EFOMP

Mammo Protocol

Mammo Working Group Protocol, March 2015

QUALITY CONTROLS IN DIGITAL MAMMOGRAPHY

PROTOCOL OF THE EFOMP MAMMO WORKING GROUP

MEMBERS

Gisella Gennaro (group leader)
Simona Avramova-Cholakova
Andrea Azzalini
Maria Luisa Chapel
Margarita Chevalier
Olivera Ciraj
Hugo de las Heras
Vesna Gershan
Bengt Hemdal
Elizabeth Keavey
Nico Lanconelli
Susanne Menhart
Màrio João Fartaria
Ana Pascoal
Kristin Pedersen
Stefano Rivetti
Ana Roda
Veronica Rossetti
Friedrich Semturs
Peter Sharp
Alberto Torresin (EFOMP Science Committe Chairman)

Preface

One of the aims of the European Federation of Organisations for Medical Physics (EFOMP), as stated in its constitution, is to encourage exchanges between the National Member Organisations and disseminate professional and scientific information through publications and meetings (**www.efomp.org**).

With the recent formation of Special Interest Groups (SIGs), it is now possible for EFOMP to publish scientific documents on topics of special interest such as scientific guidelines, quality assurance and quality control protocols, and monographs on particular topics in medical physics.

The main objective of the quality assurance protocol series is to develop a minimum set of easily implemented quality control tests on diagnostic and therapeutic systems that can be used to assure the performance of a system within a set and acceptable range. It is intended that these be implemented as part of the daily routine of medical physicists and system users throughout Europe in a harmonised way so allowing results to be compared.

The present document is the first of the series on quality control protocols and it was developed by a Task Force under the Mammo Working Group. In particular it proposes tests that can be applied by each country with the resources available locally, accepting the fact that there are different technological levels across the European countries.

Particularly with respect to image quality, the document:

1. Underlines the limitations of the so called "technical image quality", that is assessed by phantoms, namely that the acceptance criteria is arbitrary and is unable to predict the impact on clinical image quality (diagnostic performance).
2. Highlights the relationship between dose and image quality that is less and less relevant with digital mammography where an absolute value cannot be assigned to a phantom-based test and where the impact of image post-processing provides a subjective and not an objective measure of image quality that can be technically assessed.

Preface

3. Provides protocols that will allow the use of several phantoms to assess image quality so taking into consideration the fact the "minimum dose level" to provide a given threshold image quality depends both on the characteristics of the phantom and test method used.
4. Presents the results obtained with different phantoms, and compares the advantages and disadvantages of automated analysis of phantom images versus the scoring of human observers. In addition the increase of information derived from multi-detail versus single-detail phantoms is highlighted.

Results from different phantoms will be valuable in situations where a specific test object is used making people aware of its advantages and limitations.

Finally it is the intention of EFOMP to update these protocols every few years taking into consideration the experience gained through their implementation in different settings. EFOMP encourages users to report their experiences at conferences and in publications.

Table of Content

Introduction

The present document has concentrated on the need for harmonization of quality controls in (QCs) digital mammography, and proposes an alternative protocol with the intention of fulfilling the following criteria:

1. Tests included in the protocol have been selected on the basis of their relevance/priority ("Keep it simple").
2. The protocol should be applicable to any type of digital mammography system (DR and CR).
3. The protocol should be "compatible" or at least "not in contradiction" with other existing protocols.
4. The protocol should not require the use of a specific test object, and consequently, should be "tolerant" about image quality parameters and criteria.
5. The protocol should be widely applicable and adaptabl to local regulations.
6. The protocol should provide clear instructions about each test procedure, in order to minimize the potential for misinterpretation.

Figure 1: Characteristics of the EFOMP Protocol for Quality Controls in Digital Mammography.

In the following, the reasons for each characteristic summarized in Figure 1 will be further clarified.

Introduction

Relevance of tests

QC protocols often include a list of tests, generally grouped by categories (for example, dose, image quality, image detector, etc.), but usually without any reference to their relevance within the protocol. This is equivalent to assigning the same weight to any test/result, irrespective of the test importance and relevance. Moreover, as protocols on digital mammography are often derived from previous documents on screen-film mammography, they might include tests which were significant in the "analogue world", but are less and less important in the digital one. An example may help to clarify this.

> Many QC protocols require the measurement of tube voltage "accuracy" and precision, i.e. evaluating the relative difference between the peak voltage value (kV_p) set on the system and the measured one, and the reproducibility of the measured values. This test originates from the early days of screen-film mammography, when the automatic exposure control (AEC) performance was limited, as well as the dynamic range of films. At that time, the radiographer was supposed to set the peak voltage value manually according to compressed breast thickness and the AEC just determined the mAs level necessary to achieve the appropriate film exposure. Given the limited dynamic range of films, miscalibration of peak voltage setting could lead to either under- or over-exposures, compromising clinical image quality. For this reason, it was important to verify that the effective tube voltage was accurate (i.e. measured value consistent with the set value) and reproducible. In digital mammography, dynamic ranges of detectors are much wider, and high voltage generators much more accurate and precise than they were in screen-film mammography. Possible miscalibration of peak voltage, however unlikely, would have a very modest effect on the final processed images. This is why kV_p measurements have been discarded from this protocol.

In this document, it was decided to select and describe only "high-priority" tests for each component of the mammography imaging chain (x-ray source, automatic exposure control, detector, image quality, and monitor). Moreover, the harmonization process necessarily requires a simple, as well as a very widely applicable approach. The list of "highly relevant tests" included in the EFOMP protocol is shown in table 1.

Introduction

Table 1 - List of tests included in the EFOMP Protocol for Quality Controls in Digital Mammography, divided by category, with related objectives.

Category	Test	Objective(s)
X-ray source	• Tube Output • Half Value Layer (HVL)	• To determine parameters necessary for average glandular dose estimation
Automatic Exposure Control (AEC)	• AEC reproducibility • Signal-difference-to-noise ratio (SDNR) compensation and average glandular dose (AGD)	• To evaluate the AEC reproducibility and its capability to adjust exposure parameters as a function of the object absorption. • To verify that radiation is below the maximum levels considered as acceptable in screen-film mammography.
Detector	• Response function • Noise evaluation • "Uniformity" • Artifact search • Inter-plate variability (CR systems only)	• To verify that the detector behavior is consistent with that expected (linear/non-linear response function, quantum limited, signal non-uniformity compatible with heel effect when not corrected). • To verify the absence of artifacts. • To check that possible differences among different imaging plates used with CR systems are small.
Image Quality	• One or more image quality parameters, depending on the phantom used	• To assess image quality acceptability and define the baseline level from phantom images. • To test system reproducibility.
Monitors	• Grayscale DICOM display function • Luminance uniformity	- To verify monitor calibration according to the DICOM standard - To verify monitor uniformity.

Introduction

Applicability to any technology

There are several different technologies used for image detectors in digital mammography, with significant differences between them, as illustrated in Figure 2.

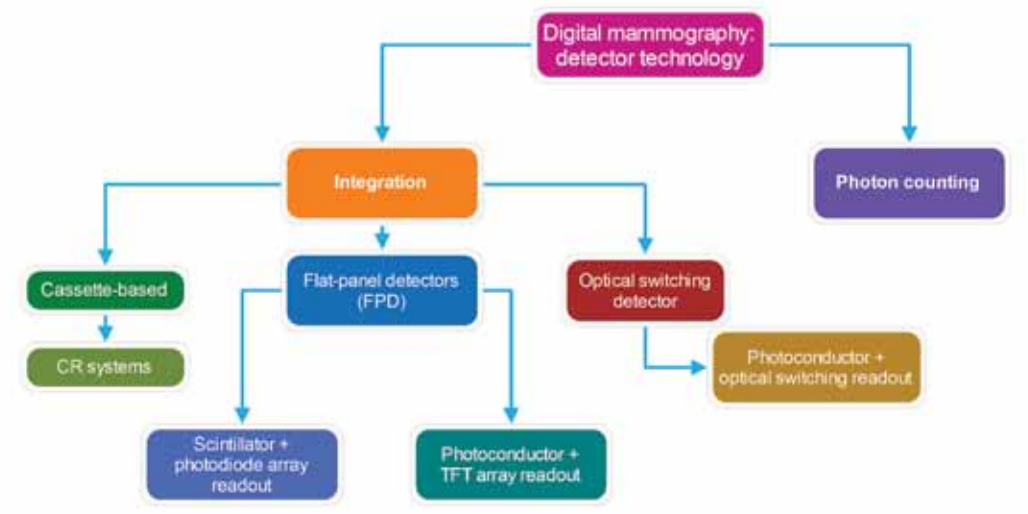


Figure 2 - Detector technologies used in digital mammography.

Moreover, digital mammography systems proposed by different manufacturers can differ not only in the detector technology, but also in the x-ray tube characteristics, the operation of the automatic exposure control (AEC), and the post-processing algorithm. The EFOMP protocol for Quality Controls in Digital Mammography is not specific for a given type of system, but has generally applicability, taking into account differences between systems.
The only exception is for the inter-plate variability test, which is specific for CR systems only.

Introduction

Compatibility with other QC documents
It is not the intention that the EFOMP protocol for Quality Controls in Digital Mammography conflicts with other QC documents/protocols. Compared to other documents, it includes a smaller number of tests, under the general principle of keeping the protocol simple in order to favour harmonization, but this does not prevent the protocol users from performing additional tests as required by other protocols or local regulations. In particular, the EFOMP protocol does not require any specific phantom for image quality assessment, allowing a range of possible choices which may depend on a number of factors (country Regulatory process, resources, phantoms already available, etc.).

Multiple phantoms for image quality assessment
The EFOMP protocol does not require the use of a given phantom for image quality assessment, leaving users free to adopt different phantoms, according to their needs/constraints. For this reason the "Image Quality" chapter is the only one in this protocol that does not provide precise instructions about the test procedures. In this chapter the limitations of the so-called "technical image quality" (i.e. image quality based on phantom images) in predicting clinical performance of any mammography system have been noted, and a list of different types of phantoms and quality features which can be obtained from phantom images have been provided. A description of several commercial phantoms and related parameters/scores able to reflect image quality was given, and this should allow users to adjust the protocol according to each country's requirement and define a precise test procedure for image quality assessment. Phantoms different from those identified in the "Image Quality" chapter could also be used.

Adapting to local regulations
Quality controls are mandatory in all European countries, but there are considerable differences in the obligations that each country's law applies to quality control. In some countries, the regulations are not specific about the protocol which should be applied, in others it is, and may even describe in details the individual tests required. This is why the need for harmonization has to take into account existing differences between countries, and leave the opportunity of adapting the EFOMP protocol to their specific needs. The characteristics of flexibility in the EFOMP protocol for Quality Controls in Digital Mammography is expected to allow its wide application, within and eventually outside Europe.

Introduction

Detailed description of test procedures

Harmonization of quality controls and the wide applicability of the EFOMP protocol also has to deal with text comprehensibility by many people with different experience in mammography measurements. In general, any experimental method is judged to be a "good method" if it is reproducible, i.e. if it is able to produce the same results when conducted by different people under the same conditions. This target is more likely to be achievable if each test procedure is described in detail, with step-by-step instructions, limiting as much as possible misinterpretations or variations by the person who is performing the test itself.

In order to do that, each chapter of this EFOMP protocol (all but the "Image Quality" one, as previously highlighted) consists of:

- A short introduction regarding each group of tests ("X-ray Source", "Automatic Exposure Control", "Detector", and "Monitors"), giving reasons why these tests have been selected.
- Test purpose(s).
- Description of the equipment necessary to perform each specific test, including pictures when necessary/useful.
- Test frequency.
- Step-by-step procedures, including settings of different types of digital systems, input and output tables corresponding to those reported in the Excel template provided for data collection and analysis, and numeric examples. All the instructions assume that the software tools used are the Excel template and the freeware software ImageJ. However, the instructions provided are detailed enough to allow development of other tools by users.
- Limiting values, when applicable.
- Some references for those who want to know more about a certain topic.

Each group of tests is characterized by a chapter color: red for the "X-ray source" chapter, violet for the "Automatic Exposure Control" chapter, orange for the "Detector chapter", magenta for the "Image Quality" chapter, and green for the "Monitors" chapter. The same style was used for the tables in the text and in the Excel template provided to facilitate data collection and analysis.

X-Ray Source

X-Ray Source

As previously mentioned, one purpose of this document is to limit the number of QC tests to the most relevant ones. Regarding the x-ray source, some tests traditionally included in mammography QC protocols have been removed, like kV_p accuracy and reproducibility, and focal spot size. The evaluation of kV_p accuracy and reproducibility was not included because modern X-ray generators used in digital mammography are highly stable and once calibrated, seldom drift. Moreover, the focal spot is no longer a limiting factor affecting spatial resolution in digital mammography [1].

The only tests on the X-ray source retained as necessary in this protocol are those related to radiation dose estimation, i.e. the X-ray tube output measured at known distance from the source, and the half value layer (HVL) or beam quality, which characterizes penetration capability of a polychromatic X-ray beam. Both those parameters, tube output and HVL, are measured with the breast compression paddle in the beam.

In more detail, the X-ray tube output, $Y(d)$, is the quotient of the air kerma at a specific distance from the X-ray tube focus, d, by the tube current-exposure time product, also called tube loading and usually expressed in units of mGy/mAs or µGy/mAs. The X-ray tube output permits the calculation of the incident air kerma, K_i, if tube loading (mAs) and compressed breast thickness are known. As clarified by the IAEA Technical Report no. 457, the incident air kerma is the kerma to air from an incident X-ray beam measured on the central beam axis at the position of the patient or phantom surface. It includes only the incident, and not the backscattered radiation. Conversely, the entrance surface air kerma, K_e, is obtained from the incident air kerma multiplied by the backscatter factor [2]. Organ dose in mammography is represented by the radiation dose absorbed by the glandular tissue of the breast, called Average Glandular Dose (AGD) or Mean Glandular Dose (MGD). AGD can be estimated by multiplying either K_i or K_e by a conversion factor which depends on HVL and breast thickness, while assuming that the breast is composed by a given ratio of fat/glandular components. The factor to derive the AGD from the incident air kerma or entrance surface air kerma depends on the Monte Carlo model used to obtain the conversion factor. In the following, the incident air kerma, K_i, is used, the term AGD will be reported (see AEC chapter), and the conversion factors used for AGD estimation will be those provided by Dance and colleagues in their publications [3,4].

X-Ray Source

Since both tube output and HVL depend on the selected X-ray beam (anode/filter combination and kVp value), in principle, it would be necessary to measure both parameters for each X-ray beam used in the clinical practice. Nevertheless, it is possible to drastically reduce the number of measurements by using a power law relationship between K_i and kV_p,

$$K_i = A\,(kV_p)^n \tag{1}$$

which becomes linear by a logarithmic transformation, and a second order polynomial relationship between HVL and kV_p

$$HVL = a\,(kV_p)^2 + b\,(kV_p) + c \tag{2}$$

as published by Robson [5].

X-Ray Source

DETECTORS FOR AIR-KERMA MEASUREMENTS

Instruments for dose measurements commonly used are ionization chambers and solid state detectors. Ionization chambers normally have weak energy dependence in the limited energy range of X-ray spectra used in mammography, while the opposite occurs with solid state detectors. For this reason, when solid state detectors are used, verifying the existence of appropriate calibration curves for the X-ray beams of interest is recommended.

For relatively old instruments, built to perform measurements with beams produced by typical anode/filter combinations used in screen-film mammography, such as Mo/Mo and Mo/Rh, calibration curves might not be available for combinations like Rh/Rh, W/Rh, W/Ag, and W/Al, more recently introduced in digital mammography. For some dose meters the displayed value can be used directly, but for others the use of calibration curves might be necessary.

According to the European Protocol on Dosimetry in Mammography, total uncertainty of the dose meters should be less than ±10% (including the uncertainty in energy response), precision less than ±5%, dynamic range at least 0.5-100 mGy [6]. However, photon counting systems require a dynamic range starting at least from 0.05 mGy.

These demands are valid also for HVL. However, other parameters should also be considered in order to reduce the uncertainty in HVL measurements, such as geometry, and possible variability of aluminum filters used to attenuate the X-ray beam [6].

X-Ray Source

Some DR systems allow disabling image acquisition, thereby speeding up tube output and HVL measurements. To do this, the presence of the equipment Field Engineer is required.

Otherwise, in the case where a system forces image acquisition, it is not necessary to store the images acquired during dose measurements into the image archive (Picture Archiving and Communication System, PACS). The automatic sending can be temporarily disabled (and images deleted). DO NOT FORGET TO TURN IT ON WHEN YOU HAVE FINISHED.

During dose measurements, TAKE CARE TO PROPERLY SHIELD THE IMAGE DETECTOR, in order to avoid image ghosting caused by the dose meter. This can be done by covering the image detector with a layer of lead or stainless steel or aluminum or other material thick enough to completely absorb the X-ray beam, and large enough to fully cover the image detector surface.

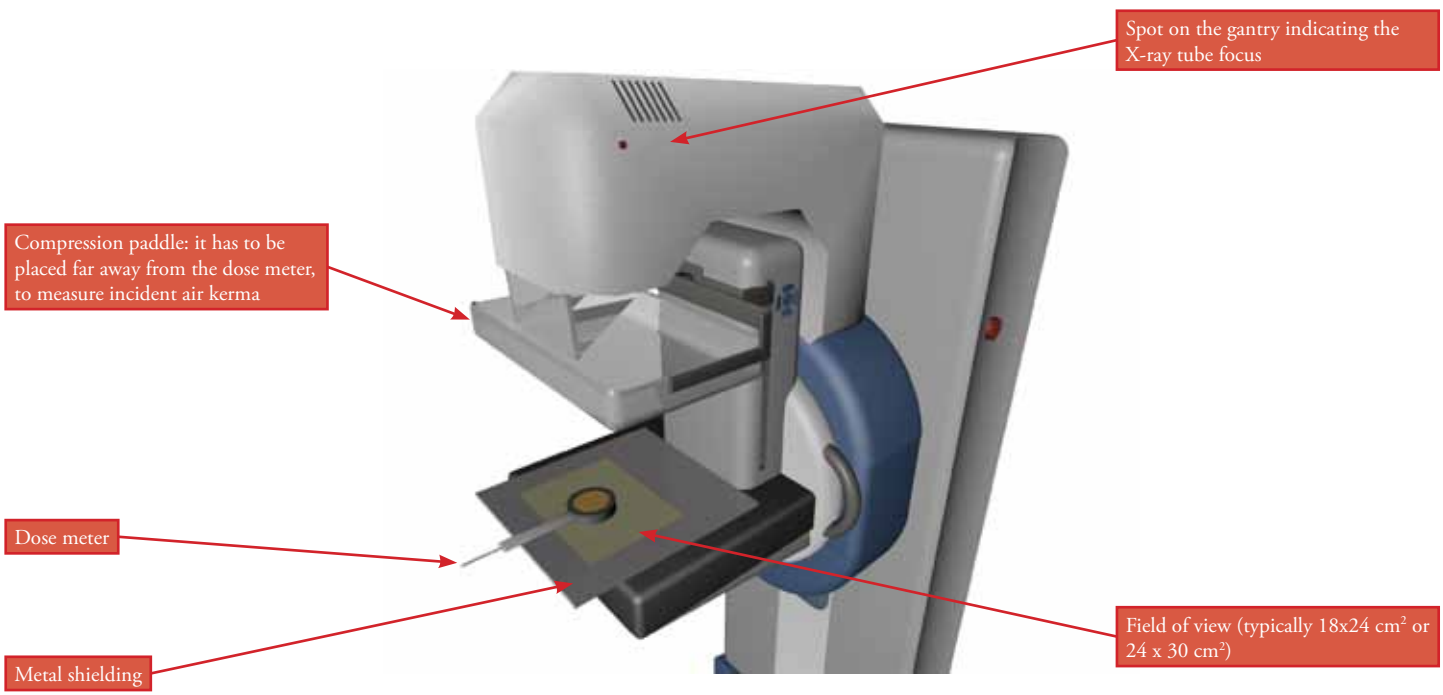
With CR systems, it may be possible to perform dose measurements without any cassette in the bucky, just disabling the sensor controlling the cassette presence, mechanically or from the mammography unit console. Otherwise, the same protection shield as used for DR systems can be used.

Tube Output

PURPOSE

MEASURE THE TUBE OUTPUT IN UNITS OF µGy/mAs, AT KNOWN DISTANCE FROM THE X-RAY TUBE FOCUS AND AFTER THE COMPRESSION PADDLE, FOR ALL THE X-RAY SPECTRA CLINICALLY USED.

EQUIPMENT



Tube Output

TEST FREQUENCY

- Acceptance/Commissioning.
- After possible replacement of X-ray source or filter block.
- Annual (tube output is expected to slowly decay in time).

Tube Output

PROCEDURE

1 Geometry

1. In principle, the incident air kerma, K_i, should be measured, as for an ionizing chamber, in contact with the compression paddle [4]. However, various dose meters have different sensitivities to forward scattered radiation from the compression paddle. To overcome this problem and, at the same time for simplicity, in this document the same experimental setup is proposed for air kerma measurements, independently of the dose meter type. The compression paddle is placed in the beam far away from the dose meter, so the scatter contribution to the dose meter is insignificant. A constant forward-scatter factor, FSF = 1.076, is used to correct the measured values, K_{im} [8].
2. Shield the image detector with a metal plate and fix it with tape. Fix the dose meter on the metal shield with tape, ensuring that the sensitive area is located at 6 cm from the chest wall edge of the breast support and laterally centered. Such distance from the chest wall might be slightly different in other QC protocols, but this has a modest impact on the measured air kerma, at least in the range 5-6 cm from the chest wall edge.
3. Measure the distance between the X-ray focus (usually indicated with a cross or a spot on the gantry) and the dose meter (the sensitive part is usually indicated with a line or other mark), SDD, and enter its value in the first page of the "Tube Output" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file.
4. Keep the collimator aperture compatible with one of the clinically used fields of view (FOV) for contact views, typically 18 x 24 cm^2, or 24 x 30 cm^2.
5. Verify that the compression paddle is in the beam, but keep it as far as possible from the dose meter surface to exclude the forward scatter from the paddle.

2 Exposure mode

1. For DR systems, enter a new patient at the acquisition workstation and start the examination. For CR systems, skip this step.
2. Select the manual exposure mode (independent setting of anode/filter combination, kV_p and mAs values) from the mammography unit console.
3. Depending on the X-ray source used by your system, acquire incident air kerma measurements for each anode/filter combination clinically used at four different kV_p values in the range of those clinically selected. In Table 1, some possible X-ray tube voltage settings are suggested for different anode/filter combination, to cover the typical range used in mammography. Different settings are allowed.

kV_p setting	**Anode/Filter**					
Voltage	Mo/Mo	Mo/Rh	Rh/Rh	W/Rh	W/Ag	W/Al
kV_{p1}	24	26	26	26	26	29
kV_{p2}	26	28	28	28	28	32
kV_{p3}	28	30	30	30	30	35
kV_{p4}	30	32	32	32	32	38

Table 1 - Suggested peak voltage settings for tube output measurements. Choose the column(s) corresponding to your mammography unit.

Tube Output

3 mAs values

1. As the linearity between tube loading (mAs) and measured dose is normally very good, two data points (mAs values) per each combination of anode/filter and kV_p are sufficient to derive fitting coefficients accurately. The following values are suggested:
   a. 50 and 100 mAs for systems using integration detectors
   b. 5 and 10 mAs for photon counting systems.
2. Enter the two mAs values in the first page of the "Tube Output" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file.
3. One shot per mAs value is sufficient, the tube output reproducibility being normally very high (if you want to be more confident about this, test the output repeatability by taking multiple air kerma measurements for one spectrum and mAs level).

4 Data input

1. For each anode/filter combination clinically used by your mammography unit, enter the setting data (SDD, anode/filter combination, and the four kV_p values) in each table of the first page of the "Tube Output" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file. The worksheet is designed to host up to four anode/filter combinations.
2. For each setting, enter the corresponding K_i value at SDD. One sample table for data input is shown below (Table 2). The gray cells should be filled with the setting data, the area with the red rectangle should contain the corresponding measured, K_{im}, incident air kerma values.

	Anode/Filter		
	Voltage	$K_{im}(mAs_1)$	$K_{im}(mAs_2)$
kV_{p1}			
kV_{p2}			
kV_{p3}			
kV_{p4}			

Table 2 - Table for data input in tube output measurements.

Incident air kerma values (mGy) measured @ each (kV_p, mAs) pair

Tube Output

5 Example of data input

	Anode/Filter	W/Ag	
	Voltage	$K_{im}(mAs_1)$	$K_{im}(mAs_2)$
kV_{p1}	26	1.826	3.696
kV_{p2}	28	2.348	4.736
kV_{p3}	30	2.849	5.747
kV_{p4}	32	3.364	6.757

Table 3 - Example of data input for tube output measurements for a DR system (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

6 Tube output calculation

1. Each incident air kerma value is multiplied by 1.076 to take account of the scatter from the compression paddle (FSF) to obtain the incident air kerma (K_i)
2. Equation (1) can be linearized by taking logarithms of both sides of the equation

$$\log_{10}(K_i) = \log_{10}(A) + n \cdot \log_{10}(kV_p) \quad (3)$$

3. Coefficients log A and n, for each of the two mAs values, can be obtained from Equation (3) by a linear fit, with the constraint of passing through zero.
4. As suggested by Robson [5], those coefficients can be used to feed Eq. (1) to determine incident air kerma (K_i) at any kV_p value.
5. Finally, the tube output at any kV_p can be obtained from the linear relationship between K_i values recalculated using the coefficients previously derived and related mAs values.

Tube Output

7 Data output

The "Tube Output" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file automatically provides the tube output in units of µGy/mAs @SDD for all anode/filter combinations and peak voltage in the range 22-40 kV_p. The square of the correlation coefficient (R^2) obtained from the linear fit is displayed to detect possible errors in entered data, the relationship between K_i and mAs being expected to be highly linear (R^2 should be very close to 1). Furthermore, another value of the tube output is recalculated at the common distance 650 mm, to allow possible comparison among different mammography systems.

SDD	Anode/Filter	
kV_p	µGy/mAs @SDD	R^2
22		
.....		
30		
.....		
40		

Table 4 - Table for data output in tube output measurements. For each anode/filter combination, the tube output is calculated in the range 22-40 kV_p.

8 Example of data output

SDD (620 mm)	Anode/Filter	W/Ag
kV_p	µGy/mAs @SDD	R^2
22	24.89	0.99997
.....	.....	.....
30	61.11	1.00000
.....	.....	.....
40	140.58	0.99999

Table 5 - Example of tube output obtained by a DR system (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

Tube Output

Knowledge of tube output is necessary to determine the air kerma incident to the breast or phantom entrance for dosimetric purpose.

Suppose we want to calculate the K_i value for a 5 cm breast, for which the technical factors selected by the automatic exposure control were W/Ag, 30 kV_p, and 63 mAs. The distance between the focal spot and the breast support is 620 mm, approximately equal to the SDD, the dose meter probe being very thin. K_i at the breast entrance can be obtained by multiplying the tube output for the selected X-ray beam by the mAs value, and correcting the resulting number to take into account the effective source-to-breast entrance distance, i.e.

$$K_i = \left.\frac{\mu Gy}{mAs}\right|_{(W/Ag,30kV_p)} \times \frac{mAs}{1000} \times f_{distance}$$

where $f_{distance}$ counts for the inverse square dependence of air kerma from distance.
From Table 5, the tube output for the W/Ag combination at 30 kV_p is 61.11 µGy/mAs. The breast thickness is 50 mm, thereby, applying the previous formula, the K_i at the breast entrance for an SSD of 620 mm becomes

$$K_i = 61.11 \times \frac{63}{1000} \times \frac{620^2}{(620-50)^2} = 4.55mGy$$

LIMITING VALUES

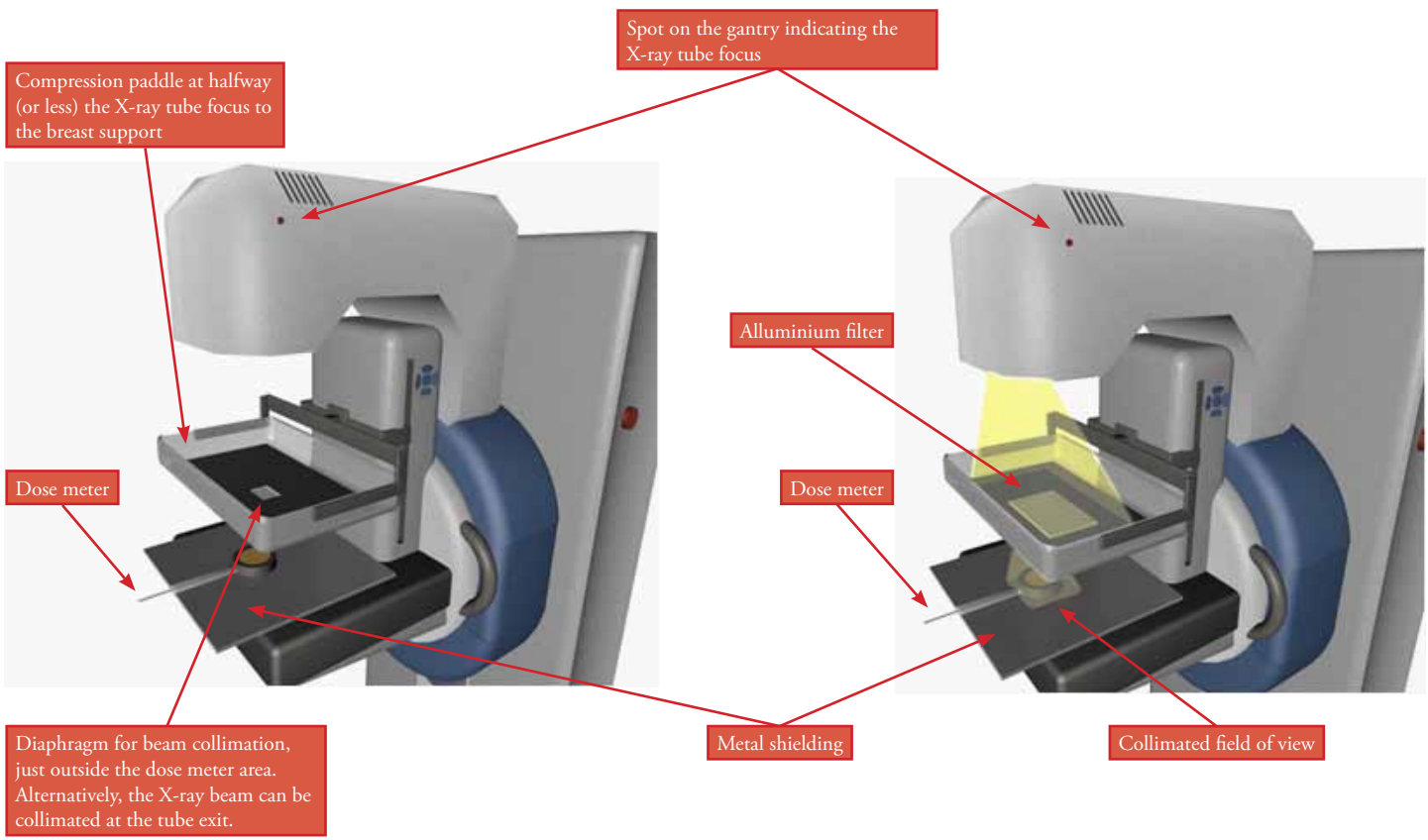
THERE IS NO LIMITING VALUE FOR THE TUBE OUTPUT. The value 30 µGy/mAs at the distance of 1 m was reported in old protocols, for the typical X-ray beam used in screen-film mammography, the anode/filter combination Mo/Mo at 28 kV_p. Since beams used by modern digital mammography systems are totally different, such value is no longer relevant.
However, a very low tube output can result in a long exposure time, with significant risk of motion artifacts. The typical exposure time for a breast thickness of 50 mm, should be well below 1 second, with the exception of scanning systems.

Half Value Layer (HVL)

PURPOSE

MEASURE THE HALF VALUE LAYER (HVL) IN UNITS OF mmAl AFTER THE COMPRESSION PADDLE, FOR ALL THE X-RAY SPECTRA CLINICALLY USED.

EQUIPMENT



Half Value Layer (HVL)

TEST FREQUENCY

- Acceptance/Commissioning.
- After possible replacement of X-ray source or filter block.
- As the HVL should not change significantly in time, unless the anode or filter as degraded, it is not necessary to repeat the whole test annually. Reproducibility can be checked for a single spectrum, frequently selected by your system.

Half Value Layer (HVL)

PROCEDURE

1 Geometry

1. The air kerma should be measured in scatter-free conditions, with and without aluminum filters in the beam in order to measure the thickness of aluminum needed to reduce the air kerma to 50% of the value measured without aluminum. In fact, it is not necessary to determine K_i. Any dose meter reading that is properly corrected for beam quality dependence would be sufficient.
   Anyhow, in the following text, the quantity incident air kerma, K_i, is used, with no correction for the forward-scatter fraction.
2. Shield the image detector with a metal plate and fix it with tape. Fix the dose meter on the metal shield with tape, ensuring that the sensitive area is located at 6 cm from the chest wall edge of the breast support and laterally centered. Such distance from the chest wall might be slightly different in other QC protocols, but this has a modest impact on the measured HVL, at least in the range 5-6 cm from the chest wall edge.
3. Reduce the collimation aperture to have only the dose meter in the exposed field. This can be done by either adjusting system collimators or using an external diaphragm placed on the compression paddle.
4. Position the compression paddle halfway between the X-ray source and the breast support , or closer to the X-ray source [6]. Use the compression paddle as support for the aluminum filters.
5. If possible, disable the "decompress after exposure" function (which automatically repositions the compression paddle after each exposure at a preset height) from the mammography unit console. Do not forget to reset the default condition when you have finished.

2 Exposure mode

1. For DR systems, enter a new patient at the acquisition workstation and start the examination. For CR systems, skip this step.
2. Select the manual exposure mode (independent setting of anode/filter combination, kV_p and mAs values) from the mammography unit console.
3. Depending on the X-ray source used by your system, acquire incident air-kerma (K_i) measurements for each anode/filter combination clinically used at four different kV_p values, according to Table 6.

	Anode/Filter					
Voltge	**Mo/Mo**	**Mo/Rh**	**Rh/Rh**	**W/Rh**	**W/Ag**	**W/Al**
kV_{p1}	26	26	28	26	28	29
kV_{p2}	28	28	30	28	30	32
kV_{p3}	30	30	32	30	32	35

Table 6 – Suggested peak voltage settings for HVL measurements. Choose the column(s) corresponding to your mammography unit.

Half Value Layer (HVL)

4. For each selected beam (anode/filter, kV_p), measure three K_i values:
   a. without any additional filter (T_0)
   b. with the aluminum filter T_1 as indicated in Table 7
   c. with the aluminum filter T_2 as indicated in Table 7
   The aluminum foils must be placed on the compression paddle to cover the dose meter area

	Anode/Filter					
Aluminium Thickness (mm)	**Mo/Mo**	**Mo/Rh**	**Rh/Rh**	**W/Rh**	**W/Ag**	**W/Al**
T_0	0.0	0.0	0.0	0.0	0.0	0.0
T_1	0.2	0.2	0.2	0.3	0.3	0.3
T_2	0.5	0.5	0.5	0.7	0.7	0.7

Table 7 – Suggested thickness (in mm) of the aluminum filtes used to experimentally determine the HVLs for the X-ray spectra listed in Table 6.
Choose the column(s) corresponding to your mammography unit.

3 mAs value

» For each anode/filter combination allowed by your mammography unit, keep the mAs value constant, to ensure the dose intensity is high enough even when the thicker aluminum filter is used. The mAs value should be in the range 80-150 mAs for any system using integration detectors, around 10 mAs for photon counting systems. Enter the selected mAs in the first page of the "HVL" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file.

» For each anode/filter combination allowed by your mammography unit, enter the setting data (anode/filter combination, and the four kV_p values) in each table of the first page of the "Tube Output" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file.

4 Data input

» For DR systems, enter a new patient at the acquisition workstation and start the examination. For CR systems, skip this step.

» For each anode/filter combination allowed by your mammography unit, enter the three kV_p values indicated in Table 6 and the attenuator thicknesses as in Table 7 in one of the tables of the first page of the "HVL" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file. The worksheet is designed to host up to four anode/filter combinations.

» Select the manual exposure mode (independent setting of anode/filter combination, kV_p and mAs values) from the mammography unit console.

» For each setting, enter in the corresponding table for data input the K_i value corresponding to each attenuator thickness (T_0, T_1, and T_2).

Half Value Layer (HVL)

Anode / Filter							
Al thickness (mm)		K_i @ kV_p		K_i @ kV_p		K_i @ kV_p	
T_0			mGy		mGy		mGy
T_1			mGy		mGy		mGy
T_2			mGy		mGy		mGy

K_i value (mGy) @ each setting (kV_p, mAs) pair and aluminum thickness

Table 8 - Table for data input in HVL measurements.

5 Example of data input

Anode / Filter					W/Ag		
Al thickness (mm)		K_i @ kV_p	28	K_i @ kV_p	30	K_i @ kV_p	32
T_0	0.0	5.007	mGy	6.073	mGy	7.173	mGy
T_1	0.3	3.317	mGy	4.089	mGy	4.901	mGy
T_2	0.7	2.023	mGy	2.540	mGy	3.094	mGy

Table 9 - Example of data input for HVL measurements for a DR system (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

6 Experimental HVL values

1. For each selected beam (anode/filter combination and kV_p pair), the experimental HVL value can be obtained from the exponential attenuation law. It can be obtained by computing slope and intercept from the log-log linear relationship between K_i and aluminum thickness:

$$\log(K_i) = \log(K_i)_{0mm} - \mu \cdot T \quad (4)$$

where $K_{i\,0.0mm}$ is the K_i measured without additional attenuation material, μ is the linear attenuation coefficient of aluminum (which is a function of mean photon energy), and T is the aluminum thickness.

2. Experimental HVL is calculated from coefficients determined by linear fitting (Eq. 4), for air kerma equal to $K_{i\,0.0mm}/2$. Resulting HVLs for measurements in Table 9 are reported in Table 10.

Half Value Layer (HVL)

Anode/Filter		W/Ag
kV_p	Ln (K_{i0}/2)	Experimental HVL (mmAl)
28	0.918	**0.530**
30	1.111	**0.551**
32	1.277	**0.572**

Table 10 - Example of experimental HVLs determined from measurements in Table 9 (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

7 Data output

1. For each anode/filter combination, the three experimental HVL values obtained at three different kV_p values are used in Equation (2) to determine the polynomial fit coefficients, a, b, and c. The coefficients from the previous example are shown in Table 11.

W/Ag	
Fit parameters	
a	-9.75×10^{-5}
b	+ 0.016
c	+ 0.148

Table 11 - Example of polynomial fitting coefficients obtained from the experimental HVLs in Table 10 (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

2. Finally, the fitting coefficients are replaced in Equation (2) to recalculate HVLs at any kV_p value. The tables with data output in the worksheet "HVL" of the "Template_EFOMP_MammoWG_DR/CR" Excel file, show HVLs in the range 22-40 kV_p. Table 12 reproduces the table for calculated HVLs.
3. As the HVL measurements are affected by a large overall uncertainty, the fitting accuracy should always be verified. By definition, the HVL values should increase with the tube voltage, for each given anode/filter combination. If you notice that the HVL decreases at a higher kV_p, add another experimental HVL value, obtained at higher peak voltage values than those indicated in Table 6, before applying the polynomial fitting.

Half Value Layer (HVL)

Calculated HVL (mm Al)				
kV_p	A/F_1	A/F_2	A/F_3	A/F_4
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				

Table 12 - Table for data output of HVL measurements, providing calculated HVLs in the range 22-40 kV_p.

Half Value Layer (HVL)

8 Example of data output

Calculated HVL (mm Al)				
kV_p	W/Ag	A/F_2	A/F_3	A/F_4
22	0.461			
23	0.473			
24	0.485			
25	0.496			
26	0.508			
27	0.519			
28	0.530			
29	0.541			
30	0.551			
31	0.562			
32	0.572			
33	0.582			
34	0.592			
35	0.601			
36	0.611			
37	0.620			
38	0.629			
39	0.638			
40	0.647			

Table 13 - Example of data output for calculated HVLs from previous measurements (IMS Giotto Image 3DL, selenium-based, and W/Ag combination).

Half Value Layer (HVL)

LIMITING VALUES

THERE IS NO LIMITING VALUE FOR THE HVL. The value 0.33 mmAl was reported in old protocols for the typical X-ray beam used in screen-film mammography, the anode/filter combination Mo/Mo at 28 kV_p. Since beams used by modern digital mammography systems are totally different, this value is no longer relevant.
However, typical HVL values are provided in the European guidelines for quality assurance in breast cancer screening and diagnosis, 4th edition [7], and reported here in Table 14. If your HVL results are far from those values, this may indicate a problem with either measurements performed or with the mammography system itself, and requires further investigation.

kV_p	Mo + 30µm Mo	Mo + 25µm Rh	Rh + 25µm Rh	W + 50µm Rh	W + 45µm Al
25	0.33 ± 0.02	0.40 ± 0.02	0.38 ± 0.02	0.52 ± 0.03	0.31 ± 0.03
28	0.36 ± 0.02	0.42 ± 0.02	0.44 ± 0.02	0.54 ± 0.03	0.37 ± 0.03
31	0.39 ± 0.02	0.44 ± 0.02	0.48 ± 0.02	0.56 ± 0.03	0.42 ± 0.03
34		0.47 ± 0.02		0.59 ± 0.03	0.47 ± 0.03
37		0.50 ± 0.02			0.51 ± 0.03

Table 14 – Typical HVL values for different tube voltage and anode/filter combinations. Attenuation from a PMMA compression plate is included.

REFERENCES

1. *M Yaffe et al. A harmonized quality control program for digital mammography. LNCS 4046:585-592, 2006.*
2. *International Atomic Energy Agency, "Dosimetry in diagnostic radiology: an international code of practice", Technical Reports Series No. 457, IAEA, Vienna 2007.*
3. *D R Dance, C L Skinner, K C Young, J R Beckett, C J Kotre. Additional factors for the estimation of mean glandular dose using the UK mammography dosimetry protocol. Phys Med Biol 45:3225-3240 (2000).*
4. *D R Dance, K C Young, RE van Engen. Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA and breast dosimetry protocols. Phys Med Biol 54:4361-4372 (2009).*
5. *K J Robson. A parametric method for determining mammographic X-ray tube output and half value layer. BJR 74:335-340, 2001.*
6. *J Zoetelief, M Fitzgerald, W Leitz, M Säbel, "European Protocol on Dosimetry in Mammography", EUR 16263, 1996.*
7. *N Perry et al. "European guidelines for quality assurance in breast cancer screening and diagnosis", 4th edition, 2006.*
8. *P Toroi et al. Aspects of forward scattering from compression paddle in the dosimetry of mammofraphy. Rad Prot Dosim 154:439-445 (2013).*

Automatic Exposure Control (AEC)

Automatic Exposure Control (AEC)

The automatic exposure control (AEC) is the device designed to select all the parameters which control the detector dose, i.e. anode/filter combination, kV_p, and mAs, according to the effective absorption by each individual breast. The AEC aim is to keep the image quality sufficiently high for diagnostic purposes and constant as well, while delivering the appropriate dose level.

In screen-film mammography, due to the limited dynamic range of film, the AEC was designed to keep the radiation dose at the film constant; the AEC correct operation was indirectly verified by varying the object thickness and checking the constancy of the mean optical density produced on the film.

In digital mammography, detector dynamic range is no more a limiting factor, and, for given detector characteristics like response function and spatial resolution, image quality is degraded by noise presence. Therefore, in order to preserve image quality, the automatic exposure control should take noise into account. This means that, ideally, to preserve image quality, the AEC should adjust the detector dose to keep the image noise unchanged. Moreover, in digital mammography, image acquisition and display are separated processes, and image processing plays an important role. Contrary to screen-film mammography, both unprocessed and processed images can be available; usually, image quality vs. dose optimization is performed using unprocessed images of given test objects, regardless of the impact of the processing algorithms on the final image quality.

Depending on the system type, dose control by the AEC is obtained either by means of one or more dose detectors placed at the exit of the image detector, or, alternatively, by the image detector itself or a portion of it, used as "exposure meter". All CR systems, which use an analog mammography unit to expose the imaging plates, belong to the first category, whilst DR systems normally belong to the second one.

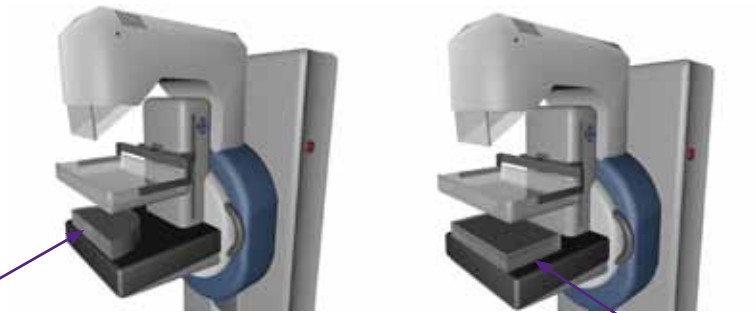
In general, any AEC is expected to "compensate" the dose level at the image detector (adjusting the technique factors, i.e. anode/filter combination, kV_p, mAs) as a function of the breast absorption, to keep image quality above a given minimum level. Breast "property" which mainly affects X-ray absorption is its compressed thickness. In mammography, breast is always compressed to reduce its thickness, increasing image quality while reducing scattered radiation and

Automatic Exposure Control (AEC)

dose. Most of the AECs use the compressed breast thickness measured by the height of the compression paddle to select the X-ray spectrum (anode/filter and kV_p value). The accuracy of the displayed thickness value versus the measured one should be verified.

DISPLAYED BREAST THICKNESS

When quality controls are performed, and rigid objects like polymethyl methacrylate (PMMA) slabs are used, the displayed thickness value can significantly deviate from reality because of the compression paddle tilt occurring when a robust force is applied to an incompressible object. This might induce the AEC to select anode/filter combinations and/or kV_p values different from those expected for an object of this thickness. In order to obtain thickness displayed values consistent with the measured thickness, the use of PMMA slabs with a surface smaller than the compression paddle surface is recommended, as well as the application of a minimum compression force (in fact, some DR systems require a minimum compression force to enable the automatic exposure control; ususally 3-5 daN is sufficient). For example, semicircular PMMA slabs as provided with some image quality phantoms are more "compatible" in this respect than large rectangular slabs covering the whole breast support.



Semicircular PMMA slabs: their shape and size limit the tilt of the compression paddle, making the displayed thickness value more consistent with the measured one

Rectangular PMMA slabs: their shape and size might induce a significant tilt of the compression paddle, making the displayed thickness value different from the measured one

Automatic Exposure Control (AEC)

In the case where one or more additional dose detector, other than the image detector, are used as the exposure meter, the only parameter which can be controlled is the dose to the exposure meter, which is proportional to the dose to the image detector. Similarly to what was done with screen-film mammography, the AEC compensation capability of such systems could be evaluated by measuring the mean output signal (mean pixel value, MPV) variation versus phantom thickness for systems with a linear response function or, for those with a non-linear response function and which normalize the output signal (CR systems), the exposure index variation versus phantom thickness.

For DR systems using a portion of the image detector as exposure meter, different choices could, in principle, be made by manufacturers about the "control variable", and parameters like the signal-to-noise ratio (SNR) or other quantities including noise could be compensated by the AEC.

A philosophical question which should be answered before designing an AEC or evaluating its performance is "What is the radiation dose necessary to produce the appropriate image quality?". Normally, image quality "appropriateness" is assessed by physicists using test objects (normally called phantoms), including "details"; X-ray images of such details can be either "rated" according to given criteria, or analyzed to measure objective parameters. "Optimization" is intended as the process which defines the radiation dose level necessary to obtain a minimum image quality, with image quality obtained according to the method associated with each given phantom. However, the "optimal AEC setting" might depend on the specific phantom used and the criteria/method applied to evaluate image quality.

In this document, the decision was taken that no particular phantom or method for image quality evaluation has been recommended. This is explained in more detail in the "Image Quality" chapter. Consequently, several different phantoms could be used according, for instance, to national legislation.

Automatic Exposure Control (AEC)

A general, simple method is proposed to test the AEC performance of any digital system, according to the following principles:

1. The AEC should deliver the radiation dose appropriate to achieve a given image quality for a defined test object. The image quality level and the parameter used to represent it are arbitrarily defined and depend on the phantom and the measurement method, as explained in more detail in the Image Quality chapter.
2. The AEC should adjust the radiation dose according to the object absorption, in order to preserve image quality as much as possible.
3. The AEC should be repeatable, i.e. able to select the same technique factors for multiple exposures of the same test object and produce the same image quality.

Independently of the phantom and method used for image quality evaluation, the AEC performance is tested by means of a simple test object made by poly-methyl methacrylate (PMMA) slabs (45 mm total thickness) to simulate breasts with different absorption properties, and a relatively large aluminum (Al) detail (area: 15x15 mm^2; thickness 0.2 mm) to produce a signal difference on the images.

The physical parameter proposed to evaluate both AEC repeatability and compensation capability is the signal-difference-to-noise ratio (SDNR), while the dose parameter is the average or mean glandular dose (AGD in the following). AGD limiting values had been established for screen-film mammography [1]; digital mammography is expected to deliver radiation dose lower or equal than screen-film mammography [2,3].
Before describing the testing of the AEC performance in practice, a more precise definition of the two parameters SDNR and AGD is necessary

Automatic Exposure Control (AEC)

Signal-difference-to-noise ratio (SDNR)

SDNR (Signal Difference-to-Noise Ratio) is obtained from the difference between the mean pixel values of background (MPV_{bkg}) and aluminum detail (MPVAl), divided by the standard deviation of the background (SD_{bkg}) where the detail is embedded, according to the following formula:

$$\mathrm{SDNR} = \frac{\mathrm{MPV}_{bkg} - \mathrm{MPV}_{Al}}{\mathrm{SD}_{bkg}} \qquad \text{Eq. (5)}$$

Historically, the SDNR is related to human capability of detecting "large" objects with given contrast in a noisy background, and it was first formulated by Albert Rose [4]. Rose observed that humans cannot perceive an object unless its SDNR is above a given threshold, and Rose's model is the basis of any contrast-detail phantom and related methods to evaluate image quality in medical imaging [4-8].

The SDNR is easy to measure and can be used to test the AEC performance, but, unfortunately, cannot be used as an absolute image quality index. In fact, there are significant differences in the SDNR produced by different mammography systems [9] which can be justified by difference of x-ray source, detector technology, read-out electronics, and, in some cases, by the presence of some type of pre-processing. For all these reasons, it is impossible to assess quality of any given digital mammography system on the basis of only the SDNR, and other image quality parameters are necessary to perform inter-system comparison.

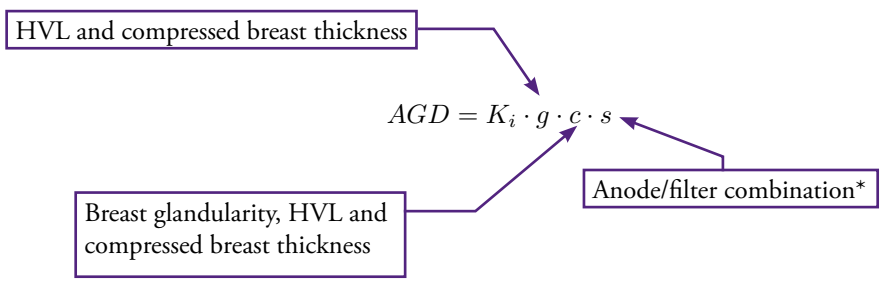
Automatic Exposure Control (AEC)

Average glandular dose (AGD)

Average or mean glandular dose (AGD or MGD) if defined as the dose absorbed by the glandular tissue within the breast, which is considered to be the most sensitive to radiation induced cancer [15]. The central part of the breast is modeled as a mixture of glandular tissue (called glandularity) and fat. This is surrounded by a layer of fat representing the skin.

For a long time, the "typical" breast was different in various countries in Europe although a common protocol was introduced in 1996 [16] proposing a standard breast with 50% glandularity, represented by a 45 mm thick PMMA phantom. That protocol is still used in some countries, however it does not reflect the large variation in thickness and composition of real breasts.

The AGD is computed from the incident air kerma (K_i) (at the object entrance), by applying conversion factors which take into account the object absorption dependence on its thickness and composition, and characteristics of the X-ray beam used. Determination of K_i is described in chapter X-Ray Source. There are a few different models for the AGD calculation, typically based on Monte Carlo techniques [11-13]. In this document, we will refer to the model by Dance and colleagues [12,14], which converts K_i at the breast or phantom entrance to AGD using the three factors g, c and s in the following equation:



Eq. (6)

**The s factor depends also on breast thickness for the W/Al combination*

Automatic Exposure Control (AEC)

AGD may be calculated for standard breasts simulated with PMMA or for real patients and different values of *g* and *c* factors are provided for those cases. The factor *g* gives the AGD for a breast of glandularity 50%. It is HVL and compressed breast thickness dependent. The factor *c* corrects for any difference in breast composition from 50% glandularity. It is given for the age groups 40-49 and 50-64 years for measurements with patients and for the age group 50-64 for PMMA measurements. This factor is determined for the UK population and may be different for other populations. The factor *c* is also HVL and compressed breast thickness dependent. Since the first two factors were initially calculated for a limited number of target/filter combinations, the factor *s* was introduced to take into account the different types of spectra used in the contemporary digital systems. It is target/filter combination dependent. The values of *g*, *c* and *s* factors are presented in Annex. AGD values evaluated after following this document should not differ significantly from AGD values according to the European Guidelines [1] being updated [14]. AGD is usually given in units of milligrays (mGy).

Automatic Exposure Control (AEC)

EXPOSURE MODES

1. **MANUAL:** Anode/filter combination, kV_p and mAs are manually selected by the operator from a discrete number of values. Manual exposure mode does not require any compression force, and can be used for some test procedures.
2. **SEMI-AUTOMATIC:** Anode/filter combination and kV_p are manually selected by the operator, while the mAs value is automatically determined when the signal recorded by the exposure meter (AEC) achieves a given threshold. Semi-automatic mode is still available with analogue equipment used with CRs and also with some DRs. Sometimes, this exposure mode requires a minimum compression force (typically 3-5 daN) to be allowed.
3. **(FULL) AUTOMATIC:** All the technique factors (anode/filter combination, kV_p and mAs) are automatically selected after a short pre-exposure used to determine the object absorption peak. Some systems use preset anode/filter combination and kV_p values on the basis of the object thickness measured by the compression arm, and adjust the mAs value according to the absorption peak. Other systems use the effective object absorption determined by the pre-exposure to adjust all the exposure parameters, including anode/filter and kV_p. Sometimes, this exposure mode requires a minimum compression force (typically 3-5 daN) to be enabled. The automatic mode is the AEC mode normally used with all DRs and most of the analog units coupled to CRs.

DO NOT FORGET ...

That some DR systems permit one to choose between two or three different AEC modes, usually associated to different dose levels to the image detector. With some systems, the choice is made during the AEC calibration (for example, Hologic, IMS, or Siemens) by the system field-engineer; with other systems, the user is allowed to choose among multiple exposure modes (for example, with GE systems, three automatic exposure modes, STD, CNT, and DOSE are available).

AEC Reproducibility

PURPOSE

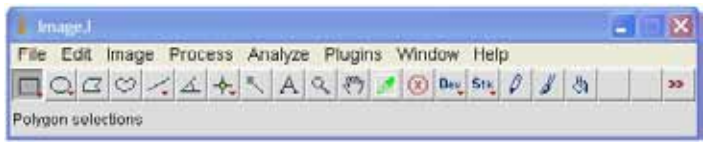
1. **TO EVALUATE THE REPEATABILITY IN SELECTING THE TECHNIQUE FACTORS (i.e. parameters controlling radiation dose) BY THE AUTOMATIC EXPOSURE CONTROL.**
2. **TO CHECK SDNR REPRODUCIBILITY.**

EQUIPMENT

1 Poly-methyl methacrylate (PMMA) block: 45 mm with an aluminum square, 0.2 mm thick (side 10-15 mm).

Some notes on the dependence of the AEC behavior from the PMMA shape and size are provided in the FAQ chapter.

2 Software for image analysis (in the following, all instructions are provided for the freeware package ImageJ, but any other equivalent software can be used.)



AEC Reproducibility

TEST FREQUENCY

» Acceptance/Commissioning.
» Annual.
» After possible replacement of components having an impact on the AEC (X-ray source, or filter block or image detector).
» After possible changes of the AEC setting.

PROCEDURE

1 Geometry

1. Place the PMMA block (45 mm) on the breast support aligned with the chest wall edge and laterally centered.
2. Place the aluminum detail (side 10-15 mm, thickness 0.2 mm) on the top of the PMMA block, at 5-6 cm from the chest wall edge, and laterally centered.
3. Mount the compression paddle compatible with the test object size, and check that the collimator aperture is appropriate.
4. Lower the compression paddle and apply a minimum compression force by 3-5 daN, if this is required to enable the automatic exposure mode.

2 Exposure mode

1. Ensure that the acquisition workstation is set to save “FOR PROCESSING” images and any possible pre-processing algorithm is disabled. For example, with GE DR systems, a pre-processing named “Fine View” is applied, acting on image noise; it can be disabled for quality controls from the “Medical preferences/Image Quality” menu on the acquisition workstation. Do not forget to enable it after the QC tests have been completed.
2. If possible, disable the “automatic decompression after exposure” from the mammography unit console. Otherwise, take care to keep the compression force unchanged for the five exposures and verify the thickness value displayed before each acquisition.
3. Record a new patient on the acquisition workstation.
4. Select the automatic exposure mode under test from the mammography unit console.
5. With CR systems, check that the AEC sensor is placed under the aluminum detail and the test cassette is in the bucky. For DR systems with multiple external exposure sensors, verify that only the sensor(s) covered by the phantom is/are enabled.
6. Expose the phantom in automatic exposure mode and acquire five images.
7. With CR systems, always use the same imaging plate. Scan the imaging plate using the test menu indicated in Table 15.

AEC Reproducibility

CR system	Test menu	Output
Agfa	System diagnostic/ flat field	Pixel gray level
Carestream	Test	Pixel gray level
Fuji	1. Ave2 (preferred) 2.Linear (EDR fix, S= 120, L = 2)	1. Exposure Index (S) 2. Pixel gray level
Konica-Minolta	Test	Exposure Index (S)

Table 15: Test menu to be used for quality controls and associated output for each CR manufacturer.

3 Exposure data collection

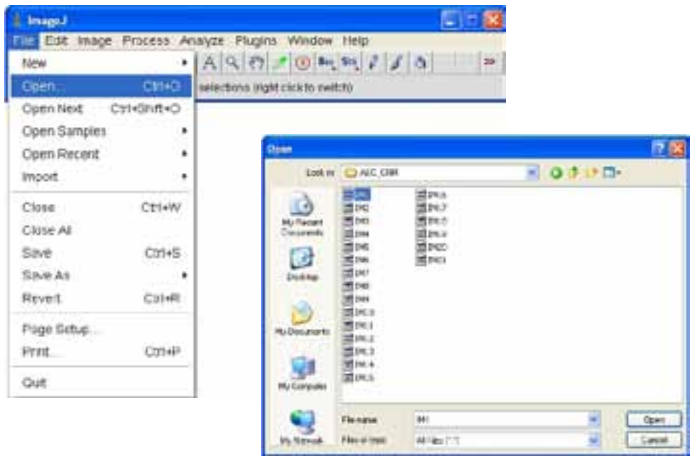
» **DR systems:** all exposure data are saved in the DICOM header of images and can be accessed at any time.
» **CR systems:** EXPOSURE DATA MUST BE ANNOTATED DURING IMAGE ACQUISITION, OTHERWISE THEY ARE LOST, unless some type of applications have been developed to communicate technique factors from the X-ray unit to the CR reader.

In the following, it is assumed that quality control measurements on the images are performed on an independent computer. This can be typically done by storing unprocessed images into the PACS and retrieving them from another PC. However, depending on each specific mammography systems, some types of measurements can be performed directly from the acquisition workstation.

AEC Reproducibility

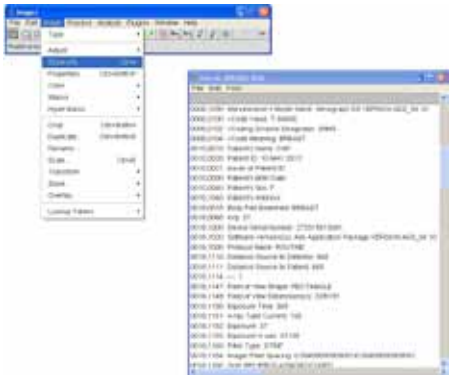
4 Exposure data extraction

1. Open the first image



2. As the image is "FOR PROCESSING", it will be displayed with the full dynamic range

3. Display the information included in the DICOM header



4. For DR systems, extract the following DICOM tags and insert them in the input Table__ of the worksheet "AGD" in the "Template_EFOMP_MammoWG" file
   - (0018,1191) "Anode Target Material"
   - (0018,7050) "Filter Material LT"
   - (0018,0060) "kV_p"
   - (0018,1152) "Exposure"
   - (0018,11A0) "Body Part Thickness"

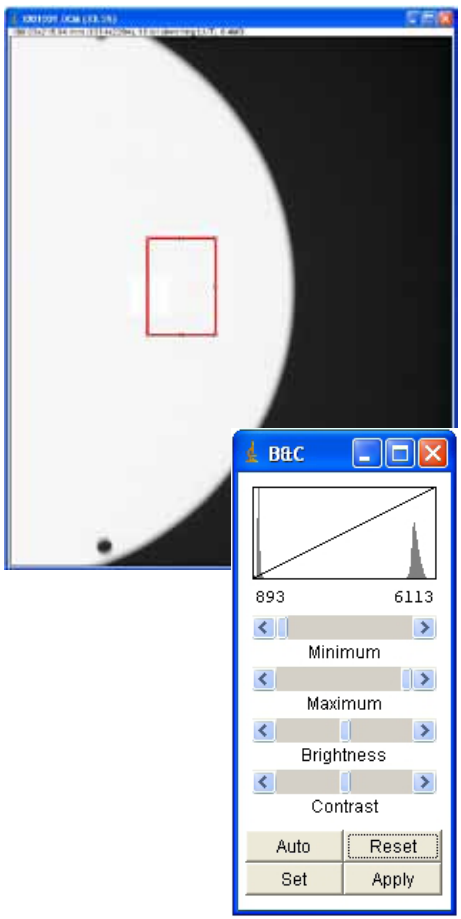
5. For CR systems, exposure data must be annotated during image acquisition.

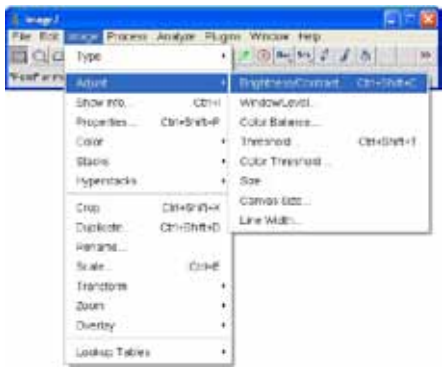
AEC Reproducibility

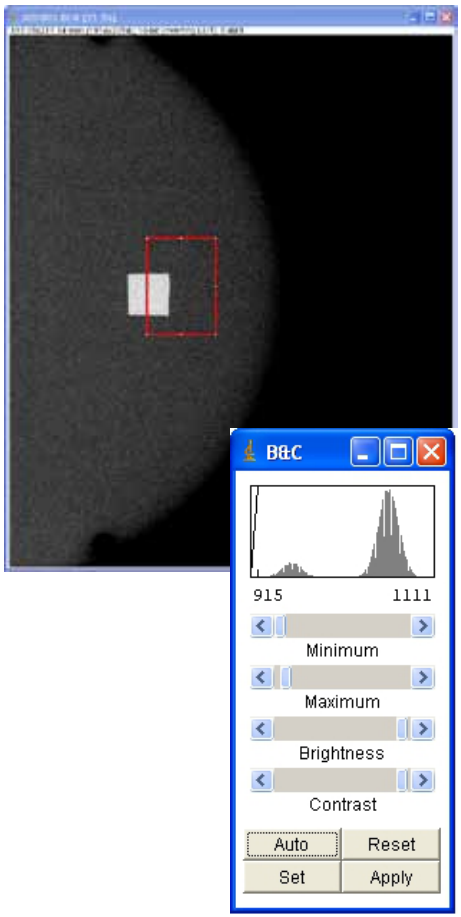
PROCEDURE

5 SDNR measurement

1. Draw a region of interest (ROI) inside the phantom and adjust window width and level (WW/WL) from the "Image/Adjust/ Brightness/Contrast…" menu by pressing the "Auto" button to display the aluminum detail.

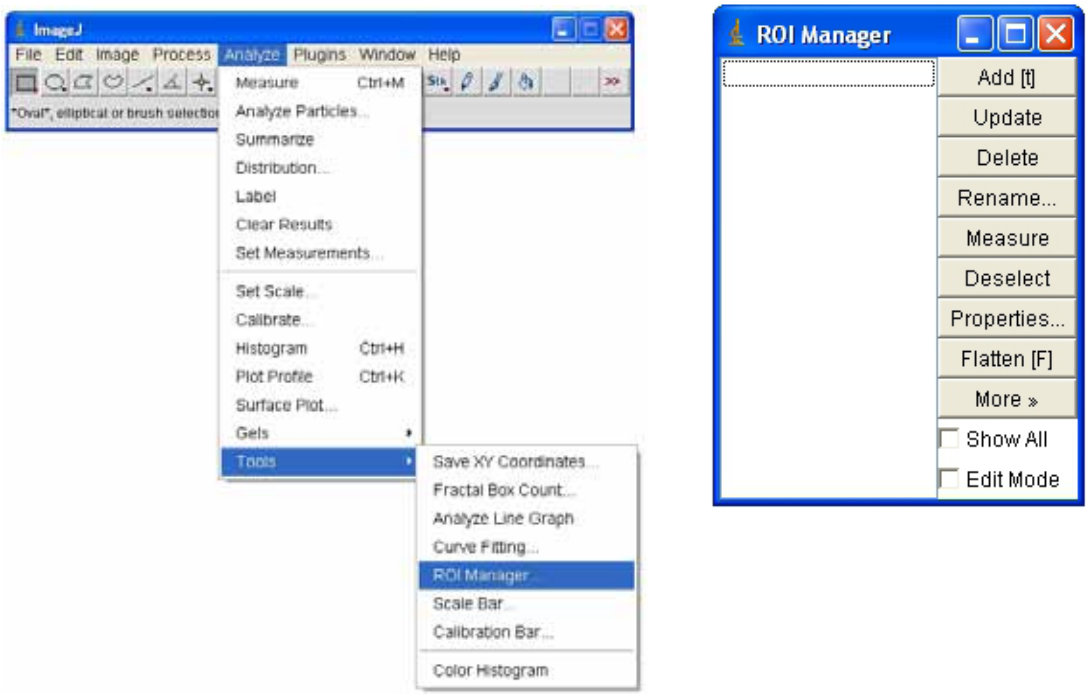






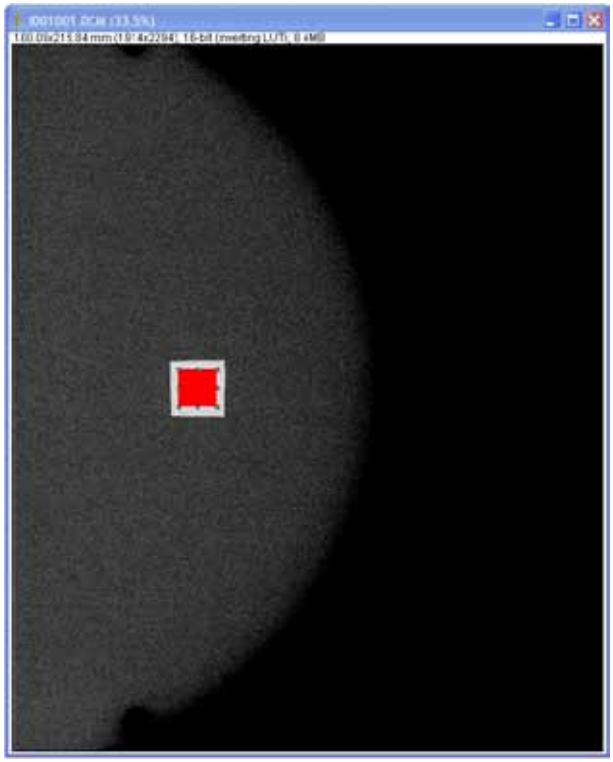
AEC Reproducibility

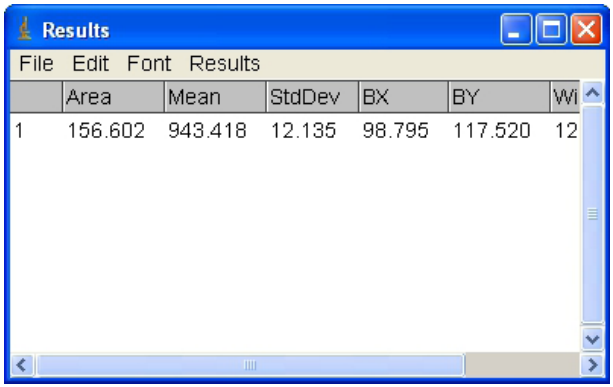
2. Click on "Analyze/Tools/ROI Manager…" to open the ROI Manager window. This tool allows the size and position of one or more ROIs to be recorded, so that the same ROIs can be used an different images, hence improving the consistency of measurements. The ROI set can also be saved and reused during annual tests. For a detailed description of all ROI Manger commands, click the following link http://imagejdocu.tudor.lu/doku.php?id=gui:analyze:tools#roi_manager. At any time one uses ROIs saved by ROI Manager to repeat measurements, the visual check of correctness of the ROI positioning versus the real position of a given object (the aluminum foil in this case) is required. In case the object is shifted compared to the ROI position, the ROI can be dragged in the new position and its coordinates can be saved by clicking the "Update" button.



AEC Reproducibility

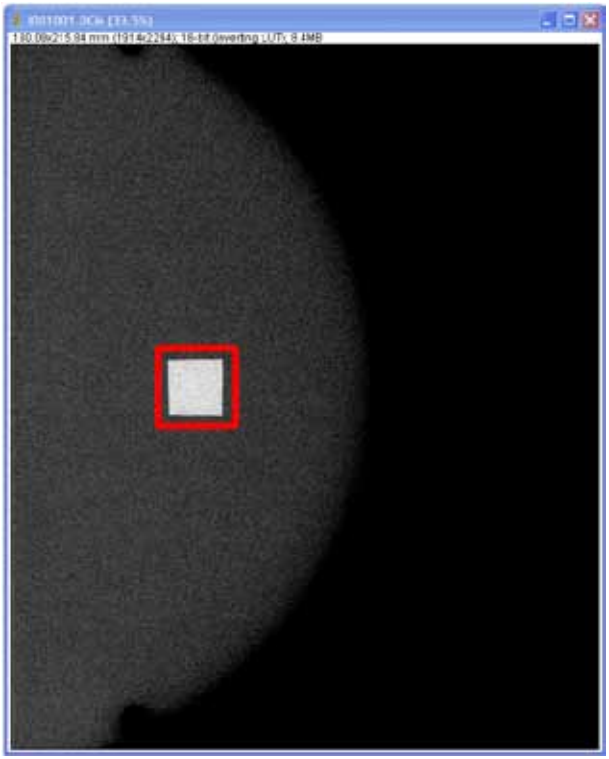
3. Select the rectangular ROI tool and keep the Shift key pressed to draw a ROI smaller than the aluminum detail and put it in the center of the detail.
4. Press the button “Add” of ROI Manager menu to record the ROI properties.
5. Click on the ROI and press the button “Rename” of ROI Manager menu to give it a proper name (optional)
6. To obtain the mean pixel value in the ROI, press either the button “Measure” of ROI Manager menu or use the command CTRL+M (equivalent to “Analyze/Measure” of the ImageJ menu. The Result window will appear with the measurements. If the mean value and/or standard deviation do not appear in the Result window, flag those data in the setting window opened by the “Analyze/Set Measurement…” command.





AEC Reproducibility

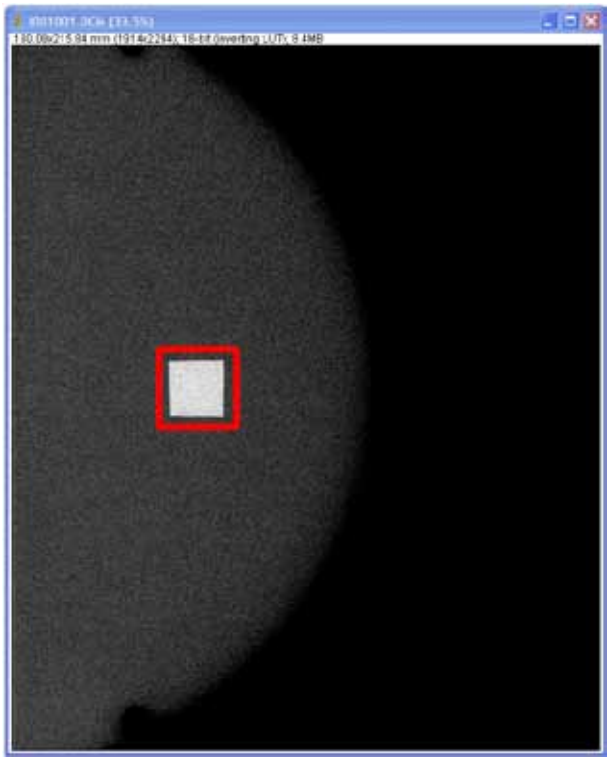
7. Draw a square ROI bigger than the aluminum detail and circumscribe the detail with it.

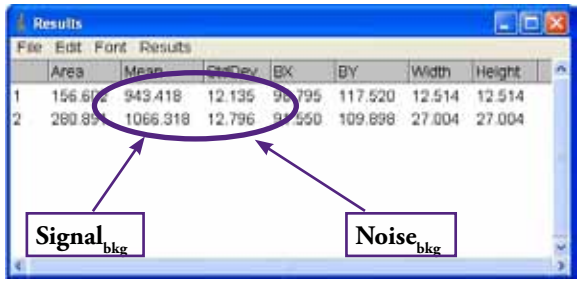


8. Define a "square band" around the aluminum detail using the "Edit Selection Make Band…" menu.

AEC Reproducibility

9. Add the new ROI to ROI Manager and eventually rename it properly.
10. Measure mean pixel value and standard deviation of pixels belonging to the band using one of the two options used for the ROI in the detail.





11. **REPEAT THE SAME MEASUREMENTS FOR ALL THE IMAGES IN THE AEC REPRODUCIBILITY DATASET. IT IS SUFFICIENT TO OPEN THE IMAGES ONE-BY-ONE AND SELECT THE TWO ROIs IN ROI Manager TO PERFORM ALL MEASUREMENTS WITH THE SAME ROI SIZE AND POSITION.**
12. Insert the values obtained for the mean pixel value of Al and PMMA and standard deviation of PMMA in Table 16.

AEC Reproducibility

6 Data input

Phantom image #	A/F	kV_p	mAs	MPV_{Al}	MPV_{PMMA}	SD_{PMMA}
1						
2						
3						
4						
5						

Table 16: Data input for the AEC reproducibility test. Exposure parameters, MPV of aluminum detail and PMMA, and SD of PMMA are collected.

7 Example of data input

Phantom image #	A/F	kV_p	mAs	MPV_{Al}	MPV_{PMMA}	SD_{PMMA}
1	Mo/Rh	28	66.8	578.34	686.60	5.73
2	Mo/Rh	28	66.7	577.36	685.34	5.66
3	Mo/Rh	28	69.3	599.96	712.06	5.68
4	Mo/Rh	28	69.3	600.39	712.52	5.79
5	Mo/Rh	28	69.3	600.61	712.79	5.81

Table 17: Example of data input for the AEC reproducibility test (GE Senographe DS, AOP/STD).

AEC Reproducibility

8 Data output

1. The incident air kerma values, K_i, are automatically calculated from the tube output measurements and shown in Table 18.
2. Short-term reproducibility of the automatic exposure control can be derived by the coefficient of variation (COV) of K_i and SDNR. The COV can be defined by the equation

$$COV = \frac{Standard\ Deviation}{Mean} \times 100 \qquad \text{Eq. (7)}$$

Table 18: Data output for the AEC reproducibility test. K_i is used to check radiation dose repeatability, while SDNR is for image quality repeatability.

Phantom image #	K_i (mGy)	SDNR
1		
2		
3		
4		
5		
Mean		
Standard deviation		
COV		

AEC Reproducibility

9

Example of data output

Phantom image #	K_i (mGy)	SDNR
1	5.904	18.90
2	5.891	19.08
3	6.109	19.72
4	6.109	19.37
5	6.109	19.32
Mean	6.024	19.278
Standard deviation	0.116	0.311
COV	1.92%	1.61%

Table 19: Example of data output for the AEC reproducibility test (GE Senographe DS, AOP/STD).

LIMITING VALUES

Index	COV
K_i	≤ 5 %
SDNR	≤ 5 %

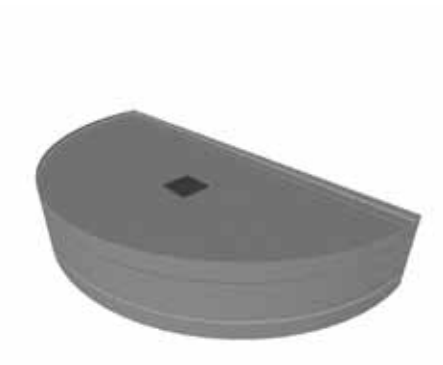
SDNR compensation and AGD

PURPOSES

1. **TO DETERMINE SDNR VS. OBJECT ABSORPTION**
2. **TO EVALUATE AGD VS. OBJECT ABSORPTION**

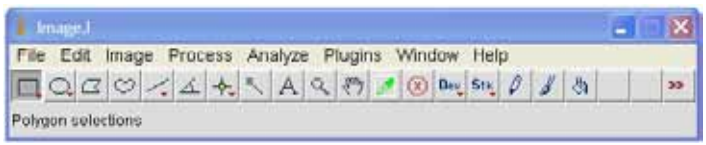
EQUIPMENT

1 Poly-methyl methacrylate (PMMA) block: 45 mm with an aluminum square, 0.2 mm thick (side 10-15 mm).



Some notes on the dependence of the AEC behavior from the PMMA shape and size are provided in the FAQ chapter.

2 Software for image analysis (in the following, all instructions are provided for the freeware package ImageJ, but any other equivalent software can be used.)



SDNR compensation and AGD

TEST FREQUENCY

» Acceptance/Commissioning.
» Annual.
» After possible replacement of components having an impact on the AEC (X-ray source, or filter block or image detector).
» After possible changes of the AEC setting.

PROCEDURE

1 Geometry

1. Place the first PMMA phantom (20 mm) on the breast support aligned with the chest wall edge and laterally centered.
2. Place the aluminum detail (side 10-15 mm, thickness 0.2 mm) on the top of the PMMA block, at 5-6 cm from the chest wall edge, and laterally centered.
3. Mount the compression paddle compatible with the test object size, and check that the collimator aperture is appropriate.
4. Lower the compression paddle and apply a minimum compression force by 3-5 daN, if this is required to enable the automatic exposure mode.

2 Exposure mode

1. Ensure that the acquisition workstation is set to save "FOR PROCESSING" images and any possible pre-processing algorithm is disabled. For example, with GE DR systems, a pre-processing named "Fine View" is applied, acting on image noise; it can be disabled for quality control from the "Medical preferences/Image Quality" menu on the acquisition workstation. Do not forget to enable it after the term of QC tests.
2. If possible, disable the "automatic decompression after exposure" from the mammography unit console. Otherwise, take care to keep the compression force unchanged for the five exposures and verify the thickness value displayed before each acquisition.
3. Record a new patient on the acquisition workstation.
4. Select the automatic exposure mode under test from the mammography unit console.
5. With CR systems, check that the AEC sensor is placed under the aluminum detail and the test cassette is in the bucky. For DR systems with multiple external exposure sensors, verify that only the sensor(s) covered by the phantom is/are enabled.
6. Expose the phantom in automatic exposure mode and acquire one image.
7. Without moving the aluminum detail, increase progressively the phantom thickness by adding PMMA slabs to obtain: 30, 40, 45, 50, 60 mm. Acquire one image per phantom thickness in automatic exposure mode.
8. With CR systems, always use the same imaging plate. Scan the imaging plate using the test menu indicated in Table 15 and reported here.

SDNR compensation and AGD

3 Exposure data collection

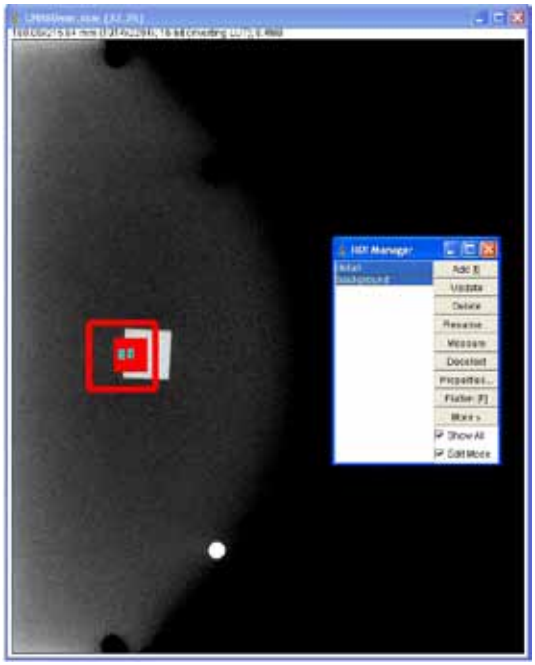
» **DR systems:** all exposure data are saved in the DICOM header of images and can be accessed at any time.
» **CR systems:** EXPOSURE DATA MUST BE ANNOTATED DURING IMAGE ACQUISITION, OTHERWISE THEY ARE LOST, unless some type of applications have been developed to communicate technique factors from the X-ray unit to the CR reader.

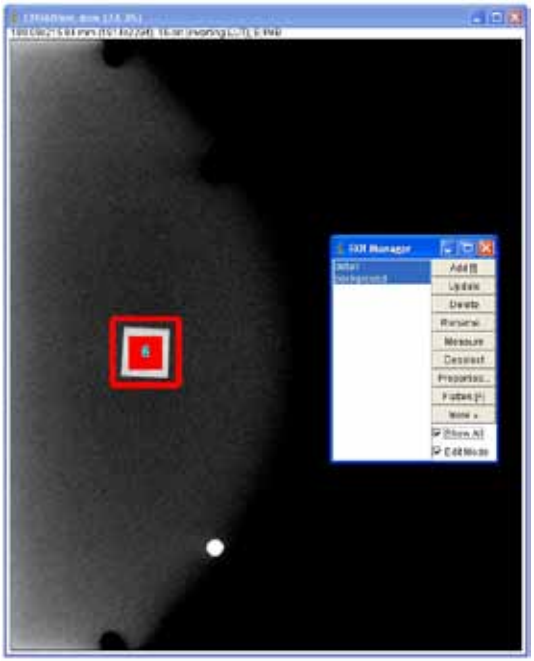
In the following, it is assumed that the quality control measurements on the images are performed on an independent computer. This can be typically done by storing unprocessed images into the PACS and retrieving them from another PC. However, depending on each specific mammography systems, some types of measurements can be performed directly from the acquisition workstation.

SDNR compensation and AGD

4 SDNR measurements

1. For the SDNR measurements versus PMMA thickness, follow the same procedure as done for the AEC reproducibility.
2. If ROI Manager has been used to record the two ROIs inside and outside the aluminum detail, the same ROIs can be used for this test, taking care to adjust their position in case the aluminum detail appears slightly shifted from one image to another one. Do not forget to click the "Update" button once the ROI positioning has been adjusted before pressing the "Measure" button.





3. Enter values obtained for mean pixel value of aluminum and PMMA and standard deviation of PMMA in Table 20.

SDNR compensation and AGD

5 Data input

1. For each phantom thickness, enter the exposure data "Anode Target Material", "Filter Material LT", "kV_p", "Exposure", "Body Part Thickness" (extracted from the DICOM header for DRs or annotated during acquisition for CRs) in Table 20.
2. Enter in the same table the MPV and SD measured from the images.

PMMA thickness (mm)	Body part thickness (mm)	Anode Target Material	Filter Material LT	kV_p	Exposure mAs	MPV_{Al}	MPV_{PMMA}	SD_{PMMA}
20								
30								
40								
45								
50								
60								

Table 20: Data input for SDNR and AGD calculation.

6 Example of data input

PMMA thickness (mm)	Body part thickness (mm)	Anode Target Material	Filter Material LT	kV_p	Exposure mAs	MPV_{Al}	MPV_{PMMA}	SD_{PMMA}
20	19	Mo	Mo	26	23.3	362.16	438.01	6.23
30	30	Mo	Rh	26	37.4	419.44	492.46	6.87
40	39	Mo	Rh	27	53.7	424.01	484.10	6.86
45	44	Mo	Rh	28	57.5	438.87	496.44	6.99
50	48	Rh	Rh	29	54.1	450.44	506.11	7.17
60	58	Rh	Rh	30	63.7	474.64	528.11	7.55

Table 21 - AEC: SDNR compensation and AGD. Example of data input for a DR system (GE Senographe DS, AOP/STD).

SDNR compensation and AGD

7 SDNR data output

1. For each phantom thickness, the SDNR is calculated according to Eq.(5) and shown in Table 22.
2. Moreover, the SDNR variation with respect to the reference thickness 45 mm, $\Delta SDNR_{45mm}$, is obtained according to the following equation

$$\Delta SDNR_{45mm} = \frac{(SDNR_i - SDNR_{45mm})}{SDNR_{45mm}} \times 100 \qquad \text{Eq. (8)}$$

where $SDNR_i$ is the SDNR measured at the *i*-th thickness and $SDNR_{45mm}$ the SDNR measured at 45 mm.

3. The AEC compensation performance is evaluated in terms of $\Delta SDNR_{45mm}$. The limiting values are reported in Table 22.

PMMA thickness(mm)	SDNR	$\Delta SDNR_{45mm}$	Limiting values $\Delta SDNR_{45mm}$	Test Status
20			≥ 0%	Pass/Fail
30			≥ 0%	Pass/Fail
40			≥ 0%	Pass/Fail
45			0 %	Pass/Fail
50			≥ - 15%	Pass/Fail
60			≥ - 30%	Pass/Fail

Table 22 Data output for SDNR values and variations.

SDNR compensation and AGD

8 SDNR example of data output

PMMA thickness(mm)	SDNR	$\Delta SDNR_{45mm}$	Limiting values $\Delta SDNR_{45mm}$	Test Status
20	19.57	+ 87.5 %	≥ 0%	Pass/Fail
30	15.40	+ 47.5 %	≥ 0%	Pass/Fail
40	12.14	+ 16.3 %	≥ 0%	Pass/Fail
45	10.44	0.0 %	0 %	Pass/Fail
50	10.44	0.0 %	≥ - 15%	Pass/Fail
60	7.87	- 24.6 %	≥ - 30%	Pass/Fail

Table 23: Example of SDNR values and variations with a DR system (GE Senographe DS, AOP/STD).

9 AGD data output

1. For each phantom thickness, the AGD is calculated according to Eq.(6) using exposure data in Table 20 and tube output and HVL measured as reported in the X-ray source chapter.
2. Table 24 shows HVL and entrance air-kerma calculated from tube output. The percentage of glandular tissue is assumed to be fix for PMMA and conversion coefficients, g-fact, s-fact, c-fact, determined from the work by Dance and colleagues.

PMMA thickness (mm)	HVL mmAl	K_i (mGy)	% glandular tissue	g-fact	s-fact	c-fact	AGD (mGy)	AGD EU Limits (mGy)	Test Status
20			97 %					1.0	Pass/Fail
30			67 %					1.5	Pass/Fail
40			41 %					2.0	Pass/Fail
45			29 %					2.5	Pass/Fail
50			20 %					3.0	Pass/Fail
60			9 %					4.5	Pass/Fail

Table 24 Data output for AGD estimation. HVL and K_i are obtained from the tests on the x-ray source.

SDNR compensation and AGD

10 Example of AGD data output

PMMA thickness (mm)	HVL mmAl	K_i (mGy)	% glandular tissue	g-fact	s-fact	c-fact	AGD (mGy)	AGD EU Limits (mGy)	Test Status
20	0.334	1.80	97 %	0.407	1.000	0.873	0.65	< 1.0	Pass
30	0.384	2.45	67 %	0.316	1.017	0.944	0.74	< 1.5	Pass
40	0.393	4.10	41 %	0.229	1.017	1.040	0.99	< 2.0	Pass
45	0.402	5.02	29 %	0.199	1.017	1.102	1.12	< 2.5	Pass
50	0.419	4.85	20 %	0.180	1.061	1.151	1.07	< 3.0	Pass
60	0.431	6.59	9 %	0.146	1.061	1.233	1.26	< 4.5	Pass

Table 25 Example of data output for AGD estimation with a DR system (GE Senographe DS, AOP/STD).

LIMITING VALUES

PMMA thickness(mm)	$\Delta SDNR_{45mm}$	AGD (mGy)
20	≥ 0%	< 1.0
30	≥ 0%	< 1.5
40	≥ 0%	< 2.0
45	0 %	< 2.5
50	≥ - 15%	< 3.0
60	≥ - 30%	< 4.5

SDNR compensation and AGD

REFERENCES

1. *European Guidelines for Breast Cancer Screening and Diagnosis, 4th edition, 2006.*
2. *E R Hendrick, E D Pisano, A Averbukh et al, Digital mammography and screen-film mammography in the Ametican College of Radiology Imaging Network Digital Mammographic Imaging Screening Trial, Am J Roen AJR 194:362-369 (2010).*
3. *G Gennaro, L Katz, H Souchay et al. Are phantoms useful for predicting the potential of dose reduction in full-field digital mammography ? Phys Med Biol 50:1851-1870 (2005).*
4. *A Rose. The sensitivity performance of the human eye on an absolute scale, JOSA 38:196-208 (1948).*
5. *R Aufrichtig. Comparison of low contrast detectability between a digital amorphous silicon and a screen-film based imaging system for thoracic radiography, Med Phys 26:1349-1358 (1999).*
6. *M S Chesters, G A Hay. Quantitative relation between detectability and noise power. Phys Med Biol 28:1113-1125 (1983).*
7. *K C Young, J J H Cook, J M Odoku, H Bosmans. Comparison of software and human observers in reading images of the CDMAM test object to assess digital mammography systems. Proceedings of SPIE vol 6142, 614206 (2006).*
8. *A E Burgess. The Rose model, revisited. JOSA A, 16:633-646 (1999).*
9. *IAEA Human Health Series No.17 "Quality Assurance Programme to Digital Mammography" (2011).*
10. *ICRU Report 74 "Patient dosimetry for X-ray used in medical imaging" (2005).*
11. *W T Sobol, X Wu, Parametrization of mammography normalized average glandular dose tables, Med Phys 24:547-554 (1997).*
12. *D R Dance, C L Skinner, K C Young, J R Beckett, C J Kotre, Additional factors for the estimation of mean glandular dose using the UK mammography dosimetry protocol , Phys Med Biol 45:3225-3240 (2000).*
13. *J M Boone, Normalized glandular dose (DgN) coefficients for arbitrary x-ray spectra in mammography: Computer-fit values of Monte Carlo derived data, Med Phys 29: 869-875 (2002).*
14. *D R Dance, K C Young, R E van Engen, Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA breast dosimetry protocol, Phys Med Biol 54: 4361-4372 (2009).*
15. *Statement from teh 1987 Como meeting of the ICRP Publication 52. Ann ICRP, 17(4) Oxford and New York Pergamon Press, 1987.*
16. *J Zoetelief, M Fietzgerald, W Leitz, M Sabel, "European Protocol on Dosimetry in Mammography", EUR 16262 (1996).*

Quality Controls

Detector

Image Detector

Introduction

Image detector performance could be fully characterized by applying the linear shift-invariant (LSI) system theory [1]. This theory describes how a system (image detector) acts on the input signal to produce the output (image), using the concept of "transfer function". The only two constraints are that the detector response is linear (output proportional to input) and shift-invariant (response not dependent on what position of the detector the response is measured at). Not all image detectors are linear and shift-invariant but most of them can be assumed to be LSI if limited response intervals and limited areas are considered. Ideal detectors are noise-free and capable of fully transferring the input signal into the output image, with no information loss. Real detectors introduce noise and transfer only a portion of the input signal into the output image. The LSI system theory permits characterization of transfer properties of an image detector in the spatial frequency domain, using the modulation transfer function (MTF) for the signal, and the noise power spectrum (NPS) for the noise. A global figure-of-merit for the performance of image detectors is the Detective Quantum Efficiency (DQE), which describes how effectively an imaging system produces an image with high signal-to-noise ratio (SNR) compared to an ideal detector. DQE can be derived from MTF and NPS measurements.

While methods to standardize DQE measurements have been developed, in this document DQE is not included because:

1. Neither the image detectors used by DR mammography systems nor the imaging plates used with CR systems, are provided with any type of "certification" by manufacturers. Therefore, if a user performs DQE measurements, results cannot be compared with any reference value or curve.
2. Detector DQE significantly affects the image formation process, but final image quality depends on many other factors. For this reason, mammography systems with quite large differences in terms of DQE are clinically accepted.
3. DQE measurements are not trivial and not easily applicable by non-experts in a routine quality control setting.

Image Detector

Quality control of the image detector only aims to compare its actual and expected behavior. Here, "expected behavior" means the following:

1. The response function is the curve which relates the output signal (typically the mean pixel value measured in a region of the image) with the input signal (radiation dose). Usually, but not always, the response function is linear for DR systems, while can be a logarithmic or square root function of dose for CR systems.

2. Quantum noise is the dominant component of the image noise. In general, there are multiple noise components in digital images (some of them dependent on detector technology), and quantum noise, related to the statistical process of interaction between incident X-ray photons and image detector, can be separated from structural and electronic noise, which depend on manufacturing characteristics of any given type of detector. It is important that the imaging system is set to work within a dose range where quantum noise, which is under "user control" through selection of technique factors, is the main component of the image noise

Image Detector

3. The detector response is “shift-invariant”, i.e. the same mean signal is expected from any detector area, given the same input signal.

There is an inherent issue in the verification of the detector shift-invariance in mammography, caused by the geometrical conditions used. In fact, the short source-to-breast distance (about 60 cm) and the angled anode target designed to produce a small X-ray spot which secures high spatial resolution images, make the X-ray beam heavily non-uniform, especially in the cathode-anode direction. Such X-ray beam non-uniformity prevents to test image detector shift-invariance, and possible “uniformity tests” assume different meanings, as remarked later in this chapter.

4. If more than one image detector is used (CR systems), all detectors are supposed to have similar responses/sensitivities.

Each imaging plate used with CR systems is an individual image detector. Nevertheless, as it would be too time consuming to perform the full set of tests for all the imaging plates, all quality control tests should be performed using the same imaging plate. However, possible differences in sensitivity among plates must be also evaluated (see “Inter-plate variability” test). Sensitivity differences among different imaging plates are supposed to be limited.

Image Detector

"Uniformity" with DR systems:
DR mammography systems apply a "bad pixel correction" to remove the effect of possible defective elements, and a so-called "flat-field" or "gain" correction to any image acquired, in order to correct possible differences in sensitivity among detector elements. Both corrections are applied before the "FOR PROCESSING" image becomes accessible to users. However, both corrections (bad pixel and flat fielding process) compensate for any type of non-uniformities, either produced by the image detector or by the X-ray beam. As a consequence of the flat fielding, despite the significant non-uniformity of the photon beam (heel effect), the image produced by a uniform object by a DR system will show homogeneous pixel value, while inhomogeneity shows up if any parameter related to image noise is calculated.
"Uniformity" with CR systems:
CR systems are equipped with multiple image detectors, i.e. cassettes containing imaging plates, based on the properties of photostimulable phosphors they are made with. When an imaging plate is scanned after exposure by the CR scanner to produce the digital image, no correction is applied. Thereby, the image of a uniform object produced by a CR system includes possible phosphor inhomogeneity as well as the spatial variation of the X-ray beam.
CR systems are expected to produce less uniform images than DR systems.

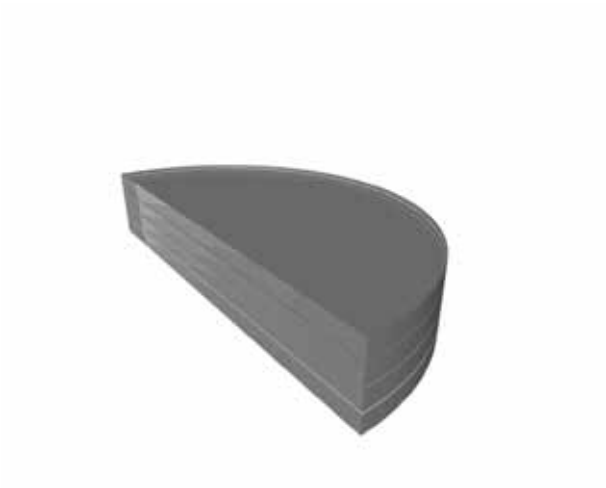
Response function and noise evaluation

PURPOSES

1. **CHECK IF THE RESPONSE FUNCTION OF MEAN PIXEL VALUE (MPV) VERSUS DOSE IS CONSISTENT WITH THE EXPECTED TREND.**
2. **CHECK IF QUANTUM NOISE IS THE MAIN COMPONENT OF IMAGE NOISE.**

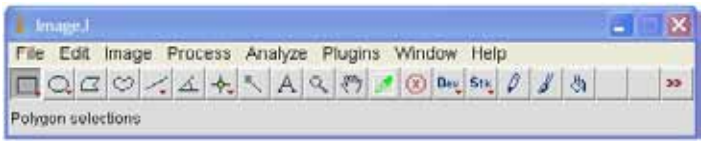
EQUIPMENT

1 Poly-methyl methacrylate (PMMA) block: 45mm



Same remarks as done for the SDNR measurements in the AEC chapter about PMMA shape versus "AEC compatibility"..

2 Software for image analysis (in the following, all instructions are provided for the freeware package ImageJ, but any other equivalent software can be used.)



Response function and noise evaluation

TEST FREQUENCY

» Acceptance/Commissioning.
» After possible replacement of image detector or X-ray source.
» Annual (tube output is expected to slowly decay in time).

PROCEDURE

1 Geometry

1. Place the PMMA block on the breast support, aligned with the chest wall edge and laterally centered.
2. Mount the compression paddle compatible with the test object size, and check that the collimator aperture is appropriate.
3. Lower the compression paddle, and apply a minimum compression force by 3-5 daN, if this is required to enable the automatic exposure mode.

2 Exposure mode

1. Ensure that the acquisition workstation is set to save "FOR PROCESSING" images and any possible pre-processing algorithm is disabled.
2. If possible, disable the "automatic decompression after exposure" from the mammography unit console.
3. Record a new patient on the acquisition workstation.
4. Expose and acquire one phantom image. The aim of this exposure is only to identify the anode/filter combination and kV_p value chosen by the AEC for the test object. Such parameters will be kept unchanged during the acquisition of the full dataset. With CR systems it is not necessary to scan the plate after exposure. Just remember to perform a "primary cancellation" of the imaging plate before re-using it.
5. Switch to the manual exposure mode from the mammography unit console and set the same anode/filter combination and kV_p value as previously chosen by the AEC.
6. Acquire a set of 8-10 images at increasing mAs values (typically, from 10 to 200 for integration systems; from 0.5 to 15 for photon counting systems).
7. With CR systems, always use the same imaging plate.

Response function and noise evaluation

3 Exposure data collection

» **DR systems:** all exposure data are saved in the DICOM header of images and can be accessed at any time.
» **CR systems:** EXPOSURE DATA MUST BE ANNOTATED DURING IMAGE ACQUISITION, OTHERWISE THEY ARE LOST, unless some type of applications have been developed to communicate technique factors from the X-ray unit to the CR reader.

4 CR systems: IP reading mode

1. Read the imaging plate using the same test menu (or reading mode) as indicated in Table 15 of the AEC chapter and reported here.

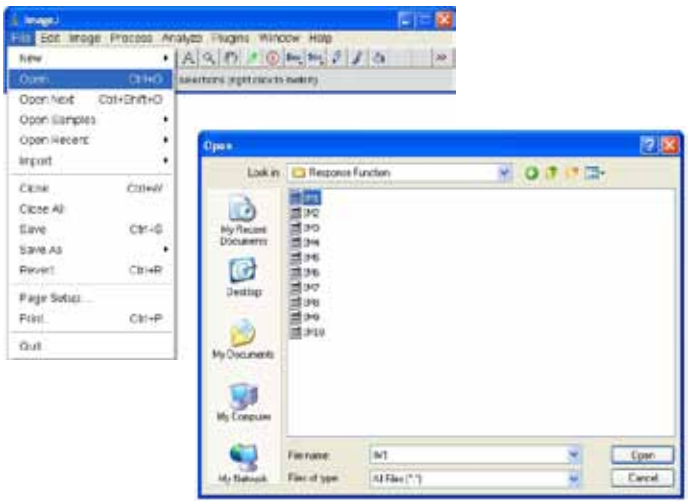
CR system	Test menu	Output
Agfa	System diagnostic/flat field	Pixel gray level
Carestream	Test	Pixel gray level
Fuji	1. Ave2 (preferred) 2. Linear (EDR fix, S=120, L=2)	1. Exposure Index (S) 2. Pixel gray level
Konica-Minolta	Test	Exposure Index (S)

In the following, it is assumed that the quality control measurements on the images are performed on an independent computer. This can be typically done by storing unprocessed images into the PACS and retrieving them from another PC. However, depending on each specific mammography systems, some types of measurements can be performed directly from the acquisition workstation.

Response function and noise evaluation

5 Exposure data extraction

1. Open an image.



2. As the image is "FOR PROCESSING", it will be displayed with the full dynamic range

3. Display the information included in the DICOM header

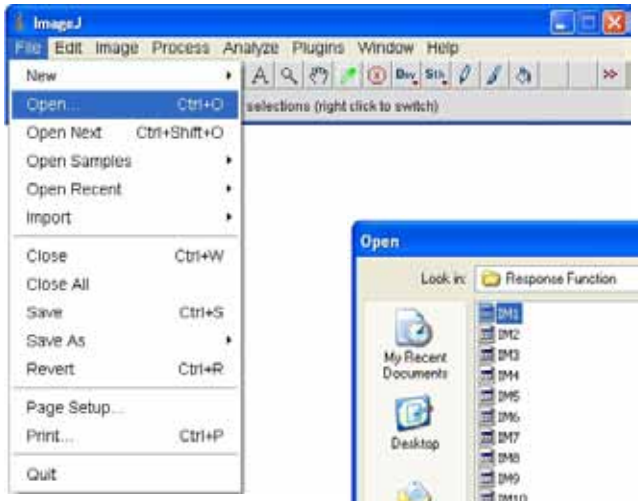
4. For DR systems, extract the following DICOM tags and enter them in the worksheet "Response Function" in the "Template_EFOMP_MammoWG" file:
   - o (0018,1191) "Anode Target Material"
   - o (0018,7050) "Filter Material LT"
   - o (0018,0060) "kV_p"
   - o (0018,1152) "Exposure"

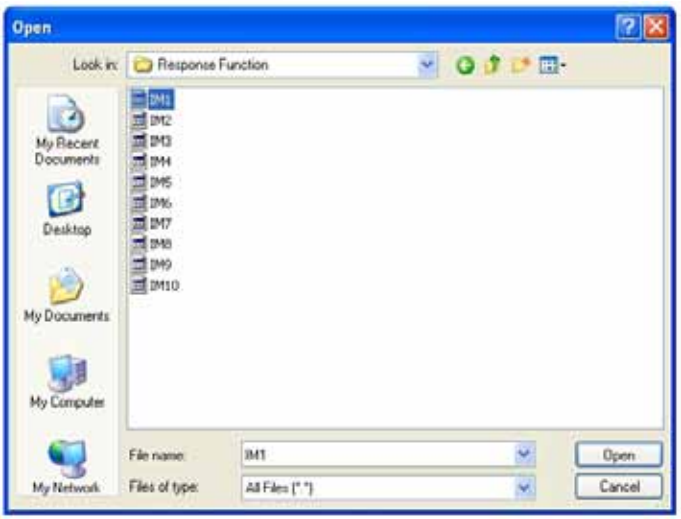
5. Enter the mAs value corresponding to each image in Table 26 of the worksheet "Response Function" in the "Template_EFOMP_MammoWG" file in ascending order.

Response function and noise evaluation

6 Response function and noise measurements

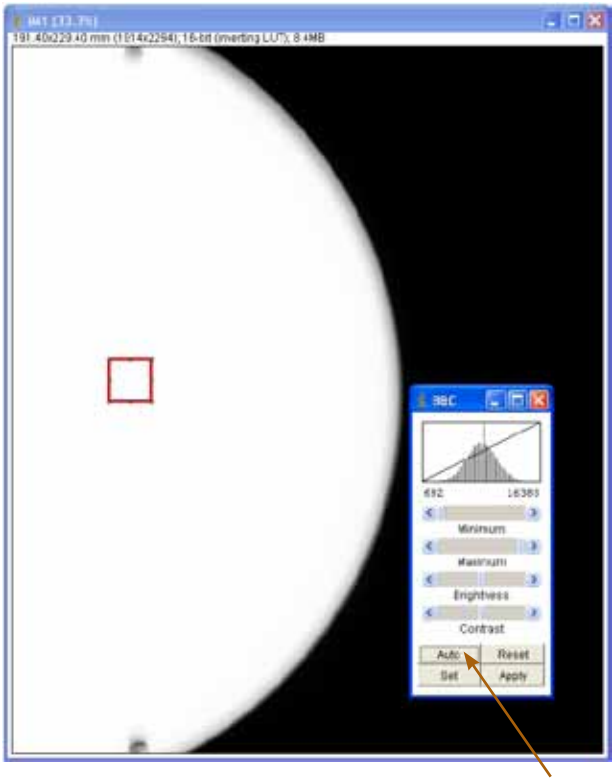
1. Open the first image acquired (lowest mAs value)

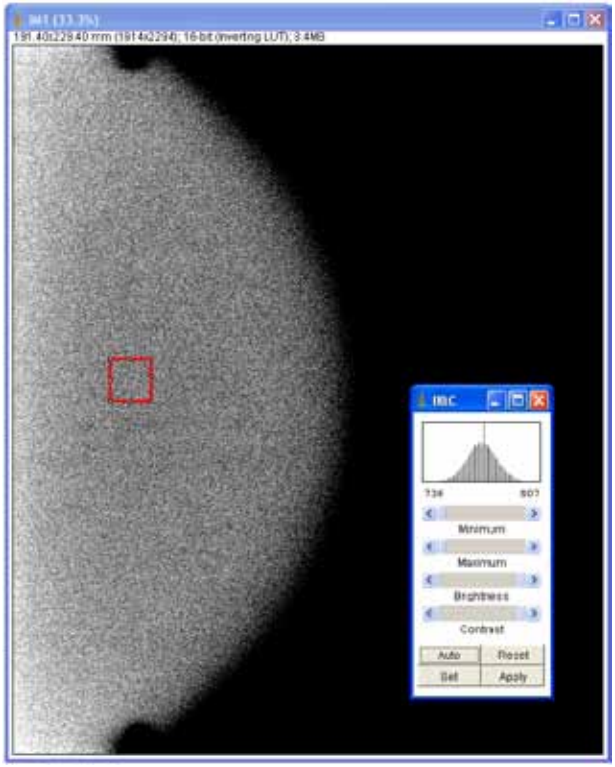




Response function and noise evaluation

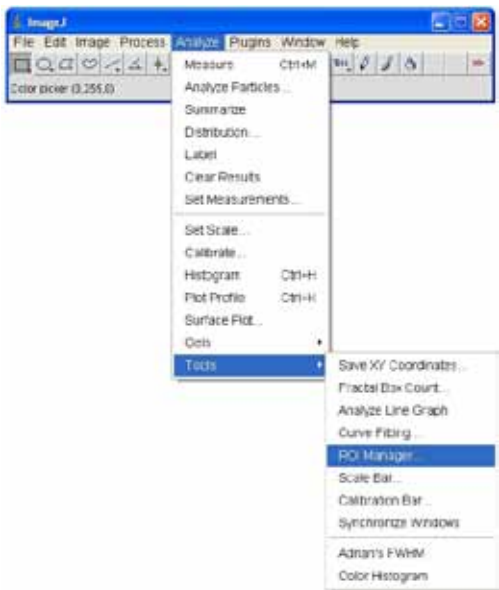
2. Select the rectangular ROI tool, keep the Shift key pressed to draw a square ROI (side between 10 and 15 mm), and put it at 5-6 cm from the chest wall edge and laterally centered.
3. Brightness and contrast adjustment is unnecessary in this case, the phantom being uniform.
4. This ROI definition is necessary only if either this is the first measurement or if a previously recorded ROI is not available (see the following pages to learn about ROI Manager use).

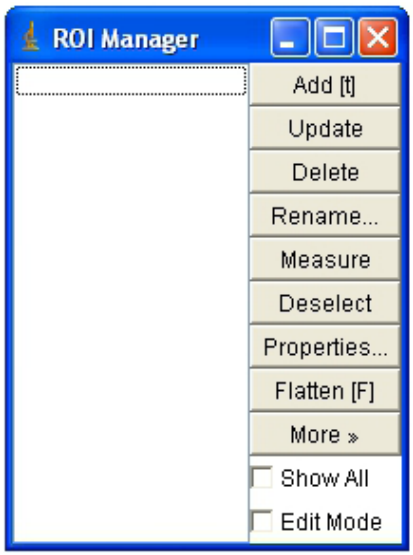




Response function and noise evaluation

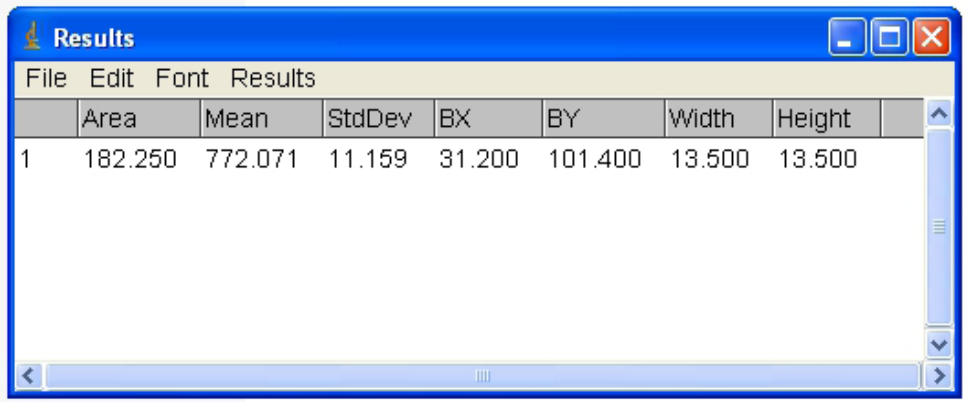
5. Click on "Analyze/Tools/ROI Manager…" to open the ROI Manager window. This tool allows the size and position of one or more ROIs to be recorded, so that the same ROIs can be used on different images, hence improving the consistency of measurements. The ROI set can also be saved and reused during annual tests. For a detailed description of all ROI Manger commands, click the following link http://imagejdocu.tudor.lu/doku.php?id=gui:analyze:tools#roi_manager .





6. Press the button "Add" in the ROI Manager menu to record the ROI.
7. Click on the ROI and press the button "Rename" in the ROI Manager menu to give it a proper name (optional)
8. To obtain the mean pixel value in the ROI, press either the button "Measure" in the ROI Manager menu or use the command CTRL+M (equivalent to "Analyze/Measure" of the ImageJ menu). The result window will appear with the resulting measurements.

Response function and noise evaluation



9. Enter the mean pixel values (MPV) and standard deviation (SD) in Table 26 of the worksheet "Response Function" in the "Template_EFOMP_MammoWG".
10. **REPEAT THE SAME MEASUREMENT FOR ALL THE IMAGES ACQUIRED AT DIFFERENT mAs LEVEL. IT IS ENOUGH TO OPEN THE IMAGES ONE-BY-ONE AND SELECT THE ROI OPENED IN THE ROI Manager TO PERFORM ALL MEASUREMENTS WITH THE SAME ROI SIZE AND POSITION.**
11. Insert MPVs and SDs obtained in Table 26 of the worksheet "Response Function" in the "Template_EFOMP_MammoWG".
12. If the same measurement was previously performed and the ROI already created and saved, just load the ROI by the command "More/Open" in the ROI Manager and adjust and save its position on the new image by using the command "Update".

Response function and noise evaluation

7 Data input for a DR system

mAs	K_i (mGy)	MPV	SD

K_i is automatically calculated from the tube output measurements

Table 26: Data input for the response function and noise evaluation tests of DR systems.

8 Example of data input for a DR system

mAs	K_i (mGy)	MPV	SD
10	0.863	141.45	5.66
16	1.379	230.40	6.80
25	2.154	363.69	8.21
36	3.101	525.61	9.57
50	4.307	732.65	11.10
80	6.889	1177.01	13.94
100	8.611	1473.42	15.78
140	12.055	2066.93	18.73
200	17.221	2955.69	22.60
250	21.525	3695.84	25.21

Table 27: Example of data input for the response function and noise evaluation tests of a DR system (GE Senographe Essential).

Response function and noise evaluation

9 Data input for a CR system

mAs	K_i (mGy)	MPV	SD	EI

Table 28: Data input for the response function and noise evaluation tests of a CR system.

K_i is automatically calculated from the tube output measurements

The Exposure Index (EI) is a parameter related to the radiation dose absorbed by the imaging plate. Its scale is different for different types of CR systems. The EI value is also a function of the menu selected on the CR reader. For Agfa and Carestream systems, the EI is fixed for the test menu, the same holds for Fuji if the "Linear"mode is selected. Otherwise, the EI changes with mAs values for both Fuji with menus different form "Linear", and Konica test mode.

Response function

10 Example of data input for a CR system

mAs	K_i (mGy)	MPV	SD	EI
10	0.81	566.68	22.21	1077
25	2.02	573.71	14.14	439
36	2.91	570.81	11.78	303
50	4.05	571.98	10.13	220
80	6.48	564.77	8.15	136
100	8.10	570.51	7.47	110
125	10.12	568.21	6.78	88
140	11.33	566.20	6.57	79
160	12.95	572.91	6.18	70
200	16.19	569.79	5.75	55

Table 29: Example of data input for the response function and noise evaluation tests of a CR system (Fuji FCR Profect, HR-BD imaging plate).

Response function

11 Data output for a DR system: response function

1. Most DR systems show LINEAR response function, as shown in the list below. The linearity of the response function can be tested by plotting MPV versus K_i or mAs, and performing a linear fit [1,2].

Linear response function	Non-linear response function
GE DRs (2000D, DS, Essential)	Fuji DR (Amulet)
Hologic DRs (Selenia, Dimensions)	
IMS DR (Giotto Image)	
Philips DRs (Mammo Diagnost, Microdose)	
Planmed DR (Nuance)	
Siemens DRs (Novation, Inspiration)	

2. Table 30 reports data output for DR systems.
3. The mean distance of experimental points from the fitted straight line is also evaluated (COV is the coefficient of variation)

$MPV = A + B \cdot x$		
MPV	**vs.**	**x**
A =		
B =		
R2 =		**≥ 0.95**
$COV = \left(\frac{MPV_i - A}{x_i}\right)$		**≤ 10 %**

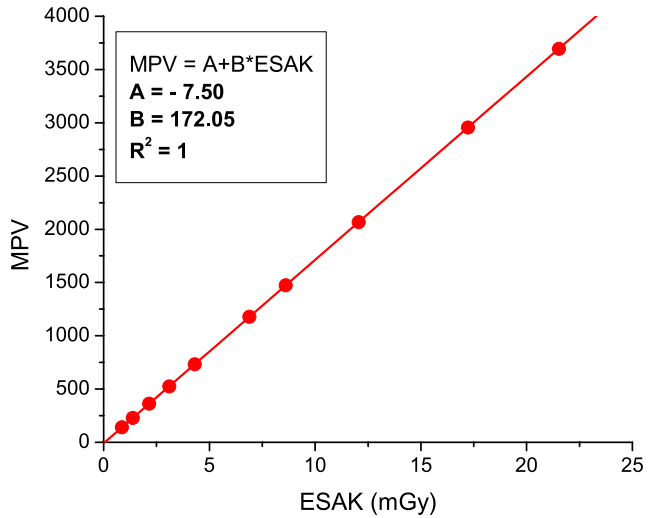
Options for x:
- » mAs
- » K_i
- » Ln(mAs)
- » Ln(K_i)

Table 30: Data output for the response function test of a DR system.

Response function

12

Example of data output for a DR system: response function



$\mathrm{MPV} = A + B \cdot x$		
MPV	**vs.**	**x**
A =	-7.50	
B =	172.05	
R^2 =	1	≥ 0.95
$\mathrm{COV} = \left(\frac{\mathrm{MPV}_i - A}{x_i}\right)$	5.85 %	≤ 10 %

Table 31: Example of data output for the response function test of a DR system (GE Senographe Essential).

Response function

13 Data output for a CR system: response function

1. All CR systems show a NON-LINEAR response function, typically logarithmic or power functions. CR response can be tested by plotting either CR output (exposure index EI or mean pixel values MPV, depending on the screen processing and the individual CR system) or its logarithm versus the logarithm of K_i or mAs, and performing a log-linear or log-log-linear fit [1,2].
2. Table 32 shows data output for CR systems.
3. The mean distance of experimental points from the fitted straight line is also evaluated (COV is the coefficient of variation). The formula is generic, and the x and y variables can change among different CR systems and test menu used to process the imaging plate.

$y = A + B \cdot x$		
y	**vs.**	**x**
A =		
B =		
R^2 =		**≥ 0.95**
$\mathrm{COV} = \left(\frac{y_i - A}{x_i}\right)$		**≤ 10 %**

Table 32: Data output for the response function test of a CR system.

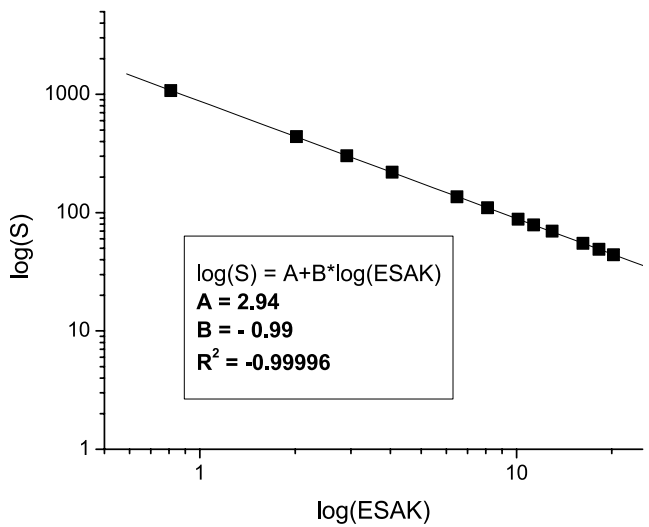
Options for y:
- EI
- Ln(EI)
- MPV
- Ln(MPV)

Options for x:
- mAs
- K_i
- Ln(mAs)
- Ln(K_i)

Noise evaluation

14

Example of data output for a CR system: response function



$y = A + B \cdot x$		
log(S)	**vs.**	**log(mAs)**
A =	**4.03**	
B =	**-0.99**	
R^2 =	**0.99998**	**≥ 0.95**
$\mathrm{COV} = \left(\frac{y_i - A}{x_i}\right)$	**0.15 %**	**≤ 10 %**

Table 33: Example of data output for the response function test of a CR system (Fuji FCR Profect, HR-BD imaging plate).

Noise evaluation

15 Noise evaluation

» An imaging system is called "quantum limited" when quantum noise (i.e. the noise component related to the statistical processes of interaction between X-ray photons and image detector) is the main component of image noise. Image noise can be derived from the standard deviation (SD) of pixel values in a region of the image. For quantum limited systems with a linear response function, the standard deviation is proportional to the square root of the detector dose. As the detector dose is difficult to estimate, dose at the test object (K_i or mAs) are used.

» Image noise can be evaluated from the same input tables used for the response function, by plotting variance (SD^2) against the dose parameter (mAs or K_i) and applying a linear fit [1,2].

» For CR systems, image pixel values are normalized, and the image noise measured through the standard deviation inside a ROI is inversely proportional to detector dose. Thereby, with CR systems, the main contribution of quantum noise to the overall image noise can be verified by plotting variance (SD^2) against the reciprocal of the dose parameter (mAs^{-1} or K_i^{-1}) and applying a linear fit [2,3].

16 Data output for a DR system: noise evaluation

$SD^2 = a + b \cdot x$		
SD^2	**vs.**	**x**
a =		
b =		
R^2 =		**≥ 0.95**

Table 34: Data output for the noise evaluation test of DR systems.

Options for x:
» mAs
» K_i

Noise evaluation

17 Example of data output for a DR system: noise evaluation

$SD^2 = a + b \cdot x$		
SD^2	vs.	K_i
a =	0.95	
b =	29.39	
R^2 =	0.99942	≥ 0.95

Table 35: Example of data output for the noise evaluation test of a DR system (GE Senographe Essential).

18 Data output for a CR system: noise evaluation

$SD^2 = a + b \cdot x$		
SD^2	vs.	x
a =		
b =		
R^2 =		≥ 0.95

Options for x:
» mAs^{-1}
» K_i^{-1}

Table 36: Data output for the noise evaluation test of CR systems.

Noise evaluation

19

Example of data output for a CR system: noise evaluation

$SD^2 = a + b \cdot x$		
SD^2	vs.	mAs^{-1}
a =	7.60	
b =	4839.97	
R^2 =	0.9981	≥ 0.95

Options for x:
» mAs^{-1}
» K_i^{-1}

Table 37: Example of data output for the noise evaluation test of a CR systems (Fuji FCR Profect, HR-BD imaging plate).

Noise evaluation

LIMITING VALUES

Response function [2]:

Type of response	Relationship with dose	Variables	Limiting value
Linear	$\mathrm{MPV} = A + B \cdot x$	x: mAs K_i y: MPV	$R^2 \geq 0.95$
	$\mathrm{COV} = \left(\frac{\mathrm{MPV}_i - A}{x_i} \right)$		≤ 10 %
Non-linear	$y = A + B \cdot x$	x: ln(mAs) ln (K_i) y: EI ln(EI) MPV ln (MPV)	$R^2 \geq 0.95$
	$\mathrm{COV} = \left(\frac{y_i - A}{x_i} \right)$		≤ 10 %

Noise evaluation [2]:

Type of response	Relationship with dose	Variables	Limiting value
Linear	$SD^2 = a + b \cdot x$	x: mAs K_i	$R^2 \geq 0.95$
Non-linear		x: mAs^{-1} K_i^{-1}	

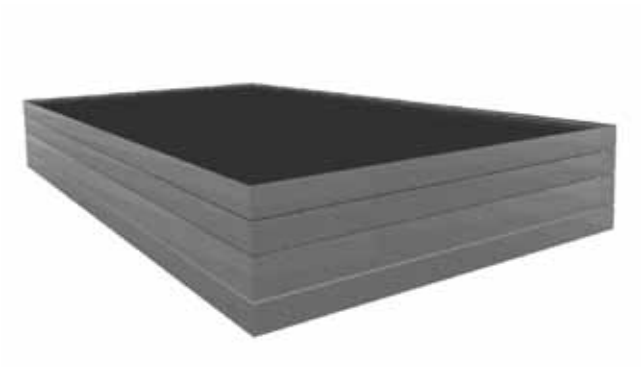
Uniformity

PURPOSES

1. **CHECK THAT THE MEAN SIGNAL PRODUCED BY THE IMAGE DETECTOR WHEN A UNIFORM TEST OBJECT IS EXPOSED, IS LOCALLY UNIFORM**
2. **TEST THE OVERALL EFFECTIVENESS OF THE "FLAT-FIELD" OR "GAIN" CORRECTION (DR systems only)**

EQUIPMENT

1 Uniform PMMA rectangular block covering the whole detector area. Thickness between 25 and 45 mm.

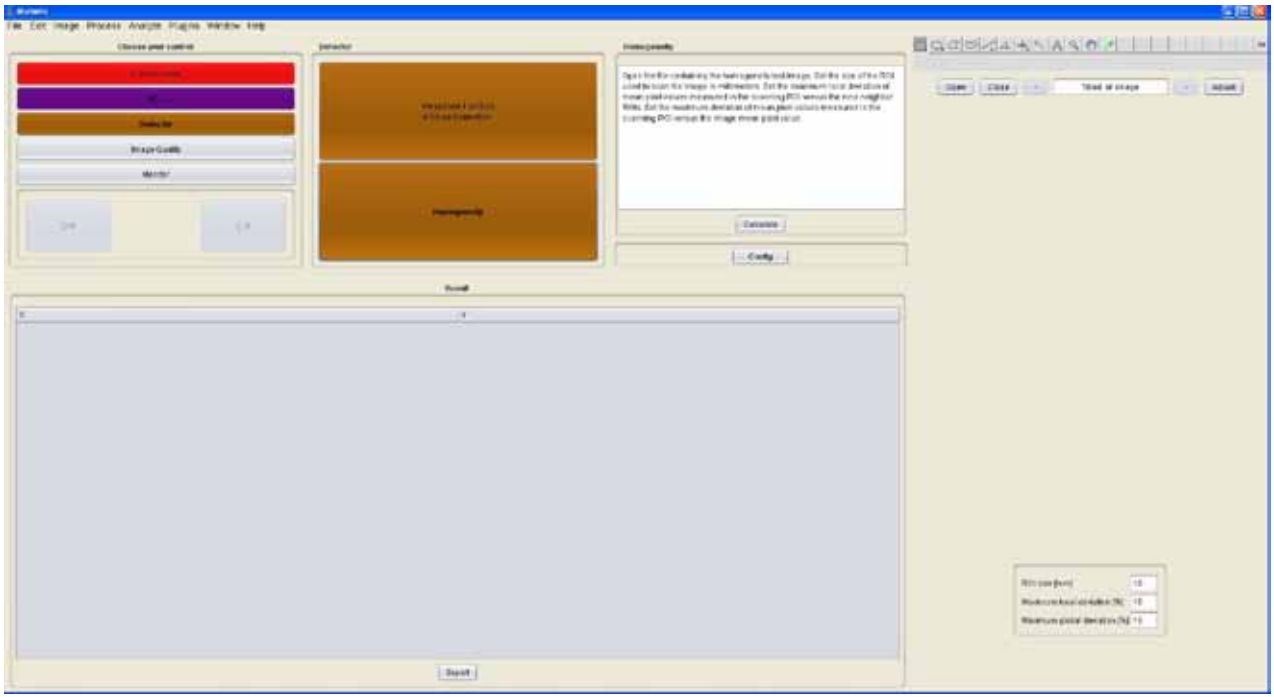


If a uniform PMMA block covering the full detector has been provided by the manufacturer, it can be used for the uniformity test.
BEFORE PERFORMING THIS TEST, IT IS VERY IMPORTANT TO VERIFY THAT THE PMMA BLOCK:

- COVERS THE FULL DETECTOR AREA
- IS REALLY UNIFORM (NO DEEP SCRATCHES)

Uniformity

2 Software to analyze the test image



Uniformity

TEST FREQUENCY

» Acceptance/Commissioning.
» After possible replacement of image detector or X-ray source.
» Annual (tube output is expected to slowly decay in time).

PROCEDURE

1 Geometry

1. Place the uniform PMMA block on the breast support, taking care to completely cover the image detector.
2. Mount the largest compression paddle and check that the collimator aperture is appropriate.
3. Lower the compression paddle and apply a minimum compression force by 3-5 daN, if this is required to enable the automatic exposure mode.

2 Exposure mode

1. Ensure that the acquisition workstation is set to save "FOR PROCESSING" images and any possible pre-processing algorithm is disabled.
2. Record a new patient on the acquisition workstation.
3. Acquire one image in AEC mode. For CR system, ensure you are using the same imaging plate as used for the response function/noise evaluation tests.
4. Take note of technique factors (A/F, kV_p, mAs) selected by the AEC. The same images and data can be used for the artifact an inter-plate variability tests.
5. For CR systems, the uniformity test should be performed for any imaging plate clinically used.

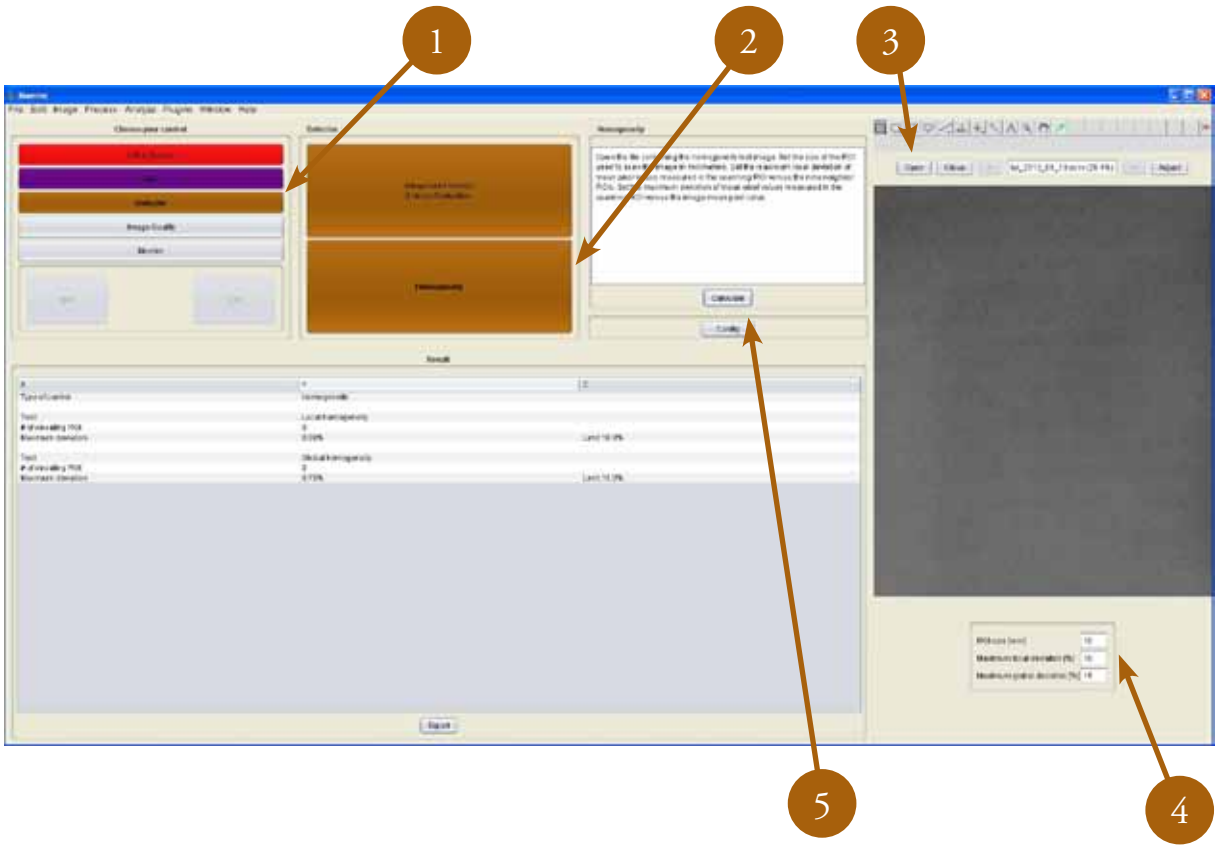
Uniformity

3 Image analysis

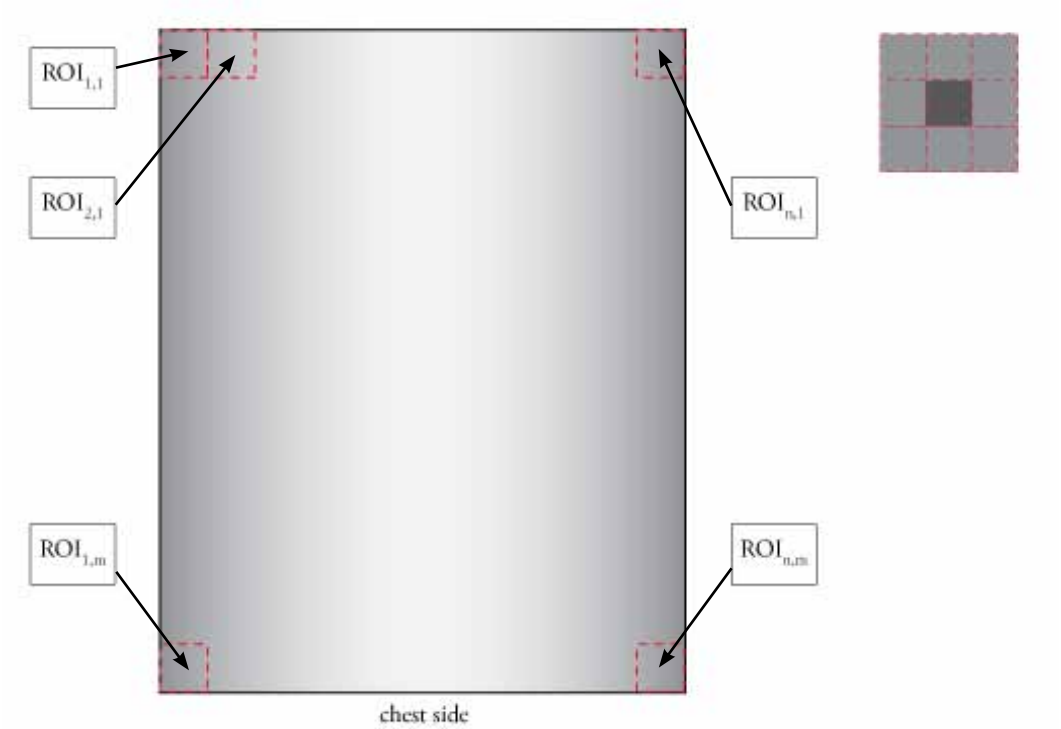
1. Launch the “COQ Mammo” plugin from ImageJ

2. Select “Detector” test type and “Uniformity” test.
3. Open the file containing the uniform image
4. Set the size of the ROI used to scan the image in units of millimeters. The default value is 10 mm.
5. Set the maximum local deviation of mean pixel values measured in the scanning ROI versus the nine neighbouring ROIs. The default value is 5 %. Only for DR systems, set the maximum deviation of mean pixel values measured in the scanning ROI versus the image mean pixel value. The default value is 5%.

Uniformity



Uniformity



Local uniformity is determined by calculating the local difference between each **MPV**$_{i,j}$ and the average across the eight neighbouring MPV's, **MPV**$_{neighbour}$:

$$\mathrm{LU} = \max\left(\frac{|\mathrm{MPV}_{i,j} - \mathrm{MPV}_{neighbour}|}{\mathrm{MPV}_{neighbour}}\right) \leq 0.05$$

Global uniformity is only required for DR systems, where flat-field correction is performed. It can be obtained as the maximum deviation between **MPV**$_{i,j}$ and the mean pixel value measured from the entire image, **MPV**$_{image}$:

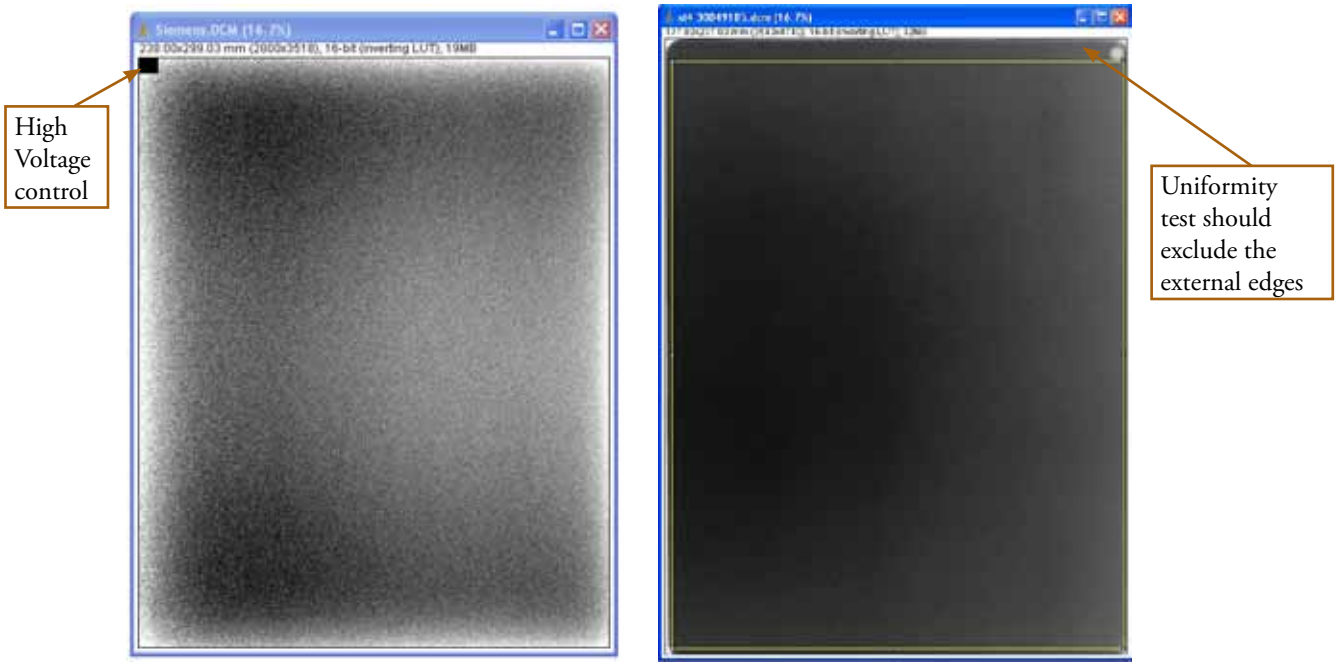
$$\mathrm{GU} = \max\left(\frac{|\mathrm{MPV}_{i,j} - \mathrm{MPV}_{image}|}{\mathrm{MPV}_{image}}\right) \leq 0.10$$

Uniformity

NATURAL NON-UNIFORMITIES

It should be remarked that some detectors have "natural non-uniformities" which should be excluded from the uniformity test:

1. DR systems using selenium flat panel detectors have one or more small areas not sensitive to radiation, used to bring the high voltage to the panel.
2. Imaging plates used by CR systems do not fit the cassette perfectly, and a white edge is visible on the image borders.



Uniformity

4 Data output: DR systems

1. The GU index with DR systems also verifies the effectiveness of flat-field (or gain) correction.
2. The output table shows how many ROIs deviated from the liming values for both LU and GU, and the maximum deviation.

Local uniformity (LU)		
N° of deviating ROI	_____	
Maximum deviation	_____	Limit ≤ 5%
Global uniformity (GU)		
N°of deviating ROI	_____	
Maximum deviation	_____	Limit ≤ 10%

Table 38: Data output for the uniformity test of DR systems.

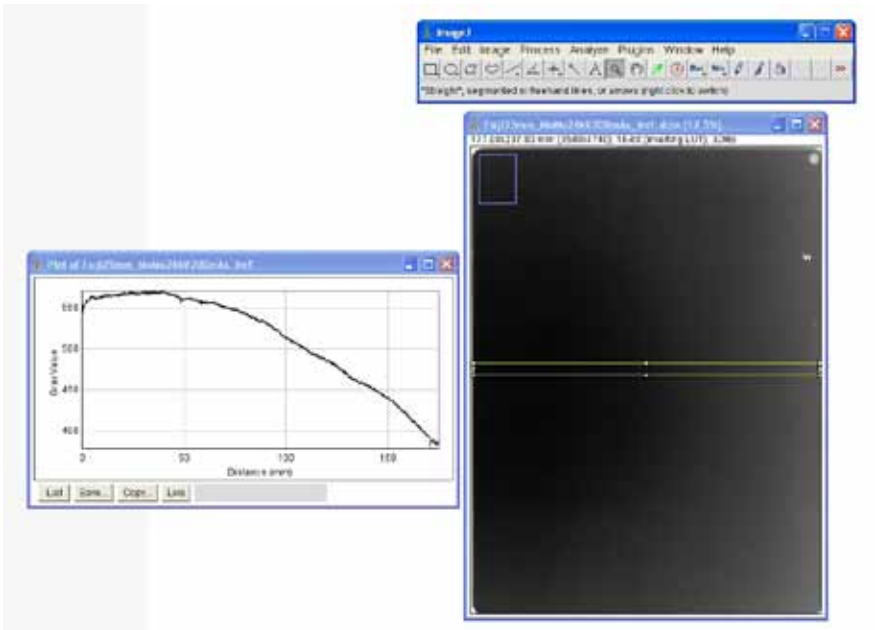
5 Example of data output: DR systems

Local uniformity (LU)		
N° of deviating ROI	0	
Maximum deviation	0.13 %	Limit ≤ 5%
Global uniformity (GU)		
N°of deviating ROI	0	
Maximum deviation	1.92 %	Limit ≤ 10%

Table 39: Example of data output for the uniformity test of a DR systems (GE Senographe Essential).

Uniformity

Why global homogeneity cannot be required for CR systems



In the figure above, a profile was obtained in the cathode-anode direction from the image obtained by exposing a uniform block of PMMA with a CR system. The output signal ranges from 550 down to about 390, showing a relative difference of 30%. Since CR systems do not correct for X-ray beam non-uniformity, there is no reason to expect signal variation compared to the whole image within 15%.

Uniformity

6 Data output: CR systems

1. The uniformity test is limited to local uniformity (LU) for CR systems.
2. The output table shows how many ROIs deviated from the liming values for LU, and the maximum deviation.

Local uniformity (LU)		
N° of deviating ROI	_____	
Maximum deviation	_____	Limit ≤ 5%

Table 40: Data output for the uniformity test of CR systems.

7 Example of data output: CR systems

Local uniformity (LU)		
N° of deviating ROI	0	
Maximum deviation	1.22 %	Limit ≤ 5%

Table 41: Example of data output for the uniformity test of CR systems (FCR Fuji Profect, HR-BD imaging plate).

LIMITING VALUES

Local uniformity (LU): ≤ 5% (any system).
Global unifotmity (GU): ≤ 10% (DR systems).

Artifacts

Artifacts can be defined as image "features" not produced by the imaged object [4]. They might be lines or streaks or dots, clearly not belonging to any imaged breast. However, in some cases, they might also bear a very close resemblance to clinical features, like microcalcification clusters or small masses. While artifacts in the first group are not desirable and should be removed, but do not have any significant impact on the clinical outcome, those belonging to the second group are really dangerous and could lead to unnecessary biopsies. Image artifacts can have different causes, for example uncorrected defective elements or degraded gain correction map in DR systems, or streaks due to the imaging plate dragging mechanism or dust on the imaging plate itself. There is no general "recipe" to distinguish between potentially dangerous and harmless artifacts, but, in general, artifacts should be detected and removed as soon as possible. The method proposed here is visual and very easy to apply, but remedial actions should be decided from case to case, according to the hazard of misinterpretation associated with each given artifact. Some artifacts could take weeks or months to get removed without any risk of adverse consequence, while others could require an immediate suspension of the use of the mammography equipment.

In principle, artifact search should be one of the daily checks performed by radiographers. It consists in the exposure of a uniform PMMA block (the same used for the uniformity test), and in the visual detection of possible artifacts in the resulting image, after having stressed the contrast [2]. If one or more artifacts are detected through this routine check, the radiographer should involve the medical physicist, who can then analyze the image according to the procedure describe in the following session.

Artifacts

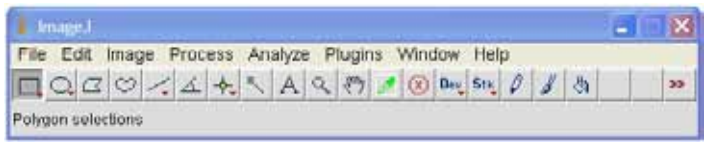
PURPOSES

IDENTIFY ANY ARTIFACT ABLE TO DISTURB/COMPROMISE DIAGNOSIS

EQUIPMENT

1 Uniform PMMA rectangular block covering the full field of view, as used for the uniformity test. Overall thickness between 25 and 45 mm

2 Software for image analysis: in the following, all instructions are provided for the freeware package ImageJ, but any other equivalent software can be used



TEST FREQUENCY

» Daily check by radiographers + analysis by medical physicist in case of artifacts.
» Acceptance/Commissioning.
» Annual.

Artifacts

PROCEDURE

1 Geometry

1. Place the uniform PMMA block on the breast support, taking care to completely cover the detector.
2. Mount the compression paddle compatible with the test object size.

2 Exposure mode

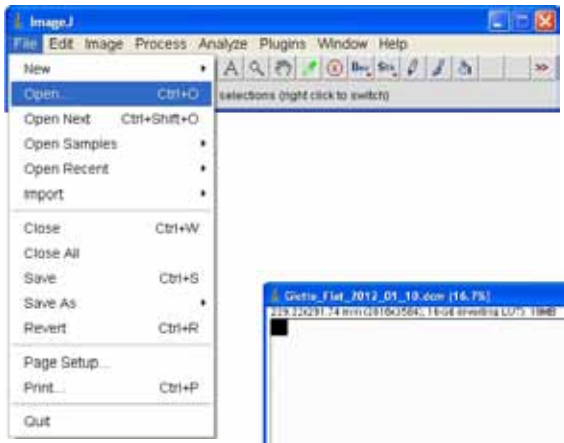
1. Record a new patient on the acquisition workstation.
2. Ensure that the acquisition workstation is set to save “FOR PROCESSING” images and any possible pre-processing algorithm is disabled.
3. Acquire an image in automatic exposure mode. On annual quality controls, all target/filter combinations should be checked, than the one routinely selected by the AEC.

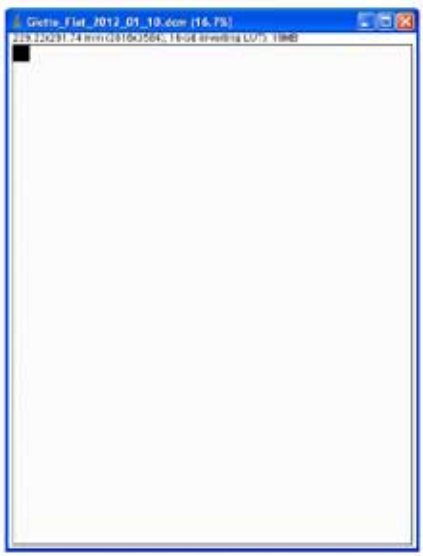
For CR systems, the search for artifacts should be performed for each imaging plate in clinical use.

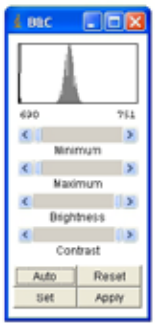
Artifacts

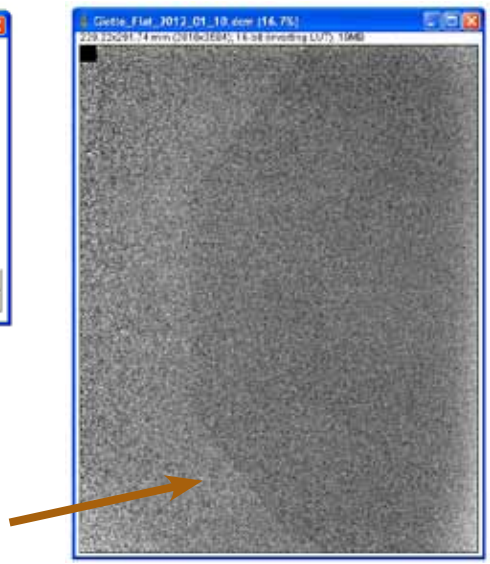
3 Artifact enhancement

1. Open the image obtained by the uniform PMMA block exposure and adjust WW and WL by the "B&C" tool.





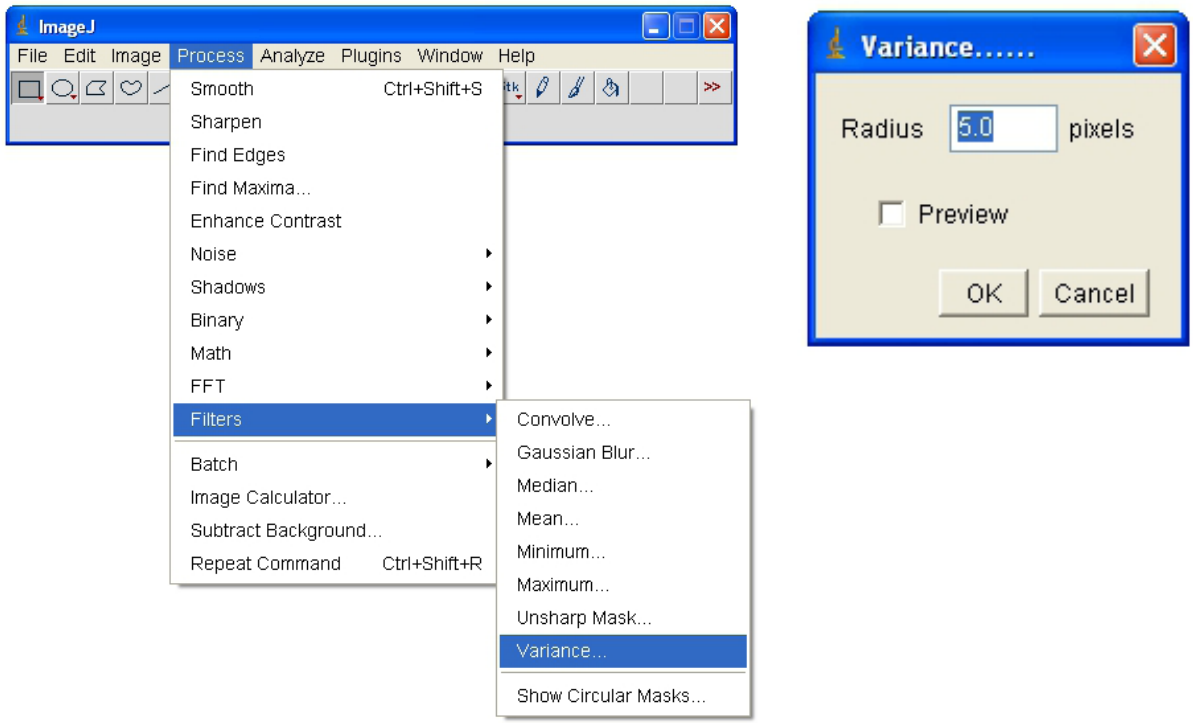




Notice that the automatic adjustment of WW and WL stresses the grey level histogram and might produce "false artifacts". Here the absolute differences between pixel values inside and outside the trapezoidal shape (compression paddle shadow) are within few units.

Artifacts

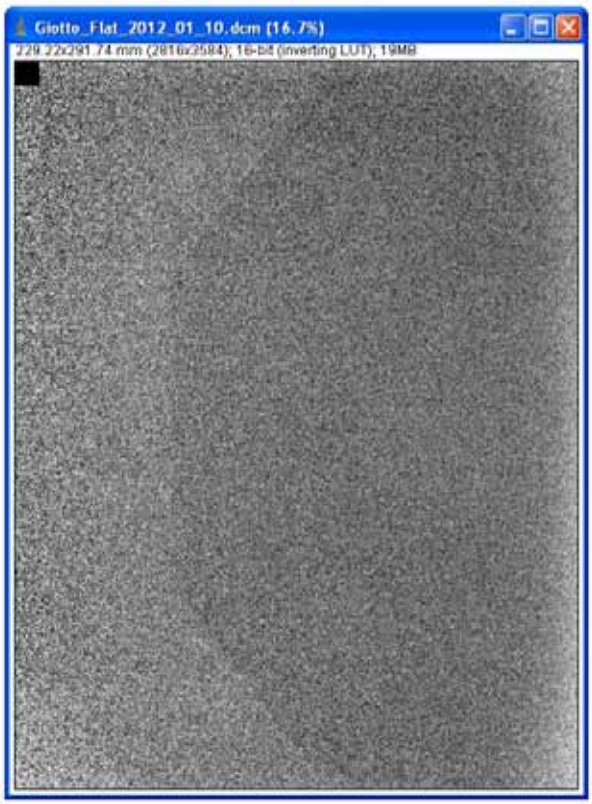
2. From the menu Process/Filters, choose the option "Variance…". This filter highlights edges by replacing each pixel with the neighborhood variance [5]. A radius of 5 pixels is suggested.



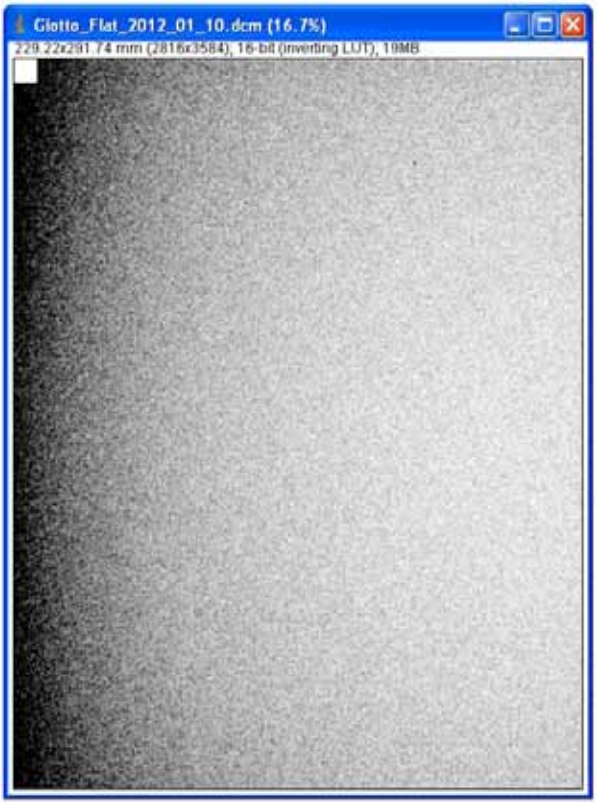
Artifacts

3. In the variance map, the “trapezoidal artifact” shown in the original image after WW/WL adjustment has disappeared.

Original image



Variance map

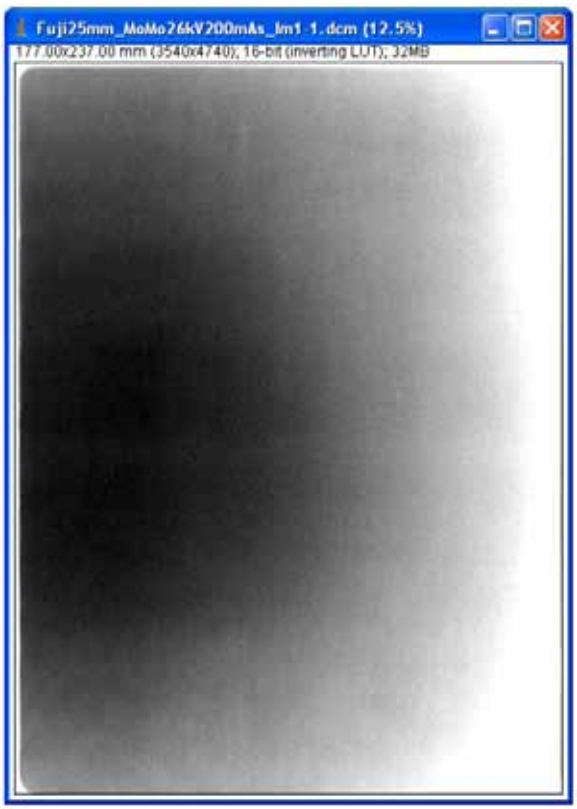


Artifacts

4. A few examples of real artifacts are provided, showing both the original image of the uniform test object and the corresponding variance map.

Example 1: CR system: scratches due to mechanical friction of rolls driving the imaging plate inside the CR scanner)

Original image



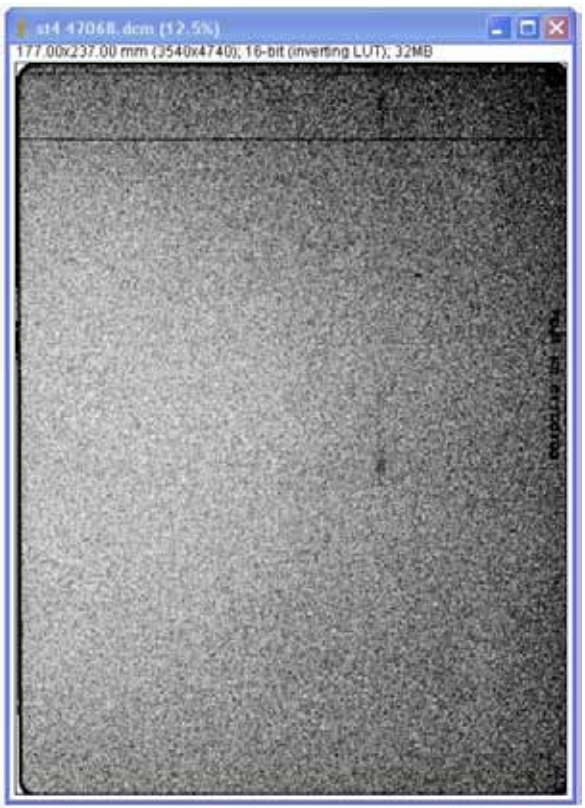
Variance map

Artifacts

Example 2: CR system: line due to the CR scanner problems. The same line was visible for all the imaging plates used with the same CR systems.

Original image

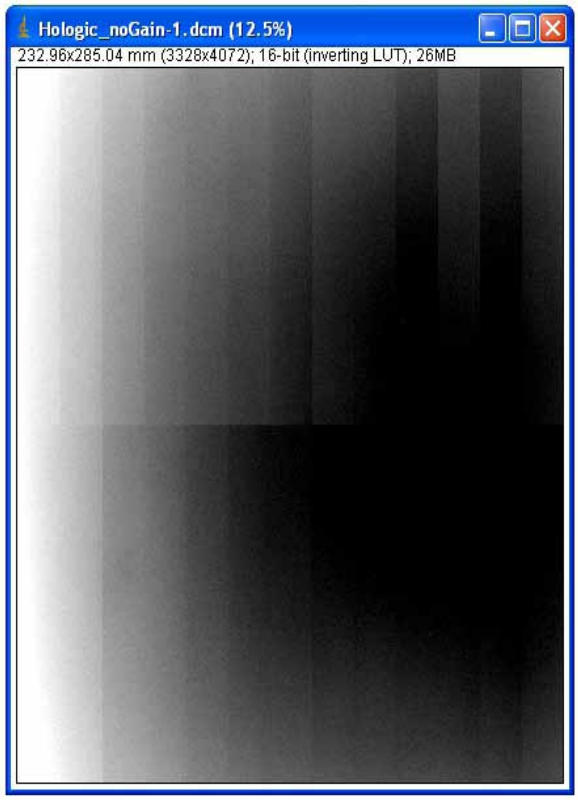
Variance map



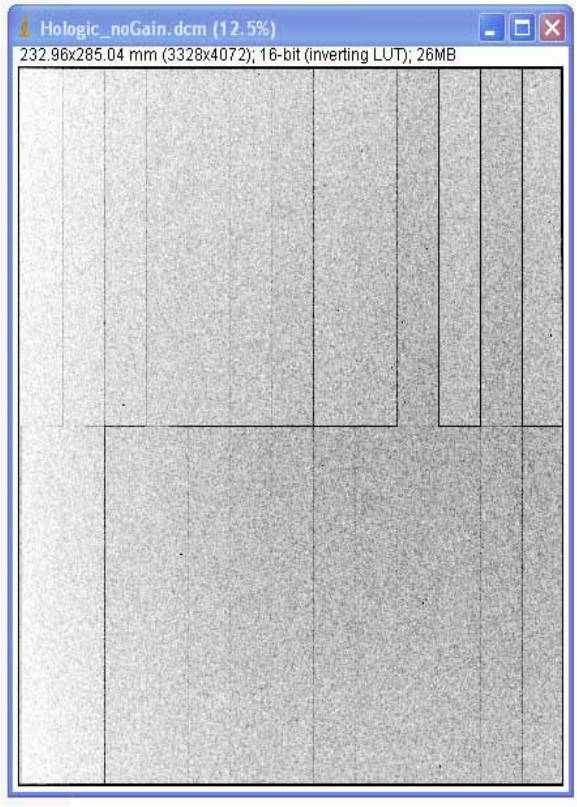
Artifacts

Example 3: DR system: uncorrected image. The image was acquired without corrections on purpose, but this mimics a case for which the gain correction map is missing.

Original image



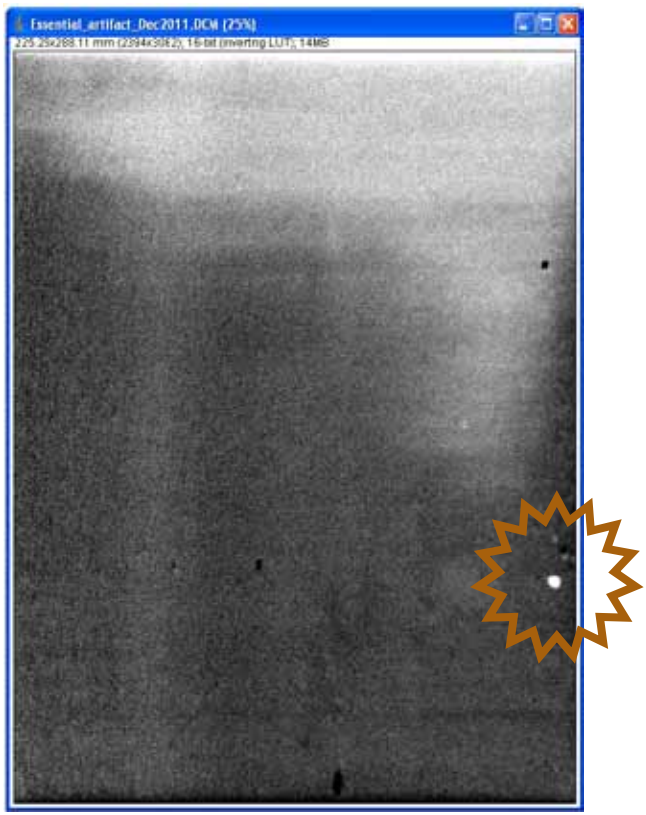
Variance map



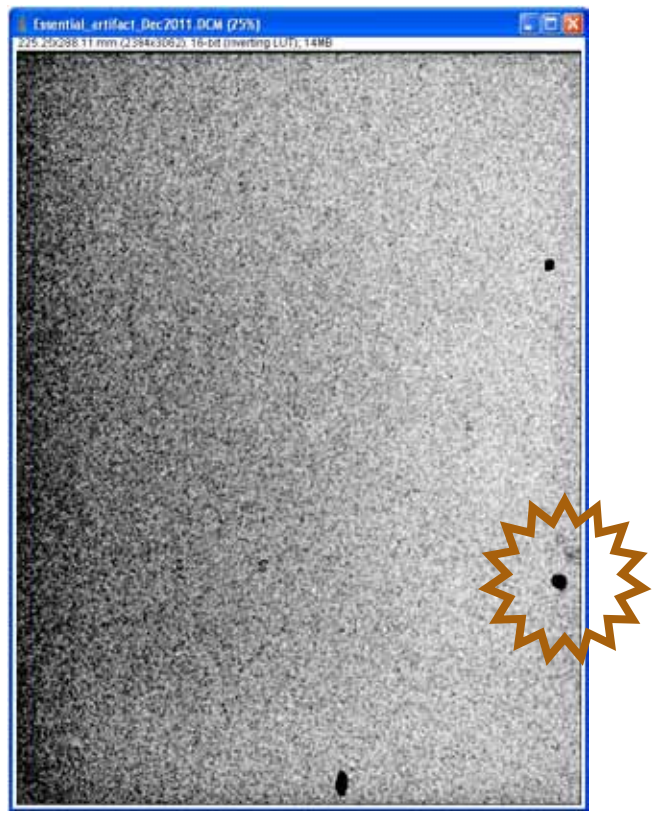
Artifacts

Example 4: DR system: artifact due to dust on the exit window of the x-ray tube. It was visible in clinical images when overlapped to fatty tissue. It caused some spot views, fortunately no biopsy.

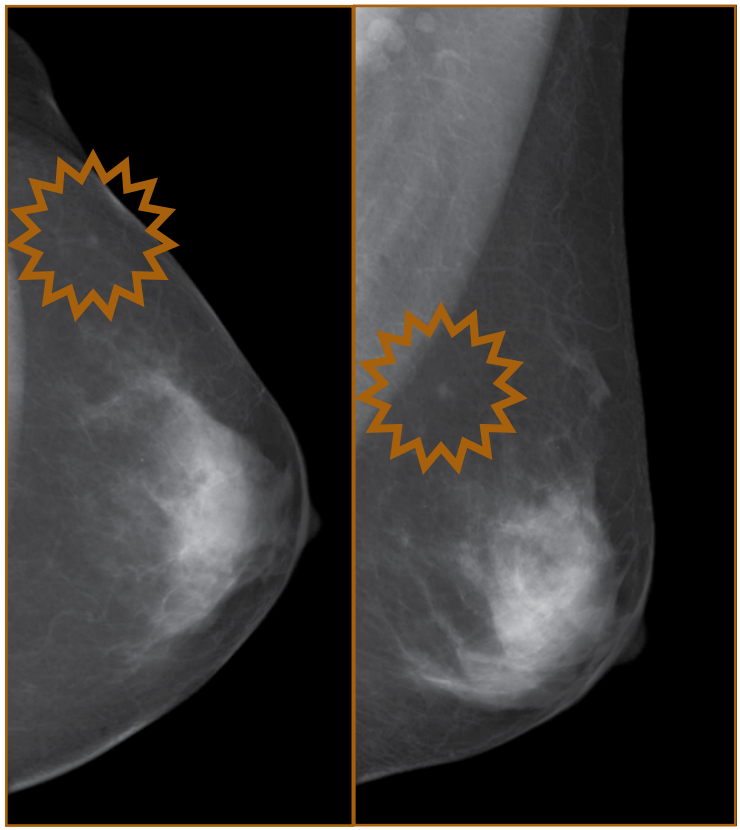
Original image



Variance map

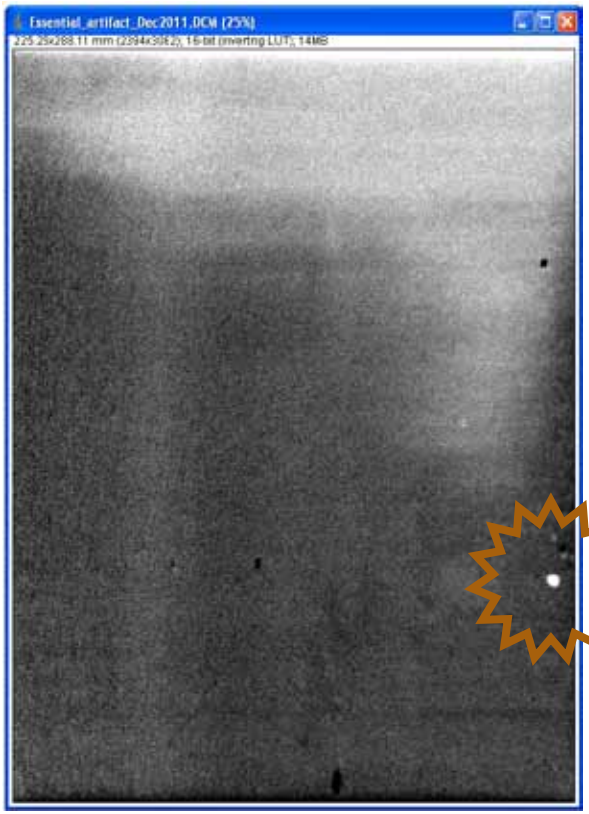


Artifacts



LCC LLMLO

Looks like a mass of size 5mm !!!



The distance from the chest wall was different in test images (Mo anode) versus clinical images (Rh anode), because of the different target angles and beam incidence.

Artifacts

4 Artifact recording

1. Insert data in Table 42 of the worksheet “Artifacts” in the “Template_EFOMP_MammoWG”.

ARTIFACTS			YES/NO
Fill data below only in case of artifact:			
Number:			1,2,3,>3
Type:	Defective row/column		Put an “X” for each artifact type. In case of “Other”, please specify
	Defective area		
	Defective stripe		
	Calcification cluster-like		
	Mass-like		
	Other		
Localization			Provide spatial localization (use the sketch in the Excel file)
Visible in			
Ghost			
Ghost cause			Image, Variance map or both

Table 42: Table for artifact recording. For CR systems, it should be multiplied by the number of imaging plates clinically used.

Artifacts

LIMITING VALUES

Basically, artifacts are not acceptable. However, as previously remarked, remedial actions in case of artifact detection can be more or less urgent, depending on the artifact itself. In general, subtle artifacts mimicking clinical features, potentially producing unnecessary biopsies could require the suspension of the equipment use and should be immediately removed.

Inter-plate variability (CR Only)

Inter-plate variability test applies to CR systems, and counts for possible differences among different imaging plates in sensitivity/response. In fact, as each IP is an independent image detector, all the tests required in this chapter should be repeated for all the IPs used with a given CR scanner. However, as this is difficult to do in practice, it is considered a good practice to use a single IP for all the QC tests, and verify that the differences among IP sensitivities are small.

In principle, this test should be done during commissioning, to ensure that all the IPs of the same "package" have similar sensitivities. However, if the QC protocol is applied for a CR system already in use, this test can be used to detect a group of IPs/cassettes with comparable sensitivities, which should be used together in clinical practice.

It is recommended to use no more than 5-6 cassettes per CR scanner, in order not to mix groups of cassettes belonging to different series, with possible significant differences in their response.

In case there is a need to speed up periodical QC tests for time constraints, the use of a group of IPs with comparable sensitivities is tolerated.

Inter-plate variability (CR Only)

PURPOSES

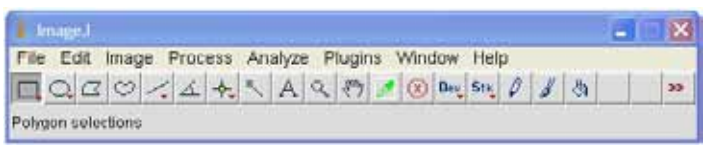
EVALUATE INTER-PLATE VARIABILITY IN ORDER TO:

1. **EXCLUDE IPs WITH MARKED DIFFERENCES IN SENSITIVITY COMPARED TO ALL THE OTHERS**
2. **SELECT GROUPS OF CASSETTES WITH COMPARABLE SENSITIVITIES**

EQUIPMENT

1 Poly-methyl methacrylate (PMMA) block: 45 mm

2 Software for image analysis: in the following, all instructions are provided for the freeware package ImageJ, but any other equivalent software can be used



TEST FREQUENCY

» Acceptance/Commissioning
» Annual
» In case new cassettes are put in use

Inter-plate variability (CR Only)

PROCEDURE

1 Geometry

1. Place the 45 mm PMMA block on the breast support
2. Lower the compression paddle and apply a minimum compression force by 3-5 daN, if this is required to enable the automatic exposure mode.

2 Exposure mode ad plate scanning

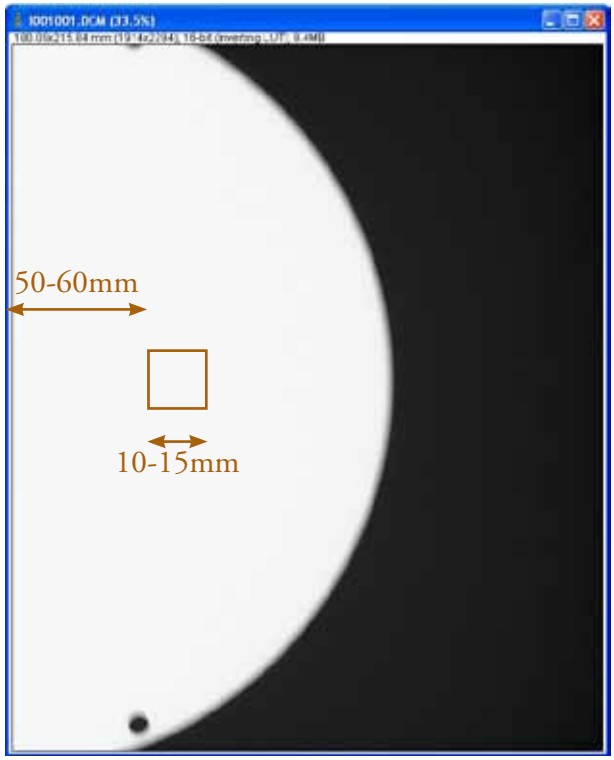
1. Record a new patient on the acquisition workstation.
2. Select the automatic exposure mode used in clinical practice from the mammography console, and position the AEC sensor close to the chest wall, verifying that it is covered by the test object. The position of the AEC sensor should be the same for the whole test.
3. Define the "Reference IP" (which ideally should be used for all the QC tests).
4. For each imaging plate clinically used, acquire an image of the PMMA block in automatic exposure mode.
5. Enter the technique factors in Table 41 (Sheet "IP_variability", excel file "Template_EFOMP_MammoWG_CR".
6. Insert the cassette in the CR scanner and read the imaging plate using the same test menu as used for the response function (Table).
7. Enter the exposure index (EI) in Table 41 (Sheet "IP_variability", excel file "Template_EFOMP_MammoWG_CR".

IT IS KNOWN THAT THERE IS A FADING OF THE LATENT IMAGE TRAPPED BY THE PHOTOSTIMULABLE PHOSPHORS WHICH CONSTITUTE THE IMAGING PLATE. FOR THIS REASON, TO PREVENT BIAS IN THE INTER-PLATE VARIABILITY TEST, TAKE CARE TO KEEP THE SAME TIME DELAY BETWEEN PLATE EXPOSURE AND READING FOR ALL THE DETECTORS USED.

Inter-plate variability (CR Only)

3 Imaging analysis

1. Open the image of the reference IP, draw a square ROI (side 10-15 mm) and center it on the phantom image at 50 mm from the chest wall, as shown below.



2. Measure the MPV and the SD using CRLT+M or the menu "Analyze/Measure" and insert the values in Table 41.
3. Use ROI Manager, as previously described to keep the ROI size unchanged for the other images
4. Repeat the same operation for the images associated to the other imaging plates.

Inter-plate variability (CR Only)

4 Data input

Test menu						
Imaging plate ID	Anode/Filter	kV_p	mAs	EI	MPV	SD

Table 43: Data input for the inter-plate variability test of CR systems.

5 Example of data input

Test menu			QC test / Ave2			
Imaging plate ID	Anode/Filter	kV_p	mAs	EI	MPV	SD
15562c (ref)	Mo/Rh	28	69	90	404.9	4.7
49178c	Mo/Rh	28	65	94	399.5	4.9
47068c	Mo/Rh	28	61	105	407.2	4.7
15851c	Mo/Rh	28	70	92	412.6	4.5
49215c	Mo/Rh	28	60	108	405.8	4.9

Table 44: Example of data input for the inter-plate variability test of a FCR Fuji Profect system (HR-BD plates).

Inter-plate variability (CR Only)

6 Inter-plate variability

1. Inter-plate variability is evaluated in terms of K_i, EI (for those systems using a test menu which adjusts the EI with the absorbed dose), and SNR.
2. For each of the three parameter, inter-plate variability (Δ) is calculated as the maximum deviation from the Reference IP, and as maximum deviation from the mean value across all the IPs, according to the following formulas

$$\Delta_{ref} = \max\left(\frac{|X_i - X_{ref}|}{X_{ref}}\right) \times 100$$

$$\Delta_{mean} = \max\left(\frac{|X_i - X_{mean}|}{X_{mean}}\right) \times 100$$

where X variable can be K_i, EI, and SNR, respectively.

3. Maximum deviations by 20% are admitted for Δ_{ref}, while 15% is proposed for Δ_{mean}.

7 Data output

	Imaging plate ID	Ki	EI	SNR
Reference IP				
	$\Delta_{reference}$			
	Δ_{mean}			

Table 45: Data output for the inter-plate variability test of CR systems.

Inter-plate variability (CR Only)

8

Example of data output

	Imaging plate ID	K_i	EI	SNR
	15562c (ref)		90	86.1
	49178c		94	81.5
	47068c		105	86.6
	15851c		92	91.7
	49215c		108	82.8
Reference IP	15562c			
	$\Delta_{reference}$			
	Δ_{mean}			

Table 45: Data output for the inter-plate variability test of CR systems.

LIMITING VALUES

Maximum deviation of EI and/or K_i compared to the reference IP: 20%
Maximum deviation of EI and/or K_i compared to the mean across IPs: 15%
Maximum deviation of SNR compared to the reference IP: 20%
Maximum deviation of SNR compared to the mean across IPs: 15%

REFERENCES

1. *N Perry et al. "European guidelines for quality assurance in breast cancer screening and diagnosis", 4th edition, 2006.*
2. *IAEA Human Health Series No. 17 "Quality Assurance Programme for Digital Mammography", Vienna, International Atomic Energy Agency (2011).*
3. *E G Christodoulou, M M Goodsitt, H-P Chan. Phototimer setup for CR imaging, Med Phys 27:2652-2658 (2000).*
4. *R. S. Ayyala, M. Chorlton, R. H. Behrman, P. J. Kornguth, P. J. Slanetz. Digital mammography artifacts on full-field systems: what are they and how do I fix them ? Radiographics 28:1999-2008 (2008).*
5. *N. W. Marshall. Retrospective analysis of a detector fault for a full field digital mammography system, Phys Med Biol 51:5655-5673 (2006).*

Image Quality

Image Quality

In this chapter the term image quality (IQ) will indicate the so-called technical image quality (TIQ), i.e. image quality parameters measured from images obtained by exposing reproducible and known objects, commonly designated as test objects (TO) or phantoms. Conversely, clinical image quality (CIQ) is usually evaluated within studies properly designed, including many clinical images and several radiologists, to count for wide variability in human perception and decision criteria. The relationship between TIQ and CIQ is complex and not easy to characterise. If it is reasonable to assume that improvements in technical image quality are likely to lead to improvements in clinical image quality, it is very difficult (virtually impossible) to establish an acceptable level of diagnostic image quality based on absolute or relative thresholds for TIQ parameters.

However, despite the known limitations of TIQ methodologies they can be useful. They provide a practical approach to characterise the image quality performance when commissioning a new piece of equipment and in routine assessment of equipment performance over time.

The evaluation of image quality using test objects and phantoms has three main objectives:

1. Verify that, under normal conditions of use of the mammography unit (e.g. at the dose level provided by the automatic exposure mode) the technical image quality is at a minimum acceptable level as defined per the applied QC protocol. This type of test is normally done during acceptance/commissioning of new equipment. Another level of IQ can be defined above the limiting value and used as a reference for monitoring the performance of systems that perform noticeably above the minimum level considered acceptable (Figure 1).

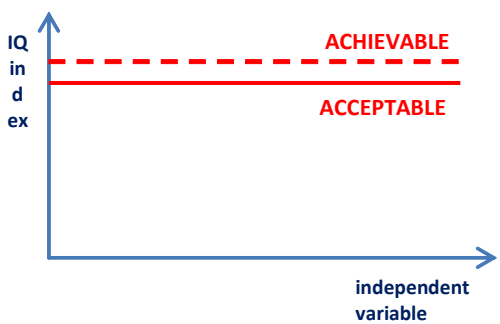
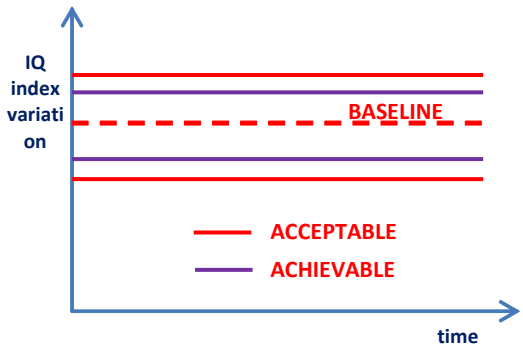


Image Quality

2. Define a baseline for one or more IQ indices and a "control interval" to be used as reference thereafter for long-term reproducibility tests. This is also done as acceptance/commissioning test. The typical acceptable control interval is calculated as baseline ± 3 standard deviations. A narrower range can be defined as "achievable control interval" to account for systems exhibiting outstanding reproducibility (Figure 2).



3. Monitor the consistency of the system´s response over time by evaluating time variations of one or more IQ indices against their baseline and associated control intervals. Those tests are often reported as long-term reproducibility tests or "routine" tests.

Image Quality

It should be noted that a number of assumptions are made in the assessment of image quality using test objects and phantoms as described above:

- A relationship exists between TIQ and CIQ and the acceptable level of phantom-based IQ is a fundamental step for clinical acceptability.
- All phantoms commercially available of a same type are assumed "equal", i.e. if image quality for the same equipment would be assessed by two phantoms of the same type, the results obtained should be very similar.
- The parameter used to assess absolute image quality can be used for "optimization" purpose, i.e. radiation dose can be "tuned" to move the IQ level closer to the acceptable/achievable lines, and the AEC behavior can be modified accordingly.

All those assumptions are questionable. As mentioned above, ensuring that image quality is above a given threshold using phantoms, cannot guarantee appropriate clinical performance. It should be acknowledged that any phantom is a simplified model of clinical reality, and the information provided by phantoms should not be used to predict performance of the imaging system in clinical practice. In Table 1 a list of characteristics of TIQ and CIQ are shown, which aims to illustrate the difficulties in using image quality based on phantom images to derive direct conclusions about clinical performance.

Image Quality

Table 1: Differences between technical image quality (phantom based) and clinical image quality.

Technical Image Quality (phantoms)	Clinical Image Quality (breasts)
Most phantoms are simplified models: provide images with uniform background displaying details in know locations for simple measurements. A few more complex phantoms mimic anatomical tissue however they still fail to reproduce the natural variability in human anatomy.	Wide variability of clinical images: different breast patterns and lesion types/characteristics/difficulties in both detection and characterization.
Physical evaluation, based on a priori knowledge of phantom contents: measurable parameters (like SDNR, MTF or many others indices) or assessement based on "visibility" of details. The mechanism of detecting an expected object whose spatial localization is known is significantly different than clinical detection.	Clinical evaluation: includes detection and characterization steps. Detection is necessary to localize a clinical feature; characterization is the decision made after the interpretation of clinical features. There is no a priori knowledge in routine clinical evaluation, and this change detection performance.
Detectability of image features in a uniform background is strongly correlated with radiation dose. Structured backgrounds, such as that generated by human anatomy due to the superimposition of normal tissue, pose a significantly different challenge and one that cannot be eliminated by increasing dose.	Detectability is mainly restricted by normal tissue superimposition, which produces structured background. Such "anatomical noise" cannot be reduced by adjusting dose. This is the main reason why "subtraction techniques" (tomosynthesis, dual energy, etc) have been introduced.
Technical image quality is normally evaluated by unprocessed images, the processing effect not being related to the physics of image formation.	Clinical image quality is normally evaluated on processed images. Radiologists do not use unprocessed images, which are usually not archived in the PACS.
Possible optimization methods are applied using unprocessed images, under the general principle that acceptable/achievable TIQ should be obtained while keeping radiation dose as low as possible.	Clinical image quality is normally evaluated using processed images. Post-processing algorithms might nullify differences between unprocessed images obtained at slightly different dose levels as determined by the optimization process.

Image Quality

For all those reasons, it is suggested that the limiting values for TIQ be set according to widely accepted, very general criteria, like non-inferiority versus previous technology (screen-film mammography). It should be remembered that the purpose of acceptance/commissioning tests is to reject only clearly problematic systems. Setting the threshold of acceptability close to the normal operation of mammography systems might lead to unwanted "border line" situations, for which a given system passes or fails the test depending on the number of images used for the TIQ evaluation or on the specific phantom used (i.e. with one phantom the systems passes the test, with another phantom of the same type, it fails the test), or on the method applied to process information obtained by phantom images.

Non-inferiority

The term *non-inferiority* is normally used to statistically prove the "equivalence" between two systems. In clinical performance studies, non-inferiority is the **minimum set requirement** before accepting a new technology. **The underlying principle is that, given the statistical nature of any type of measurement, and of clinical performance measurements in particular, the comparison between two "equivalent" systems will very unlikely produce exactly** the same results; therefore, equivalence should be considered as a difference small enough to be considered negligible or, even better, a difference that is significantly smaller than an arbitrarily set tolerance limit, called *non-inferiority margin* [1].

The same concept of non-inferiority is also applicable to physics measurements, including image quality evaluation phantom-based, assuming that the related statistics is well known. In other words, the non-inferiority margin should be statistically well grounded, considering any possible source of variability and what is actually achievable, given the state of the art of a given technology. Unfortunately, in the few cases image quality phantoms are provided with limiting values, the method used to define limiting values and related non-inferiority margins is not supplied.

Image Quality

The purpose of this document is not to recommend the use of specific test objects or phantoms to assess IQ but to provide a general guidance on the use of these types of tools for assessment of IQ. There are wide differences between phantoms regarding their contents and consequently possible parameters/attributes to evaluate IQ from phantom images. Scores can be assigned by human observers according to "visibility" of the details embedded in the phantom, but software packages can be implemented reproducing the detection process, with the benefit of eliminating intra- and inter-observer variability. Quantitative IQ parameters can also be measured from phantom images, by means of more or less automatic software tools.

It should be noticed that national regulatory requirements may vary across countries as well as the resources available and those differences should be taken into considerations in the planning and implementation of IQ evaluations within quality assurance programs.

In the first part of the chapter a list of some of the most common phantoms and test objects commercially available is provided with a brief description of each one and the recommended method of scoring. The list presented does not include all the existing phantoms and does not aim to be exhaustive. It aims to support users in the selection and use of a suitable tool for quality controls and users should be aware of other test objects and phantoms as well as ones under development.

The second part of this chapter describes various established methods of image quality assessment. The methods presented are classified as "subjective" (or "human-based" because they include an element of human perception and decision criteria) or "objective". Examples are provided of the most common human-based methods to score phantom images and of some objective IQ indices for some typical categories of details contained in several phantoms.

Phantoms

Phantoms for IQ evaluation "replace" the anatomical part, the breast in this case, and, as such, should have comparable attenuation properties and scatter behavior as a typical breast. Almost all test objects used in 2D imaging have uniform background, and include one or more physical objects to produce features in the image that challenge the performance of the mammography system and provide images suitable for assessment of image quality.

In that sense a uniform test object containing no details is not appropriate to be used as IQ phantom, while it can be useful to investigate the presence or image artifacts that may have impact on image quality (as discussed in the "Detector" chapter). In the following, only phantoms including one or more details producing signal variations on the final image will be considered for IQ evaluation.

Commercial IQ phantoms used in 2D mammography are usually made of rigid plastic materials. The test object thickness typically varies between 40-50 mm to mimic the attenuation characteristics equivalent to 45-60 mm breast tissue with average glandularity [2]. The phantom details can either reproduce clinically relevant features found by radiologists in clinical images, like small masses, spicules or microcalcifications, or include specific objects, like spatial resolution patterns or step wedges, to allow measurements of physical parameters. The composition and size of the physical objects inserted in the phantom to mimic structures of clinical interest are normally provided by the phantom manufacturer. The phantom details can be either confined in a plane (thin slab), or distributed over the phantom volume. When the phantom objects are all contained in a same slab it may be possible to produce phantom with varying thickness by stacking on to the test object uniform plastic slabs to produce a range of intended thickness.

Phantoms

IQ phantoms are usually provided with a user manual that contains information about the intended use of the tool, technical specifications and recommended criteria for image quality evaluation. A general assumption by phantom manufacturers is that all phantoms of the same type have the same characteristics and are "equivalent". However, the current phantom manufacturing processes are not very sophisticated and this is expected to cause variability between items. When acquiring a phantom it may be useful making effort to gather information about the phantom reproducibility. This would be useful especially before using a test object for "absolute" image quality evaluation.

In this chapter seven different types of phantoms are presented together with information about their technical specification, radiographic image appearance, and proposed methodology for image quality scoring method where provided by the manufacturer. As previously stated, such a list is not intended to be exhaustive, but aims to cover a wide range of different phantom types and suggest different possible methods to assess image quality. Medical physics experts should adapt the procedure for IQ evaluation to the specific requirements of their country and their local needs/resources.

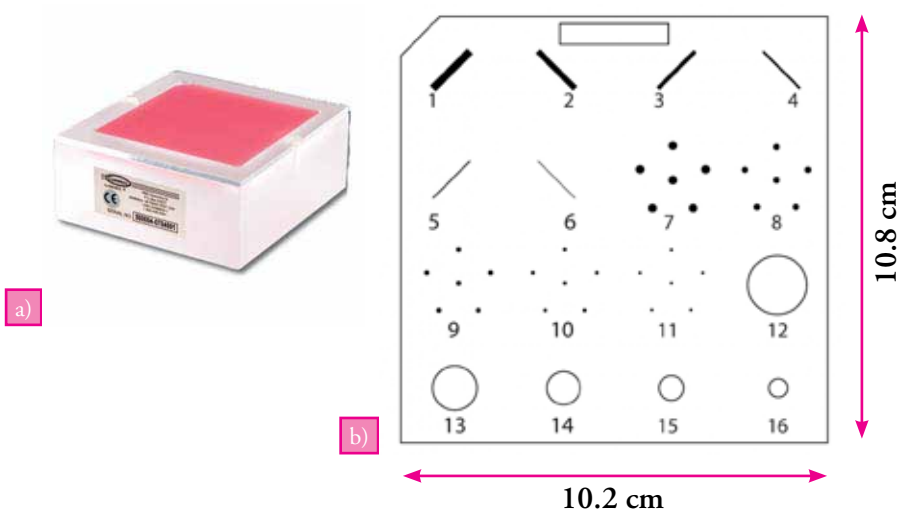
Phantoms

ACR Mammographic Accreditation Phantom

The American College of Radiology (ACR) phantom is the IQ assessment tool proposed by the ACR for the accreditation of mammography units. It is frequently identified as the CIRS model 015 or the RMI/Gammex 156.

It consists of a wax box, containing 16 details, placed on an acrylic base and with a plastic cover. All of this together approximates a 42 mm compressed breast. Six different size nylon fibers simulate fibrous structures, five groups of specks simulate different size calcification clusters, and five different size spheres simulate tumor-like masses. In Figure 3 a picture of the ACR phantom is shown, together with the location of the phantom details in the wax insert and a table with the description of the detail properties.

Figure 3: ACR Mammographic Accreditation Phantom.
(a) Phantom picture;
(b) Phantom sketch;
(c) Description of the phantom details.



Phantoms

Nylon Fibers		Groups of specks		Masses	
1	ϕ = 1.56 mm	7	0.54 mm	12	ϕ = 2.00 mm
2	ϕ = 1.12 mm	8	0.40 mm	13	ϕ = 1.00 mm
3	ϕ = 0.89 mm	9	0.30 mm	14	ϕ = 0.75 mm
4	ϕ = 0.75 mm	10	0.24 mm	15	ϕ = 0.50 mm
5	ϕ = 0.54 mm	11	0.16 mm	16	ϕ = 0.25 mm
6	ϕ = 0.40 mm	-	-	-	

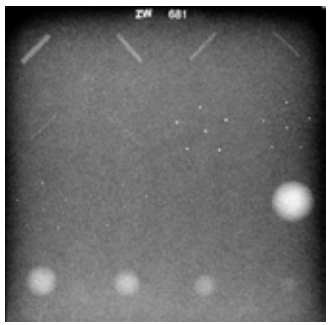
c)

The use of an ACR accredited phantom is mandatory in the US where all mammography facilities need to be accredited by an authorized accreditation body. To receive accredited status a facility needs to fulfill established criteria, which include the submission of an ACR phantom image of acceptable quality produced with the mammography equipment. To maintain accreditation an image of the ACR phantom must be imaged weekly and records of these images maintained as evidence.

A radiographic image of the ACR phantom obtained by a digital mammography system in automatic exposure mode is shown below (Figure 4) together with the criteria applied to rate visibility of each type of detail from phantom images. The total score must be at least 4 for fibers, 3 for groups of specks, and 3 for masses, leading to a minimum grand total minimum score of 10. Image quality assessment is supposed to be performed by trained persons. Further details about the rating method will be provided in the second part of this chapter.

Phantoms

a)



b)

Detail type	Visibility criteria	Limiting value
Fibers	• Full=1 • Partial=0.5 • None=0	*4*
Calcs	• Five calcs=1 • 1 < calcs < 5=0.5 • None=0	*3 (groups)*
Masses	• Full=1 • Partial=0.5 • None=0	*3*

Figure 4:
(a) X-ray image of the ACR phantom showing how the details appear;
(b) Table illustrating standard visibility criteria used to evaluate ACR phantom images, and limiting values for accreditation in the U.S.

The total minimum acceptable score of 4 fibers + 3 groups of specks + 3 masses = 10 details is derived from screen-film mammography and its validity confirmed for digital mammography under the general concept of non-inferiority. The scoring method is human-based, even if there is some research showing potential advantages of automation of image analysis and scoring, applying the same limiting values [3-4]. Some manufacturers have noticed that phantom scoring for DR systems provides noticeable higher scores and suggest the use of higher threshold values accordingly (this is allowed by the ACR). Work is in progress to "update" the ACR phantom to improve its suitability to the characteristics of digital mammography systems.

Phantoms

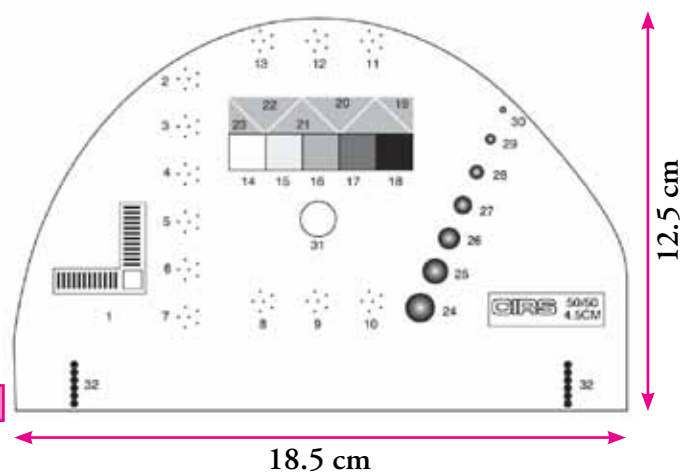
CIRS Phantom, Model 011A

The CIRS Model 011A is breast shaped phantom, made of an epoxy resin equivalent to a 45 mm breast (mixture 50% glandular – 50% fatty), plus a 5 mm shell simulating the skin, for a total 50 mm thickness. Groups of objects within the phantom mimic clinical details. Six nylon fibers simulate fibrous structures of different sizes in the breast, five groups of specks mimic clusters of microcalcifications of various sizes and five spheres mimic low contrast of various diameters.

A picture of the CIRS Model 011A is shown (Figure 5) together with a schematic of the location of the phantom details. The detail properties are described in the table.

Figure 5: CIRS Model 11 A Phantom.
(a) Phantom picture;
(b) Phantom sketch;
(c) Description of the phantom details.

a)



b)

Phantoms

Line pair target	$CaCO_3$ specs grain size (mm)	Step Wedge 1 cm thick	Nylon Fibers diameter size (mm)	Hemispheric Masses 75% glandular/ 25% adipose Thickness (mm)	Optical Density 31. Reference zone Edge of Beam 32. Localization target
1. 20 lp/mm	2. 0.130	14. 100% gland	19. 1.25	24. 4.76	
	3. 0.165	15. 70% gland	20. 0.83	25. 3.16	
	4. 0.196	16. 50% gland	21. 0.71	26. 2.38	
	5. 0.230	17. 30% gland	22. 0.53	27. 1.98	
	6. 0.275	18. 0% gland	23. 0.30	28. 1.59	
	7. 0.400			29. 1.19	
	8. 0.230			30. 0.90	
	9. 0.196				
	10. 0.165				
	11. 0.230				
	12. 0.196				
	13. 0.165				

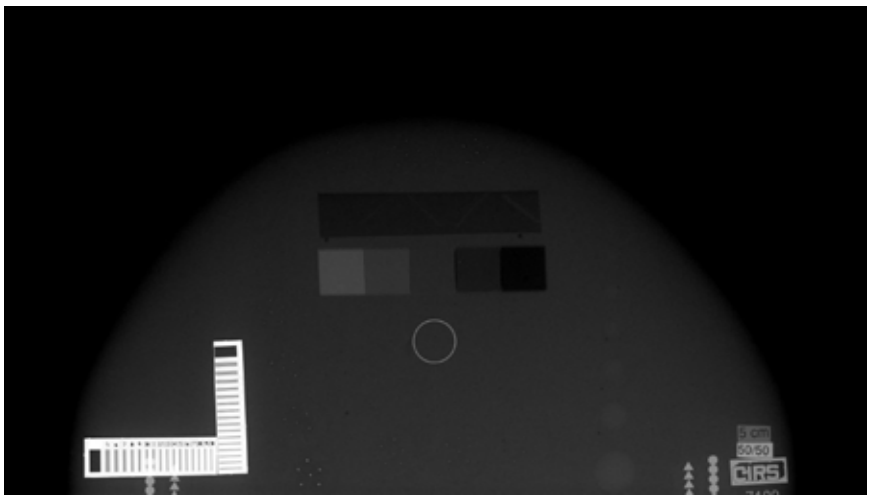
c)

Phantoms

CIRS Model 011A contains objects similar to those contained in the ACR phantom to simulate the clinical features but more of them. This gives a larger number of steps levels for the detail sizes so potentially improving the scoring sensitivity. In addition, it includes other details, like two orthogonal bar patterns, a step wedge and a reference area for quantitative measurements. It is also a phantom originally designed for screen-film mammography, for which scoring was supposed to be done on the basis of human assessment [5]. The manufacturer does not provide an indication of acceptable scores or limiting values. The general principle of non-inferiority to screen-film can be applied.

A radiographic image of the CIRS phantom obtained by a digital mammography system in automatic exposure mode is shown in Figure 6.

Figure 6: X-ray image of the CIRS Model 11 A Phantom showing how the details actually appear.



Phantoms

Leeds TORMAS / TORMAX Phantoms

Both Leeds TORMAS and TORMAX test objects have been designed for screen-film mammography. They are provided with variable thickness attenuator stack (poly-methyl methacrylate, 5-70 mm). The phantoms themselves include a ten-step step wedge, one (TORMAS) or two (TORMAX) high-contrast resolution patterns (1.0 to 20.0 lp/mm), a low-contrast resolution pattern, twelve low-contrast large details (5.6 mm diameter, decreasing thickness), 22 high-contrast small details (0.5 mm and 0.25 mm diameter, decreasing thickness, 11 details per each diameter), microparticles representing microcalcifications on a step wedge (three different grain sizes). A picture of the phantom and the acrylic stack is shown in Figure 7, together with a table including a description of physical properties of the details included in those phantoms.

Figure 7: Leeds TORMAS and TORMAX phantoms.
(a) Phantom picture;
(b) acrylic stack allowing modulation of total phantom thickness;
(c) Detail physical properties.

Phantoms

Type of detail	Physical properties
High contrast resolution gratings	Spatial frequencies (lp/mm): 1.0, 1.12, 1.25, 1.4, 1.6, 1.8, 2.0, 2.24, 2.5, 2.8, 3.15, 3.55, 4.0, 4.5, 5.0, 5.6, 6.3, 7.1, 8.0, 8.9, 10.0, 11.1, 12.5, 14.3, 16.6, 20.0
Low contrast circular details (12 details)	Size: 5.6 mm (diameter) Relative contrast (%) for Mo/Mo 28 kVp: 8.3, 5.6, 3.9, 2.8, 2.0, 1.4, 1.0, 0.7, 0.5, 0.35, 0.25, 0.15
High contrast circular details (11 + 11 details)	Size: 0.5 mm and 0.25 mm(diameter) Relative contrast (%) for Mo/Mo 28 kVp: 41%, 30%, 21% 16%, 11%, 8.3%, 5.6, 3.9, 2.8, 2.0, 1.4
Grey scale step wedge (10 steps)	Size: 8 mm square Relative contrast: to cover a typical film characteristic curve
Micro-particle step wedge (5 + 5 +5 steps, different grain size)	Grade 1: median size: 125 µm, max size: 212 µm Grade 2: median size: 234 µm, max size: 425 µm Grade 3: median size: 328 µm, max size: 500 µm

c)

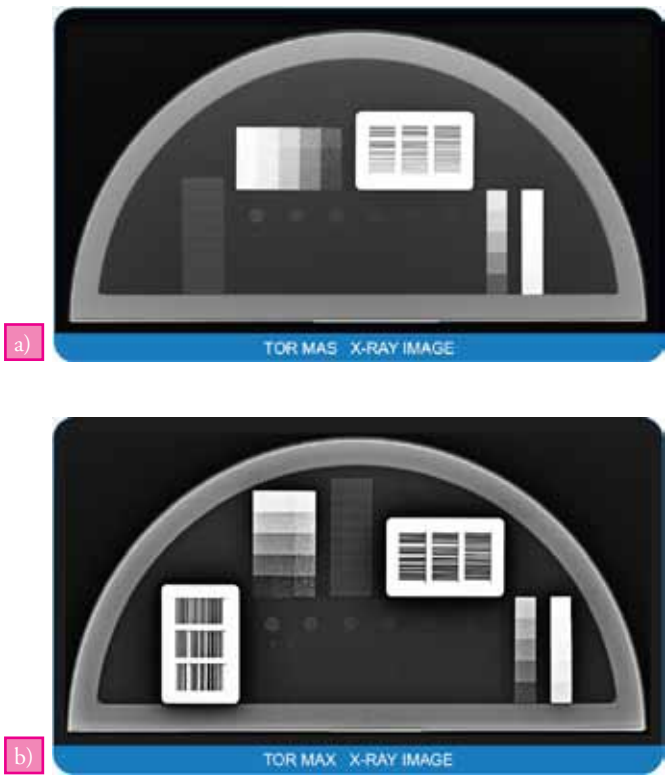
For routine tests, it is suggested to put the TORMAS/X on the top of 35 mm PMMA stack to obtain an object 45 mm thick. However, if a thinner or thicker test object is needed so as to be more representative of the breast population under study, the overall thickness can be adjusted accordingly.

Image quality assessment is supposed to be human-based applying "visibility" criteria similar to those used for other phantoms. No recommended scores or limiting values are provided by the manufacturer for those phantoms, but the general principle of non-inferiority versus screen-film can always be applied.

A radiographic image of both TORMAS and TORMAX is shown in Figure 8.

Phantoms

Figure 8: X-ray images of the Leeds TORMAS (a) and TORMAX (b) phantoms. The two test objects differ for the presence of an extra high contrast resolution pattern in the TORMAX. Some details are also differently distributed in the two phantoms.



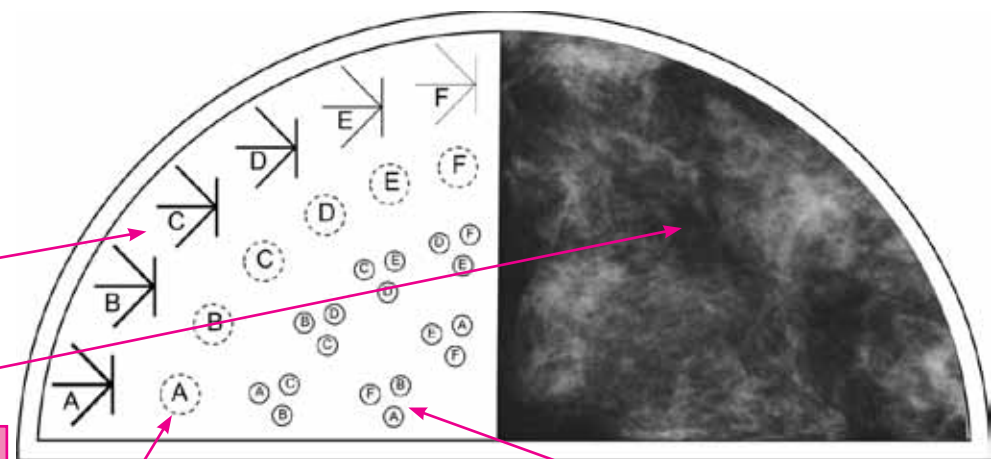
A commercial software package for automatic analysis of images produced by TORMAS/X phantoms exists; it provides different types of image quality indices, both "visible" detail counts (similar to human IQ assessment) and objective measurements [6]. More detailed information about the software can be obtained from the web site **http://autopia.cyberqual.it/index.php/AutoPIA**.

Phantoms

Leeds TORMAM Phantom

TORMAM is a test object designed for quick and easy use on a routine basis to provide an ongoing check of imaging performance. The regular record of the TORMAM scores allows users to notice possible deterioration in imaging performance. TORMAM is recommended by the manufacturer as supplementary to TORMAS or MAX and provides a more "realistic" breast image which may be preferred by radiographers and radiologists. TORMAM is composed of 2 parts (Figure 9). The left half contains a range of 6 groups of multi-directional filaments, 6 groups of micro-calcification in the range 300-100 µm and 6 groups of 3 low-contrast detail subgroups , that aim to mimic clinical features. The limiting visibility of the details depends on the image dynamic range contrast, spatial resolution and noise properties of the image, and can be used to obtain an image-quality "score" which can be used as an IQ attribute of performance of the mammography system. The right half mimics the appearance of breast tissue and contains micro-calcification clusters in addition to fibrous and nodular details.

Figure 9: Leeds TORMAM phantom.
(a) Sketch of the left part with countable features and X-ray image of the right, anthropomorphic part, including microcalcification clusters;
(b) Table describing physical properties of the details.



Phantoms

	A	B	C	D	E	F
Filaments (length 10 mm) Diameters [mm]	**0.40**	**0.35**	**0.30**	**0.25**	**0.225**	**0.20**
Particles Size ranges [μm]	354-224	283-180	226-150	177-106	141-90	106-63
Circular details (diam 3 mm) Nominal contrasts (%)	**4.0**	**3.0 (?)**	**2.0**	**1.5**	**1.0**	**0.5**

b)

TORMAM is provided with a variable thickness attenuator stack (poly-methyl methacrylate, 5-70 mm). For routine testing, the manufacturer suggests imaging TORMAM on top of 35 mm PMMA stack to obtain an object 50 mm thick (TORMAM plate with the details is thicker than TORMAS/X plates). However, if a thinner or thicker test object is needed so as to be more representative of the breast population, the overall thickness can be adjusted accordingly by adding or removing PMMA layers.

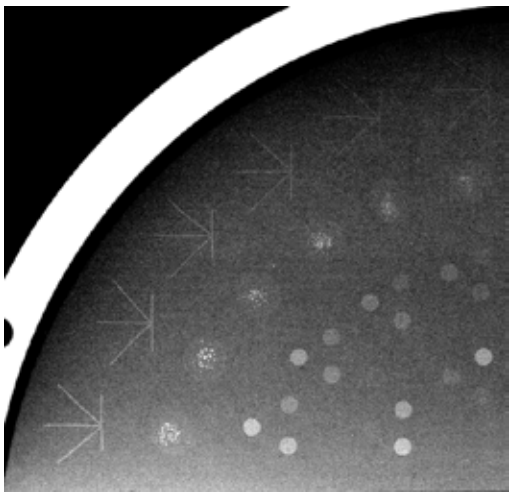
The manufacturer does not recommend exposure settings to produce the image with TORMAM. Recommendations of experienced users are to produce an image using exposure settings comparable to those used in the clinical practice and to acquire images in clinical acquisition mode (e.g. AEC exposure) so image quality will exhibit the effects of image processing.

Image quality assessment is supposed to be human-based applying "visibility" criteria similar to those used for other phantoms. For example, the following criteria can be applied: a 0-3 scale of visibility can be used to assign IQ score to each given detail [7].

A high dose X-ray image of TORMAM is depicted in Figure 10, together with the assessment scale for scoring detail visibility.

Phantoms

Figure 10: Radiographic image of a Leeds TORMAM phantom acquired at high dose level to show all the details, and assessment scale for scoring detail visibility.



Detail appearance	Detail score
Detail seen unambiguously with clear edges	3
Detail seen unambiguously but edges not entirely seen	2
Detail possibly seen but may have been artifact	1
Detail not seen	0

Since TORMAM was originally designed as reproducibility test object, there is no established threshold score for this phantom, but the general principle of non-inferiority (e.g. compared to screen-film or previous image quality evaluations) can be applied.

Phantoms

Artinis CDMAM Phantom

The CDMAM Phantom is manufactured by Artinis (Nijmegen, The Netherlands) and is required for evaluating technical image quality as described in the European guidelines for quality assurance in breast cancer screening and diagnosis, fourth edition since 2006. The description provided hereafter refers to CDMAM 3.4, which was modified from previous version 3.2 to make it more suitable for digital mammography [8].

This phantom (162.5 x 240 mm^2) consists of an aluminum base 0.05 mm thick with gold discs (99.99% pure gold) of varying thicknesses and diameters, which is attached to a Plexiglas® cover. The assembly (PMMA and aluminum) has a Plexiglas® equivalent thickness of 10 mm, under standard mammography-exposure conditions. A matrix of square cells with golden disks whose diameters range from 0.06 to 2.0 mm and whose contrast is varied by their different thickness of gold, ranging from 0.03 to 2 µm, is attached to the base by means of evaporation process. In each of the 205 cells of the matrix, one dot is always placed at the center of the cell and another one, identical, is positioned in a randomly selected corner within the cells.

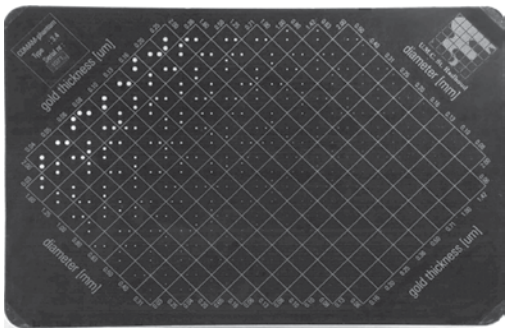
CDMAM v. 3.2 previously developed for screen-film mammography had the same structure, but the gold disks ranged in diameter from 0.1 to 3.2 mm and in thickness from 0.05 to 1.6 µm.

The CDMAM is provided with four PMMA slabs (10 mm per each), and is recommended to be exposed sandwiched between 20 + 20 mm PMMA, as shown in Figure 11.

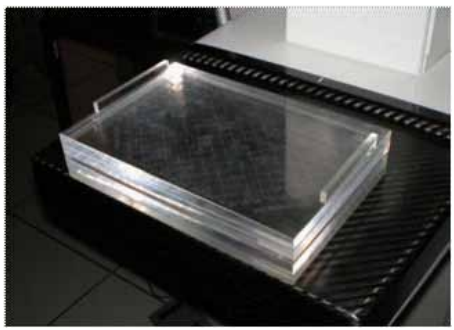
Phantoms

Figure 11: Artinis CDMAM 3.4 phantom.
(a) Picture of the test object;
(b) picture to show how the CDMAM should be exposed, ensandwitched between 20 + 20 mm PMMA.

a)



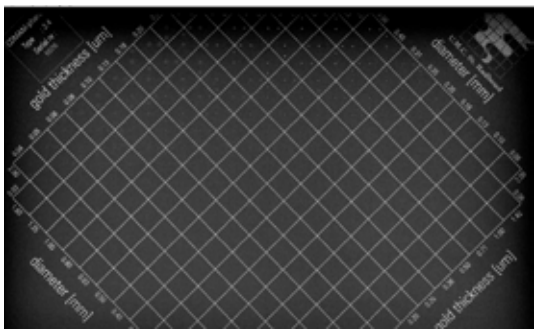
b)



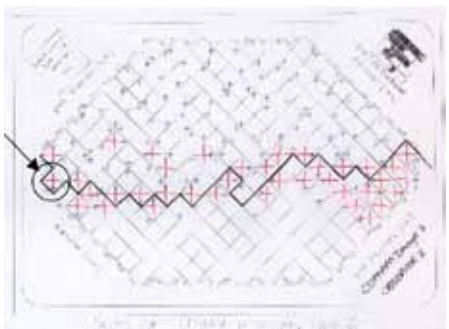
The X-ray image of the CDMAM shows 16 rows and 16 columns forming squares, in which two disks are projected, one in the center and one in one of the four corners (Figure 12 a). The image evaluation is based on indicating for each square in what corner the second disk is visible (Figure 12 b). The CDMAM phantom is a special kind of contrast detail test object. Further considerations on the general philosophy behind contrast-detail phantoms will be provided in the section on image evaluation. For the moment, it is enough to say that, after comparing image evaluation done by one or more observers with the truth (real positions of gold disks are known), a "threshold curve" can be obtained, named "contrast-detail" curve, showing, for each contrast in the phantom, the smallest detail correctly localized.

Phantoms

Figure 12:
(a) X-ray image of an Artinis CDMAM 3.4 phantom;
(b) scoring sheet used for human-based IQ evaluation of CDMAM images.



a)



b)

Software packages for automatic analysis of CDMAM images and calculation of threshold contrast for different disk diameters while removing inter-observer variability have been developed (CDCOM by Artinis http://www.euref.org/downloads/software-physico-technical-protocol/cdmam-readout; Erica2 by University of Leuven http://cdmamanalysis.com/) [9].

A more detailed discussion on methods to derive contrast-detail curves will be provided in the section on image evaluation. In the EUREF protocol, limiting values are provided for human readout [9]:

Au disk diameter (mm)	Acceptable thickness (µm)	Achievable thickness (µm)
0.1	1.68	1.10
0.25	0.35	0.24
0.5	0.15	0.10
1.0	0.09	0.06
2.0	0.07	0.04

Phantoms

Quart MAM/DIGI phantom

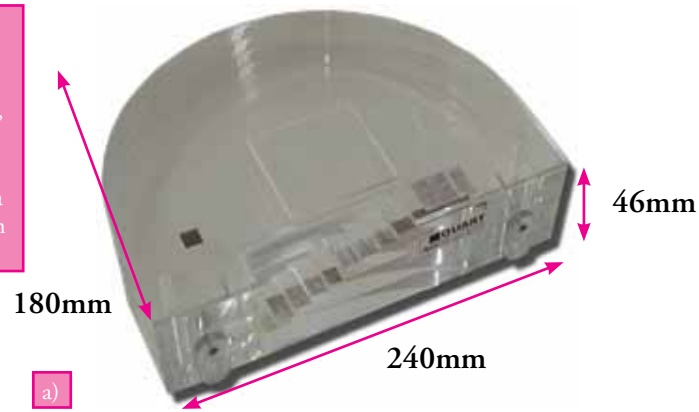
TThe Quart MAM/DIGI phantom is a semicircular, 46 mm-block of PMMA, which includes multiple details, some of them designed for observer-based evaluation (subjective) and others for objective measurements of image quality. Some details are "in-plane", others are distributed at different planes along a step-wedge.

The step-wedge consists of 12 steps dug in the PMMA block (5 mm intervals), with the last three steps including additional aluminum layers to cover breast equivalent thickness range between 5 and 82.5 mm.

A uniform titanium strip below half of each step produces an area (at different signal levels) for signal-difference-to-noise ratio (SDNR) measurements. On the other half of each step a group of six "Landolt rings", i.e. C-shaped rings with a small gap in one of the four perpendicular directions, constitute a new type of contrast-detail test object. In fact, object size is varied by changing the gap size, while contrast is varied by placing equal ring groups on each step of the step-wedge. Detection capability of Landolt rings at different contrast levels is evaluated by human assessment.

Other details in the MAM/DIGI phantom are a brass square (1 cm^2) to allow modulation transfer function (MTF) measurement, and 2 mm steel balls close to the chest wall, used for phantom alignment and for testing the possible missed tissue on the chest wall [11]. A picture of the phantom and a sketch of the details is provided in Figure 13.

Phantoms



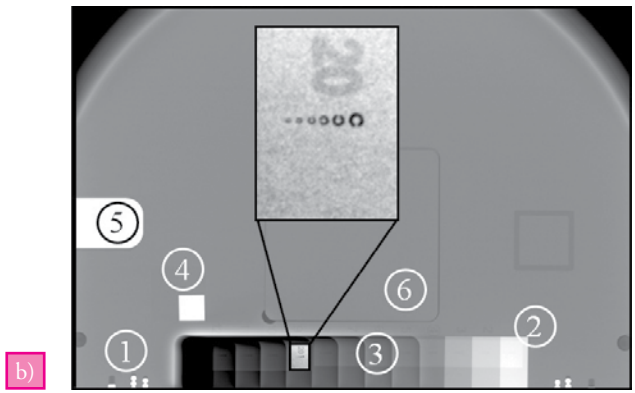


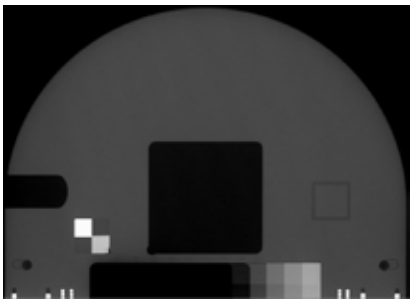
Figure 13: Picture (a) and sketch (b) of the Quart MAM/DIGI phantom. (1) are the steel balls for phantom alignment and missing tissue evaluation, (2) is the step-wedge with the Landolt rings, (3) the uniform step-wedge, (4) the brass square for MTF, (5) a slot for a dosemeter, and (6) a from carved section for optional additional inserts.

A commercial software package for automatic analysis of images produced by MAM/DIGI phantom exists. It provides the Nyquist frequency and the MTF of the system, and the SDNR for each one of the steps within the step wedge.

More detailed information about the phantom and software can be obtained in [11].
An X-ray image of the Quart MAM/DIGI phantom is shown in Figure 14.

Phantoms

Figure 14: Radiographic image of the Quart MAM/DIGI phantom.



The Landolt ring test:

Twelve groups of six rings (three of them are depicted in Figure 15) are inserted into the twelve-step wedge to cover a wide range of contrasts using a single exposure. The task with the Landolt test consists of detecting the gap direction (four-alternative choices) within rings of diameters between 260 and 800 µm and gaps of lengths between 52 and 160 µm. The observer results are then compared with a known truth. The number of correctly determined gaps provides a score for each equivalent breast thickness. An overall measure of image quality can be obtained by totaling scores from the seven thickest steps. A Landolt7score of 19 and 8.5 has been found to be equivalent to the acceptability thresholds suggested for the CDMAM and ACR phantoms respectively [11]. The evaluation of the Landolt rings can be automated using software like the one developed by University of Mainz [12].

Figure 15: Example of Landolt ring groups for three different steps of the QUART mam/digi phantom. A close-up view of the left group is shown for clarity.

Image quality evaluation

Image quality is a generic term that can be applied to any type of image. Image quality evaluation requires the definition of one or more parameters/attributes that can be measured/assessed to describe the quality characteristics of an image. The usual rule applied (with few exceptions) is that the higher the value of such parameter/index, the better the image quality. In the following (technical) image quality evaluation methods used in mammography QC have been grouped in two main categories: "human-based" methods and objective methods.

Human-based methods include the attribution of IQ scores on the basis of the degree of perceived "visibility" of each feature produced by phantom details, and the search for a "visibility threshold curve" obtainable by means of contrast-detail phantoms. The vision process is typically human, but automatic software packages can be developed including some type of criteria which reproduce the detection mechanism and provide as output detail count or a threshold.

The alternative methods (so-called "objective") consist in the extraction of quantitative parameters from image matrices; this is obtained by placing regions of interest (ROIs) inside/outside/across the features produced by the details embedded in the phantom, extracting signal and noise information and applying suitable formulas. Objective measurements can be performed either by human observers using software tools able to handle ROIs and their contents, or automatic software packages capable to localize phantom details, define the ROIs, and automatically calculate the relevant parameters.

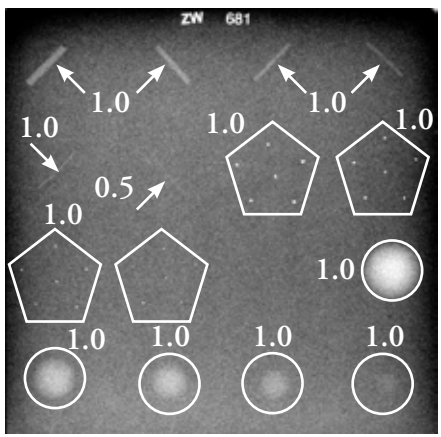
In this section, several examples of image quality evaluation by human-based and objective methods will be illustrated, using images/details produced by phantoms previously described. Same or similar approach/principles can be followed for evaluating image quality by images produced with other phantoms, different from those reported in this document.

Image quality evaluation

Human-based scores – Counting "visible" details

A general approach to image quality assessment based on phantom images comes from the "analog world" and consists in counting which details in the phantom image are "visible", according to different "visibility" criteria. This can be done by asking a certain number of observers (to include inter-observer variability) experienced in evaluating phantom images to give a score per each type of detail depicted by the phantom image, with reference to their confidence in presence of each expected detail. For each type of detail simulating clinical features, a score of 1 can be assigned if the detail is well/totally visible, 0.5 if it is not-well/partially visible, and 0 if is not visible at all, as usually done with the ACR accreditation phantom or with the CIRS Model 011A. The sum of scores associated to all detail visibility is taken as overall IQ index, and a minimum threshold can be set for the total score and, separately, per each type of detail [13]. An example is reported in Figure 16.

Figure 16: Example of scoring of an ACR phantom image using the standard rating.



IQ SCORE	
Fibers = 1.0+1.0+1.0+1.0+0.5 =	4.5
Specks = 1.0+1.0+1.0+1.0 =	4.0
Masses = 1.0+1.0+1.0+1.0+0.5 =	4.5
TOTAL =	13.0

Detail count is a very simple and general method, which is applicable, in principle, to any phantom containing multiple details of the same type; detail "visibility" for an image obtained at a given dose level is associated to the physical properties of each feature, its thickness (influencing perceived contrast) and size; however, the decision about the amount of visibility is subjective and can lead to variability when performed by human observers.

Image quality evaluation

The categorical scale used to score phantom images can be slightly extended, in order to differentiate between more levels of visibility and improve the sensitivity of IQ evaluation. An example is given for the TORMAM phantom, for which four different scores between 0 and 3 are admitted per each type of feature produced by the non-anthropomorphic portion of the test object. Figure 17 summarizes visibility criteria proposed for the TORMAM phantom and shows a few scoring examples.

Figure 17: Example of 0-3 scale application to score TORMAM images

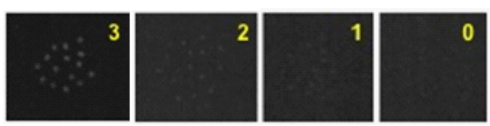
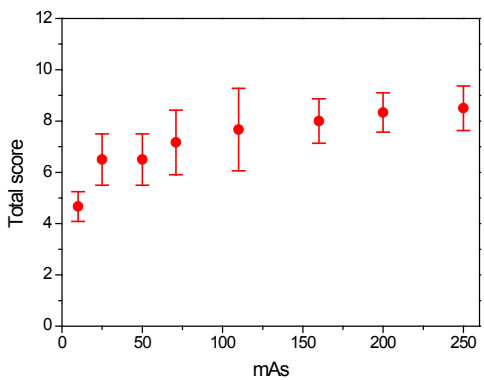
Detail group	Visibility criteria	Examples of scoring
Multi-directional filaments	Filament complete/easily detected = 3 Filament complete/less visible = 2 Filament partly complete/threshold = 1 Filament not complete/not seen = 0	
Microcalcs (MC)	MC group easily seen = 3 MC group less visible/faint = 2 MC group barely visible/threshold = 1 MC not seen/ background noise = 0	
Low-contrast detail	Disc easily seen/contour defined = 3 Disc visible/contour not defined = 2 Disc partly visible/contour not defined = 1 Disc indistinguishable from background noise = 0	

Image quality evaluation

Independently of the scale used to score phantom images, these human-based methods have inherent limitations. In particular, intra- and inter-observer variability associated respectively to the lack of reproducibility of the individual decision about the visibility threshold and to differences among different observers, unavoidably affects the image quality evaluation process. An example is reported in Figure 18 for scores assigned by four observers to the visibility of the low-contrast circular details of TORMAS images acquired at different dose levels and using a rating scale similar to that illustrated for the ACR phantom; the total count is plotted versus dose (detector dose is proportional to selected mAs for a given spectrum), and the error bars represent variability across readers [6]. Similar results can be found in the literature for different phantoms [7, 14-17].

Figure 18: Low-contrast circular details of a TORMAS phantom: total score versus mAs level (dose). Three step scale (score 1 = well visible; 0.5 = partially visible; 0 = not visible). Red dots are the mean scores across the four observers and the error bars represent the inter-observer variability.

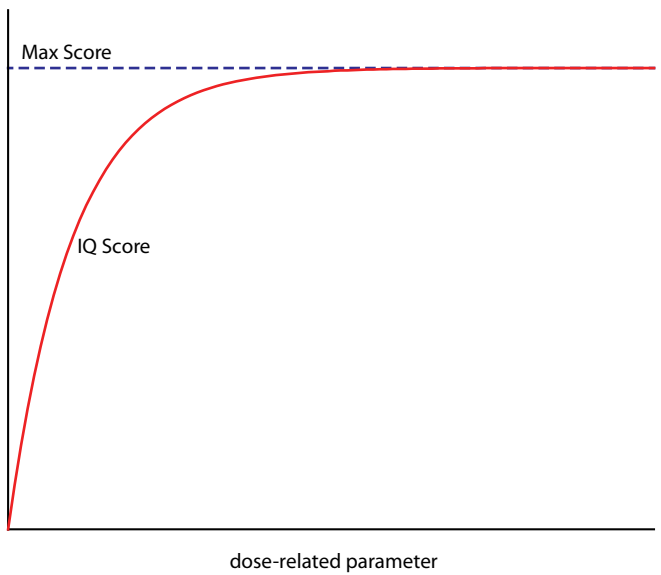


Variability across the observers may reduce capability of discriminating real differences between image quality levels. Such variability can partially be reduced through training sessions, where image scores are discussed in consensus and observers try to harmonize their decision criteria. However, some degree of variability has to be accepted as "natural noise" when image quality is assessed using human-based methods. Statistical uncertainties associated with variations in human decisions about visibility threshold should be evaluated before using those methods in assessing phantom image quality for quality controls.

Image quality evaluation

In general, whatever the phantom and the number of its details are, image score versus dose is well fitted by the function $y = A \cdot (1 - e^{-B \cdot x})$ where A is a coefficient equal to the total number of details in the phantom which could be visible only at very high dose level and B is another coefficient representing the growth rate [15]. The curve shape is represented in Figure 19.

Figure 19: Typical model which can be used to fit the IQ scores obtained by human observers. The IQ score increases with the dose level at which the phantom image was acquired. There is a superior limit, corresponding to the maximum score associated to the full visibility of all the details included in the phantom. Even if such limit is not achieved in the dose range normally explored, there is a dose threshold above which a further increase of radiation dose has a minimum effect on the IQ score.



Alternatively, the process of counting details or assessing visibility scores can be performed automatically by software tools, with the primary benefit of eliminating the uncertainty associated to intra- and inter-observer variability. However, it should be noticed that the decision about detail presence and degree of visibility from software tools are made by applying threshold criteria to objective parameters (e.g. signal intensity produced by the detail, size, etc.) extracted from the image [3,4,6,9,18-20], not using observer models able to mimic the human visual system; this is the main reason why software tools often count or "see more" than the "mean observer" [6,9].

Image quality evaluation

Human-based scores – "Visibility" threshold

A different approach, originally developed for screen-film imaging and still classified as "human-based", consists in the determination of the "visibility threshold" for a given type of detail, with known physical properties. This is usually done by contrast-detail (C-D) phantoms, which include multiple details (typically discs) with variable contrast and size. Normally, detail contrast is varied in one direction (by varying detail thickness) while keeping the same size, and in the other direction detail size is varied while keeping the contrast unchanged. C-D phantoms can be obtained either superimposing details with variable thickness (i.e. image contrast) to a uniform background, or digging holes with variable depth in the uniform background. This type of test objects comes from experiments on human visual perception performed in the 40's [21]. Humans can perceive objects on the basis of their size, shape, and contrast relative to their surroundings, as illustrated by the simulated image in Figure 20.

Figure 20: Simulated image as obtained from a C-D phantom. Details change in size in one direction (left-to-right) and in contrast in the other direction (up-to-down).

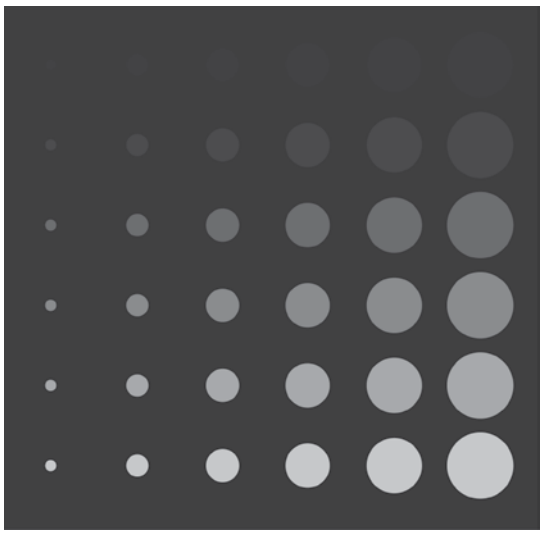
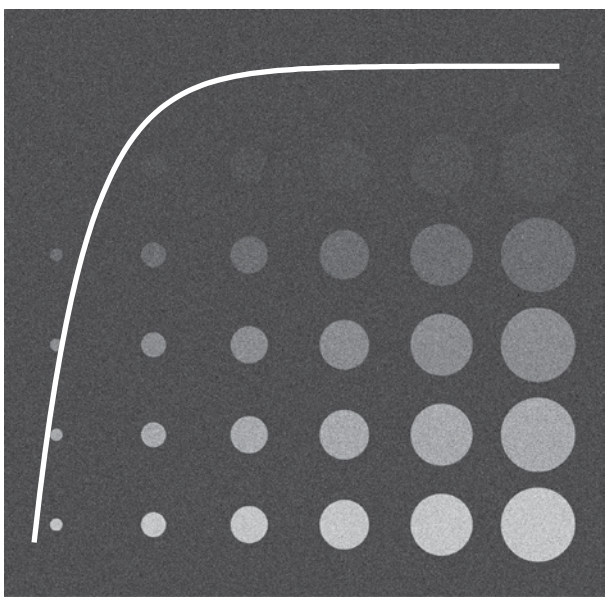


Image quality evaluation

It has been demonstrated that in the presence of noise, human ability to perceive details in an image is further obscured. Therefore, a trade-off exists between object size, contrast, and noise. In general, given the unavoidable presence of noise on images produced by medical imaging equipment, the contrast necessary to detect a small object is higher than the contrast to perceive a large one [22]. The negative effect of noise on the detection of small objects is represented in Figure 21.

Figure 21: Evaluation of medical images always occurs in presence of noise, which reduces human capability of detecting low-contrast object in uniform background. A "contrast-detail" curve can be defined, corresponding to the "visibility threshold" for different size of the details.



A curve, named contrast-detail curve, can be obtained by asking human observers to detect the visibility threshold (minimum contrast) for each detail size, assuming the phantom location of the details is known.

Image quality evaluation

The capability of the human eye to detect low-contrast objects in a uniform background in presence of noise was described by Albert Rose in his work to determine the sensitivity of the human eye on an absolute scale [23]. The Rose's model provides a simple description of the relationship between noise, contrast, and spatial resolution. It formulates the human visibility threshold in terms of "signal-to-noise ratio", where the "signal" is the net difference between the signal produced by the target detail and the background, while the "noise" can be calculated by the square root of the background signal, assuming that quantum noise (uncorrelated) is the only (at least the main) source of noise in the image, following Poisson statistics. The SNR definition by Rose corresponds to what we now call contrast-to-noise ratio (CNR) or signal-difference-to-noise ratio (SDNR):

$$K = \frac{signal}{noise} = C \times \sqrt{\Phi \times A}$$

where C is the relative contrast of a given detail, Φ is the photon fluence (number of photons per unit area), and A is the area of the smallest object we wish to detect. Rose observed that humans cannot perceive a low-contrast disk in a noise unless the value of K is at least 5-7.

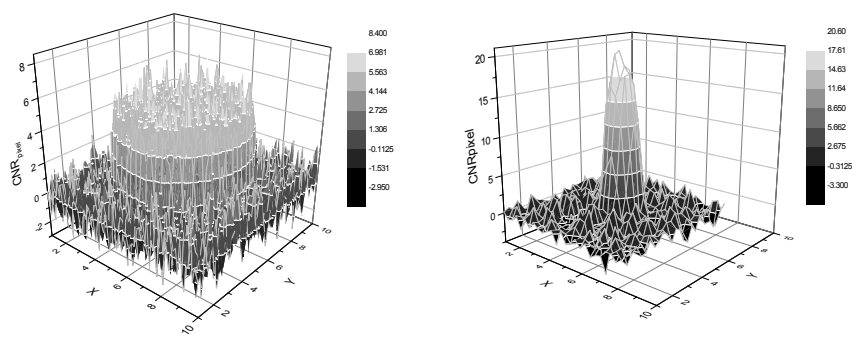
As the object size decreases, either the contrast or the number of photons (radiation dose in X-ray imaging) need to increase to allow the object detection [22].

Image quality evaluation

Rose's model applicability

Rose's model has the merit of using a simple formalism to define "visibility threshold" by human observers. Experiments by Rose aimed to determine the absolute human contrast threshold. Its extension to medical images requires awareness by users of Rose's model. The simple task addressed by a contrast-detail phantom is the detection of a signal with known parameters (size, shape, location) in additive, uncorrelated Gaussian noise on a known background. This is usually referred to as the signal-known-exactly (SKE), background-known-exactly (BKE) detection task. Rose's model is a useful approximation for SKE and BKE detection of a limited subset of signals and noise. In particular, the Rose's model neglects the fact that noise in the potential signal location has unequal variances for the signal-present and signal-absent cases; such approximation is acceptable only when the variance produced by the detail is very small compared to the variance of the background, and this is true only for low-contrast details, which are a portion of details in a contrast-detail test object. Moreover, Rose assumed "flat-topped" signals with sharp edges; this is approximately true for large details produced by flat disks, but is very far from the signal produced by small details, as shown in Figure 22).

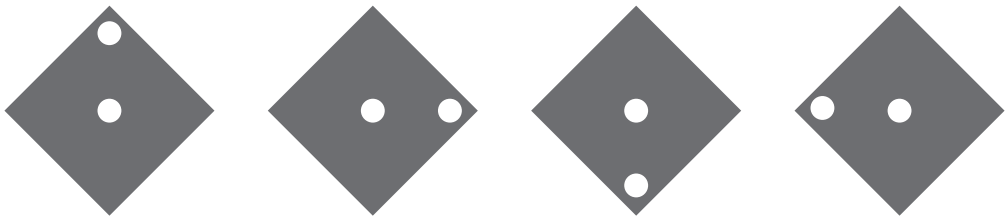
Figure 22: Signal differences produced by a large circular detail and a small one in a noisy background; the large detail produces a signal reasonably "flat topped", while this does not happen for the small detail .



In other words, the Rose's model is applied with a significant number of approximations when contrast-detail phantoms are used to assess image quality [24].

Image quality evaluation

A special class of C-D phantoms is that requiring the application of "multiple-alternative forced choices" (MAFCs). Those phantoms are designed to have each detail of given contrast and size located in one of multiple possible alternative positions (usually two or four) [25]. The CDMAM previously presented belongs to this class of phantoms and requires a choice among four alternatives. In fact, each diamond-shaped cell in the CDMAM contains a disk of given thickness and size in the center and an identical disk in one of the four vertices. Detection of each detail pair is required, together with the identification of the correct localization of the vertex detail.



The observer completes a each scoring sheet reporting the positions of the vertex details he/she guessed. After comparison with the "mask of truth" to verify correctness of detail localization, the contrast-detail curve is obtained by plotting the contrast (or nominal thickness) of the last visible disc against the disc size.

A psychometric curve is often applied to the discontinuous experimental data points produced by image scoring MAFC to obtain a smooth C-D curve [10]. Human readings of images produced by C-D phantoms have the same limitations of intra- and inter-observer variability as previously reported for other types of phantom for which different scoring criteria are used to correlate image quality with the visible details. For this reason, contrast-detail curves can be derived from human readings only involving multiple phantom images and multiple readers. An example showing error bars illustrating the significant variability across four readers is provided in Figure 23 [26].

Image quality evaluation

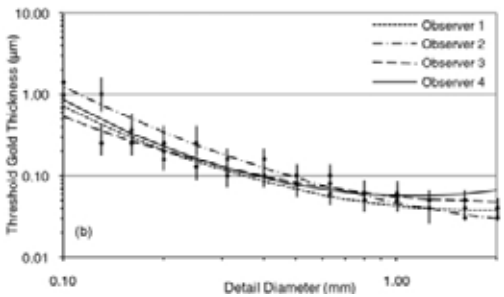


Figure 23: Contrast-detail curves obtained by four human observers from CDMAM (version 3.4) image score. Error bars represent variability across the readers [26].

The use of software analysis can overcome such problem while saving a large amount of time necessary to score those phantom images by human observers. The basic software (freeware) developed by the University of Njimegen [9] for automatically identifying discs on digital images of the CDMAM is called CDCOM (current version 1.6) and is available for downloading at the EUREF website (http://www.euref.org/downloads). The CDCOM program locates the position of the gold discs on a single DICOM image, but does not combine the data from more than one image or determine the threshold contrasts.

Other tools exist to combine multiple CDCOM outputs and derive the contrast-detail curve and the visibility thresholds for multiple disk diameters by fitting experimental data with a psychometric curve. One free tool is called CDMAM Analysing Software, developed by the UK National Health Service; it can be downloaded at EUREF website. Two other applications are commercial, CDMAM Analyser by Artinis (Nijmegen, NL) (infos at http://www.artinis.com/additional/download), and Erica2 by Qaelum (Leuven, BE) (infos at http://www.qaelum.com/products/quality-control-tools/erica.html).

Software tools are usually more sensitive than human observers in detecting low contrast gold discs; therefore, the results of CDCOM output are multiplied by a certain factor (e.g. 1.5 for 0.1 mm diameter) to make threshold values comparable to those obtained by human readout [16]. The CDMAM Analysing Software from NHS-UK has implemented such factors, and results already adjusted for the human threshold are provided.

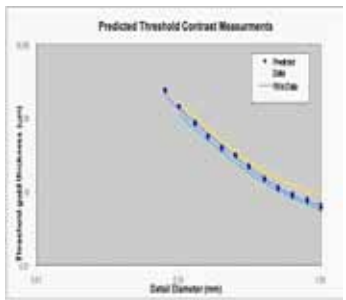
Image quality evaluation

Figure 24 shows the output of the three tools mentioned above.

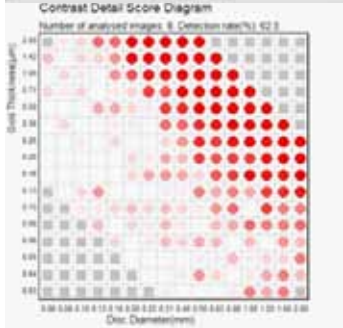
Figure 24: (a) Table showing the threshold provided by the CDCOM software tool and the predicted human threshold [16]; (b) Output of the CDMAM Analysing Software; (c) Output of the Artinis CDMAM Analyser; (d) Output of the Qaelum Erica² software.

ϕ (mm)	Software threshold (µm Au)	Predicted human threshold (µm Au)	Fit to predicted human threshold (µm Au)
0.08	1.17	1.74	1.52
0.10	0.67	1.05	1.05
0.13	0.43	0.71	0.70
0.16	0.29	0.50	0.52
0.20	0.21	0.37	0.38
0.25	0.16	0.29	0.29
0.31	0.12	0.23	0.22
0.40	0.08	0.16	0.17
0.50	0.06	0.13	0.13
0.63	0.05	0.11	0.10
0.80	0.04	0.08	0.08

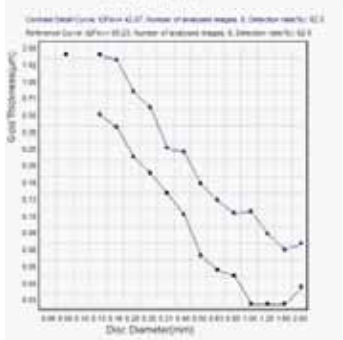
a)



b)



c)



d)

Image quality evaluation

Quantitative image quality evaluation

Electronic images produced by digital mammography systems allow the calculation of a large number of quantitative parameters, reported in the following as Image Quality Indices (IQIs). It was much more complicated to obtain such indices from film images, as the film had to be first scanned in order to obtain a digital image. This is the main reason why the simple criteria related to detail visibility, such as discussed above, were used. Conversely, digital images allow direct access to any information recorded by the image itself; the only issue is to define appropriate IQIs for each given phantom, in order to show differences in image quality.

Quantitative IQIs present several advantages compared to human-based IQ assessment:

- They are objective, i.e. independent of subjective decision (causing intra- inter-observer variability when human observers assess image quality), and independent of arbitrary threshold setting (when image quality is assessed by automatic software, but still using "human-based" methods).
- They are reproducible, i.e. different experimenters or software tools, measuring the same parameters would obtain the same results. For this reason, quantitative image quality can be assessed using a limited number of images under each exposure condition, once system reproducibility has been verified.
- They are more sensitive than human-based methods in assessing IQ differences. In fact, quantitative IQIs are usually defined in continuous space, and, as such, they can provide a continuous function of different independent parameters (dose, photon energy, object thickness, etc). This is particularly useful to monitor long-term reproducibility and to try to assess possible causes of variations.

The general approach taken by this document is to be independent of the phantom used. So, in the following examples are given of possible IQIs measurable by the various phantoms. As before, the list of examples does not aim to be exhaustive.

Image quality evaluation

Phantoms including discs or spheres to mimic low-contrast masses

Several phantoms include discs or spheres to reproduce low-contrast masses. Usually discs are used keeping the object size unchanged and varying the disc thickness to produce contrast variations, while with spheres, changes in sphere diameter produces a change of both object size and contrast. Examples for three of the phantoms previously described are illustrated in Figure 25: ACR and CIRS Model 011A phantoms including spheres, and Leeds TORMAS including discs, respectively.

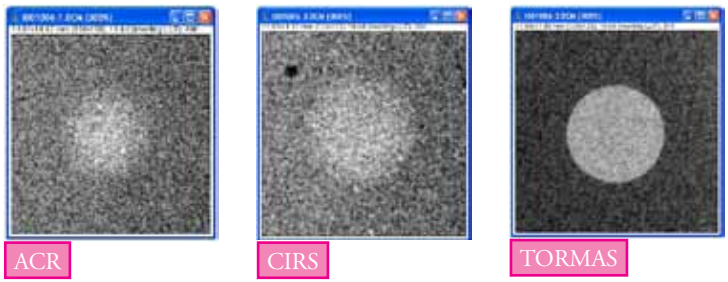


Figure 25: Low-contrast circular details included in three different phantoms: ACR RMI 156, CIRS 011A, and Leeds TORMAS. For the first two acrylic spheres are used, fort the third one thin metal discs.

For "large", flat circular details, the signal-difference-to-noise ratio (SDNR) is something which is always measurable according to the Rose's model. The classic definition of SDNR assumes the signal produced by the detail is uniform in the detail area and can be measured by means of a circular ROI inside the detail and a cicular ring including the overall background around the detail, as represented in Figure 26.

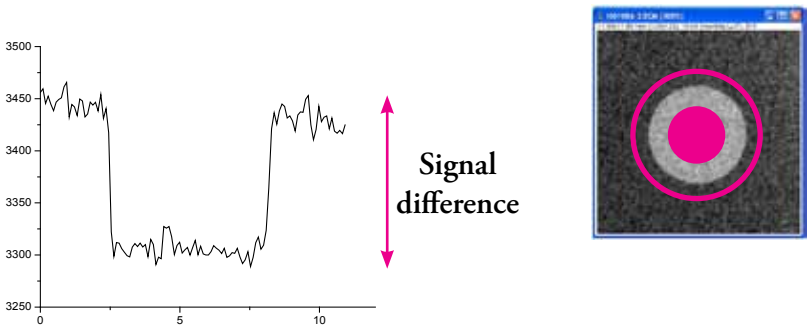


Figure 26: Signal profile and ROIs for SDNR measurement with low-contrast circular details obtained from discs as present in some QC phantoms.

Image quality evaluation

With phantoms using spheres to simulate masses, the profile shape is different, and differences in the selected ROIs, even small, can change significantly the resulting SDNR value. In this case, the SDNR can be obtained by determining the peak height and using the external ROI only for noise evaluation (Figure 27).

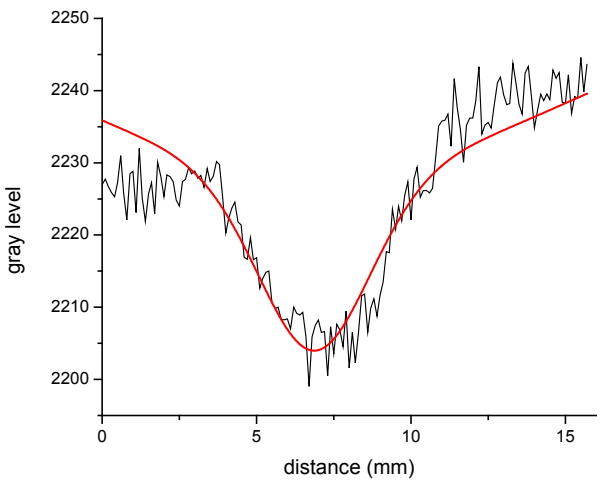


Figure 27: Signal profile and ROIs for SDNR measurement with low-contrast circular details obtained from discs as present in some QC phantoms.

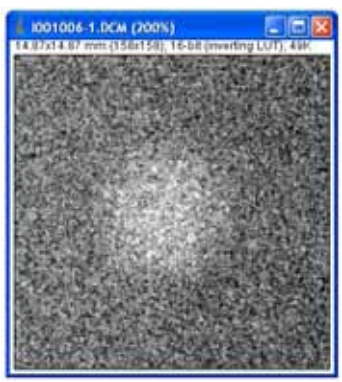
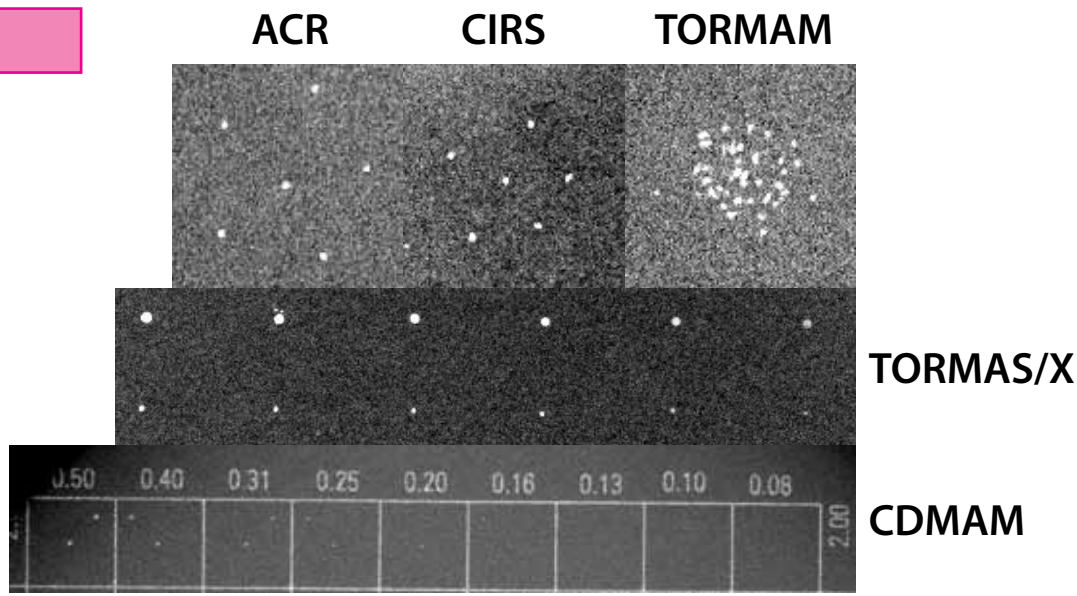


Image quality evaluation

Phantoms including small high-contrast objects to mimic microcalcifications

Several phantoms include small high contrast objects (either individual or in clusters) to reproduce microcalcifications. Some examples are illustrated in Figure 28.

Figure 28: Small high contrast objects included in many phantoms.



Detection of small objects is limited by spatial resolution properties (pixel size / sampling interval), as well as image noise. Quantitative parameters can be derived from "peak analysis". An example is provided in Figure 29 for the details 0.5 mm in diameter of the TORMAS phantom: out of the 11 details.

Image quality evaluation

Peak parameter	Detail 01	Detail 02	Detail 03	Detail 04	Detail 05	Detail 06
Amplitude	150.2	125.5	88.5	87.8	43.2	38.0
σ	0.41	0.42	0.40	0.39	0.36	0.34
FWHM ~ 2.35 σ	0.96	0.99	0.94	0.92	0.85	0.80

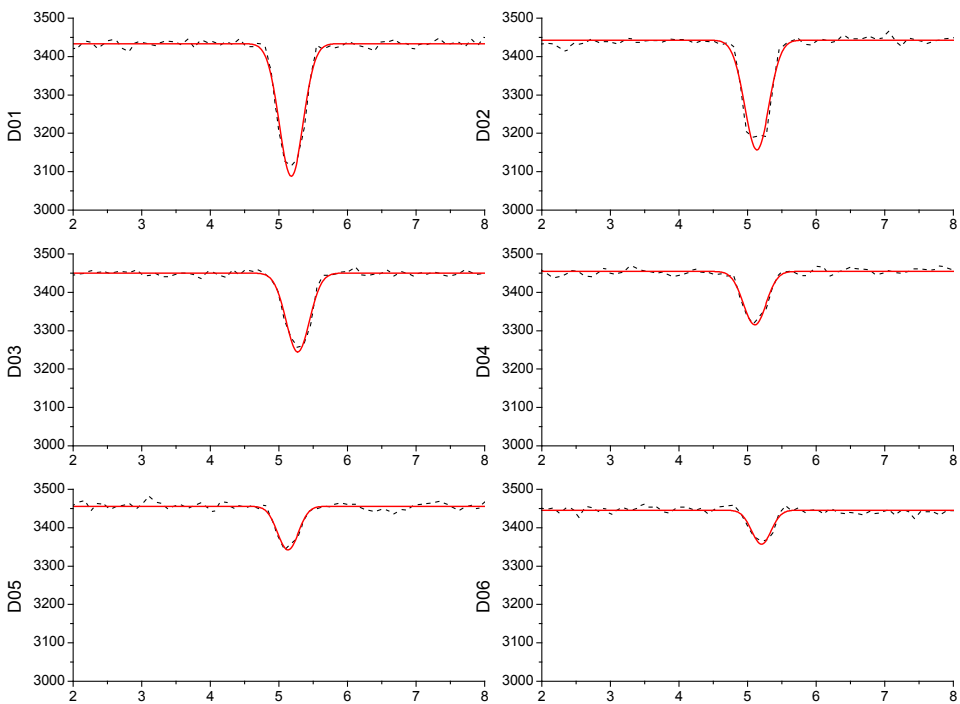


Figure 29: Peak parameters and profiles of the first 6 small high contrast details included in TORMAS.

Image quality evaluation

Phantoms including fiber-shaped objects

Several phantoms include fiber-shaped objects to reproduce low contrast fibers with different spatial orientation or spicules in clinical images. Some examples are illustrated in Figure 30.

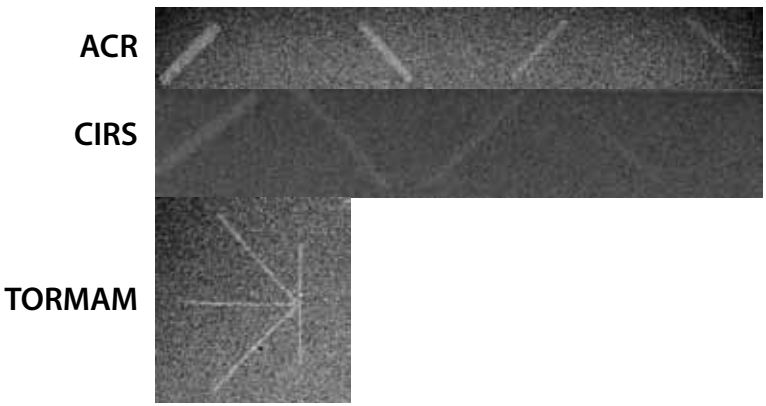
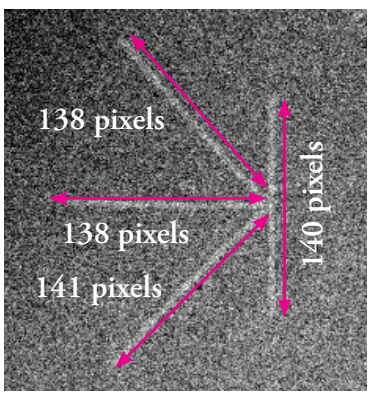


Figure 30: Low contrast fibers with different spatial orientation included in three QC phantoms (ACR, CIRS, TORMAM)

Fibers can be characterized by the filament thickness and length, both quantitative parameters. An example is shown in Figure 31 for the Leeds TORMAM phantom.



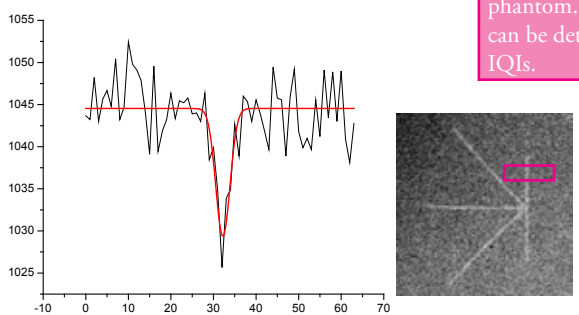


Figure 31: Group of fibers in the TORMAM phantom. For each fiber, length and thickness can be determined and used as quantitative IQIs.

Image quality evaluation

Phantoms including step-wedges

Several phantoms include step wedges, for relative contrast or SDNR measurements. Some examples of step wedge images are depicted in Figure 32.

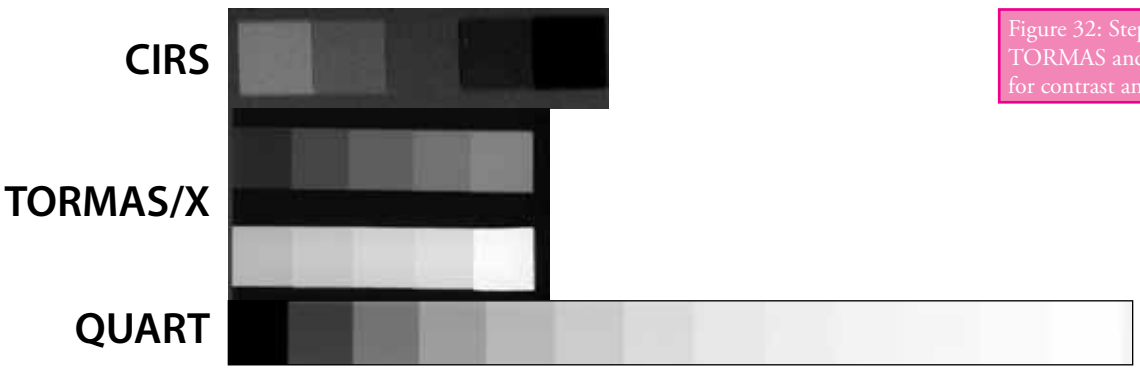
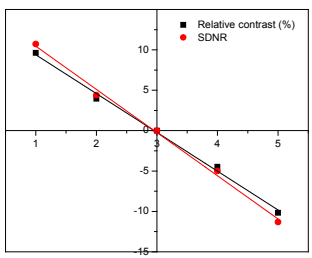


Figure 32: Step wedges contained in CIRS, TORMAS and QUART phantoms. Useful for contrast and/or SDNR meadurement.

Relative contrast and signal-difference-to-noise ratio from the step wedge of the CIRS image (the central step has the same absorption as the phantom background) are shown in Figure 33.



Step	Relative contrast (%)	SDNR
1	9.6	10.72
2	3.9	4.40
3	0.0	0.00
4	-4.5	-4.99
5	-10.2	-11.32

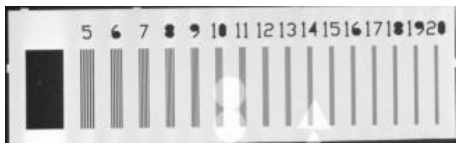
Figure 33: Relative contrast (%) and SDNR from the step wedge of CIRS phantom.

Image quality evaluation

Phantoms including spatial resolution patterns

Some phantoms include high contrast spatial resolution patterns, often bar patterns if phantoms were designed for screen-film mammography, or sharp edges for phantoms developed for digital mammography; such objects can be used to determine the modulation transfer function (MTF) or other parameters related to spatial resolution. Some examples are illustrated in Figure 34.

CIRS



TORMAS/X

$$MTF(f) = \begin{cases} \dfrac{\pi\sqrt{2}}{4} \cdot \dfrac{M(f)}{M_0}, & \text{if } f \geq f_c/3 \\ \dfrac{\pi\sqrt{2}}{4M_0} \cdot \sqrt{M^2(f) - \dfrac{M^2(3f)}{9} - \dfrac{M^2(5f)}{25} - \dfrac{M^2(7f)}{49}}, & \text{if } f \geq f_c/11 \end{cases}$$

QUART

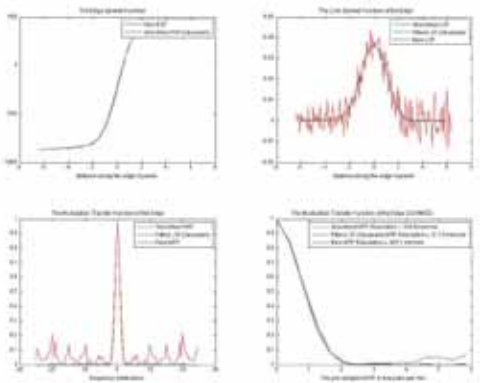


Figure 34: Objects for the evaluation of spatial resolution properties included in some phantoms. Bar pattern are still present in phantoms developed for screen-film mammography, while an edge is usually inserted in new phantoms developed for digital.

Image quality evaluation

Resolution patterns included in IQ phantoms should not be used for absolute measurements of modulation transfer function (MTF), but IQI related to the overall spatial resolution can be derived. For example, bar patterns do not allow the calculation of modulation with digital images, due to the limitation in sampling interval, but the MTF can be derived from measurements of standard deviations calculated within ROIs including groups of bars, as explained by Droege and applied to old computed tomography (CT) systems [27,28]. The evaluation of an index associated with resolution properties can be helpful in long-term reproducibility to check, for instance, that geometrical conditions (overall phantom thickness) are those expected by the test procedure.

Phantoms and AEC

IQ phantoms and automatic exposure control

In principle, quality controls should be performed using the imaging equipment in conditions as close as possible to to clinical practice. This means, for example, that if in clinical practice the automatic exposure control is always used (with very few exceptions, like implants or extremely thick breasts), the AEC should be used also to expose the IQ phantoms for the evaluation of technical image quality. However, it is necessary to remark that IQ phantoms can be more or less "compatible" with the AEC, depending on the type of objects they include. Objects "disturbing" some types of AECs are highly absorbing object, like the spatial resolution patterns contained in some phantoms originally developed for screen-film mammography. In fact, most of the AECs of DR systems use a short pre-exposure to detect a sub-area of the image detector corresponding to the absorption peak, and use such information to adjust the mAs level (anode/filter combination and kV_p are mostly established on the basis of the object thickness) to avoid under-exposed areas in the final image. Presence of highly absorbing objects in the field-of-view may induce the AEC to select long exposure time, with mAs values much higher than those selected in clinical applications, with some risks of producing ghost images. There are a few options to address this potential issue:

1. Choose another type of phantom, not including high density objects, which is compatible with the AEC.
2. Given a phantom including one or more high density object, try to expose it using the standard AEC mode used for breasts and verify if the mAs value is comparable or much higher than typical values found for breasts. If it is much higher, ask to your DR vendor if there is a way to exclude the area covered by the high intensity object from the AEC search. In fact, some AECs give the possibility to specify an area of the digital detector used as exposure meter, but the procedures are different per each specific system.
3. Given a phantom including one or more high density object, expose the phantom in manual exposure mode, setting anode/filter combination and kV_p as chosen by the AEC and mAs derived from the exposure of a phantom having similar absorption but no high density object inside (for example acrylic alone, same thickness). This solution has the limitation that if one wants to test system reproducibility, possible variability of the AEC is not included.
4. Given a phantom including one or more high density object and an analog mammography unit used with CRs, take care to choose a position of the AEC sensor (in this case external to the image detector) which excludes the high density object.

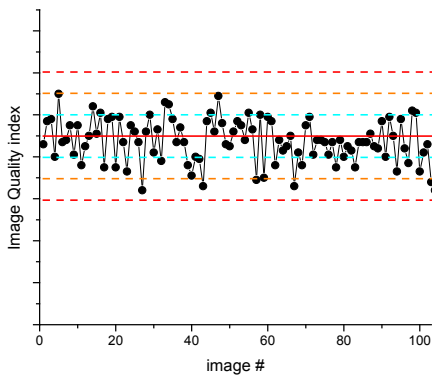
Reproducibility tests

A correct picture of such "natural variability" of an imaging system needs to be taken in order to properly define the baseline against which the following variations are calculated. This is usually done by acquiring multiple phantom images during the equipment commissioning, and over a short period after it (daily for one or few weeks). As there is no reason to suppose that the distribution of short-term data is not Gaussian, the standard estimators used to describe Gaussian distributions, mean and standard deviation, can be used to determine the natural variability of the imaging system. The baseline or central line is calculated by averaging the values of the IQ parameter extracted from the images acquired during the commissioning and eventually in the following short term. The standard deviation (SD) of the distribution is used to calculate the so-called control interval, equal to mean ± 3 SD, within which the values of the IQ parameter measured in the following days/weeks/months are expected to fall with very high probability (three standard deviations correspond to probability by 99.9% to belong to the Gaussian distribution). Variations beyond the control interval are considered "abnormal" and need to be investigated in order to find explanations or possible causes.

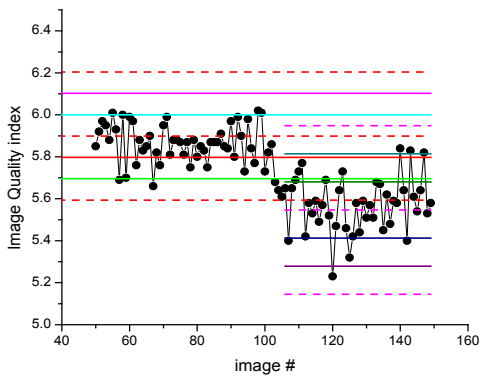
Monitoring of IQ performance can be done by means of control charts, a graphic technique introduced in industry in 1924 (by Shewart) to study temporal evolution of productive processes and with the aim of controlling the input variables to limit as much as possible causes of possible defects [30]. Each new value of the IQ parameter is represented as a point in a scatter plot as a function of time; the lines depicting the baseline and the limits of the control interval are also plotted to allow the reader to detect by a glance possible points exceeding such limits. The imaging system is considered to be reproducible over time if the IQ parameter fluctuates around the baseline value, with most values within ±1 SD and less values towards the distribution tails; otherwise, the imaging system shows an abnormal behavior if several points occur beyond the limits of the control interval, or if systematic effects are observed, like trends or sudden shift of points far from the baseline previously defined. In Figure 35 two examples are provided for (a) a reproducible system and (b) for the same system in a partially different time interval, which showed an abnormal change due to detector calibration.

Reproducibility tests

Figure 35: Example of long-term reproducibility test. (a) The image quality index normally fluctuates around the baseline value, within the control interval; (b) There is a systematic shift of the image quality index, which was found to be associated to system recalibration.



a)



b)

From the description above, it can be noticed that the definition of baseline and related control interval is critical and needs to be accurately performed.

Reproducibility tests

One possible problem is when the control interval is "overestimated", i.e. too large compared to the normal system fluctuations. In this case, the data points obtained from the following reproducibility tests would be distributed systematically within the control interval, but far from the control limits (Figure 36). This effect might occur, for instance, if during the acquisition of the baseline images some phantom exposures are performed with exposure mode/parameters different than that/those which should be used for the baseline, causing outliers in resulting IQI values. This situation would require the baseline retake or at least the baseline recalculation after removal of outliers (if in limited number).

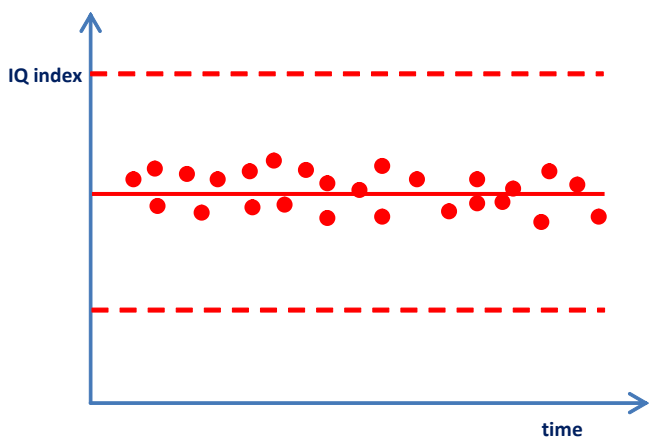


Figure 36: In case of overestimation of the control interval (caused by variability during baseline definition superior to the real system variability), reproducibility data points will be well within the control limits, very far from them. Such situation would require the baseline retake or the baseline recalculation after removal the outliers (if only few).

Reproducibility tests

The opposite problem could occur when the control interval is "underestimated", i.e. too narrow compared to the normal system fluctuations. In this case, the data points obtained from the following reproducibility tests would be very often out of the control interval (Figure 37). This effect might occur if the acquisition of the baseline images is performed in more reproducible conditions than the following tests, resulting in a small value of the IQI standard deviation. This situation would require the baseline retake or at least the baseline recalculation after removal of outliers (if in limited number).

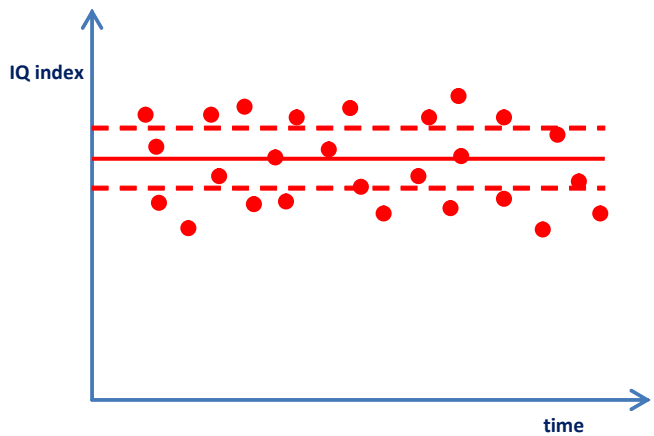


Figure 37: In case of underestimation of the control interval (caused by variability during baseline definition lower than the real system variability), many points will be abnormally out of the calculated control interval.

Reproducibility tests

A possible criticism of this method is that both a very stable system and a very variable one could be accepted, the acceptance criteria being relative to the statistics of each individual mammography system. In other words, a very stable system would produce a narrow control interval, while variable one would determine a large control interval, but the following reproducibility tests could produce data consistent with both such different situations. Further information about possible differences in constancy across systems of the same type can be obtained if the same test is applied, with the same type of phantom and image quality index measured, to several systems. An example is shown in Figure 38, where the baselines obtained from four digital systems (same brand and model) for an IQ index extracted from 25 phantom images (same type of phantom and 5 images acquired in automatic exposure mode for 5 consecutive days).

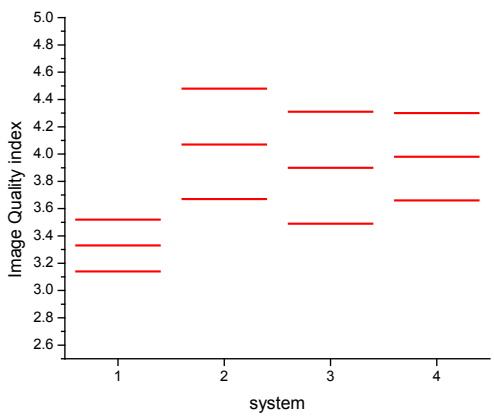


Figure 38: Baseline values and related control intervals for the image quality index measured by the same type of phantom images and four mammography systems of the same type.

Baseline value is very different for the first system compared to the others, and the control interval for system 1 is much narrower than systems 2-4. Dose parameters should be analyzed to explain reasons for differences in absolute values, especially for system 1.

Reproducibility tests

Using processed vs. unprocessed images

Physics tests are always performed using unprocessed images, i.e. images for which the pixel values are proportional to the absorbed dose; the relationship between absorbed dose and output signal depends on the detector response function. Image post-processing is absolutely necessary to make clinical images appropriate for the human vision, in terms of contrast and dynamic range, but it removes or reduces the physical relationship between input and output (Figure 39).

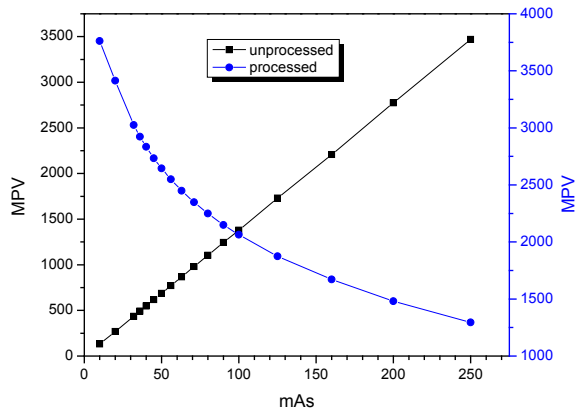


Figure 39: Mean pixel value (MPV) versus dose (mAs) for unprocessed and processed images obtained by a linear digital system. It can be noticed that system linearity can be evaluated only from unprocessed images.

However, for reproducibility tests, the use of post-processed images could be acceptable, the objective being the analysis of variations of image quality indices, parameters which usually include the contribution of both signal and noise. An example is illustrated in Figure 40 for the SDNR of a low-contrast, large detail, using both unprocessed and processed images: SDNR values are systematically higher for processed images, as the post-processing algorithm reduces the dynamic range (increasing contrast), but does not change significantly noise.

Reproducibility tests

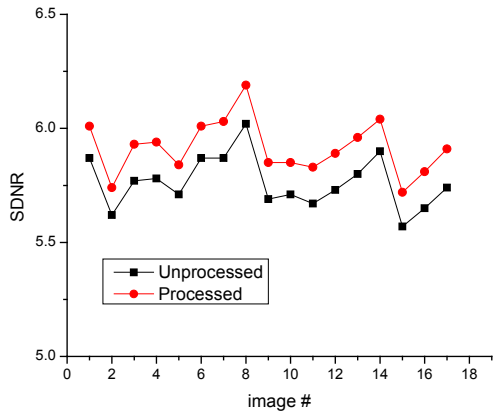


Figure 40: SDNR measured from unprocessed (black) and processed (red) images from a low-contrast detail, in a time series of phantom images acquired with daily frequency.

Image Quality and CR systems

"Philosophy" of CR systems

Computed radiography (CR) is the first technology used to produce digital radiography images, which has been developed in the early eighties from those manufacturers who produced screen-film systems at that time [31]. CRs have been widely used for conventional radiology applications, while their use for mammography was significantly delayed, due to their limitations in spatial resolution and quantum efficiency at energy ranges used in mammography. The first mammographic CR system appeared in 2004, a few years after the introduction of direct digital (DR) mammography systems [32]. CR systems were well tolerated because they left the general workflow of image production unchanged. The same X-ray unit as used with screen-film was maintained, but the cassettes were replaced with similar objects containing imaging plates (IPs), a special type of screens whose sensitive material was able to trap the latent image until an external source of energy intervened to take it out; such source of energy was provided by a laser beam included into a scanner together with a light detector and a digitalization stage. After the exposure, the imaging plate was scanned and a digital image was produced, processed and displayed, as illustrated in Figure 41. Possible residual signal is cancelled from the imaging plate, and the imaging plate reused for other exposures.

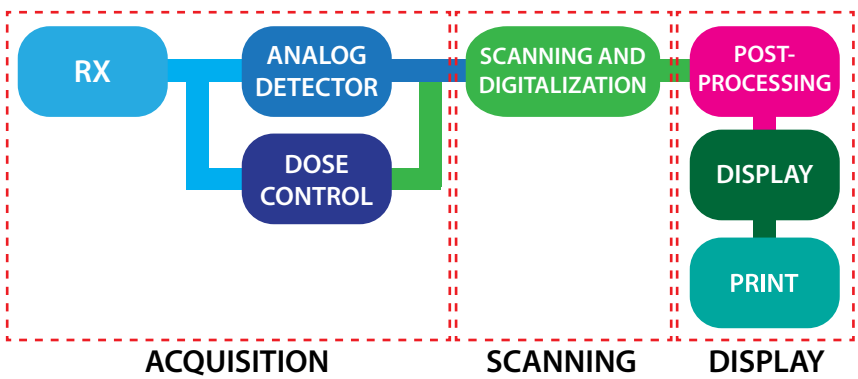


Figure 41: Image formation with CR systems. Image acquisition is analog, exposing a cassette containing an imaging plate by a conventional X-ray unit, the imaging plate is scanned after the exposure, and a digital image is produced, processed, and displayed.

Image Quality and CR systems

The general philosophy followed by CR manufacturers was to propose a general purpose technology, able to provide digital images of any anatomical part with an "adequate" image quality while avoiding under- and over-exposures. This was done by adjusting both scanning conditions and post-processing parameters, as a function of the anatomical part, and keeping the image contrast at a preset level, assumed to be consistent with what radiologists expect to obtain. In fact, normally, CR manufacturers have no control on the X-ray unit, and, as such, on the technique factors; if no work is done during system commissioning to match the CR output with the automatic exposure control of the mammography equipment, the two system components act independently one to each other, and either under- or over-exposures can very likely occur. To address this lack of control on the exposure parameters, CR manufacturers developed IP scanners able to adjust scanning parameter according to the overall radiation dose absorbed by the imaging plate during the exposure; in case of under-exposure they amplify the IP signal more, in case of over-exposure they amplify it less, with a final image which appears almost unchanged, even for very different dose levels; moreover, they developed post-processing algorithms applying multiple filters, with the final target of providing an image which is "apparently good" for radiologists, with several margins of customization. The high number of degrees of freedom offered by such "tools" (adaptive scanning and multi-parametric post-processing), reduces the user's control on the CR systems, and induces large variability among settings of different systems of the same type. There is a finite risk that the CR field engineer installs the CR system without communicating with his X-ray unit colleague, with the result of keeping the same calibration of the AEC as used for the previous screen-film combination, and consequent possible systematic under- or over-exposure.

Image Quality and CR systems

Mammography clinical performance with CR systems

Recently, some concerns about clinical performance of CR mammography have been arisen, some results having shown a decrease in cancer detection rate by screening mammography with CRs, compared to the cancer detection rate obtained by screen-film combinations and with integrated digital systems (DRs). Results published from the Ontario Breast Screening Program showed that the cancer detection rate was similar for DR (0.49%) and screen-film (0.47%), while it was significantly lower (0.34 %) for CRs, and the authors concluded that screening programs should monitor the performance of computed radiography separately and may consider informing lower cancer detection rates [33]. Differences between DRs and CRs in favour of DRs can be easily obtained from physics measurements, in particular in terms of spatial resolution, and noise transfer properties, as well illustrated by Martin Yaffe and colleagues in a recent publication [34]. This means that, at the same dose levels and using physical parameters like modulation transfer function (MTF) and detective quantum efficiency (DQE), tested CRs perform worse than tested DRs [34].

A different point of view is that, given the inherent physical differences between CR and DR technology, those two types of mammography systems need different dose levels to achieve comparable results. Results published by Hilde Bosmans and colleagues about the Flemish breast screening program showed that screening performance indicators, i.e. cancer detection rate and recall rate, were comparable for CRs and DRs, but CRs needed on the average 60% extra-dose. The authors conclude that both CRs and DRs can be employed in screening mammography, taking care to calibrate the automatic exposure control of mammography units used with CR systems to achieve an image quality level similar to that obtained by DR systems [35].

Image Quality and CR systems

Things one should know to use CR mammography

- CRs are not fully digital: in fact, they use analog detectors (imaging plates, IPs) and the digital image is obtained after the exposure by digitizing the signal stimulated from the IP by laser scanning.
- CRs do not use one detector only, but a number of cassettes, each with one IP are used, reproducing the same workflow as used in conventional mammography. Care should be taken in limiting the number of IPs associated to a same CR scanner, securing a periodical test of each individual IP in terms of efficiency and absence of defects/artifacts. IP cassettes cannot be used forever, but they need to be replace in case their sensitivity becomes too low (or need too high dose to provide a given signal level), or in case of relevant artifacts.
- CRs are not integrated systems: in fact, contrary to DRs, IPs are exposed by a conventional X-ray unit (the same used with screen-film combinations), and the automatic exposure control is the one available on such unit. This means that, in absence of proper actions, the probability of having the "optimal exposure conditions" to achieve the target image quality is very low.
- There are risks of misuse of CRs: possible suboptimal exposure conditions are not easily detectable at a glance because of the multiple tools implemented by CR manufacturers to improve the "apparent image quality", the clinical information in each breast being unknown a priori.
- Results derived from screening experience show that CRs need more radiation dose than DRs to achieve comparable clinical results, and this is consistent with physical differences between the two technologies.
- Care should be taken by both Medical Physics Experts (MPEs) and Radiologists when CR mammography systems are commissioned. The appropriateness of the automatic exposure control setting should be periodically checked and constancy of image quality over time verified.

Image Quality

REFERENCES

1. C Pater (2004) Equivalence and non-inferiority trials – are they viable alternatives for registration of new drigs ? (III), Curr Control Trial Cardiovasc Med 5:6-14.
2. D R Dance, C L Skinner, K C Young, J R Beckett, C J Kotre (2000) Additional factors for the estimation of mean glandular dose using the UK mammography dosimetry protocol , Phys Med Biol 45:3225-3240.
3. Y Lee, S-Y Tsai, N Shinohara (2010), Computerized quantitative evaluation of mammographic accreditation phantom images, Med Phys 37:6323-6331.
4. M Asahara, Y Kodera (2012) Computerized scheme for evaluating mammographic phantom images, Med Phys 39:1609-1617.
5. R Sharma, S D Sharma, Y S Mayya (2012) A survey on performance status of mammography machines: image quality and dosimetry studies using a standard mammography imaging phantom, Rad Prot Dosim 150:325-333.
6. G Gennaro, F Ferro, G Contento, F Fornasin, C di Maggio (2007) Automated analysis of phantom images for the evaluation of long-term reproducibility in digital mammography, Phys Med Biol 52:1387-1407.
7. P E Udrill, A D O'Kane, F J Gilbert (2000) A comparison of digital and screen-film mammography using quality control phantoms, Clin Radiol 55:782-790.
8. K R Bijerk, M A O Thijssen, TH J M Arnoldssen (2000), Modification of the CDMAM contrast-detail phantom for image quality evaluation of full-field digital mammography systems, Proceedings of the 5th International Workshop on Digital Mammography (IWDM), Toronto, Canada, Jun 11-14, 2000.
9. W J H Veldkamp, N Karssmeijer (1998), Accurate segmentation and contrast measurement of microcalcifications in mammograms: a phantom study, Med Phys 25:1102-1110.
10. K C Young, A Alsager, J M Odoku, H Bosmans, B Verbrugge, T Geertse, R van Engen (2008) Evaluation of softare for reading images of the CDMAM test object to assess digital mammography systems, Proc. of SPIE vol 6913, Medical Imaging 2008: Physics of Medical Imaging.
11. H de las Heras Galas, F Schöefer, B Tiller, M Chevalier, G Zwettler, F Semturs (2013), A phantom using titanium and Landolt rings for image quality evaluation in mammography, Phys Med Biol 58:L17-L30.
12. S Fränkel, K Wunder, U Heil, D Groß, R Schulze, U Schwanecke, C Düber, E Schömer, O Weinheimer (2013), Total variation regularization in digital breast tomosynthesis, Bildverarbeitung für die Medizin 2013, 356-361.
13. R. E. Hendrick et al. (1999) "Mammography quality control manual" American College of Radiology, Revised edition,

REFERENCES

1999.

14. *W Huda, A M Sajewicz, K M Ogden, E M Scalzetti, D R dance (2002), How good is the ACR accreditation phantom for assessing image quality in digital mammography ?, Acad Radiol 9:764-772.*
15. *G Gennaro, L Katz, H Souchay, C Alberelli, C di Maggio (2005), Are phantoms useful for rpedicting the potential of dose reduction in full-field digital mammography ?, Phys med Biol 50:1851-1870.*
16. *K C Young, J H Cook, J M Odoku, H Bosmans (2006), Comparison of software and human observers in reading images of the CDMAM test object to assess digital mammography systems, roc. of SPIE vol 6142, Medical Imaging 2006: Physics of Medical Imaging.*
17. *M Figl, F Semturs, M Kaar, R Hoffmann, H Kaldarar, P Homolka, G Mostbeck, B Scholz, J Hummel (2013), Dose sensitivity of three phantoms used for quality assurance in digital mammography, Phys Med Biol 58:N13-N23.*
18. *K W Brooks, J H Trueblood, K J Kearfott, D T Lawton (1997), Automated analysis of the American College of Radiology mammographic accreditation phantom images, Med Phys 24:709-723.*
19. *L Blot, A Davis, M Holubinka, R Martí, R Zwiggelar (2002), Automated quality assurance applies to mammographic imaging, EURASIP J App Sign Proc 7:736-745.*
20. *P Mayo, F Rodenas, G Verdú, J I Villaescusa, J M Campayo (2004), Automatic evaluation of the image quality of a mammographic phantom, Comp Meth Prog Biomed 73:115-128.*
21. *H R Blackwell (1946), Contrast threshold of human eye, J Opt Soc Amer 36:624-643.*
22. *M S Rzeszotarski (1999), Counting statistics, Radiographics 19:765-782.*
23. *A Rose (1948), The sensitivity performance of the human eye on an bsolute scale, J Opt Soc Amer 38:196-208.*
24. *A E Burgess (1998), The Rose model revisited, J Opt Soc Amer A 16:633-646.*
25. *A E Burgess (1995), Comparison of receiver operating characteristic and forced choice observer performance measurement methods, Med Phys 22:643-655.*
26. *H Urbanczyk, E McDonagh, NW Marshall, I Castellano (2012), Characterization of the effects of the Fine View algorithm for full field digital mammography, Phys Med Biol 57:1987-2003.*
27. *R T Droege (1983), A practical method to routinely monitor resolution in digital images, Med Phys 10:337-343.*
28. *R T Droege, M S Rzeszotarski (1985), An MTF method immune to aliasing, Med Phys 12:721-725.*
29. *European Guidelines for Breast Cancer Screening and Diagnosis, 4th edition, 2006.*

Image Quality

REFERENCES

30. *W A Schewart, "Statistical Method from the Viewpoint of Qaultiy Control", Dover Publications, New York, 1939, 1986.*
31. *M Sonoda, M Takano, J Miyahara, H Kato (1983), Computed radiography utilizing scannig laser stimulated luminescence, Radiology 148:833838.*
32. *K A Fetterly, B A Schueler (2003), Performance evaluation of a "dual-side read" dedicated mammography computed radiography system, Med Phys 30:1743-1854.*
33. *A M Chiarelli, , S A Edwards, M V Prummel, D Muradali, V Majpruz, S J Done, P Brown, R S Shumak, M J Yaffe (2013), Digital compared with screen-film mammography: performance measures in concurrent cohorts within an organized breast screening program, Radiology 268:684-693.*
34. *M J Yaffe, A K Bloomquist, D M Hunter, G E Mawdsley, A M Chiarelli, D Muradali, J G Mainprize (2013), Comparative performance of modern digital mammography systems in a large breast screening program, Med Phys 40: 121915; doi: 10.1118/1.4829516.*
35. *H Bosmans, A De Hauwere, K Lemmens, F Zanca, H Thierens, C Van Ongeval, K Van Herck, A Van Steen, P Martens, L Bleyen, G Vande Putte, E Kellen, G Mortier, E Van Limbergen (2103), Technical and clinical breast cancer screening performance indicators for computed radiography versus direct digital radiography, Eur Radiol 23:2891-2898.*

Quality Controls

Monitor

Monitor

In digital mammography, the two processes of image acquisition and display are physically separated. Such separation constitutes one of the main general benefits of digital mammography, as the acquired images can be further processed and their display mode optimized and adjusted according to individual preferences. However, the separation between image acquisition and display also implies that it is not sufficient to evaluate the performance of the acquisition system and make efforts to optimize acquired image quality, without accounting for the display system. Digital mammograms can be either displayed on monitors with appropriate technical specifications or printed on films; however, monitors will be considered the "natural output" of digital images and all future discussion will be focused on them, with no reference to film printers.

Tests described in this chapter are applicable to both liquid crystal displays (LCD) and cathode ray tube (CRT) displays used for the interpretation of medical images ('diagnostic' display devices) with the review workstation (RWS), although CRT monitors are no longer used in new systems. The tests are not mandatory for other display devices like acquisition workstation (AWS) monitors.

Displays for mammography images have special requirements in terms of spatial and contrast resolution compared to displays for other imaging modalities. As a standard mammography examination is done by two views per breast (a cranio-caudal, CC, and a medio-lateral oblique, MLO), mammography displays are used in pairs, with portrait format, allowing the evaluation of the current examination as well as comparison with the previous mammograms.

Screen size and spatial resolution should be chosen to match image inherent resolution, allowing for it being viewed at the standard distance which is typically 30 cm. In mammography, display size is typically 53 cm (diagonal) with an active area by 33x42 cm^2, and spatial resolution about 2000x2500 pixel, often indicated as "5MP". The other important features are monitor contrast and contrast resolution capability. They depend on monitor performance in producing and modulating a light signal (obtained by a cathode ray tube or modulating the light emitted by a lamp for LCDs) which determines multiple different gray tones on its surface. The signal level at the monitor surface is measured by a parameter called "luminance" (L), measured in SI units of candela per square meter (cd/m2). The higher the brightness the higher the luminance measured. In mammography, high luminance displays are required in order to maximize the displayed contrast of breast images, which are inherently limited in contrast.

However, since the human vision is an adaptive system, i.e. is able to adjust its sensitivity to each specific luminance condition, the luminance ratio (ratio between maximum and minimum luminance) is a quality factor which

Monitor

characterized displays better than the maximum luminance. In mammography, the luminance ratio should be at least 250.

As medical images are interpreted by human observers, the choice of monitor characteristics requires consideration of both technique factors and performance of the human visual system. The response of the human visual system is measured in terms of the smallest perceivable difference, or just noticeable difference (JND), defined as the perceivable change in luminance for a given luminance level (ΔL/L or dL/L). Threshold contrast (dL/L) is nonlinear, according to Barten's model [1-3], and decreases while luminance decreases (Figure 1); for this reason, high-luminance monitors are required in mammography.

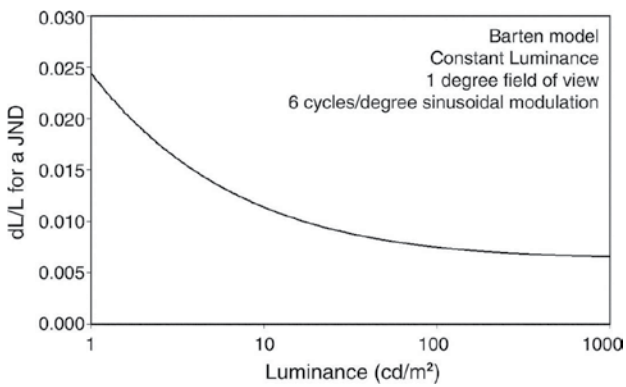


Figure 1 : Response of the average human visual system as a function of the background luminance.

Ideally, the luminance response of a display device should match this nonlinear response of the human visual system such that image values are displayed in equally perceptible luminance increments. The DICOM grayscale standard display function (GSDF) offers a way to approach this goal by applying a specific look-up-table to the display values, such that the display values present equally discriminable levels of brightness [2]. GSDF is the standard calibration curve for medical displays and its verification is important, especially in mammography, to ensure image consistency when breast images are displayed on different workstations and monitors, for example moving from the mammography RWS to a PACS WS.

Monitor

Tests proposed in this chapter are extracted from the report published in 2005 by the Task Group 18 (TG18) of the American Association of Medical Physics (AAPM), "Assessment of display performance for medical imaging systems" (https://www.aapm.org/pubs/reports/OR_03.pdf). They are limited to the contrast factor, the verification of calibration according to the GSDF, and the uniformity.
Monitor tests are performed displaying test patterns (electronic phantoms), which can be downloaded from the TG18 website http://deckard.mc.duke.edu/~samei/tg18#_DOWNLOAD_THE_TG18. The 2k resolution version is good for mammography displays. If the RWS with the monitors you are testing belongs to a hospital network equipped with an archiving system (PACS), it is strongly suggested to save the test patterns in the image archive, to be reused whenever you need.

- **Prior to carrying out tests, monitors should be warmed up for approximately 30 minutes. Dirt and smudges on the face plate of the monitor may lead to incorrect results due to the emitted light being absorbed, reflected of refracted, therefore monitors should be cleaned according to manufacturers' instructions, if necessary, before testing.**
- **The quality tests in this chapter are highly sensitive to ambient light, therefore all of them should be performed under "optimal clinical conditions" (room light low, external intense light sources emitting in the monitor direction should be avoided). The ambient light level should be low (illuminance < 20 lux) for an optimal perception particularly with regard to low contrast objects.**
- **All test patterns should be displayed at full resolution 1:1 (exactly one display pixel for each pixel in the digital image).**
- **Contrast and brightness buttons on the monitors should be disabled after monitor calibration, to avoid any arbitrary change to the monitor response. If this is not possible, then personnel should be asked not to touch those buttons.**
- **The hospital cleaning staff should be instructed only to clean the surface of the devices with the proper cleaning agent and not touch any of the adjusting buttons to avoid changes in contrast or brightness of the monitor.**

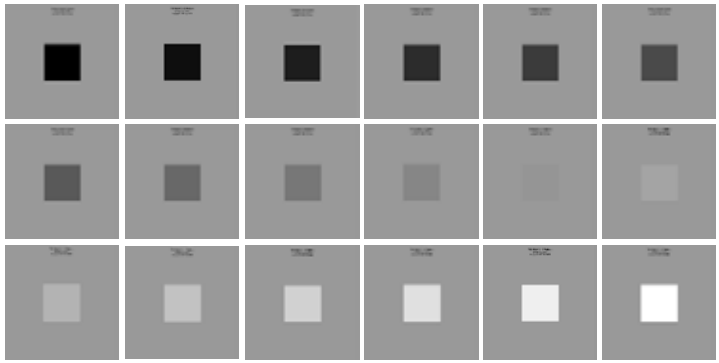
Calibration

PURPOSE

TO VERIFY THE COMPLIANCE OF MONITOR CALIBRATION WITH THE DICOM GREYSCALE STANDARD DISPLAY FUNCTION (GSDF) AND CALCULATE THE CONTRAST RATIO.

EQUIPMENT

CALIBRATED LUMINANCE METER
(range 0.05-1000cd/m^2 and accuracy <5%)



TG18LN test patterns

TEST FREQUENCY

- Acceptance/Commissioning.
- After possible monitor replacement.
- Annual

Calibration

PROCEDURE

Luminance measurements and data input

1. Load the TG18-LN12-i (with i = 1, 2, …18) test patterns.
2. If possible, display the test patterns one-by-one using the same software as for displaying clinical images.
3. Otherwise, some RWSs are equipped with specific software packages also used by manufacturers for monitor calibration, allowing use of the AAPM test patterns. They can also be used for calibration verification.
4. Display the first image of the TG18-LN test patterns, TG18-LN12-01, showing a black square in the middle on the left monitor.
5. Place the luminance meter in the center of the square and measure luminance. Care should be taken, especially with the first image (minimum luminance), to wait until the luminance meter provides a "stable" number.
6. Enter the luminance value in the third column of Table 1 of the "Monitor – GSDF" worksheet in the "Template_ EFOMP_MammoWG_DR/CR" Excel file.
7. Do the same for the following images TG18-LN12-02 - TG18-LN12-18.
8. Repeat the procedure displaying the TG18-LN test patterns on the right monitor and fill the fourth column of Table 1.

Calibration

Test Pattern	Grey Level (12 bit)	Left monitor luminance (cd/m^2)	Right monitor luminance (cd/m^2)
TG18-LN12-01	0		
TG18-LN12-02	240		
TG18-LN12-03	480		
TG18-LN12-04	720		
TG18-LN12-05	960		
TG18-LN12-06	1200		
TG18-LN12-07	1440		
TG18-LN12-08	1680		
TG18-LN12-09	1920		
TG18-LN12-10	2160		
TG18-LN12-11	2400		
TG18-LN12-12	2640		
TG18-LN12-13	2880		
TG18-LN12-14	3120		
TG18-LN12-15	3360		
TG18-LN12-16	3600		
TG18-LN12-17	3840		
TG18-LN12-18	4080		

Table 1 – Table for data input of monitor luminance measurements

Calibration

Example of data input

Test Pattern	Grey Level (12 bit)	Left monitor luminance (cd/m^2)	Right monitor luminance (cd/m^2)
TG18-LN12-01	0	0.57	0.55
TG18-LN12-02	240	1.61	1.31
TG18-LN12-03	480	3.28	2.65
TG18-LN12-04	720	5.53	4.45
TG18-LN12-05	960	8.67	7.01
TG18-LN12-06	1200	12.75	10.58
TG18-LN12-07	1440	18.16	15.30
TG18-LN12-08	1680	24.78	21.53
TG18-LN12-09	1920	33.42	29.64
TG18-LN12-10	2160	44.24	39.83
TG18-LN12-11	2400	57.81	52.75
TG18-LN12-12	2640	74.70	69.59
TG18-LN12-13	2880	95.83	90.25
TG18-LN12-14	3120	122.0	116.6
TG18-LN12-15	3360	154.5	148.6
TG18-LN12-16	3600	193.3	188.3
TG18-LN12-17	3840	241.1	239.0
TG18-LN12-18	4080	299.0	298.2

Table 2 – Example of data input for monitor luminance measurements

Calibration

Data output

1. The contrast factor is calculated in the "Monitor – GSDF" worksheet of the "Template_EFOMP_MammoWG_DR/CR" Excel file, as the ratio of the maximum and the minimum luminance values measured, corresponding to the TG18-LN12-18 and TG18-LN12-01 images, respectively. The test is passed if the contrast factor is above 250.
2. From the relationship between measured luminance and gray levels of the luminance test patterns, the monitor response in terms of "just noticeable difference curve" (JND) is calculated [REF DICOM sup 28], according to the following equation:

$$\begin{aligned} j(L) = & A + B \cdot \log_{10}(L) + C \cdot [\log_{10}(L)]^2 + D \cdot [\log_{10}(L)]^3 + \\ & + E \cdot [\log_{10}(L)]^4 + F \cdot [\log_{10}(L)]^5 + G \cdot [\log_{10}(L)]^6 + \\ & + G \cdot [\log_{10}(L)]^7 + I \cdot [\log_{10}(L)]^8 \end{aligned}$$

   with log(L) decimal logarithm of luminance and values of A, B, C, D, E, F, G, H and I coefficients reported in the document "DICOM Part 14: Greyscale standard display function" [REF]
3. Maximum and minimum values calculated from the previous equation, JND_{max} and JND_{min}, respectively, are used to compute the JND variation ΔJND and recalculate the response curve as luminance variations dL/L (displayed contrast) vs. gray levels.
4. Measured monitor response is compared with the GSDF.
5. A ± 10% difference interval versus the GSDF is tolerated.

In Table 3 all parameters necessary to compare the actual calibration curve to the GSDF are reported. The first two columns are a copy of the input data, the third column is the response function in JNDs obtained from the second column using the equation above, the fourth column is the same recalculated at regular JND intervals, the fifth column is the displayed contrast, the sixth one the luminance response for the DICOM GSDF, and the last three columns the DICOM GSDF and the curves at ± 10% of the DICOM GSDF.

Calibration

Grey Level	L_{means} (cd/m^2)	JND_{meas}	JND_{calc}	$(dL/L)_{meas}$	L_{DICOM} (cd/m^2)	GSDF	GSDF-10%	GSDF+10%

Table 3 – Table for data output of monitor luminance calibration

Calibration

Example of data output

Grey Level	L_{means} (cd/m^2)	JND_{meas}	JND_{calc}	$(dL/L)_{meas}$	L_{DICOM} (cd/m^2)	GSDF	GSDF-10%	GSDF+10%
0	0.33	34.62	34.62		0.33			
240	0.97	70.25	69.74	**0.985**	0.96	**0.975**	**0.877**	**1.072**
480	2.07	105.88	104.86	**0.724**	2.03	**0.718**	**0.646**	**0.789**
720	3.80	142.30	139.98	**0.589**	3.67	**0.575**	**0.517**	**0.632**
960	6.17	177.00	175.10	**0.475**	6.02	**0.486**	**0.437**	**0.534**
1200	9.68	214.02	210.22	**0.443**	9.27	**0.425**	**0.383**	**0.468**
1440	14.15	248.84	245.34	**0.375**	13.64	**0.382**	**0.344**	**0.420**
1680	20.21	284.53	280.46	**0.353**	19.44	**0.350**	**0.315**	**0.385**
1920	28.02	319.69	31.58	**0.324**	27.00	**0.326**	**0.293**	**0.358**
2460	38.46	355.90	350.70	**0.314**	36.78	**0.307**	**0.276**	**0.337**
2400	51.35	390.67	385.82	**0.287**	49.35	**0.292**	**0.263**	**0.321**
2640	67.24	424.46	420.94	**0.268**	65.40	**0.280**	**0.252**	**0.308**
2880	86.94	457.79	456.06	**0.256**	85.79	**0.270**	**0.243**	**0.297**
3120	113.50	493.41	491.18	**0.265**	111.64	**0.262**	**0.236**	**0.288**
3360	145.40	527.37	526.31	**0.246**	144.29	**0.255**	**0.230**	**0.281**
3600	186.10	561.94	561.43	**0.246**	185.45	**0.250**	**0.225**	**0.275**
3840	236.50	596.12	596.55	**0.239**	237.25	**0.245**	**0.221**	**0.270**
4080	302.30	631.67	631.67	**0.244**	302.35	**0.241**	**0.217**	**0.265**

Table 4 – Example of data output for monitor luminance calibration

Calibration

The calibration curve is plotted in figure 2, together with the DICOM GSDF and the tolerance interval, showing that the monitor is properly calibrated.

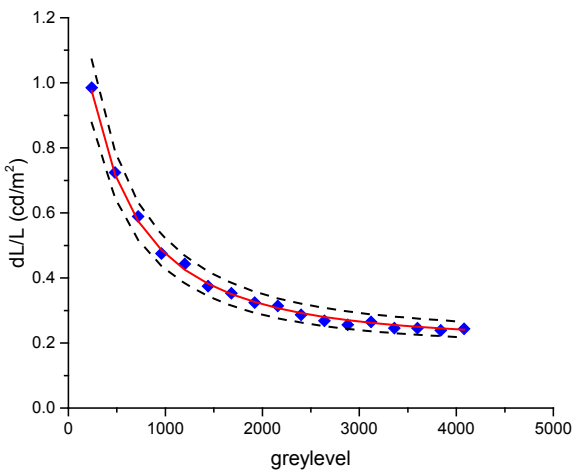


Figure 2 –Example of results of monitor calibration test.

LIMITING VALUES

Consistency with the DICOM GSDF should keep within ±10%.

In some cases, it might become difficult to adjust the curve to have the first point (black, very low luminance) inside the interval.
If this is the only "slightly uncalibrated" point then the display can be used; if there are multiple points then the manufacturer should be asked to perform a recalibration.

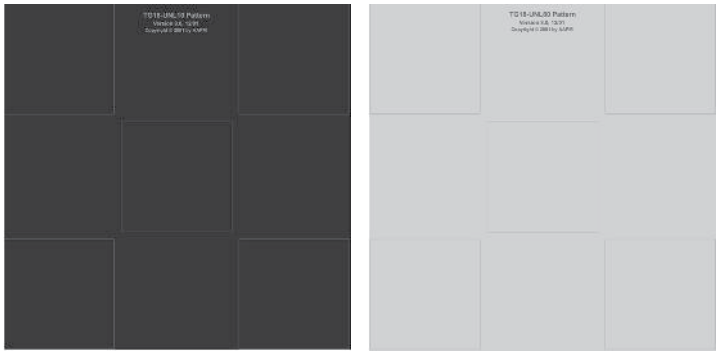
Uniformity

PURPOSE

TO VERIFY THAT THE DISPLAY LUMINANCE RESPONSE IS UNIFORM OVER THE ENTIRE MONITOR SURFACE.

EQUIPMENT

CALIBRATED LUMINANCE METER
(range 0.05-1000cd/m^2 and accuracy <5%)



TG18UNL test patterns

TEST FREQUENCY

- Acceptance/Commissioning.
- After possible monitor replacement.
- Annual

Uniformity

Luminance measurements and data input

1. Load the TG18-UNL10 and TG18-UNL80 test patterns.
2. If possible, display the test patterns one-by-one using the same software as for displaying clinical images.
3. Otherwise, some RWSs are equipped with specific software packages also used by manufacturers for monitor calibration, allowing use of the AAPM test patterns.
4. Display the TG18-UNL10 on the left monitor, showing a uniform dark luminance level with five squares in the middle and at the four corners, respectively.
5. Place the luminance meter in the center of each square and measure luminance. Care should be taken to wait until the luminance meter provides a "stable" number.
6. Enter the luminance values in Table 5 of the "Monitor – Uniformity" worksheet in the "Template_EFOMP_MammoWG_DR/CR" Excel file.
7. Do the same for TG18-UNL80.
8. Repeat the procedure displaying the TG18-UNL test patterns on the right monitor.

Uniformity

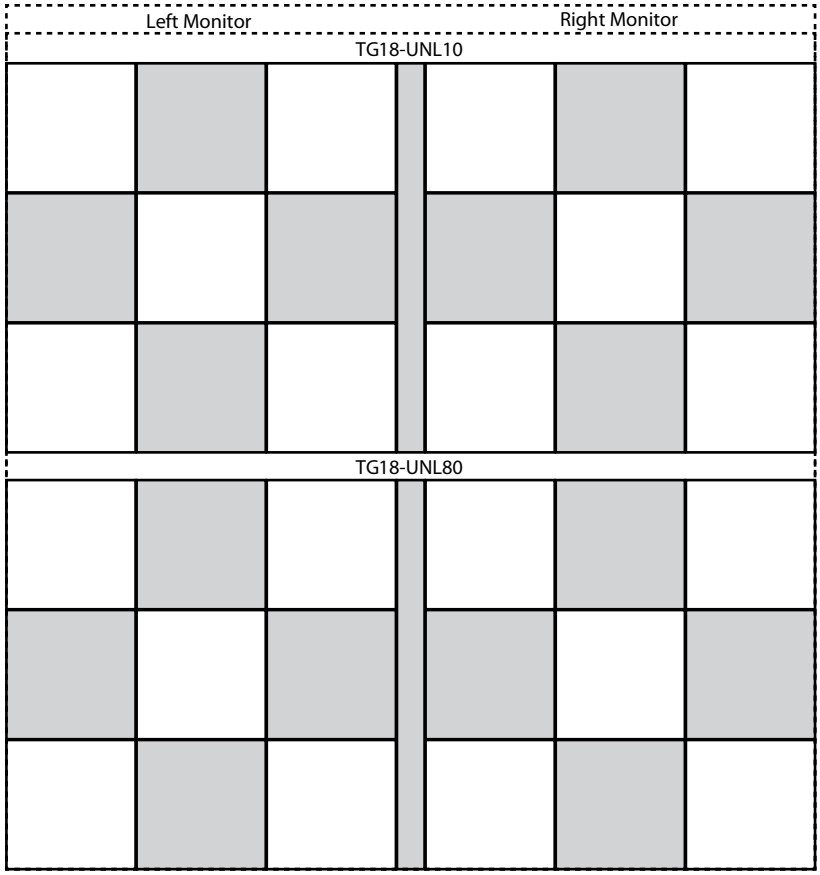


Table 5 – Table for data input of the monitor uniformity test

Uniformity

Example of data input

Left Monitor			Right Monitor		
TG18-UNL10					
1.67		1.64	1.64		1.67
	1.74			1.85	
1.61		1.61	1.62		1.64
TG18-UNL80					
124.4		121.1	115.1		115.5
	132.2			134.2	
118.0		114.0	111.2		116.1

Table 6 - Example of data input for the monitor uniformity test. Luminance measurements are in units of cd/m^2.

Uniformity

Data output

1. Uniformity is calculated as the relative difference between the central and each peripheral luminance value and is reported in Table 7.
2. The maximum deviation from the central luminance is taken as the uniformity index.

Left Monitor				Right Monitor	
TG18-UNL10					
%Δ		%Δ	%Δ		%Δ
%Δ		%Δ	%Δ		%Δ
TG18-UNL80					
%Δ		%Δ	%Δ		%Δ
%Δ		%Δ	%Δ		%Δ

Table 7 – Table for data output of the monitor uniformity test.

Uniformity

Example of data input

Left Monitor				Right Monitor	
TG18-UNL10					
4.02%		5.75%	11.35%		9.73%
7.47%		7.47%	12.43%		11.35%
TG18-UNL80					
5.90%		8.40%	14.23%		13.93%
10.74%		13.77%	17.14%		13.49%
Left	Δmax=	13.77%	Right	Δmax=	17.14%

Table 8 – Example of data output for the monitor uniformity test.

LIMITING VALUES

The maximum deviation from the central luminance should be below 30%.

Uniformity

REFERENCES

1. *P G Barten (1992) Physical model for the contrast sensitivity of the humaneye, Proc SPIE1666:57-72.*
2. *E Samei (2005) Technological and psychophysical considerations for digital mammographic displays, Radiographics 25:491-501.*
3. *K A Fetterly, H R Blume, M J Flynn, E Samei (2008) Introduction to greyscale calibration and related aspects of medical imaging grade liquid crystal displays, J Digit Imaging 21:193-207.*

Definitions of terms

TERM	DEFINITION
Absorbed dose	Absorbed dose, is the quotient of dE by dm, D = dE/dm, where dE is the mean energy imparted by ionizing radiation to matter in a volume element and dm is the mass of matter in the volume element. The measurement unit for absorbed dose in the SI system is the gray (Gy).
Acceptance test	The process of verifying that the contractor has supplied all of the equipment specified and has performed adequate tests to demonstrate that the specified requirements in the contract have been met.
Accuracy	Term that indicates how close the measured value of a quantity is to the true value. Typically, accuracy is used to check the correspondence between nominal and measured values of a parameter. The nominal value is taken as true value. The accuracy is calculated as relative difference between measured and true value.
Air kerma	Air kerma, is the quotient of dE_{tr} by dm, $K_a = dE_{tr}/dm$, where dE_{tr} is the sum of initial kinetic energies of all the charged ionizing particles liberated by uncharged ionizing particles in a mass of air dm. The measurement unit for air kerma in the SI system is the gray (Gy). Air kerma measures, employing an ionization chamber or another dose detector calibrated in mammography energy range, can be used to evaluate the Incident air kerma (K_i) or Entrance Surface Air Kerma (ESAK).
Antiscatter grid	Device positioned close to the entrance surface of an image receptor for reducing the amount of scattered radiation reaching the receptor.
Artifacts	Artifacts are image "features" not produced by the imaged object. They might be lines or streaks or dots, clearly not belonging to any imaged breast, but, in some cases, they might also resemble very closely clinical features, like microcalcification clusters or small masses.

Definitions of terms

TERM	DEFINITION
Automatic Exposure Control (AEC)	Device designed to select parameters which control detector dose, i.e. anode/filter combination, kV_p, and mAs, according to the beam attenuation by each individual breast, with the aim of keeping image quality high and constant.
Average (Mean) Glandular Dose (AGD or MGD)	The mean absorbed dose in the glandular tissue (excluding skin) in a uniformly compressed breast of, for example, 50% adipose, 50% glandular tissue composition. The reference breast thickness should be specified.
Baseline (level or value)	Value of a parameter defined on the basis of many repeated measurements (at least 10), that can be considered typical for a system. Generally, the baseline level is used when absolute limits for a parameter do not exist.
Bit depth	The number of bits used to digitize the detector signal, giving rise to the number of digital signal levels (grey scale). For a bit depth of n, the number of possible grey levels is equal to 2^n (i.e. 12 bits = 4096 shades of grey). Bit depth cannot be changed after equipment is purchased and is a vendor specific system characteristic.
Brightness	Quantity of light emitted by or reflected from any surface.
Clinical image quality (CIQ)	Image quality evaluated using clinical images, usually interpreted by clinicians.
Coefficient of variation (COV)	Standard deviation (SD) divided by the absolute value (mean value) of the measured quantity value. COV = SD/mean value.
Collimation	The process of restricting and confining an X-ray beam to a given area.

Definitions of terms

TERM	DEFINITION
Commissioning	The set of tests carried out by the customer's representative to ensure that the equipment is ready for clinical use and to establish baseline values against which the results of subsequent routine tests can be compared.
Compression paddle	Thin, rectangular or oval shaped device, made of plastic material (typical PMMA or polycarbonate) that can be positioned parallel to and above the breast table of a mammography apparatus. It is intended to compress the breast during X-ray examination.
Computed radiography (CR)	Digital radiology technology using photostimulable phosphor plates.
Contrast-detail phantom	A test object used in the assessment of imaging systems, which employs details of different sizes and contrasts.
Contrast resolution	The smallest relative exposure change that can be usefully imaged by a system. Ultimately, contrast resolution is limited by the dynamic range and the quantization (number of bits per pixel) of the digital detector.
Cranio-caudal projection	A routine view for mammography. The detector system is placed caudal (below the breast) and the vertical X-ray beam is directed from cranial to caudal direction (downward through the breast).
Del	Discrete element in a detector employed in Digital Radiology (DR) systems.
Detector correction	Correction in DR systems in which the pixel value of any defective detector elements is reconstructed and in which the detector is corrected for individual detector element sensitivity variations and electronic gain of the read-out.
Detector dose index	A parameter associated with CR and DR systems that is intended to provide an indication of the absorbed dose at the imaging detector of CR or DR reader.

Definitions of terms

TERM	DEFINITION
DICOM (Digital imaging and Communications in Medicine)	International standard for handling, storing, printing and transferring medical image data and patient information
Digital mammography	Radiographic images of the breast performed with digital detector that samples a finite number of locations and produces an electronic signal for each location. The magnitude of each signal is related to the transmission of X rays through the breast, and is digitized and stored in computer memory.
Digital radiography (DR)	Digital radiology technology using sealed units mounted on a radiography system, which captures X-rays and produces a digital image by sampling the X-ray image.
Display	see Monitor
Display area	Display area specifies the physical size of the active image display area. The display area of a display device is measured as the diagonal length of the active display area or horizontal width and vertical height of the display area along with the diagonal dimension.
Dose detector	Instrument for dose measurements, usually ionization chambers or solid state detectors.
Display noise	Any high-frequency fluctuations/patterns (< 1 cm) that interfere with the detection of the true signal. It influences the detectability of small objects and objects of low contrast in medical images.
Display (spatial) resolution	A quantitative measure of the ability of a display system to spatially resolve different points within the object.
Dynamic range	The range of exposures over which a detector can acquire image data in a single image. Typical digital systems will respond to exposures as low as 1 μGy and as high as 50 mGy outside the breast.
Electronic noise	Noise component which originates from various electronic sources.

Definitions of terms

TERM	DEFINITION
Entrance surface air kerma (ESAK)	The air kerma at a point in a plane corresponding to the entrance surface of a specified object, e.g. a patient's breast or a standard phantom. The radiation incident on the object and the backscattered radiation are included.
Exposure	The act of initiating and producing x-radiation from an X-ray unit.
Exposure data	see Exposure parameters
Exposure index (EI)	see Detector dose index
Exposure meter	see Dose detector
Exposure mode	MANUAL: Anode/filter combination, kV_p and mAs are manually selected by the operator, among a discrete number of values. SEMI-AUTOMATIC: Anode/filter combination and kV_p are manually selected by the operator, while the mAs value is automatically determined when the signal recorded by the exposure meter (AEC) achieves a given threshold. (FULL) AUTOMATIC: All the exposure parameters (anode/filter combination, kV_p and mAs) are automatically selected (after a short pre-exposure used to determine the object absorption peak, or depending on compressed breast thickness, or using both parameters).
Exposure parameters	Set of parameters used to control X-ray spectrum. It includes the combination of anode material, filter, KV_p and mAs. Selections are made based on breast size, density , image receptor type etc., to optimize contrast sensitivity with respect to radiation dose. The selections are made either manually by the operator or automatically by some systems.
Flat-field analysis	Analysis of the homogeneity of the detector field.

Definitions of terms

TERM	DEFINITION
Focal spot	Area on the surface of an X-ray tube anode (target) that is bombarded by the high-energy electrons from the cathode and where the x-radiation is produced.
Geometric distortions	An aberrations that cause the displayed image to be geometrically dissimilar from the original image. The practical consequences of such distortions affect the relative sizes and shapes of image features, particularly for larger displays or large deflection angles.
Ghost image	Residuals of previous images visible on the current image.
Global uniformity (GU)	Maximum deviation between th mean pixel value in region of interest (i,j), $MPV_{i,j}$, and the mean pixel value measured from the entire image, MPV_{image}: $$GU = \max\left(\frac{|MPV_{i,j} - MPV_{image}|}{MPV_{image}}\right) \leq 0.10$$
Gray level	see Bit depth
Grayscale Standard Display Function in DICOM	Standardized display function for grayscale images. It is intended to allow images transferred using the DICOM standard to be displayed on any DICOM-compatible display device with a consistent grayscale appearance. The consistent appearance of images is approached through perceptual linearization, where equal changes in digital values cause equal changes in perceived brightness.
Half value layer (HVL)	Thickness of absorber which attenuates the air kerma of X-ray beams by half. The absorber normally used to evaluate HVL of low energy X-ray beams, such as mammography beams, is high purity aluminum.

Definitions of terms

TERM	DEFINITION
Heel effect	The non-uniform distribution of air-kerma rate in an X-ray beam in a direction parallel to the cathode-anode axis. It causes decreasing signal intensity measurable on a image detector in the cathode-anode direction. It is due to the geometric setup of the X-ray tube.
Illuminance	Photometric term used to describe the rate at which visible light strikes a surface. It is often used to describe the amount of ambient lighting or the light striking a display surface. The unit of illuminance is lumen per meter squared (lm/m^2), or lux (lx). The luminous flux per unit area at any point on a surface exposed to incident light. The measurement unit for illuminance in the SI system is the lux (lx).
Illuminance meter	A device used to measure illuminance for quantitative assessment of display reflection and for monitoring ambient conditions.
Image detector	A device that detects and records the distribution of x- rays in order to form an image.
Imaging plate (IP)	The unexposed photostimulable phosphor detector is commonly known as an imaging plate.
Image quality	There is no uniformly accepted definition of image quality. Any general definition must address the effectiveness with which the image can be used for its intended task. It is possible to define quality indices representing the information content of the image. Usually, image quality "appropriateness" is assessed by physicists using test objects called phantoms, including "details", whose images can be either "rated" according to given criteria, or analyzed to measure objective parameters.
Image quality criteria	Criteria which characterize a level of acceptability for radiological images which allows a specific clinical question to be answered.

Definitions of terms

TERM	DEFINITION
Incident air kerma (Ki)	The air kerma at a point in a plane corresponding to the entrance surface of a specific object, e.g. a patient's breast or a standard phantom. Only the radiation incident on the object and not the backscattered radiation is included.
Just noticeable difference (jnd)	The smallest detectable difference between a starting and secondary level of a particular sensory stimulus. Since the visual system is not linear and has varying response depending on illumination levels, the video display must be set up so that similar changes in pixel values should be equally visible in both the dark and light regions of an image.
Limiting value	Maximum or minimum limit of a possible range, considered acceptable for a given parameter
Linear shift-invariant (LSI) system theory	A system which is both linear and shift-invariant. A system is linear if it obeys the principle of superposition: the response to a weighted sum of any two inputs is the (same) weighted sum of the responses to each individual input. A system is shift-invariant (also called translation-invariant for spatial signals, or time-invariant for temporal signals) if the response to any input shifted by any amount is equal to the response to the original input shifted by amount . Such theory provides a description of how the system (image detector) acts on the input signal to produce the output (image) using the concept of "transfer function".
Linearity	System response where the output increases in direct proportion to the input signal increase.

Definitions of terms

TERM	DEFINITION
Local uniformity (LU)	Local difference between each , $MPV_{i,j}$ and the average across the eight neighbours, $MPV_{neighbours}$: $$LU = \max\left(\frac{|MPV_{i,j} - MPV_{neighbour}|}{MPV_{neighbour}}\right) \leq 0.05$$
Luminance	A photometric quantity describing the light power per unit area per unit solid angle emitted by a light source. The SI unit for luminance is candelas per square meter (cd/m^2).
Luminance meter (Photometer)	A device used to measure the luminance of the display device.
Luminance response	The luminance response of a display device refers to the relationship between displayed luminance and the input values of a standardized display system.
Luminance uniformity	Luminance uniformity refers to the maximum variation in luminance across the display area when displaying a uniform pattern.
Mammography	The X-ray examination of the female breast. This may be undertaken for health screening of a population (mammography screening) or to investigate symptoms of breast disease (symptomatic diagnosis).
mAs value (tube loading)	Product of the X-ray tube current (milliampere, mA) and the exposure time (seconds, s).
Mean pixel value (MPV)	The mean value of the pixels in a region of interest.

Definitions of terms

TERM	DEFINITION
Medio-lateral oblique projection	A routine view for mammography. The detector system is placed lateral to the breast and the horizontal X ray beam is directed from medial to lateral aspect through the breast. A routine view for mammography. The detector system is placed inferolateral to the breast, and the x-ray beam is directed from superomedial to inferolateral with an angle usually 45° (from 30° to 60°).
Modulation transfer function (MTF)	A function, which describes how the contrast of image components is transmitted as a function of their spatial frequency content.
Monitor	Device used for viewing digital images. It is the actual physical unit that generates a visible image from analog or digital video signals.
Noise (image noise)	Fluctuations in pixel values which are unrelated to the imaged object. The standard deviation in a region of interest in the output image of a homegeneous object can be taken as measure of noise.
Noise power spectrum (NPS)	Function which describes image noise as a function of spatial frequency.
Object contrast	The inherent differences in X-ray attenuation in the object being imaged.
Patient dose or dose	Generic term for a variety of radiation dose quantities applied to a patient or group of patients.
Phantom	An object used to absorb and/or scatter radiation equivalent to that of a patient and hence to aid estimation of radiation doses and test imaging systems without actually exposing a patient. It may be an anthropomorphic or a physical test object.

Definitions of terms

TERM	DEFINITION
Photostimulable phosphor (PSP)	PSP is a material that stores absorbed X-ray energy in crystal structure "traps," and is sometimes referred to as a "storage" phosphor. This trapped energy can be released if stimulated by additional light energy of the proper wavelength by the process of photostimulated luminescence.
Pixel	Picture element, the smallest unit in the image.
Pixel value	Discrete value assigned to a pixel. In mammography systems the number of pixel values range from 1024 (10-bits) to 16384 (14 bits), depending on the detector.
Pixel value offset	For some systems a constant value is added to the values of all pixels. This constant value is defined as the pixel value offset.
Processed image	The image after image processing, ready for presentation on the monitor or print-out. In the DICOM file the value of tag Pixel Intensity Relationship (0028,1040) is 'for presentation'.
Poly-methyl methacrylate (PMMA)	The synthetic material poly-methyl methacrylate. Trade names include Lucite, Perspex and Plexiglas.
Quality Assurance (QA)	As defined by the WHO (1982): All those planned and systematic actions necessary to provide adequate confidence that a structure, system or component will perform satisfactorily in service. Satisfactory performance in service implies the optimum quality of the entire diagnostic process i.e., the consistent production of adequate diagnostic information with minimum exposure of both patients and personnel.
Quality Control (QC)	Part of quality assurance. The set of operations (programming, coordinating, implementing) intended to maintain or to improve quality. It covers monitoring, evaluation and maintenance at required levels of all characteristics of performance of equipment that can be defined, measured and controlled.

Definitions of terms

TERM	DEFINITION
Quantum limited imaging system	System in which quantum noise is the main component of image noise.
Quantum noise	Noise component related to the statistical process of interaction between incident X-ray photons and image detector
Radiation output	see X-ray tube output
Radiographic contrast	The difference of signal intensity between two adjacent elements of a radiographic image.
Region of interest (ROI)	Area of interest in an image, often circular or square, in which measurements are made.
Repeatability or precision (short-term reproducibility)	The variation (usually relative standard deviation) in observed values usually for a set of measurements made at about the same time
Reproducibility	The variation (usually relative standard deviation) in observed values usually for a set of measurements made over a period of time
Response function (detector)	A function that describes a response of image detector to a stimulus (input signal).
Screen-film mammography	Radiographic imaging of the breast performed with a high detail intensifying screen that is in close contact with the film in the cassette.
Signal-difference-to-noise ratio (SDNR)	SDNR (Signal Difference-to-Noise Ratio) is obtained from the difference between the mean pixel values, of background (MPV_{bkg}) and aluminum detail (MPV_{Al}), divided by the standard deviation of the background (SD_{bkg}) where the detail is embedded (background standard deviation), according to the following formula: $$SDNR = \frac{MPV_{bkg} - MPV_{Al}}{SD_{bkg}}$$

Definitions of terms

TERM	DEFINITION
Spatial frequency	A spatial pattern such as an image can be represented as the summation of a set of spatial sinusoidal functions of appropriate amplitudes, each sinusoid covering a specific distance (e.g. millimetres) per cycle. The spatial frequency is the reciprocal of that distance and is specified in cycles/millimetre (or mm^{-1})
Spatial resolution	The fineness of spatial detail that an imaging system can demonstrate. This can be measured from the image of a resolution pattern in terms of the number of line pairs per millimetre.
Standard breast phantom	A PMMA phantom to represent approximately the average breast (although not an exact tissue-substitute) so that the X-ray machine operates correctly under automatic exposure control and the dosimeter readings may be converted into dose to glandular tissue.
Structural noise	Noise component related to the structure of a detector.
Technical image quality (TIQ)	Image quality obtained by images of reproducible and known objects called test objects or phantoms.
Test object	see Phantom
Test pattern	Patterns used to evaluate the performance of display devices.
Threshold contrast	Contrast level that produces a just visible difference between an object and the background. It is usually evaluated by contrast-detail (C-D) phantoms, which include multiple details (typically discs) with variable contrast and size.
Tube voltage (kVp)	The potential difference in units of kilovolt peak (kV_p) applied across the anode and cathode of an X-ray tube during a radiographic exposure.

Definitions of terms

TERM	DEFINITION
Unprocessed image	The image of a DR system after flat-fielding and detector corrections but before other image processing has been applied. In the unprocessed image the pixel value is, in general, linearly releated to pixel exposure. In the DICOM file the value of tag Pixel Intensity Relationship (0028,1040) is 'for processing'. International Electrotechnical Commission (IEC) Maintenance Team (MT) 31 refers to the unprocessed image as 'raw data'.
Window width (WW)	Setting defining (together with window level) a linear relationship between modality pixel values and pixel values meaningful for presentation (presentation values).
Window level (WL)	Setting defining (together with window width) a linear relationship between modality pixel values and pixel values meaningful for presentation (presentation values).
Window and levelling settings	Settings that alter way the image data is displayed. Level (brightness) controls the offset level for the blackest part of the image, while window (contrast) controls the difference between light and dark areas.
X-ray beam quality	The spectrum of radiant energy produced by a given radiation source with respect to its penetration or its suitability for a specific application. The radiation quality of X-ray beams used for medical imaging can be characterized by a combination of various parameters. These include first half-value layer (HVL; symbol HVL_1); the second HVL (HVL_2); the ratio of HVL1 to HVL_2, referred to as the homogeneity coefficient; the tube voltage and the total filtration
X-ray generator	A device that provides the high voltage supply for X-ray unit.

Definitions of terms

TERM	DEFINITION
X-ray spectrum	Distribution of photon energies in an X-ray beam. It depends on anode and filter material and tube potential, as well on all attenuators (tube output window, compression device, air gap) between anode and breast.
X-ray tube (source)	Vacuum tube designed to produce X-rays by bombardment of the anode with a beam of electrons accelerated through a potential difference.
X-ray tube output	Ratio between air kerma (mGy) measured without any test object and the tube loading (mAs), for a known distance between the X-ray source and the dosimeter and for preset exposure parameters (anode/filter combination and kVp value).
X-ray unit	Assembly comprising a high voltage supply, an X ray tube with its protective housing and high voltage electrical connections.