THE PROBLEM OF IMAGE QUALITY ASSESSMENT IN COMPUTED TOMOGRAPHY AND IN MAGNETIC RESONANCE IMAGING

S.Polvi[1] , *V. Busechian*[2]

1) Ital-TBS S.p.A., Area Science Park, Trieste, Italy
2) CRS-TBS, Area Science Park, Trieste, Italy

Abstract

The meaning and the aims of Quality Controls (QC) in Diagnostic Radiology are discussed, underlining the basic philosophy of "Quality" and the difference between the concept of Quality Controls and Quality Assurance. A brief overview of QC in Computed Tomography is given describing the main advantages. The methods are similar in Magnetic Resonance Imaging (MRI) but far more complex. Therefore only the method for the measurement of Signal-to-Noise ratio (SNR) is examined in detail while the other image quality parameters are just mentioned. As an a example of a right method for the evaluation of image quality in MRI, a new technique for the measurement of SNR based on a spatial frequencies domain analysis is proposed. Finally some guidelines for the personnel involved in QC activity are provided.

Introduction

The term "Quality" is very popular in this period, but sometimes there is some confusion about the terms.

In fact the quality of an organization is often confused with the quality of the product and the "effectiveness" of a process is considered to be the same as the "efficiency". We must pay a lot of attention to the terms we use: the quality of a product can be very good even if the quality of the organization in the factory is poor. Similarly the quality of the organization can be very good even if the quality of the product is very poor. The ISO 9000 certification, for instance, doesn't refer to the product quality but it refers just to the organizational criteria of the manufacturer.

The same attention must be paid using the terms "effectiveness" and "efficiency": if the aim of a company is selling equipment without defects and if 100% of the equipment put into the market has no defects, it is possible to speak of a good effectiveness of the process. This means that the goal has been achieved. But if the quality control service in the company throws away 50% of the manufactured equipment before putting it into the market because it is defective, it is not correct to speak of a good efficiency. In fact, *the effectiveness refers to the capacity of reaching the target while the efficiency refers to the ability of optimizing the resources.*

These concepts were born in the world of the manufacturers but can be easily extended also to the advanced services

In a radiological department for example the primary need is obtaining films with a good image quality to produce accurate diagnosis. The radiologist could be very embarrassed in giving an answer to the patient if the available images are not clear enough and he could even decide to repeat the exam to improve the image quality setting different acquisition parameters. Therefore the radiologist needs effectiveness in the process of production of the images: image quality first of all! If it is necessary to repeat the exam 10 times to obtain the image quality needed (which means also 10 times the dose to the patient and the costs) then the efficiency of that radiological department would be very poor while the effectiveness would be good.

A poor efficiency in a radiological department means an increase of the costs for the extra films, for the development of the extra films, for the extra time the radiologist and the technician must spend. Moreover it means that the patients have received an extra dose of radiation and that statistically a part of them will die for a cancer derived from the extra dose.

Therefore in a modern radiological department both effectiveness and efficiency in the process of formation of the images are required.

Quality Assurance (QA) and Quality Control (QC)

If the top management of the hospital finally realize that it is necessary to improve the effectiveness and the efficiency in the radiological department, some studies must be done to characterize the goal as a function of the patients needs.

First of all the resources must be evaluated and compared with the aim (the resources are not only money, but also time, people, etc.). The organization of the department must be checked to verify if some procedures can be improved. Some old equipment may be replaced and a long term plan must be done for the replacement of all the equipment. It is important to outline that the planned aim must be *achieved and maintained* in the time.

The program of improvements started by the management is called Quality Assurance (QA). A program of QA, in fact, improves the effectiveness and the efficiency in the department and can be defined as "a program planning all activities and functions concerned with the attainment of quality", where the "quality" is defined as "the totality of features and characteristics of a product or service that bear upon its ability to satisfy a given need" (I R F Hendra, 1986).

In this case, of course, the "given need" is that of the patient, that is: accurate diagnosis (good image quality) and low dose. Therefore the quality is attained when the patient's need is met by the service.

The program of quality assurance must cover many different fields and must consider different aspects: the personnel organization, the management of the exams, the plan of the purchase of new equipment and so on.

Quality control is a branch of the program of QA referring to the control of the performances of the equipment. The goal of QC is to ensure the best performances of the equipment, by checking and assessing the image quality and other parameters.

Of course an improvement in the service offered by the radiological department can be achieved also without starting a QA program. But the improvement would be modest, because it depends strongly on the current chief of the department. To guarantee that the QA program will continue efficiently it should be coordinated at high level with the other departments of

the hospital.

Quality controls cost money since trained personnel is required to carry on this activity (J L Williams and W Howarth, 1986), but in some particular fields, as conventional radiography, quality controls can help also to save some money. For example it can be estimated that in a radiological department without a program of quality control up to 15% of the used films must be rejected for poor image quality. How many films could be saved with the establishment of a service of quality controls? Probably up to 60%. In a large hospital these savings are more than the costs of quality controls.

In diagnostic radiology the final result is an image. Therefore the most popular techniques of QC are based on the assessment of image quality. A proper evaluation of the images can give useful indications about the calibration of the equipment allowing a better organization of the extraordinary maintenance interventions.

In the following chapters some interesting techniques of QC will be presented and some typical results will be discussed.

Quality controls in computed tomography

The assessment of CT scanners is ideally based on the ratio "image quality"/exposure. In fact there is no use in measuring the image quality without knowing anything about the exposure. Therefore for each setting of the scan parameters the image quality and the exposure must be measured.

While the exposure can be measured using tested methods and calibrated equipment, there are no norms or special requirements concerning the image quality parameters. Without exploring the details, it seems reasonable to use the following image quality parameters since they represent an exhaustive set describing the main feature of an image:

- Image uniformity
- Image noise
- Slice thickness
- Scale factor
- Spatial resolution
- Low contrast detection

Image uniformity: when a uniform phantom of water or water equivalent is scanned, ideally the resultant phantom image should everywhere show a CT number of zero. Of course, even in the best of actual images all the image pixels would not be zero because of image noise. Still the average CT number over any significantly large region should be close to zero no matter where in the image that region is located.

Image noise: CT image noise is a result of the quantum photon nature of x-rays. It is the equivalent of quantum mottle in conventional radiographic images. As a result of this noise the CT image of a uniform phantom will demonstrate pixel CT number fluctuations even in a perfectly adjusted scanner. CT image noise is usually expressed as the standard deviation of the CT numbers of a large number of pixels taken from the image of a uniform phantom.

Slice thickness: the slice thickness is set from the main console of the CT scanner. But the true slice thickness can be different from the set up. Therefore it is important to check this parameter by imaging a ramp oriented at 45° to the scan plane.

Scale factor: the CT number of an image pixel is derived from the measured x-ray linear attenuation coefficient of the corresponding point in the sample being scanned. The relationship between linear attenuation coefficient (μ) and CT number of a given substance is given by the following expression:

$$CT\ number = \frac{\mu - \mu H_2O}{\mu H_2O}[Scale\ factor]$$

For modern scanners the nominal value of the scale factor is 1000, which roughly means that the CT number of water is 0 and that the CT number of air is -1000. The actual value of the scale factor can be determined using a phantom constructed of two materials (A and B) that differ in linear attenuation coefficient of a value $\Delta\mu$.

$$Scale\ factor = \frac{|CT_B - CT_A|}{\Delta\mu\ /\ \mu H_2O}$$

Spatial resolution: it can be measured using phantoms containing patterns of parallel rods. An alternative is offered by measuring the MTF of an edge. This second method is more accurate but it requires dedicated software and the transfer of the image to a different computing system.

Low contrast detection: one of the most significant properties of CT scanners is their ability to perceive the presence of low contrast lesions that are quite invisible by any other radiographic method. Therefore monitoring low contrast detectability is an essential component of CT QC. Noise measurements usually correlate well with detectability measurements. However, noise measurements by themselves may give very little information as to how two different scanners compare in their low contrast performance.

Low contrast detectability is usually measured using low contrast phantoms (0.3 - 0.6% contrast). It is defined as the smallest detail still visible at the lowest contrast.

Measuring the exposure in computed tomography

The dose is strictly connected with the measurement of the exposure. The dose is the total amount of energy deposed in a tissue by the x-rays. The exposure is the amount of ions produced in air by the x-rays. But while the dose depends by the features of the tissue, as the thickness, the atomic properties, and so on, the exposure can be easily measured using ionization chambers and a dosimeter. Therefore people involved in QC on CT scanners should always measure the exposure with standard methods and calibrated instruments.

The image quality (noise and low contrast detectability in particular) improve increasing the dose. But the improvement is not linear and after a given level a plateau is reached. A further increase of the dose produces a small improvement of image quality and it should be avoided.

Fig. 1 gives an idea of the exposures (in Roentgen) measured in a program of QC performed in north-eastern Italy on 15 different scanners. The scan parameters were typical of each site for standard abdomen exams.

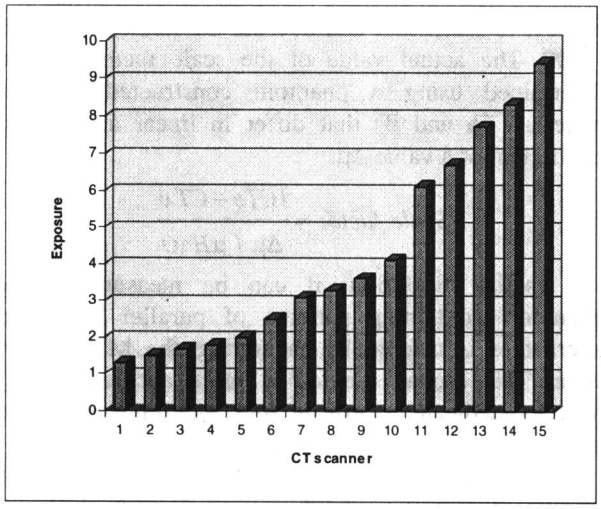

Fig. 1: exposure for standard exams measured on 15 CT scanners.

Even if the differences in image quality were not relevant, the differences of the exposures were huge and could vary of a factor 10. Moreover the wear of the x rays tubes was quite different. The tubes of the scanner n. 1 were usually replaced after 250.000 slices, while the tubes of the scanner n. 15 were usually replaced after 50-60.000 slices. It can be estimated that the cost for the extra tubes was more than 100.000 US$ per year (the cost for a CT tube is about 50.000 US$).

The previous example gives an idea of the utility of a QC program: if the QC program belongs to a QA program, the maintenance service immediately act where it is necessary to recover the best image quality and the physicians check the scan parameters with the literature to reduce the wear of the tubes.

Quality controls in MR imaging

In computed tomography Quality Controls are really very useful. It is possible to show that after a beginning period the information obtained can be used to improve the image quality and that some money can even be saved reducing the wear of the tubes and the dose to the patients. There are no doubts that a QC program is very attractive for the administration of the hospital and for the radiologist too.

In Magnetic Resonance Imaging, there are no wearing parts, the dose to the patient is not so important as in CT (there are no ionizing radiation) and finally it is difficult to assess the performances of the scanner and to compare different image qualities. Is it worthy to perform quality controls on MR scanners? The answer is *yes*. In fact, the following features of MR scanners must be noted:

1. MR scanners are more expensive than CT scanners;
2. The exams last more than CT exams;
3. The exams are more specific than CT exams;
4. The electronic is not so stable as in CT and frequent set up are required;
5. MR scanners are often used also for research studies.

It appears that MR imaging is a more advanced and specific technique than Computed Tomography. Therefore the performances of the equipment must always be the best achievable to justify the higher cost. Moreover the worsening of the image quality can be gradual and the physicians can seldom reveal it because he adapts himself to the image quality and a numeric evaluation of the image quality is the only way to detect the worsening.

A periodic plan of QC can produce the following benefits:

1. Warranty of a good image quality;
2. Warranty of the right response also for advanced research studies.
3. Early detection of breakdowns or electronic failures;
4. No necessity of repeating exams;

Of course the QC program is quite expensive and the results are not so evident as in CT, but we do believe that in MR imaging QC should not be considered as an optional but a primary need at the moment of the purchase, adding the cost to the other programmed costs, as the maintenance service or the refilling of the cryogen.

However, by comparison with CT, the problem is

orders of magnitude more complex because of the number of variables. Some millions of different scan parameters can be selected and their effect on the image quality is not easily understandable.

Image Quality parameters in MRI

The need for QC in MRI has emerged by the aim of characterize the tissues. In fact, it is well known that T1 and T2 relaxation times (two functional parameters calculated from MR images) generally vary from tissue to tissue and may exhibit significant alterations in the presence of pathological conditions. Hundreds of papers were published on T1 and T2 values of normal and pathological tissues in the seventies and eighties. But seldom the results obtained from different research studies could match. Many times the agree was very poor (EEC Concerted Research Project, 1988).

It was clear, even if the methods were accurate, the personnel was trained and the theory was right, that the calculated values of T1 and T2 could vary in a wide range due to a bad calibration of the equipment.

Therefore, a concerted research project entitled "Identification and Characterization of Biological Tissues by NMR" was actived in 1984 by the European Economic Community. One of the results of the project was the identification of appropriate test substances and phantoms for the calibration of MR imagers. Using the right test objects it was possible to check the calibration of the equipment and to get comparable results also from different institutes and centres.

Even if the phantoms design improved in the years (R. A. Lersky, D. W. McRobbie, 1992), the image quality parameters and the basic philosophy remained the same. The following image parameters are still currently used:

1. Signal to noise ratio (SNR)
2. Image uniformity
3. Slice width
4. Spatial resolution
5. Contrast to noise ratio
6. T1 and T2 measurement accuracy.

Signal to noise ratio: it is the ratio between the signal and the white noise. The signal can be measured as the mean value of the pixel intensity over a flat field image. The white noise is usually measured as the standard deviation of the pixel intensity over the same image.

Image uniformity: it is the variation of the pixel intensity over a flat field image.

Slice width: the meaning is the same as the slice thickness in CT. Also the way of measurements is about the same.

Spatial resolution: the same as in CT.

Contrast to noise ratio: It is probably the most important parameter from a clinical point of view since a high contrast allows the physician to discriminate between two different tissues. It is defined for couples of different tissues with the following formula:

$$CNR = \frac{|S1 - S2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

where S is the pixel intensity and σ is the noise. The index 1 and 2 refers to the tissue types.

T1 and T2 measurement accuracy: it is the accuracy achieved by the scanner when computing T1 and T2 relaxation times. Test substances are used to check the results.

All the image quality parameters are deeply influenced by many technical features of the scanners, as field intensity, field homogeneity, sensitivity of the coil, gain of the preamplifier, range of the ADC, etc. (R. A. Lersky, J. D. De Certaines, 1991).

The measurements of Signal to Noise Ratio

Each of the parameters mentioned above could require a paper to describe the meaning, the best way of measurement and the right interpretation of the results (R. L. Dixon, 1988). In this paper the only SNR is discussed, since it has been measured for a long period in a wrong way and the alternative measure we suggest seems attractive.

Signal-to-noise ratio has always been considered one of the most important image quality parameters, even if the methods suggested were not universally accepted (W. A. Edelstein, 1984). A further analysis of different methods for SNR measurements show clearly that the image theory does not match with the results obtained. In particular the dependence from the excitation number is not respected (O. Nalcioglu, Z. H. Cho, 1984).

SNR is usually measured (SNRm) on images obtained from a flat field test object. One or more regions of interest (ROI) are then considered and the ratio between the main pixel intensity and the standard deviation of the pixel intensity is calculated:

$$SNR_m = \frac{Main\ Pixel\ Intensity}{Standard\ Deviation}$$

It seems reasonable to define the "Signal" as the main pixel intensity in a given ROI. But is quite unreasonable to define the white noise as the standard deviation of the pixel intensity in the same ROI. In fact, if we superimpose strong artefacts to the image, the white noise should not change, since it must be the same at every spatial frequency, while the standard deviation is instead strongly affected by the

superimposition. Therefore the previous formula allows just a rough calculation of SNR if there are no problems of artefacts and if the image uniformity is good. Unfortunately in MRI the uniformity is always very poor and the artefacts are often present!

The theory of MR foresee the following behaviour for SNR as a function of the excitation number N:

$$SNR = kN^{1/2}$$

Where k is a complex function of other parameters.

If the previous definition is adopted, the behaviour of the measured SNRm is far different from the theoretical result and the values are far smaller since in the computation of the white noise also artefacts and lack of uniformity give their contribution.

It seem obvious that a right calculation of SNR must be based on the right measurement of the white noise.

An attractive tool for the computation of the white noise is the Wiener Spectrum (or Noise Power Spectrum). The main features of the Wiener Spectrum are the following (L. W. Grossman, 1986):

1. It has general applicability;
2. It can be applied to almost any imaging modalities (digital images included);
3. It is additive for independent sources of noise.

The Wiener Spectrum is a frequency representation of the correlation of the noise. It can be written as follows, as a function of the spatial frequencies:

$$\Phi(u,v) = \lim_{\substack{X\to\infty \\ Y\to\infty}} \frac{1}{4XY} \overline{|F_D(u,v)|^2}$$

where the bar over F_D means an average over an infinite ensemble of sampling functions (images), F_D is defined as:

$$F_D(u,v) = \int_{-X}^{X}\int_{-Y}^{Y} \Delta D(x,y) e^{-2\pi i(ux+vy)}\,dxdy$$

and $\Delta D(x, y)$ is the difference between the pixel intensity I in the position (x, y) and main pixel intensity I°. If an ensemble of MR images is available, the approximated Wiener Spectrum can be calculated (the exact Wiener Spectrum should be calculated over an infinite ensemble of images) (R. E. Challis, R. I. Kitney, 1990).

The Wiener-Khintchin theorem establishes that the Wiener Spectrum is the bidimensional Fourier Transform of the autocorrelation function $R(\xi,,\eta)$, where (K. Faulkner, B. M. Moores, 1984):

$$R(\xi,\eta) = \lim_{\substack{X\to\infty \\ Y\to\infty}} \frac{1}{4XY} \int_{-X}^{X}\int_{-Y}^{Y} I(x,y)I(x+\xi, y+\eta)\,dxdy$$

Moreover it can be easily obtained that the autocorrelation function calculated in (0, 0) coincides with the definition of the variance:

$$R(0,0) = \sigma^2$$

Finally, from the previous formulas, we obtain the relationship between the Wiener spectrum and the standard deviation of the pixel intensity:

$$\sigma^2 = R(0,0) = \int_{-\infty}^{+\infty}\int \Phi(u,v)\,dudv$$

The physics of MR guarantee that the lack of uniformity, the artefacts and the white noise arise from independent sources. Therefore the Wiener Spectrum can be split in three additive terms:

$$\Phi(u,v) = \Phi(u,v)_{white} + \Phi(u,v)_{artefact} + \Phi(u,v)_{uniformity}$$

Where Φ_{white} is the contribution of the only white noise, $\Phi_{artefact}$ is the contribution of only the artefacts and $\Phi_{uniformity}$ is the contribution of only the lack of uniformity (I. Pinto et al., 1990).

Moreover it is well known that the lack of uniformity is a smooth noise and it is correlated with low spatial frequencies. The artefacts can cover a wide range, from low to high spatial frequencies, while the white noise gives the same contribution at every spatial frequency.

Should we be able to evaluate $\Phi(u, v)_{white}$, it would be possible to integrate the function obtaining the standard deviation of the white noise. But actually we need much less: it is enough to know the average value of $\Phi(u, v)_{white}$, since the value is about the same at all the spatial frequencies, without knowing exactly the behaviour of Φ for each point.

At the very high spatial frequencies of $\Phi(u, v)$ only the white noise contributes to the spectrum.

Therefore the very high spatial frequencies can be used to calculate the average value of $\Phi(u, v)_{white}$.

The integration of the Wiener Spectrum gives the standard deviation of the white noise. The value can be used for a correct evaluation of SNR.

A deeper analysis of the results shows that the dependence from the scan parameters (as Excitation number, Field of View, Matrix size, etc.) is very similar to the provisions of the theory and that the accuracy of the measurements is quite good (less than 3%). However, in the method described above there are two problems which cannot be ignored:

1. The main features of the Wiener Spectrum has

been obtained for continuous functions. It must be demonstrated that the same results are still true for the discrete functions (digital images).

2. The average over an infinite number of images of course is not realistic. It must be demonstrated that the same results can be obtained without the average operation.

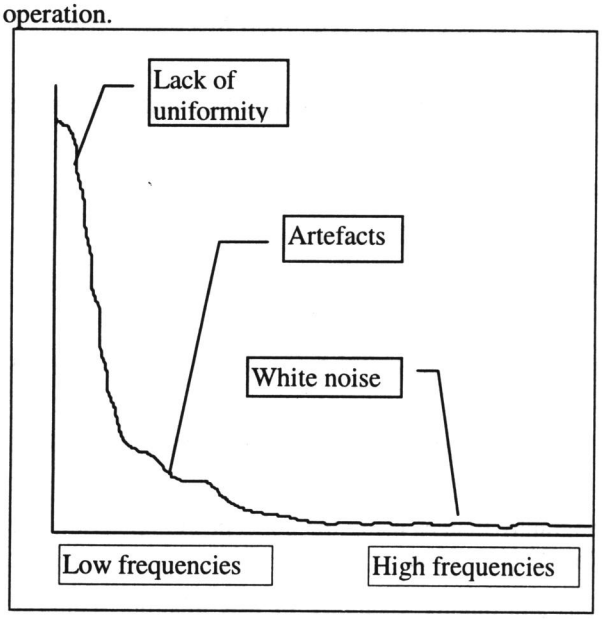

Fig. 2. Typical shape of Wiener spectrum of flat field images

The demonstration that the results are still true also for digital images is not immediate but requires some calculations (R. E. Challis, R. I. Kitney, 1990), extending the formula from the integrals to the sums. However, it can be shown that the main features of the Wiener Spectrum are maintained and, therefore, it is possible to use the method also for digital images.

About the averages over an infinite number of images, it is possible to avoid the average operation and to substitute it with a more convenient average over the spatial frequencies domain, if the sample functions (images) can be considered as an ergodic process (J. S. Bendat, A. G. Piersol, 1971). It is not true, from a general point of view, for the MR images. But is quite true for the white noise. Therefore, even if the formation of MR images is not an ergodic process, it is possible simply to average over the spatial frequencies the spectrum from a single image to obtain the value of the white noise.

Guidelines for Quality Controls in medical imaging

1. Quality Controls should always be a part of a wider program including at least the organization of the maintenance service.
2. A QC program can be successful only if physicians, technicians and engineers are informed of the aim of the program itself.
3. Only trained personnel should be involved in a QC program. The formation of product specialists should be encouraged.
4. Only standard and calibrated equipment should be used.
5. The protocols should always be written, procedures, methods and scan parameters included.
6. The final reports should always be written and immediately provided to the maintenance service, to the physicians or to the people liable of the activity.
7. If there are no enough resources to obtain the required accuracy from the measurements, it must be noted clearly, to avoid wrong interpretations of the results. The limits of the measurement should be outlined.
8. The results of the QC program should be discussed with the maintenance service and with the manufacturer.
9. A detailed calendar of the tests must be agreed with the physicians before starting the activity.
10. Always try to be updated on the recent advances in the field of QC.

Conclusions

In some countries (mainly in UK) QA and QC are an integral part in the management of the biomedical equipment. The same product specialist is used for Quality Controls, acceptance testing and purchase.

In other countries there are no constancy checks on the equipment performed by internal or independent services and the Quality Controls are entirely entrusted to the manufacturers. It means that there is no trained technical personnel in the hospital able to verify the true performances of the equipment.

A program of QC, therefore, is not just merely a verify of the performances, but also a way to mature a deeper expertise and a better knowledge of the technical features on a wide number of biomedical equipment.

It is a short-sighted policy to avoid QC because they are expensive. In a long term QC and even a QA program can produce large savings and a better service. Of course adequate resources must be available and the "culture" of the safety for the patient and for the user must be spread.

The personnel involved in the activity of QC, moreover, should be aware of the aims and of the limits of the tests. Better techniques should always be investigated, also sharing the experience with similar groups in other hospitals.

References

I. Pinto, S. Polvi, P. Giribona, T. dell'Aquila, Power spectrum analysis of MRI noise, Physica Medica, Vol. VI, N. 3-4, 1990.

I.R.F. Hendra, A systematic approach to quality assurance in medical diagnostic imaging, IOP Short Meetings Series No. 2, Institute of Physics, London, 1986.

J. L Williams, W. Howarth, Quality assurance for CT scanners and NMR imagers, IOP Short Meetings Series No. 2, Institute of Physics, London, 1986.

EEC Concerted Research Project, Identification and characterization of biological tissues by NMR, Magnetic Resonance Imaging, March-April, 1988.

J. S. Bendat, A. G. Piersol, Random data, analysis and measurements procedures, Wiley Interscience, 1971.

E. O. Brigham, The fast Fourier transform, Prentice-Hall Inc., Englewood Cliffs, NJ, 1974.

R. E. Challis, R. I Kitney, Biomedical signal processing. Part 3. The power spectrum and coherence function, Medical and Biological Engineering, Vol. 29, 1990.

R. L. Dixon, MRI: acceptance testing and quality control, Proceedings of an AAPM symposium, Medical Physics Publishing Corp., Medison, Wisconsin, 1988.

W. A. Edelstein, A signal-to-noise calibration procedure for NMR imaging systems, Medical Physics, Vol. 11(2), 1984.

K. Faulkner, B. M. Moores, Analysis of x-ray computed tomography images using the noise power spectrum and autocorrelation function. Physics in Medicine and Biology, Vol. 29 (11), 1984.

D. J. Goodenough, K. E. Weaver, Phantoms for specifications and quality assurance of MR imaging scanners, Comp. of Medical Imaging and Graphics, Vol. 12(4), 1988.

L. W. Grossman et al., Noise analysis of scintillation camera images: stochastic and non-stochastic effects, Physics in Medicine and Biology, Vol. 31 (9), 1986.

L. Kaufman et al., Measuring signal-to-noise ratios in MR imaging, Radiology, Vol. 173, 1989

R. A. Lersky, J. D. De Certaines, Performance assessment and quality control in MRI by Eurospin test objects and protocols, Eurospin n. 29/30, 1991.

R. A. Lersky, D. W. McRobbie, Eurospin II magnetic resonance quality assessment test objects, Diagnostic Sonar Ltd, 1992.

O. Nalcioglu, Z. H. Cho, Limits to signal-to-noise improvements by FID averaging in NMR imaging, Physics in Medicine and Biology, Vol. 29(8), 1984.

NEMA, Determination of signal-to-noise ratio in diagnostic magnetic resonance images.