

# **EUREF type test protocol**



**Version 1.2**

**November 2011**

**EUREF Physico-technical steering group:**

H. Bosmans  
R. van Engen  
P. Heid  
B. Lazzari  
S. Schopphoven  
M. Thijssen  
K. Young

Corresponding address:  
EUREF office  
[info@euref.org](mailto:info@euref.org)  
National Expert and Training Centre for  
Breast Cancer Screening  
Radboud University Nijmegen Medical Centre  
P.O. Box 6873  
6503 GJ Nijmegen  
The Netherlands

# 1. Introduction

The aim of the European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services (EUREF) is to improve the quality of mammography in Europe and to disseminate knowledge on high quality breast imaging. Within this context EUREF has produced European Guidelines. The current version of these guidelines is the fourth edition, published by the European Commission in 2006. A supplement of the fourth edition has been submitted to the EU in 2011.

Recently, manufacturers and users have expressed the need for type testing at a European level. For that reason EUREF decided to organise and standardise type testing for mammography systems at the European level. The standards applied in the EUREF type testing are based on chapter two of the fourth edition of the European Guidelines.

Type testing is defined as a test (1) to verify whether a type of system is able to pass the acceptability criteria of the European protocol, (2) to provide guidelines about the best practice in terms of dose and (clinical) image quality. After a successful type test, individual mammography units of the same and type brand still need to undergo an acceptance test before clinical use. Passing the type test only guarantees that the system is in principle capable of meeting the requirements of the European Protocol and the report may list suggestions for optimal use or conditions to be avoided in practice.

The physico-technical chapter of the fourth edition of the Guidelines has been written for acceptance testing and not for type testing. Therefore some small differences between the protocol for type testing and the European Guidelines are introduced. These differences are described in this document.

In this document it is described how type testing is organised and the physico-technical evaluation and the clinical evaluation are specified in detail.

Type tests are currently performed on digital mammography units (DR and CR systems). In future type testing may be performed on image processing algorithms, workstations and film digitisers.

## 2. Type testing procedure

A type test will consist of:

- Phase 1: Technical evaluation:

Two full physico-technical evaluations will be performed on different systems at different locations. The applicant for the test (i.e. the manufacturer) arranges the locations and all practicalities in co-operation with the EUREF physico-technical steering group. The system must be available for the tests for at least three days. The tests are primarily based on the measurements in chapter 2b of the European Guidelines, fourth edition and the Supplement to the European Guidelines. However, some tests are adapted and some additional tests will be performed, which are described in chapter 3. After all results are available, a meeting will be organized to discuss the results and decide whether phase 2: the clinical evaluation is allowed to start. A report with the findings of the technical evaluation will be compiled and the applicant will be asked to comment on the findings. If the system fails and it is not expected that the problems which cause the failure can be solved, the type test may be aborted by either the physico-technical steering group or the applicant. In this case only the costs of the technical evaluation(s) will be charged to the applicant.

If both physico-technical evaluations give different results, (some) tests may be repeated on a third system.

For CR systems the physico-technical tests must be performed on two X-ray units of different brand. The applicant should give guidance on the setup of the X-ray unit for their CR system and take care that the mammography unit is set up correctly for their CR plates at the sites of type testing.

- Phase 2: Clinical evaluation:

After the technical tests the digital system will be used clinically for a period of at least three months. The site at which the clinical evaluation is performed (one of the sites at which a technical evaluation was performed) must have sufficient workflow (i.e. similar to that used in European screening centres with an average of at least 50 clients/patients per system per day) and will be selected by the applicant in co-operation with EUREF. In the clinical evaluation soft copy reading is expected and all workstations must pass the European protocol for viewing conditions.

At the start of the clinical evaluation an application specialist of the applicant and a EUREF representative will be present (for two days) at the clinical site to help in starting up the clinical evaluation.

In the first days of the clinical evaluation, an initial assessment of the images will be performed by the representative of EUREF. If image quality is not satisfactory or if dose is higher than allowed, adjustments to the equipment must be made. If image quality and dose levels remain unsatisfactory the clinical evaluation may be stopped. Adjustments are only allowed after consultation with the EUREF physico-technical steering group. Some adjustments will require an additional technical survey.

In the clinical evaluation period a homogeneity image has to be acquired every day in full-automatic mode to monitor the stability of the equipment. A record must be kept

of all artifacts on clinical images and of all problems that occurred with the equipment by the radiographers/radiologists at the test site. At the end of the clinical evaluation period a bad pixel map has to be presented to EUREF to determine the number of detector elements which became defective during this test period.

When problems occur, the steering group has the right to extend the clinical evaluation period appropriately or to stop the evaluation.

In addition to the stability test a dose survey will be conducted. For this dose survey X-ray exposure data must be available and the images must be available for verification. The mean glandular dose recorded in the DICOM header will be compared with the values from the dose survey.

After the clinical test period, an evaluation of a set of clinical images will be performed, by two radiologists on invitation of EUREF with substantial experience in digital mammography and a physicist appointed by EUREF. If the results of the evaluation of clinical images are such that it is doubted whether overall image quality is sufficient, an additional evaluation of clinical images might be performed, for example with inserted simulated lesions in unprocessed images to evaluate image processing.

The images that are evaluated consist of two groups: 50 patients/clients will be selected based on breast thickness and 20 cases will be selected randomly out of referred patients/clients. If several modes of a system need to be evaluated, at least 30 images of each mode are required.

### Duration of a EUREF type test

After both technical evaluations are performed, the results of both surveys will be discussed by EUREF and permission to start the clinical evaluation may be given. The duration of the clinical evaluation is at least 3 months.

### Final report

After the clinical evaluation a final report is presented. The report will be sent to the applicant. When a system passes, this will be published on the EUREF website. When a system fails it will not be published on the website.

### Non-disclosure of confidential information

Type tests might be performed on systems using techniques which are completely different from existing systems. For these systems the current methods of measurement might be unsuitable and adaptations may be necessary. It might also be the case that new test items should be added. Therefore the physico-technical steering group will be provided with all relevant information on the system being type tested by the applicant.

Relevant information includes: basic principles of the mammography system and in specific the image receptor, philosophy and practice of the AEC system, pre-exposure parameters, reconstruction technique of bad pixels, accepted number of bad pixels etc. If requested by the applicant this information can be regarded as confidential and a non-disclosure agreement can be signed.

### Publication of results

The full report on the results of the type test will be made available to the applicant. If a system passes the type test the results will be published on the EUREF website. The applicant of the test will be able to comment on the report before publication on the EUREF website. EUREF will not use the information of the type tests in other publications without consent of the applicant of the test. It will not be mentioned on the website that a specific system has failed the type test.

### **3. EUREF type testing protocol**

By definition type testing will be performed on new types of equipment. Therefore type tests might be performed on systems using techniques which are completely different from existing systems. For these systems the current methods of measurement might be unsuitable and therefore adaptations may be necessary. It might also be the case that new tests should be added. Therefore the physico-technical steering group will be provided with all relevant information on the system being type tested by the firm applying for type testing. If measurements techniques need to be adapted this will be communicated in advance (if possible) by the physico-technical steering group and the applicant. If differences are noticed during type testing adaptations of methods of measurement will be made on the spot and discussed afterwards. This discussion will take place before the second physico-technical test is performed.

In section 3.1 an overview of all tests performed in a EUREF type test is given. Some additional tests will be performed, which are described in section 3.2.

For CR systems at least 4 cassettes of standard size (18 x 24 cm) and 4 cassettes of large size (24 x 30 cm) should be available during the type test, at the site of the clinical test at least 8 cassettes should be available. If large size cassettes are available they may also be used at the clinical site.

# EUREF type testing

Technical evaluation protocol





The following test-items described in the European Guidelines for quality assurance in breast cancer screening and diagnosis, fourth edition and its Supplement are measured in a EUREF type test technical evaluation:

- 2b.2.1.1.5     Tube output  
The method described in the Supplement to the European Guidelines is used. No limiting values are used. Measured for reference purposes and to calculate (mean glandular) dose.
  
- 2b.2.1.2        Tube voltage and beam quality
  - 2b.2.1.2.1     Tube voltage  
The method described in the Supplement to the European Guidelines is used.
  - 2b.2.1.2.2     Half value layer  
The method described in the Supplement to the European Guidelines is used.
  
- 2b.2.1.3        AEC-system
  - 2b.2.1.3.2     Back-up timer and security cut-off  
The method and limiting values of the European Guidelines are used.
  - 2b.2.1.3.3     Short term reproducibility  
The method and limiting values of the European Guidelines and its Supplement are used.
  - 2b.2.1.3.4     Long term reproducibility  
The method of the European Guidelines is used. The limiting values of the Guidelines are used as action limits for further investigation. In the final report measurements on long term stability from the clinical test period will be presented.
  - 2b.2.1.3.5     Breast thickness and composition compensation  
The method and limiting values of the European Guidelines and its Supplement are used.
  - 2b.2.1.3.6     Local dense area  
The method and limiting values of the Supplement to the European Guidelines is used.
  
- 2b.2.2          Image receptor
  - 2b.2.2.1        Image receptor response
    - 2b.2.2.1.1     Response function  
The method and limiting values of the Supplement to the European Guidelines is used.

- 2b.2.2.1.2 Noise evaluation  
The method and limiting values of the Supplement to the European Guidelines is used.
- 2b.2.2.2 Missed tissue at chest wall side  
The method and limiting values of the European Guidelines are used.
- 2b.2.2.3 Image receptor homogeneity and stability
  - 2b.2.2.3.1 Image receptor homogeneity  
The method and limiting values of the European Guidelines and its Supplement are used.
  - 2b.2.2.3.2 Detector element failure (DR systems)  
The method and limiting values of the European Guidelines are used. The bad pixel map should be easily accessible for all users. If uncorrected bad pixels are visible on the images this should be taken into account when evaluating detector element failure.
  - 2b.2.2.3.3 Uncorrected defective detector elements (DR systems)  
The method and limiting values of the European Guidelines are used.
- 2b.2.2.4 Inter plate sensitivity variations (CR systems)  
The method and limiting values of the European Guidelines and its Supplement are used. At least four cassettes of each size should be present.
- 2b.2.2.6 Fading of latent image (CR systems)  
The method of the European Guidelines is used.
- 2b.2.3 Dosimetry  
The method and limiting values of the European Guidelines and its Supplement are used.
- 2b.2.4 Image quality
  - 2b.2.4.1 Threshold contrast visibility  
The method and limiting values of the European Guidelines and its Supplement are used.

Threshold contrast visibility is determined at a number of dose levels: the clinical glandular dose level and two to four other dose levels, which will be chosen such that a large range of dose levels is covered (for example between  $\frac{1}{2}$  and 2 times the clinical dose level).

For details of 0.1, 0.25, 0.5 and 1.0 mm diameter, threshold contrast is plotted against glandular dose. The dose at which the minimum acceptable image quality standard and achievable image quality standard is achieved for the 0.1,

0.25, 0.5 and 1.0 mm diameter objects on the contrast threshold visibility phantom is calculated.

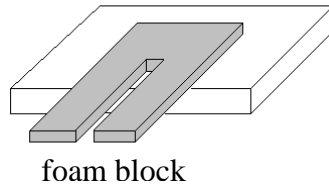
- 2b.2.4.2 Modulation Transfer Function (MTF), Noise Power Spectrum (NPS) and Detective Quantum Efficiency (DQE)  
The method described in the Supplement to the European Guidelines, appendix 7 is used.
- 2b.2.4.3 Exposure time  
The method and limiting values of the European Guidelines are used.
- 2b.2.4.4 Geometric distortion and artefact evaluation  
The method and limiting values of the European Guidelines are used.
- 2b.2.4.5 Ghost image / erasure thoroughness  
The method and limiting values of the European Guidelines are used.

## **Additional test performed in a EUREF typetest**

### **Thickness indication**

In the clinical test a dose survey will be conducted. For this dose survey the indicator of the height of the compression paddle needs to be checked.

For this measurement two foam blocks with compressed thickness of about 2 and 4 cm are used. A strip has been cut out of the foam block to allow measurement of thickness during compression (see figure below). Thickness indication can be checked when the foam blocks (18 x 24 cm) are placed on the bucky. Position the blocks such that half of the block is positioned on the bucky and half of the block is positioned over the edge of the bucky at chest wall side, see figure V.1. Apply compression (approximately 100 N), record the thickness indication and measure thickness at the reference point with an appropriate device (for example a calliper). Perform this measurement for the two foam blocks separately and together (so measurements can be done at about 2, 4 and 6 cm compressed thickness).



# EUREF type testing

## Clinical evaluation protocol



Authors: Hans Lelivelt  
Chantal van Ongeval  
Jurgen Jacobs  
Hilde Bosmans  
Ruben van Engen  
Roland Holland

### Introduction:

The imaging chain consists of a different number of components from a client who is being x-rayed to the radiologist who is reporting the images. During physico-technical type testing, all these components are checked separately. The clinical part of the type testing is mainly a global test. The quality of clinical images is determined by all components simultaneously, including the image processing which is not evaluated technically. The clinical part of the type testing should not be used separately and can only be performed on a system that passed the technical evaluation.

The clinical part of the type testing includes, next to an assessment of the image quality by a team of radiologists, also a stability test and a dose evaluation. The duration of this test is three months but when problems occur, EUREF has the right to extend the clinical evaluation period accordingly or to stop the evaluation.

The clinical test site, for obtaining the images, will be decided upon by EUREF together with the applicant. It must be a site where one of the physical tests was performed or where a successful acceptance test by EUREF had been performed. It must have a sufficient workflow of 4 images per client and softcopy reading is required. The clinical images will be sent to the EUREF office in Nijmegen and scored. The workstations, used to evaluate the images, must pass the European Guidelines criteria. If an applicant wants his images to be scored on his own workstation he needs to supply that workstation and it will be checked first whether this station passes the European Guidelines criteria.

### Procedure of acquiring the images:

The clinical test will be performed during at least three months. At the beginning of the clinical phase a EUREF representative will be present, together with a specialist of the applicant. EUREF will install a hard disk where all images from the clinical test period will be copied to. On the hard disk the patients will be anonymised. All occurring image quality issues like artefacts etc. must be recorded.

### Procedure of scoring the images:

Images of 70 women will all be scored by a team of radiologists. These 70 cases will be selected from the clinical test period collection as follows: The DICOM headers of all the images have to be correctly filled in. This allows to calculate the doses of all patients of that period. A plot of mean glandular dose as a function of compressed breast thickness will be made. The aim of this is twofold: (1) to learn about the dose distribution and (2) to make a good selection of cases for the evaluation by radiologists (all thicknesses, all types of tissues). In each thickness class, patients will be ranked as a function of their dose. From 5 thickness classes, 10 patients will be selected per thickness class; in the ranked series, the image at 10<sup>th</sup> percentile of the dose distribution, the image at 20<sup>th</sup> percentile, the image at the 30<sup>th</sup> percentile, etc will be selected. This makes in total 50 patients. The other 20 cases will be referred patients and must include opacities and microcalcifications.

Reading will be done by two radiologists and a physicist at the same time. The radiologists and the physicist that will score the image quality have experience in digital mammography and with the assessment of image quality. A third radiologist will also be appointed in case of image quality problems. This will either be a radiologist from the board of EUREF, or appointed by EUREF together with the applicant. This third radiologist has substantial experience in digital mammography.

The scoring form:

The form is based on experience with type testing in the Belgian and Dutch screening and consists of three parts. In the first part (13 questions) the radiologist has to answer questions with “yes” or “no”. This part includes questions about anatomical structures, noise and contrast in low and high pixel value areas. In the second part (4 questions) the radiologist has to score the contrast and sharpness of the images on a scale between -2 and +2. The third part (3 question) has questions about how confident the radiologists are with the image on a scale from 1 to 10. The scoring will be done at the workstation where the reading team views the images.

Procedure of evaluating the data:

A software program to analyse the data has been developed. Images are shown within this program and the score is entered directly into it. The software provides the medium score for each question and also the number of questions that were answered by “yes” or “no”. The results will be compared with other systems. Any significantly lower value on some questions for a system will be investigated more in detail together with the third radiologist.



Dose evaluation:

In the clinical evaluation period a dose survey will be conducted. For this dose survey X-ray exposure data must be available. The mean glandular dose recorded in the DICOM header will be compared with calculated values. It is desirable that the DICOM header contains the exposure values. When testing CR systems the applicant should take actions to connect the mammography unit and reader in order to fill in the DICOM header with exposure values, breast thickness and dose indicator.

Stability test:

In the clinical evaluation period an image of a 4.5 cm homogeneous block of PMMA covering the whole image receptor has to be acquired every day in full-automatic mode to monitor the stability of the equipment. A record will be kept of all artefacts on clinical images and of all problems that were noticed by the radiographers/radiologists at the test site. At the beginning and at the end of the clinical evaluation period a bad pixel map will be obtained to determine the number of detector elements which became defective during this test period. These QC images and records will be sent to the EUREF office for evaluation.

Additional comments:

Comments about potential special features, ergonomics and other remarkable points will be noted by the EUREF representative who will be present at the start of the clinical evaluation period. Besides this, comments might be asked from the radiologists and radiographers at the clinical evaluation site.

Requirements for the setup at the clinical test site (Figure 1):

1. all image data (both clinical and technical) are being sent to the normal PACS and in parallel also to a local, dedicated DICOM Service Class Provider (SCP). To enable this, the vendor has to configure a DICOM output node at the mammography unit
2. at a local computer, a software package ('Gladys') [ref] is installed. This is a DICOM SCP which can be scripted to perform specific actions based on the values of the DICOM headers. The software runs as a Windows service. All data are stored on an external hard drive, provided by EUREF
  - requirements:
    - A computer which is connected to the mammography unit via the network
    - A free USB 2.0 port
    - A fixed IP address
    - An administrator access to this computer to install Gladys as a Windows service
3. For the stability test the homogeneous images (only 'FOR PROCESSING' images) are stored by 'Gladys' on the external hard drive. The technical analysis is performed in a QC reference center. To make the distinction between patient data and the

homogeneity data, a specific patient name ('QCMAMMO') is taken by convention and action is undertaken by 'Gladys' based on this patient name

4. All clinical images (both 'FOR PRESENTATION' and 'FOR PROCESSING') are also sent to 'Gladys'. Using a one-way hash algorithm, this image data is anonymised and stored on the external hard drive in a specific folder structure. The DICOM headers which need to be anonymised can be different at each clinical test site and "Gladys" can therefore be configured. This action will be performed on all data with a PatientName that is different from 'QCMAMMO'
5. For the collection of patient dosimetry data, a second copy of the anonymised patient data (only 'FOR PRESENTATION') is stored without any image information. These DICOM files are used by EUREF to simplify the patient dosimetry study
6. For the collection of the selected 20 referred cases local radiologists are asked to note down the unique identifiers of at least 20 possible cases. They must include opacities and microcalcifications. To preserve confidentiality, the unique identifiers of the selected clinical cases have to be anonymised at the clinical test site. To do this, these identifiers have to be saved as a text file under a format similar to Fig. 2a. Gladys converts this text file using the same one-way hash algorithm as in (4). An example output file can be seen in Fig. 2b. This way we ensure that no patient related information leaves the clinical test site. This task will be done by a EUREF representative at the end of the clinical evaluation period.

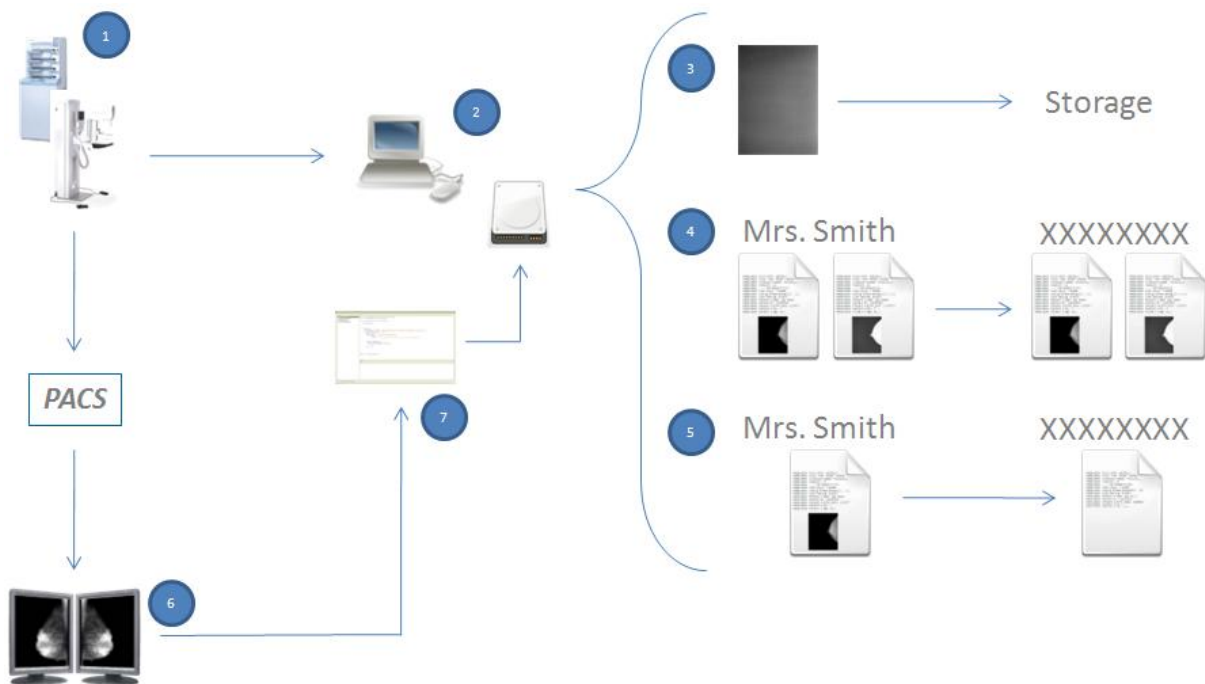


Fig. 1. Software setup at clinical test site



Fig. 2a. Example input format for selected patient cases

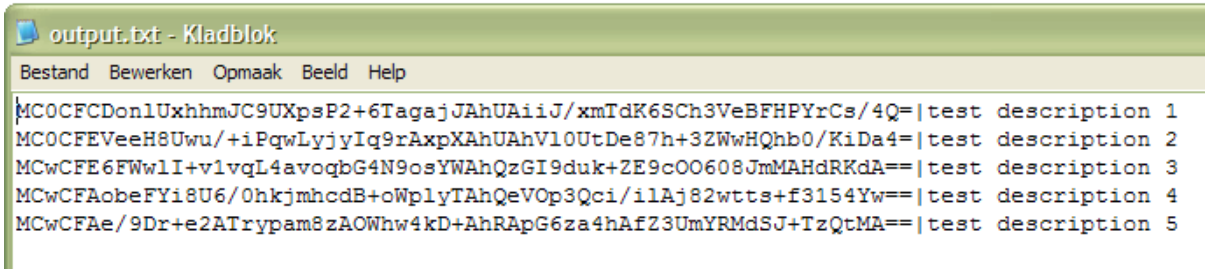


Fig. 2b. Example output format for selected patient cases

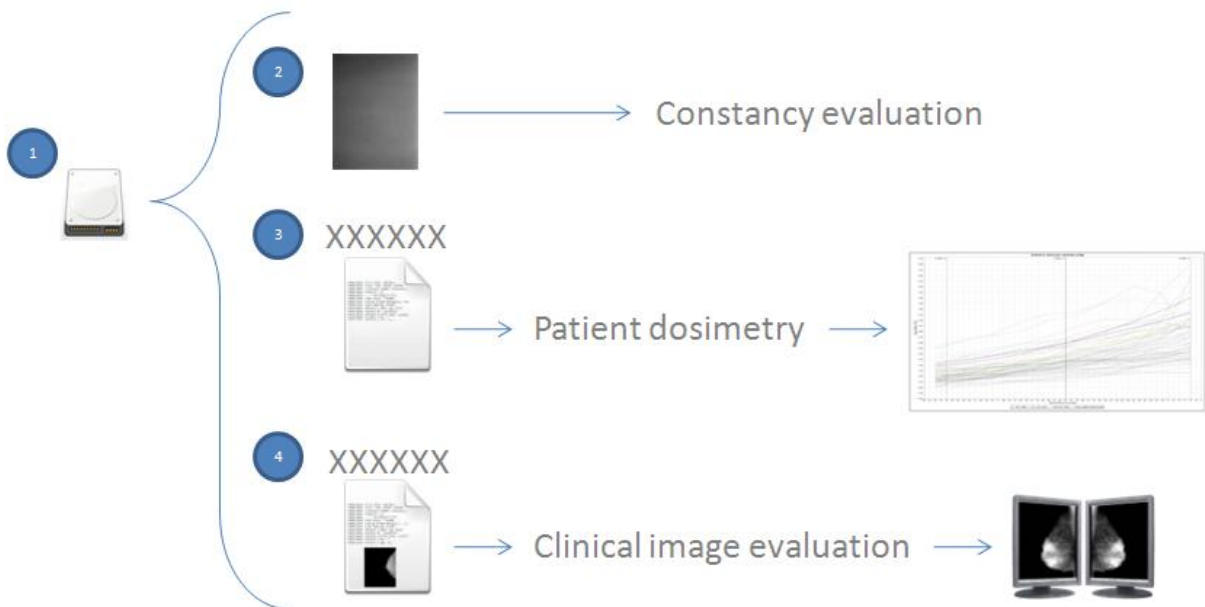


Fig. 3. Software setup at reference site

Setup at the reference site (Figure 3)

1. all data are transferred from the clinical test site to the EUREF reference site using the external hard disk provided by EUREF
2. the homogeneous acquisitions are being evaluated for artifacts and the stability over time of the mammography unit is being monitored
3. the DICOM headers without the image data are scanned for dosimetry related DICOM data. These header values will be used to calculate the patient dose using the method

- of Dance [ref]. Calculated values will be compared with DICOM header dose indications (if available) and with the current European Guidelines levels
4. the clinical image quality will be evaluated using a dedicated software platform to perform observer studies by means of a Visual Grading Analysis following the clinical evaluation form, which can be found at the end of this document [ref]. This task will be performed in a controlled environment using DICOM [ref] calibrated viewing stations. If the applicant of this clinical evaluation prefers to use a different workstation, this has to be discussed with the EUREF representative, prior to the clinical evaluation period.

References:

N. Perry et al. (ed.), European Guidelines for quality assurance in breast cancer screening and diagnosis, EU 2006.

R. van Engen et al., Supplement to the European Guidelines for quality assurance in breast cancer screening and diagnosis, to be published.

C. van Ongeval, H. Bosmans and A. van Steen, Current challenges of full field digital mammography, Radiation Protection and Dosimetry (2005), Vol. 117, No 1-3, pp. 148-153.

C. van Ongeval, A. van Steen, C. Geniets, F. Dekeyzer, H. Bosmans and G. Marchal, Clinical image quality criteria for full field digital mammography: A first practical application, Radiation Protection and Dosimetry (2008), Vol. 129, No 1-3, pp. 265-270).

J. Jacobs, T Deprez, G. Marchal and H. Bosmans, The automatic analysis of digital images for quality control purposes made easy with a generic extendable and scriptable DICOM router, Proceedings of the Scientific Assembly and Annual Meeting of the RSNA, Chicago, IL(2006).

J. Jacobs, F. Zanca, G. Marchal and H. Bosmans, Implementation of a Novel Software Framework for Increased Efficiency in Observer Performance Studies in Digital Radiology, Proceedings of the Scientific Assembly and Annual Meeting of the RSNA, Chicago, IL, (2008).

D.R. Dance, Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose, Phys. Med. Biol. 35, pp 1211-1219

D.R. Dance, C.L. Skinner, K.C. Young, J.R Becket and C.J. Kotre, Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol Phys. Med. Biol. 45, pp 3225-3240

D.R. Dance, K.C. Young and R.E. van Engen, Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA dosimetry protocols. Phys. Med. Biol. 54, pp 4361-72

D.R. Dance, K.C. Young and R.E. van Engen Estimation of mean glandular dose for breast tomosynthesis: factors for use with the UK, European and IAEA breast dosimetry protocols. Phys. Med. Biol. 56, pp 453-471.



## Clinical evaluation form for EUREF type testing

Number :

The following questions are to be answered by yes or no:

		Yes	No	N/A
1	Is there a good visualization of the skin line ?			
2	Are the vascular structures visible through the dense parenchyma ?			
3	Is there a sharp visualisation of the pectoral muscle ?			
4	Is there a good visualisation of the Coopers ligaments and vascular structures in the subcutaneous and prepectoral area ?			
5	Are the micro calcifications visualized and well outlined ?			
6	Is there sufficient contrast in the dark areas (e.g. no saturation of intensity of signals, no fully dark regions) ?			
7	Is there sufficient contrast in the white areas (e.g. no fully white regions) ?			
8	Is the glandular tissue sufficient white ?			
9	Is the background sufficient dark ?			
10	Do all images appear in the same way ? (if no, please place a remark)			
11	Is there disturbing noise in the dark areas?			
12	Is there disturbing noise in the white areas?			
13	Are there any artefacts ?			

For the following questions you score the images with a number from -5 (bad) to +5 (good). Please use the whole range

		-2	-1	0	1	2
1	Contrast in the white regions					
2	Contrast in the dark regions					
3	Overall contrast					
4	Sharpness					

For the following questions you score the images with a number from 1 (bad) to 10 (good). Please use the whole range.

		1	2	3	4	5	6	7	8	9	10
5	How satisfied are you with the representation of micro calcifications ?										
6	How satisfied are you with the representation of opacities ?										
7	How satisfied are you with the representation of the image ?										

Remarks:

Thank you very much for cooperating.