radiologist shows overall poor agreement with $\kappa = 0.20$ and most images with poor agreement were related to good, moderate and inadequate images.

Conclusion: Reliability of PGMI must be improved as current study showed its variability and subjective interpretation. Efforts must be made for improvements or modifications of PGMI in order to reduce its subjectivity and maintains its efficiency.

Prof Madya Dr Mohd Ezane Abd Aziz: Supervisor

CHAPTER ONE: INTRODUCTION

1.0 INTRODUCTION

Detection of breast lesions in mammograms were crucial in managing patients with breast symptoms. Both diagnostic and screening mammograms, therefore, remain a gold standard in the detection of breast cancer. Nevertheless, early detection was of utmost importance as it might lead to better outcome in terms of morbidity and mortality where several studies have shown screening mammography could decrease breast cancer mortality from 15% to 58 % .

The mammography has its own unique quality assurance program in order to produce a constantly high quality images. It emphasizes on quality control (QC) such as direct equipment assessment (mammogram machines) and quality audits i.e. film reject analysis and PGMI which acts as an indirect assessment tool for personnel involved; mainly – radiographers. PGMI is a part of standard quality audit for mammogram and widely practiced in many countries.

In Malaysia, College of Radiology had adapted this method as part of quality assurance program (QAP) since 2008. In HUSM, only quality control (QC) assessment is routinely done for mammogram. However, the quality audits using PGMI are never being conducted. Therefore, the status of mammographic images quality remains undetermined.

This study was performed to fulfill this gap by evaluating the quality of mammograms done in HUSM whether it is in accordance to the quality standard as set by Malaysian College of Radiology.

Hopefully, it would improve the current mammographic practice here and might serve as a baseline for quality improvement in future.

CHAPTER TWO: LITERATURE REVIEW

2.0 LITERATURE REVIEW

Quality assurance (QA) programs for mammography were first introduced in 1989 by United Kingdom Mammography Trainers Group as an instrument for analysing mammography images. The initial emphasis of QA programme was that all aspects of mammography must be of a very high quality in order to achieve the anticipated reduction in breast cancer mortality (NHSBSP, March 2000). Later, several other European countries such as Ireland, Norway, Italy, Slovenia and other countries like Australia, Hong Kong and Singapore also adapted similar QA program ((Hofvind *et al.*, 2009); (*Breast Screen* New South Wales, 2013); (The National Cancer Screening Service Board, 2008).

In Malaysia, the guidelines for QA mammography has been established since 2008 by the Malaysian College of Radiology (CoR) which was based on the guidelines for the National Health Service Breast Screening Program (NHSBSP) and European Guidelines.

The guideline mentions quality audits as an integral part of the quality management in mammography where radiographers and radiologist were part of the QA team. The radiographers or mammographers plays an important role especially to their task in handling the clients; while at the same time, performing the mammogram and doing it optimally with a good techniques in order to achieve high quality images.

It clearly states that QA assessment is to be performed at regular intervals, preferably on annual basis.

2.1 What is PGMI?

PGMI is a methodical assessment of quality of mammogram images; in which it is an acronym where P stands for a *perfect* mammogram, G is for a *good* mammogram, M is for a *moderate* mammogram and I is for an *inadequate* mammogram.

This method was adapted by Malaysian College of Radiology for quality mammogram guidelines (Malaysian College of Radiology, 2008) from the UK and Ireland and several other countries (NHSBSP, March 2000), (The National Cancer Screening Service Board, 2008), (Australia, 2008).

There are several criterias for acceptability of mammogram images using PGMI (Appendix 1). The images were classified as P, G, M or I according to these criterias.

2.1.1 Classification of Images – Craniocaudal view (CC)

P = *Perfect* images - both breast (right and left) images meet the listed criteria
(refer figure 1 & 2)

G = Good images

- All postero-medial tissue visualized (axillary portion of breast not to be included at expense of medial portion)
- Nipple in profile
- Nipple in midline of imaged breast
- Both images meet all criteria listed inclusive of b to f as listed in Appendix 1.1
- A minor degree of variation in items g to i as listed in Appendix 1.1 will be accepted for categorization as G

M = Moderate images (refer figure 3)

- Most breast tissue imaged (all breast tissue must be imaged on MLO view)
- Nipple not in profile but clearly distinguishable from surrounding breast tissue (however, nipple must be in profile on MLO view)
- Nipple not in midline of the imaged breast
- Correct film identification to workplace requirement
- Correct exposure
- Adequate compression
- Absence of movement

- Correct processing
- Artefacts which do not obscure the image
- Skin folds which do not obscure the breast tissue
- Asymmetrical images

I = *Inadequate images*

- Significant part of the breast tissue is not imaged
- Incomplete or incorrect identification
- Incorrect exposure
- Inadequate compression which hinders diagnosis
- Blurred image
- Incorrect processing
- Overlying artifacts
- Skin folds which obscure the image

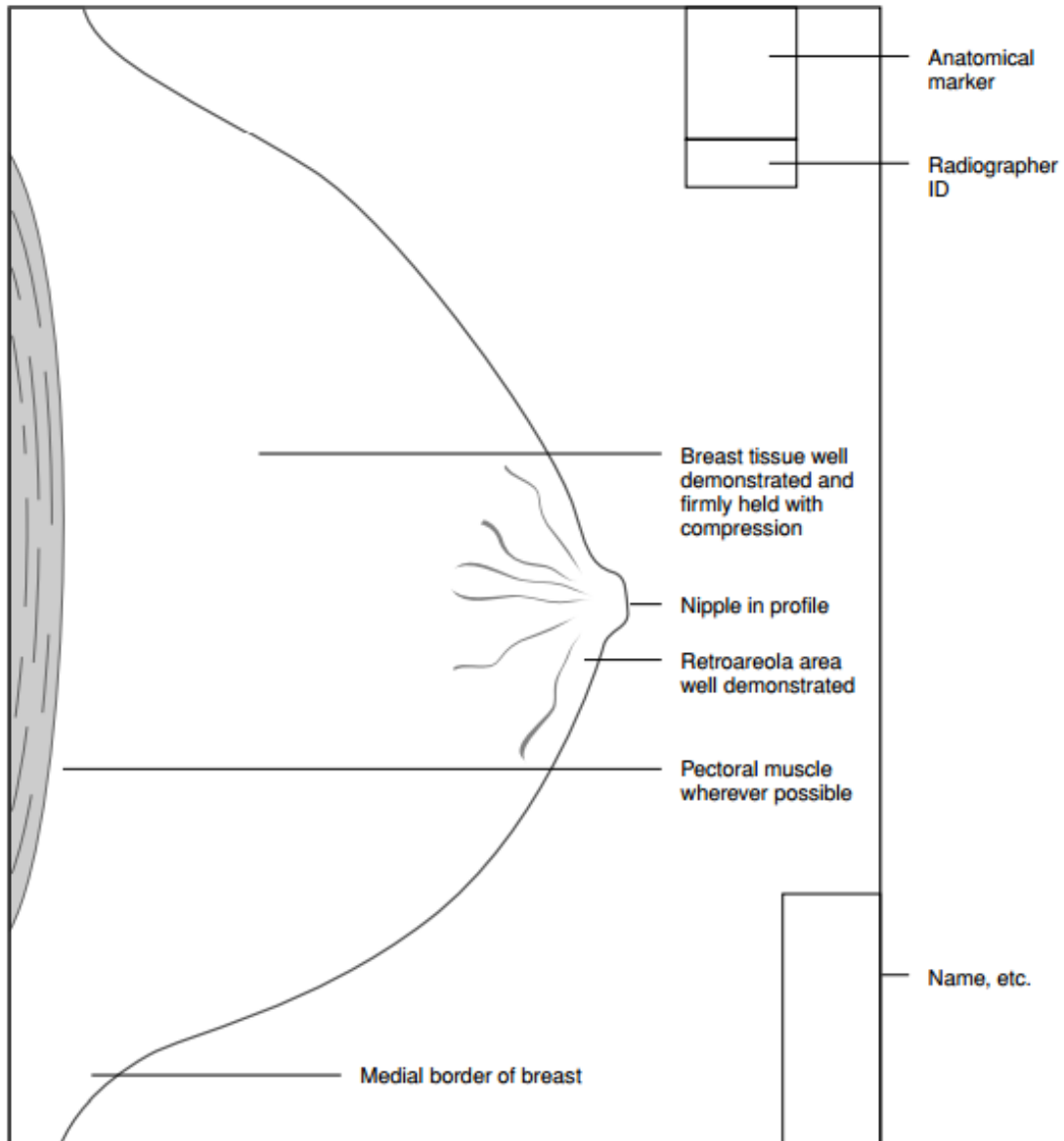


Figure 1: PGMI criteria for crano-caudal view (NHSBSP, March 2000)

SAMPLE IMAGES – NOT FOR MEDICAL USE



Figure 2: Mammogram image in left CC view (*Perfect*)

SAMPLE IMAGES – NOT FOR MEDICAL USE



Figure 3: Mammogram image in Right CC view (*Moderate*) – *nipple not in profile*

2.1.2 Classification of Images (MLO view)

P = Perfect images - both images meet all listed criteria (refer figure 4 & 5).

G = Good images

- All breast tissue imaged
- Pectoral muscle well demonstrated
- Nipple in profile
- Infra-mammary fold well demonstrated
- Both images meet all criteria listed inclusive of b to f (refer Appendix 1.2)
- A minor degree of variation in items g to i as listed in Appendix 1.2 will be accepted for categorization as G

M = Moderate images

- All breast tissue imaged
- Pectoral muscle not to nipple level but posterior breast tissue adequately shown
- Nipple not in profile but retro-areolar tissue well demonstrated
- Infra-mammary fold not clearly demonstrated but breast tissue adequately shown (refer figure 6)
- Correct film identification
- Correct exposure
- Adequate compression

- Absence of movement
- Correct processing
- Artefacts which do not obscure the image
- Skin folds which do not obscure the breast tissue
- Asymmetrical images

I = Inadequate images

- Part of the breast not imaged (refer figure 7)
- Incomplete or incorrect identification
- Incorrect or inadequate exposure
- Inadequate compression which hinders diagnosis
- Blurred image
- Incorrect processing
- Overlying artifacts
- Skin folds which obscure the image

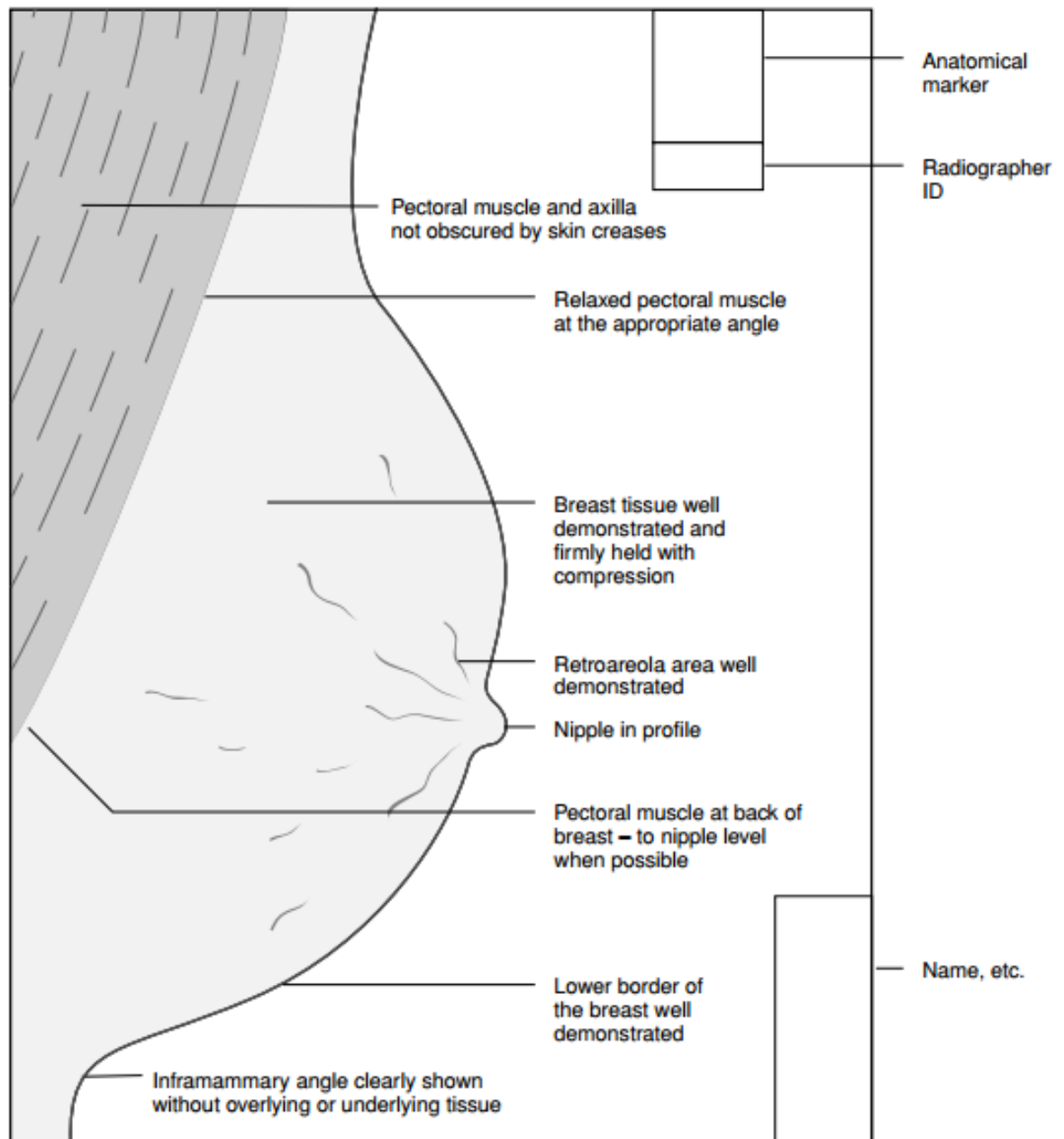


Figure 4: PGMI criteria for medio-lateral oblique view (NHSBSP, March 2000)

SAMPLE IMAGES – NOT FOR MEDICAL USE



Figure 5: Mammogram image in right MLO view (*Perfect*)

SAMPLE IMAGES – NOT FOR MEDICAL USE



Figure 6: Mammogram image in left MLO view (*Moderate*) - *Infra-mammary fold not clearly demonstrated but breast tissue is adequately shown*

SAMPLE IMAGES – NOT FOR MEDICAL USE



Figure 7: Mammogram image in Right MLO view (*Inadequate*) - Folds of skin obscuring parts breast tissue at axillary region

2.2 QUALITY STANDARDS OF MAMMOGRAM

The standard criteria of PGMI rating as set by the Malaysian College of Radiology guideline includes:

- > 97% of images to be in P, G or M categories
- overall 75% in the P & G groups is desirable with
- > 3% in the P group
- < 3% of images to be classified "Inadequate"

This guideline also mentions on film rating performance of every radiographer in which their criteria includes:

- > 75% should be in *perfect* or *good* group in PGMI rating system.
- > 97% should be in P, G, M groups.
- < 3% in *inadequate* group.

However, the aim of this study was not related to rating mammogram achievement of every radiographer.

2.3 STUDIES ON PGMI : A PERSPECTIVE

There was no published Malaysian PGMI studies known to author except for QA mammogram guideline manual on PGMI in Malaysian College of Radiology website. Apart from that, there were limited studies conducted around the world for assessment of mammogram images using PGMI criteria.

An Australian study for example, conducted by *BreastScreen* New South Wales Hunter Region and Wyong Shire by (Thompson, 2009) in which PGMI assessment were made for 343 sets of mammogram images; where majority of the mammogram images were ranked as M (moderate) (44.9%), G (good) (42.3%), I (10.5%) and P (perfect) (2.3%) .

A study conducted by Podobnik Gasper from Slovenia from 2008 until 2009 evaluated radiographers understanding of PGMI by using questionnaires related to its criteria. 600 questionnaires were distributed, had fairly good response rate (88.2%) and the results mentioned on problems pertaining to breast positioning at CC projection where *pectoral muscle were not seen* and *nipple were not in profile*. While at MLO views, the problems were related to the pectoral muscles which were not up to the nipple level and IMF were not clearly seen (Podobnik, 2008).

Several studies pertaining to inter-observer agreement using PGMI were conducted in Europe. There were two Norwegian studies done from 2009 and 2010; one study done by Hofvind, *et al* 2009 from the Norwegian Breast Cancer

Screening Programme and another by Gullien, *et al* 2010 from Oslo University Hospital.

The former had evaluated a total of 1280 mammograms from all 16 breast centres involved in breast screening programmes and these images were PGMI-rated by 2 groups comprised of local-PGMI radiographers and another group by expert radiographers. Results shown that the expert radiographer classified a higher proportion of both CC (28%) and MLO (14%) mammograms as *inadequate* than did the local-PGMI radiographers (7% and 3%, respectively; $P < 0.001$ for both) (Hofvind *et al.*, 2009).

Meanwhile in the latter study, which emphasized upon inter-observer agreement among radiographers for assessing MLO screening mammograms; each with varied PGMI experiences. One internal experienced (A, Oslo Univ.), one external experienced (B, non-Oslo Univ.) and one internal inexperienced (C, Oslo Univ.) evaluated 240 images using the PGMI of the Norwegian Breast Cancer Screening Program (NBCSP) QA manual.

The results shown varying inter-observer agreement between fair and good (as according to the κ - statistic) and that the agreement between the two internal PGMI radiographers was highest (A and C) and lowest between the two experienced PGMI radiographers (A and B) (Hofvind *et al.*, 2009).

Both studies were significant as their results were quite contradicting from each other. The former study mentioned that expert radiographers tended to grade mammograms as *inadequate* more than less experienced ones. While the

latter mentioned that work experience in the PGMI classification is actually, unimportant.

There was another important pilot study done by Boyce, *et al* which compared the use of PGMI scoring systems used in the UK and Norway by methodically assessing the technical quality of screening digital mammograms of 112 women in each center. The sample images were independently scored by four mammographers, each with ≥ 4 years' experience, using their own local PGMI. The test sets were later exchanged (Cambridge to Oslo, and vice versa) and similar process were repeated. The results shown that there is fair agreement ($\kappa = 0.38$) between centers in assigning images as acceptable overall (P, G, M) but poor inter-rater agreement within and between centers in further categorizing acceptable mammograms as P, G or M ($\kappa < 2$) criteria. Most common faults in Oslo were skin folds, and inadequate pectoralis muscle in Cambridge. Most faults overall in both centers were related to oblique views (MLO).

The poor rater agreement with differing faults due to the variation in number and interpretation of categories being used is an important point to note in this study. Apart from this, it is an important study that proved PGMI performance across countries (as in this case, of similar pan-European identity) can be very difficult and that the implementation of PGMI can be variable, subjective and interpreted locally.

With such notable variability in PGMI study, other studies were being conducted to compare PGMI with other methods of image classification.

One study in particular was done by Moreira, et al comparing the validity and reliability of PGMI with a modified classification system, EAR ; an acronym which stands for “Excellent”, “Acceptable” and “Repeat”. This study was done in New South Wales, Australia; population-based screening programmes (*BreastScreen NSW*) (Moreira *et al.*, 2005) where 30 sets of mammograms were rated by 21 radiographers and an expert panel.

The PGMI and EAR criteria were used to assign ratings to the medio-lateral oblique (MLO) and cranio-caudal (CC) views for each set of films. The results of this study shown low κ - values for both classification systems (0.01–0.17). However, PGMI produced significantly higher values than EAR with inter-observer agreement higher using PGMI than EAR for the MLO view (77% versus 74%, $p < 0.05$), but was similar for the CC view. The κ - values between raters and the reference standard were also low for both classification systems (0.05–0.15).

This study concluded that both PGMI and EAR have poor reliability and validity in evaluating mammogram quality; in which EAR was not a suitable alternative to PGMI (Moreira *et al.*, 2005)

There were several other methods recommended in quality assessment of digital mammogram images. For example;

- i. From (Bassett *et al.*, 2000) using Five step scale method in which grading were made from 1=worst to 5=best (Image quality categories used in the Clinical Image – Evaluation Process by ACR using) based on
 - a) *Positioning*: in which on a properly positioned medio-lateral oblique (MLO) view, the inferior aspect of the pectoral muscle should come to the posterior nipple line (PNL) and the pectoralis muscle should also be sufficiently wide. The breast is not sagging. Inframammary fold is open. The PNL on the CC view is within 1cm of its length on the MLO view.
 - b) *Compression*: Better compression can be identified by better spreading out of the breast markings.
 - c) *Exposure level*: Better exposure is evident from better penetration of the denser fibro-glandular tissue. Underexposure of the pectoralis muscle may prevent visualization of underlying structures in the breast.
 - d) *Contrast*: Image contrast shall permit differentiation of subtle tissue density differences.

- e) *Sharpness*: Margins of normal breast structures shall be distinct and not blurred.
 - f) *Noise*: Noise can be identified by an inhomogeneity in the background.
 - g) *Artefacts*: An artefact is any density variation on an image that does not reflect true attenuation differences in the subject.
- ii. From (Hemdal *et al.*, 2005) and (Grahn *et al.*, 2005) Revised European Union criteria using Relative grading method in which using right breast as reference, compared from left breast to right for image evaluation criteria related to positioning:
- a) Pectoral muscle at correct angle ,
 - b) Infra-mammary angle visualised ,
 - c) Nipple in profile, clear of overlying breast tissue and/or indicated by marker
 - d) No skin folds seen

Rating method used for evaluation of above criteria was 5 step scale grading method where

- 2: much worse than . . . ,
- 1: slightly worse than . . . ,
- 0: equal to . . . ,
- +1: slightly better than . . . ,
- +2: much better than . . .

Another method also used by Revised European Union criteria was Absolute grading method for image criteria related to detector performance, exposure parameters and patient movement which includes;

- (1) Visually sharp/clear reproduction of glandular tissue
- (2) Visually sharp/clear reproduction of fibrous strands in fat tissue
- (3) Visually sharp/clear reproduction of vascular structures in fat tissue
- (4) Visually sharp/clear reproduction of pectoral muscle margin
- (5) Visually sharp/clear reproduction of calcifications, when present
- (6) Acceptable noise level in the reproduction of the pectoral muscle

Rating method used: based on scoring of whether

- 1: the criterion was fulfilled,
- 0: the criterion was not fulfilled