December 2009

Medisch Spectrum △ Twente

UNIVERSITEIT TWENTE.

COMMITTEE:
DR. J.M. HUMMEL
DR. IR. S. MANOHAR
PROF. DR. M.J. IJZERMAN
DR. IR. S.R. VAARTJES

IMPLEMENTATION SCENARIOS AND PREDICTED PERFORMANCE OF THE TWENTE PHOTOACOUSTIC MAMMOSCOPE AS AN ALTERNATIVE IMAGING MODALITY IN BREAST CANCER DIAGNOSIS

FACULTY OF MANAGEMENT AND GOVERNANCE
HEALTH TECHNOLOGY AND SERVICES RESEARCH

Master Thesis Health Sciences | Marjolein Hilgerink

Implementation scenarios and predicted performance of the Twente Photoacoustic Mammoscope as an alternative imaging modality in breast cancer diagnosis

Marjolein Hilgerink
December 2009

ABSTRACT

The purpose of this study is to predict the future performance of the Twente Photoacoustic Mammoscope (PAM), and to identify different implementation scenarios based on this performance. PAM is a recently developed breast imaging device that makes use of photoacoustics. Criteria that are important in the design of a diagnostic breast cancer imaging device are identified, and these criteria are translated into design goals for further development of PAM. The Analytic Hierarchy Process (AHP) analysis has been used to identify the relative importance of design criteria, and to identify the relative performance of four different breast cancer imaging techniques: X-ray mammography, ultrasound, magnetic resonance imaging, and PAM. The currently used breast imaging techniques are used at different stages in the diagnostic track for breast cancer, two different AHP analyses have been carried out to determine the most optimal stage for PAM. House of Quality analysis is used to translate the implementation criteria into technical improvement goals for PAM. Because of the early developmental phase of PAM, there is some uncertainty about its future performance. This uncertainty has been assimilated into a sensitivity analysis from which different scenarios of the performance of PAM result.

The most important criterion in the design of a diagnostic breast imaging device is sensitivity. Factors that mainly determine the performance on sensitivity are the visualization of mass margins, mass shape and vascularization. To improve the performance of PAM with respect to these factors, the quality of the reconstruction algorithm, detector sensitivity, detector bandwidth and number of wavelengths used in PAM should be improved. From the overall performance on the costs, effectiveness, patient comfort and safety/risks criteria assessed, it can be concluded that PAM will be the most preferred alternative. This is true for all scenarios that result from the uncertainty analysis, except for the most negative scenario, in which MRI is preferred slightly over PAM. When implemented for diagnostic use, PAM can best be positioned at the start of the diagnostic track, as a substitute for the combined use of X-ray mammography and ultrasound. This research has shown that the AHP method is useful to provide guidelines for the most promising area of application for which a new technology can be implemented. It will be helpful if all participants have the same basic level of understanding about the technology that is assessed, for future research it is therefore recommended to focus on an efficient way of knowledge sharing before the AHP is carried out.

INTRODUCTION

The Netherlands has among the highest rates of breast cancer incidence in the world and the highest in Europe; about 1 in 8 women in the Netherlands will develop breast cancer at some time during their life. [Borstkanker, Visser *et al.* 2005] Breast cancer dominates all other cancers in females; with a prevalence of about 31% in Dutch women in 2006. [Cijfers]

Breast cancer screening has resulted in an increase of breast cancer detection in the Netherlands since the first half of the 90's. Screening offers the possibility of detecting cancer at an early stage, and therefore allows immediate and thus better treatment. This has resulted in a reduced mortality. The 5-year survival is 85%, when cancer is isolated to the breast this is 90-100%. [Visser *et al.* 2005]

In screening and diagnosing breast cancer, imaging techniques are used frequently and are of crucial importance. The most commonly used screening method for breast cancer is to make an X-ray image, or mammogram, of both breasts. Every year, about 800.000 women between age 50 and 75 are screened for breast cancer this way in the Netherlands. Of all breast tumors found in this age group, 40% is discovered in this screening program. This results in about a quarter of all new breast cancer patients each year. [SCK 2005] Diagnosis is often based on the results of a combination of the imaging techniques X-ray mammography,

ultrasound and MRI. With X-ray mammography, small microcalcifications can become visible which may indicate tumor growth. This method is relatively easy and reliable, but has a disadvantage of offering poor contrast in young women and a (small) risk of radiation induced tumors. Ultrasound is often used in addition to X-ray mammography, to discriminate between a cyst or other benign lesion and a (malignant) tumor. Because of its high complexity, ultrasound diagnosis can only be performed by trained radiologists, and results are strongly operator-dependent. MRI uses a contrast agent (Gadolinium), that, being a blood pool agent, helps identify angiogenesis: the increase in blood vessels and permeability of the vessel wall around the tumor. This technique has a very high sensitivity (> 95%), but also a low specificity (between 20% and 90%, strongly dependent on patient population). Besides this, MRI is more expensive than the other imaging techniques, therefore it is used only for a limited amount of indications. [Mammacarcinoom 2008, SCK 2005, Berg *et al.* 2008]

The Twente Photoacoustic Mammoscope (PAM)
The Twente Photoacoustic Mammoscope (PAM) is a new diagnostic device, developed by the Biomedical Photonic Imaging (BPI) group of the University of Twente. The device can be used to visualize breast cancer, possibly in screening as well as in diagnosis. Its working principle is based on photoacoustics. The PAM uses short pulses of Near Infrared (NIR) laser light, that are selectively absorbed by blood vessels. Upon absorption by the blood vessels, the temperature in the vessels rises and thermal expansion will occur. An ultrasound pulse (pressure wave) is generated, which can be detected by an ultrasound detector that is built into the device. After data acquisition of the signals of this detector, a (3D) image of the blood vessels in the breast can be reconstructed (figure 1). [Manohar *et al.* 2007]

An alternative to this technique would be to directly measure the light that is transmitted through the breast. The amount of scattering is however so large that it is very complex to locate the exact position of the tumor. On the other hand, using solely ultrasound would not display blood vessel details, because blood hardly reflects ultrasound. By combining light and ultrasound in photoacoustics, both disadvantages are cancelled: ultrasound is not scattered as much and the photoacoustic signal displays the interaction of light with blood vessels. This way, it is possible to identify angiogenesis, the same process that is imaged with MRI.

PAM is still early in its development, currently only a prototype of this device exists (figure 2). The prototype has been tested in a diagnostic setting on five patients that had a known breast tumor. In four out of five patients, the researchers identified areas with a higher intensity photoacoustic signal, which they refer to as vascularization of the tumor. The size of this higher photoacoustic intensity area correlated well with the pathologically determined tumor size of these patients. Photoacoustics also seems to offer additional information about the nature of the tumor: in one of the published cases the gross features in the X-ray mammogram and ultrasound image displayed benign tumor aspects, whereas the photoacoustic image displayed a ring-shaped high intensity area that is indicative of possible malignancy. [Manohar *et al.* 2007] Photoacoustics may offer some additional benefits over the existing imaging techniques. It is for example expected that this technique will be less expensive than MRI, more comfortable for the patient than the other imaging techniques, and the technique does not make use of ionizing radiation.
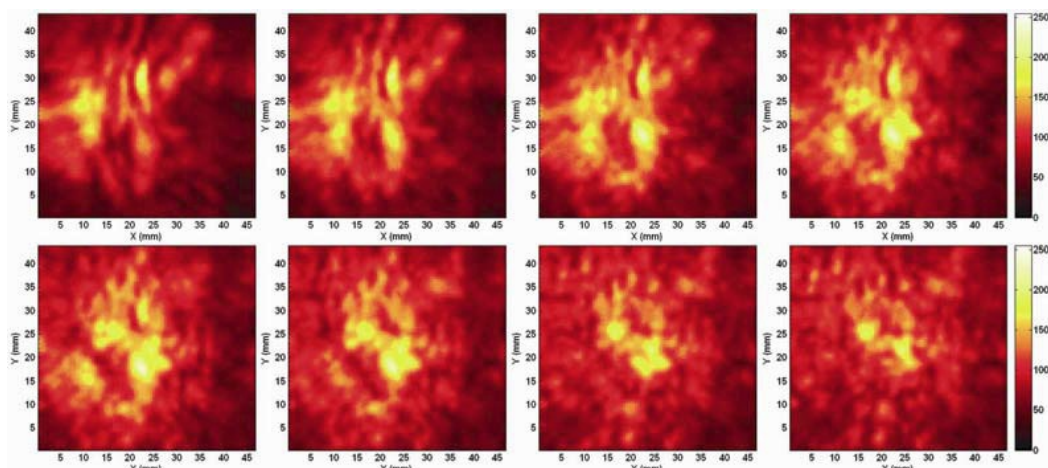

Figure 1 Selected slice images of photoacoustic reconstructed data set in craniocaudal view. The inter-slice spacing is 1 mm with the first slice 9.5 mm below the illuminated breast surface. In the top images, a ring pattern of higher intensity which depicts strong vascularization at the tumor periphery becomes evident. [Manohar et al. 2007]

Figure 2 Prototype of Twente Photoacoustic Mammoscope (PAM). a breast aperture, b ultrasound detector matrix, c glass window, d scanning system compartment, e laser, f laser safety curtain, g interface electronics between detector and computer, h positioning wheel for detector matrix (slightly pressed against breast), i laser remote controller, j laser power supply [Manohar *et al.* 2007]

Breast cancer diagnosis

Breast cancer diagnosis is defined as the use of imaging methods to evaluate a clinical or screening-detected breast abnormality. This exam includes an on-site review by a radiologist and communication of the results to the patient. The goal is to provide clinician the information to determine the future cause of action, so that the patient leaves with a final recommendation: either the extent and location of abnormalities is described or other evaluation is needed (too vague lesions). [Berg et al. 2008] The main question in diagnosis that has to be answered is: Is the detected lesion benign (fibroadenoma/cyst/…) or malignant (invasive lobular/ductal carcinoma)? [Mammacarcinoom 2008]

The Dutch Breast Cancer Guideline describes the diagnostic track for breast cancer that is used in the Netherlands. [Mammacarcinoom 2008] For Medisch Spectrum Twente, a large teaching hospital in Enschede, the Netherlands, the diagnostic track used in the Centre for Mammacare is enclosed in appendix 1. The current diagnostic track used at the Centre for Mammacare at Medisch Spectrum Twente starts with an X-ray examination together with an ultrasound-examination. This is according to the Dutch Breast Cancer Guideline. [Mammacarcinoom 2008, Koertshuis 2008] During X-ray mammography, both breasts are imaged in two directions, while the breasts are forced between two plates. The total examination takes about twenty minutes, recording the images takes several seconds. The inherent contrast between the tumor and the surrounding tissue is low, and depends on the age of the patient. Young patients have relatively denser breasts, which worsens the contrast. Contrast in fatty tissue is better. Often X-ray mammography is used to look for microcalcifications that, in a certain presentation, are indicators of malignancy. Ultrasound is used next to X-ray mammography for further evaluation to discriminate between cysts and solid tumors. Ultrasound examination is performed by a radiologist, and takes about 15 minutes. The quality of ultrasound examination depends for a large part on the experience of the radiologist. The combined use of X-ray mammography and ultrasound yields a sensitivity of about 90% and a specificity of about 90% compared to pathology results. [Mammacarcinoom 2008]

MRI is applied for breast cancer diagnosis of a select patient group, e.g. diagnosis of the postoperative breast or when axillary nodes are positive and/or X-ray mammography and ultrasound do not provide satisfactory results. [Mammacarcinoom 2008, Koertshuis 2008] During MRI examination, the patient lies in prone position in a 'tunnel' with her breast positioned in a dedicated coil. The total examination takes about 30 minutes. Breast tumors do not offer inherent MRI contrast. To be able to identify tumors in an MRI image, Gadolinium-containing (Gd) contrast medium is administered intravenously. The use of Gd contrast agent offers the possibility image angiogenesis, the increase in blood vessels and permeability of the vessel wall around the tumor. This is a complex process: also benign lesions can appear as malignant lesions. MRI is a very sensitive technique, about 86-100%, at the cost of the specificity, which is about 40-70%. [Mammacarcinoom 2008]

There is not much research carried out with respect to the future use of diagnostic breast cancer imaging techniques. In 2005, the Dutch Cancer Society has published a report about the future of imaging techniques

within cancer research. [SCK 2005] Based on a literature review, and on a Delphi-research, they conclude that currently, X-ray mammography is the most frequently used technique for breast cancer diagnosis, and ultrasound is the second most frequently used technique, which is in accordance with the Breast Cancer Guideline. The Delphi panel expects that the use of X-ray mammography will decrease and that MRI will be more frequently used, in 2015 it is expected that MRI will be the most frequently used technique. The use of ultrasound is expected to remain the same. Furthermore, the Delphi panel indicates that they expect a lot from new, especially multi-modality, imaging techniques. Photoacoustics may be one of those promising new techniques.

Research questions

Though the PAM is still early in its development, it may be interesting to prospectively evaluate if PAM would be a better alternative compared to the existing imaging techniques, and to identify factors that need more attention in the further development of photoacoustics. The choice for an imaging strategy merely depends on the diagnostic performance (e.g. likelihood of tumor if test is positive), patient friendliness, and device and operating costs to the hospital. Therefore, the goal of this research is to find criteria that are important in the design of a diagnostic breast cancer imaging device, and to translate these criteria into design goals for further development of the PAM. Furthermore, the best position within the diagnostic track for breast cancer will be determined.

The main research question that will be addressed in this paper is:

What is the appropriate position for PAM in a diagnostic track, and does PAM perform equally or better compared to existing breast imaging modalities?

In this study, it has been chosen to look at the possible success of PAM in diagnosis, instead of screening. The current focus of development of PAM is on diagnosis, and the results of this research will therefore be more directly relevant. Also, currently more is known about the performance of PAM in a diagnostic setting, and the little data that has been produced in this diagnostic setting could provide some reference in the discussion about the possible success of PAM. The most appropriate position for PAM will be the position in which its performance will be equal to or better than the currently used breast imaging techniques at that position, and a position at which many patients are examined by an imaging modality. An equal or better performance of PAM means that different experts in the breast imaging field judge, based on their expertise and experience, that they will prefer PAM compared to the existing breast imaging modalities. These judgments will be based on the relative performance of the breast imaging modalities with respect to specific characteristics (among others: costs, effectiveness, and patient comfort), and the relative importance of these characteristics. Furthermore, the following sub-question will be addressed:

How can the implementation criteria be translated into development goals for the PAM, in order to improve its position with respect to alternative imaging techniques?

The main research question will be examined using the Analytic Hierarchy Process (AHP). AHP is a quantitative technique for multi-criteria decision analysis. This technique supports decisions about different concurring alternatives, and is also applicable for the assessment of relatively new technologies, when clinical evidence is not yet available or incomplete. [Hummel *et al.* 2000]

The sub-question will be answered using House of Quality analysis. House of quality is a decision making and planning tool that provides a structure to relate the demands of customers to engineering characteristics in the product design process.

Methodological approach: Analytic Hierarchy Process and House of Quality

To evaluate the future use of new or existing health care technologies, also known as Health Technology Assessment (HTA), the Analytic Hierarchy Process (AHP) analysis will be used. (More information about HTA can be found in appendix 2). AHP analysis derives an assessment of a technology by facilitating discussions between stakeholders with diverging backgrounds with respect to the technology being assessed. Stakeholders with these diverging backgrounds often have contrasting opinions about the relevance of the criteria for technology development, that may hinder technological change. Discussions about the future social, cultural and technical context of the technology may create more awareness about relevant issues that could concern technology development. [Hummel *et al.* 2000] AHP analysis, developed by Saaty (1980), is a structured technique for dealing with complex decisions. [Saaty 1980] The technique does not prescribe a correct

decision, but provides information about the preferences and importance of certain aspects of the problem. AHP provides a framework for structuring a problem, for representing and quantifying its elements, for relating those elements with overall goals, and for evaluating alternative solutions. First, a decision problem, often the question which alternative serves a goal best, is decomposed into a hierarchy of more easily comprehended sub-problems or levels that can be analyzed independently, such as the goal, the criteria and subcriteria, and the alternatives. Once this hierarchy is built, a decision making team can systematically evaluate the criteria of the problem by pairwise comparing the criteria to one another. In making the comparisons, the decision making team can use concrete data about the criteria, or they can use their judgments about the criteria's relative importance. Also, the performance of the alternatives with respect to the criteria and subcriteria is assessed pairwise. As a result, weighting factors, reflecting the importance of the criteria and the preferences for the alternatives, are computed. The relative importance or the preferences can be appointed on a 9-point ordinal scale, in which 1 reflects equal importance or preference and 9 is extreme importance or preference. This way, diverse and incommensurable elements can be compared to each other. In table 1, the fundamental scale for pairwise comparisons is presented. In the final step of the process, numerical priorities are calculated for each of the decision alternatives. These numbers represent the alternatives' relative ability to achieve the decision goal, so they allow a straightforward consideration of the various courses of action. This way, the AHP allows for prioritizing alternatives when multiple criteria must be considered. [Liberatore *et al.* 2007, Hummel *et al.* 2000] In addition, AHP provides a measure of inconsistency to ensure that each pairwise comparison is consistent with the remainder of the comparisons. When the AHP supports a group of decision makers, it aggregates the individual pairwise comparisons by computing a geometric group average. [Liberatore *et al.* 2007, Hummel *et al.* 2000]

Scenario analysis
Because PAM is early in its development, there will be uncertainty about its performance. There are different ways to assess uncertainty, depending among others on the type of uncertainty (statistical/scenario) the nature of the uncertainty (knowledge related/variability related) and the context and data available. One type of uncertainty assessment is for example asking experts to state the extreme minimum and maximum conceivable values for the variable. Another way of assessment is to systematically change input of a system and look at the corresponding change in output, also known as sensitivity analysis. [Van der Sluijs *et al.* 2004] In this study, both the above options will be combined: a decision making team will be asked to individually write down the uncertainty about different judgments, next this uncertainty is converted into a deviation from their initial judgment. This in turn results in different implementation scenarios: negative, average and positive, for PAM.

From the AHP analysis, numerical priorities that reflect the importance of criteria and the preference for alternatives result. The ultimate goal for PAM is to perform the best on the most important criteria. To achieve the best performance, technical development goals have to be deduced from the most important criteria. This will be done using House of Quality analysis. The method of House of Quality analysis was first published by J.R. Hauser and D. Clausing (1988), as a basic design tool for quality function deployment. The House of Quality is a kind of conceptual map that provides means for interfunctional planning and communications. It is mainly used to link customers desires to engineering possibilities. [Hauser *et al.* 1988]

Table 1 The fundamental scale for pairwise comparisons		
Intensity of importance	**Definition**	**Explanation**
1	Equal importance	Two elements contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one element over another
5	Strong importance	Experience and judgment strongly favor one element over another
7	Very strong importance	One element is favored very strongly over another
9	Extreme importance	One element is extremely favored over another
Intensities of 2,4,6 and 8 can be used to express intermediate values.		

In this research, the customers are the decision making team members. They have made their judgments about the relative importance of the criteria in the AHP. Also, the relative performance of the alternatives is known from the AHP. So, it is known what is wished for, the second step is to look how to realize this. This is done by adding technical criteria to the House of Quality, that are likely to affect one or more (clinical) criteria. The technical criteria for evaluation of the PAM will be provided by technology experts. The above steps form the 'walls' of the house. Next, the 'body' is constructed by filling in a 'relationship matrix' indicating how much each technical criterion affects each clinical criterion. At the 'bottom' of the house, objective measures and target values can be added. The 'roof' of the house represents how the change of one technical criterion affects the other technical criteria. This way, a schematic representation is obtained that can help the designers of the PAM to set targets and to understand the clinical priorities and goals.

METHODS

In the present study AHP is applied. The AHP analysis will be based on the hierarchical structure depicted in figure 3. This structure is composed of a goal, alternatives and criteria. The evaluation goal, to identify the best alternative for diagnostic breast imaging, is identified by interviews with the designers of PAM, they have an idea of the feasibility of different options for PAM in the coming five years and are interested in the possible success of these options. Also, interviews with medical specialists have been carried out in order to verify if the goal set by the designers of PAM is realistic.

Implementing PAM in a diagnostic track: alternatives

The alternatives in the AHP hierarchy are the new technology, and its most important alternatives. The new technology is in this case PAM II. PAM II is the version of PAM that will be released within five years. This version produces tomographic images of the full breast. During scanning, the patient lies in prone position with her breast hanging in a cup filled with water (at body temperature). The laser and detector matrix spin around this cup. A unique feature that this device will offer is the possibility to provide information about the level of oxygen saturation of the blood around the tumor. This extra information may make a (invasive) biopsy unnecessary for certain patients. Speed-of-sound imaging is another extra feature that will make the production of ultrasound like images possible. It is assumed that photo-acoustics will replace one or more techniques, instead of being an additional technique. This assumption is made because time and money are important, and additional methods will result in longer procedure times and higher costs. Therefore, the implementation barrier is expected to be higher. For the selection of alternatives, the diagnostic track used at the mammacare department of Medisch Spectrum Twente is studied (appendix 1), together with the Dutch Breast Cancer Guideline. The diagnostic track starts with a combined examination by ultrasound and X-ray mammography. In 2008, all 764 patients that visited the mammacare department of Medisch Spectrum Twente were examined by these imaging modalities. Of these 764 patients, about 200 patients (26%) were further examined by MRI. About half of the lesions of these 200 patients were classified, based on ultrasound and X-ray mammography, as Bi-RADS 0 – III (classification not possible – probably benign), the other half of the lesions was classified as Bi-RADS IV or V (probably malignant). [Cijfers CvM 2008] In this research, it has been chosen to compare PAM to the combined use of X-ray mammography and ultrasound, and also to compare the combined use of PAM and X-ray mammography (PAM as a substitute for ultrasound only) to these alternatives. The combined use of X-ray mammography and ultrasound is the conventional start of the diagnostic track, these imaging modalities therefore are the most used. A role for PAM at the start of the track would therefore be beneficial from an economical point of view. Because a considerable amount of patients is further referred for an MRI examination, PAM will also be compared to MRI. The use of MRI is expected to increase in the coming years, and may be the most frequently used imaging modality in cancer imaging in the near future. [SCK 2005]

Identification of relevant criteria

To identify relevant criteria for the AHP, a first literature search and observations at the centre for mammacare at Medisch Spectrum Twente provided a longlist of factors influencing the quality of a breast imaging device. These longlists were discussed with different professionals involved in this research, and the most important criteria were put in a shortlist. From this shortlist, the hierarchical structure presented in figure 3 was produced. The literature search started with a visit to the website of the different cancer institutes and societies in the Netherlands. Via these websites national guidelines for breast cancer diagnosis were found. The guidelines provided the top-level criteria.
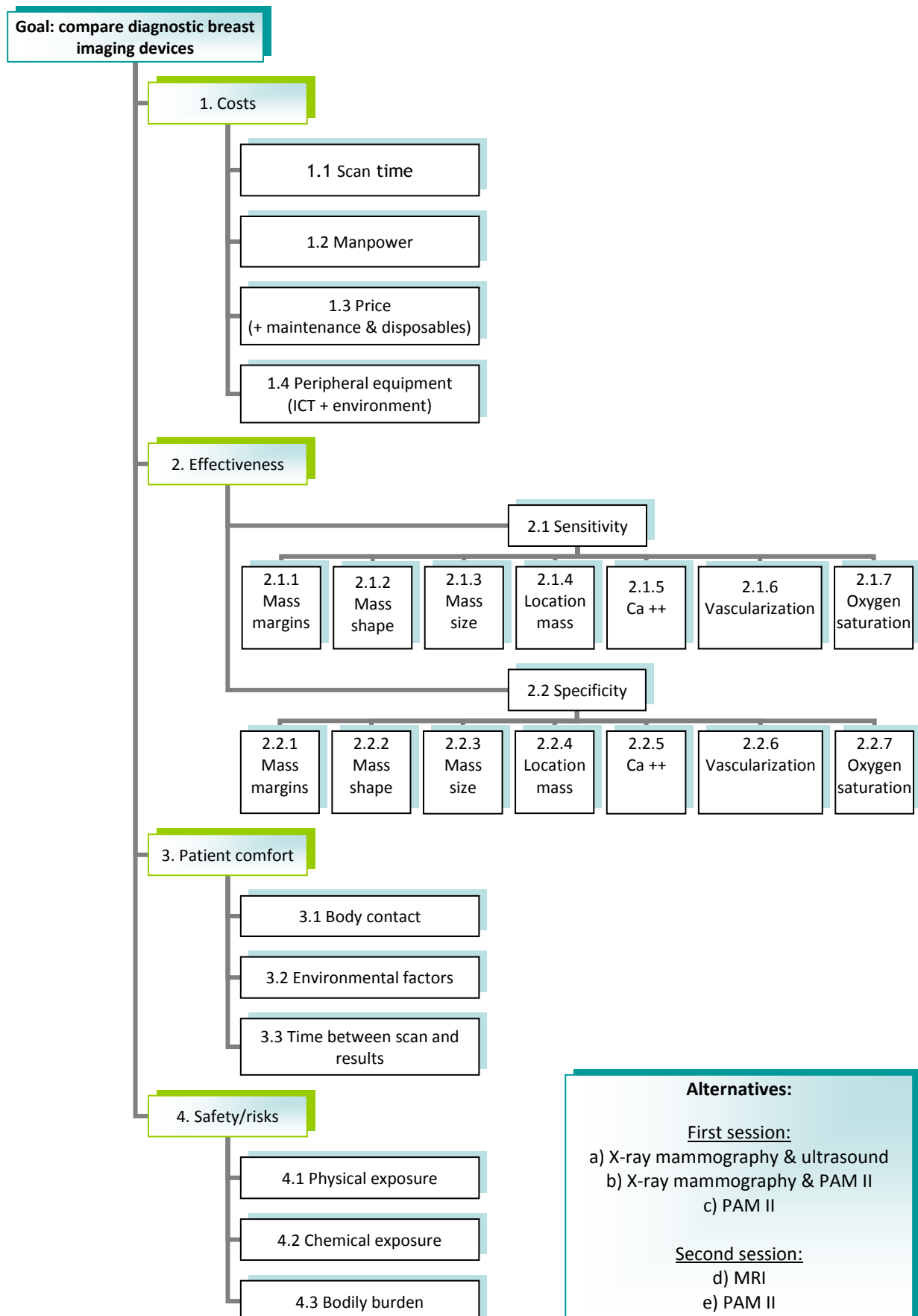
Figure 3 *AHP hierarchy diagnostic breast imaging device*

The costs subcriteria followed from an interview with the investment coordinator of Medisch Spectrum Twente. The effectiveness of a device can be evaluated both 'technically', e.g. by using criteria like spatial and temporal resolution, and 'clinically', by e.g. using criteria like ability to determine tumor size and shape. It has been chosen to use a 'clinical' approach in this AHP, and to relate the clinical factors to the technical factors afterwards using House of Quality. The clinical factors followed from clinical breast imaging literature that was provided by the radiology department of Medisch Spectrum Twente, and from the world-wide used classification standard for breast lesions (Bi-RADS). The most frequently mentioned tumor aspects in these documents are presented in the hierarchy. Furthermore, oxygen saturation is added because this is a new feature PAM offers, and the AHP results may show if this new feature will be important. Patient comfort subcriteria followed from the information on the websites of the different cancer institutes and societies, and from information on patient organization websites. Furthermore, detailed studies on the different alternatives (used for the performance matrix) sometimes revealed reasons for non-cooperation of the patients, e.g. claustrophobic patients that are afraid of MRI examination because of the small 'tunnel' in which they are situated. All patient comfort subcriteria identified could be classified under one of the three criteria presented. The safety/risks subcriteria followed from incident analyses and an interview with a medical physicist of Medisch Spectrum Twente. The definitions of all criteria and subcriteria presented in the hierarchy can be found in appendix 4. The performance of all alternatives with respect to the criteria and subcriteria defined above, is presented in a performance matrix. This performance matrix can be found in appendix 5.

AHP expert team
In interviews with the developers of PAM and with different medical experts, potential team members were identified. The aim was to create a multidisciplinary expert team, in which experts with different professional backgrounds would take place. All stakeholder groups identified by Wallner (2008) were approached. Table 2 presents the invited team members together with their professional backgrounds. The health insurance expert did not react on the invitation. Two AHP sessions were organized, for which different people were approached (table 2). Unfortunately, only during the second session a manager was present. The manager that was approached for the first session decided last minute not to join the expert group and could not be replaced at that time by another manager.

AHP feedback session
All members of the AHP expert team have received information about the AHP hierarchy and the performance matrix in advance. A day-section feedback session was organized in which the expert team discussed about the relative importance of the criteria, and about the pursued quality of each diagnostic alternative. This was supported using Team Expert Choice software (a commercially available group decision support system that incorporates the mathematical procedures of the AHP). The feedback session started with an introduction of the software and the procedures of Team Expert Choice, and the designer of the PAM explained the backgrounds of this diagnostic device. Furthermore, the AHP hierarchy was explained. Then, using hand-held radiographic keypads, the members of the expert team provided their judgments on each pairwise comparison. Individual judgments were projected on a screen, allowing the members of the expert team to discuss the rationales behind their individual scores. During the discussions, the expert team members could alter their judgments.

No.	Profession	Core PAM activity/relation
Table 2 Composition AHP expert team		
1	Health insurance expert	Reimbursement, costs
2	Manager*	Costs for health care organization
3	Medical specialist (Radiologist)*	User
4	Nurse practitioner/radiology assistant*	User, patient representative
5	Physicist	Technical design of PAM
6	Laser physicist	Technical design PAM principles
7	Physicist	Research for one of the world's leading medical device companies, main interest in optical mammography
8	Medical physicist	Safety & Quality assurance
*these professions were represented by different persons during the first and the second session		

To support the team members and to make sure every expert used the same definition of a criterion, the descriptions of the criteria that were assessed were displayed on a second screen. For each pairwise comparison, the final individual judgments were aggregated based on the geometric mean. Team Expert Choice next provided weighting factors representing the importances of the criteria and the priorities reflecting the qualities of the alternatives.

Assessment of uncertainty: scenario analysis

It is expected that the criteria regarding the effectiveness of the devices will be difficult to assess, because of the limited amount of data that has been obtained with the PAM to this date. Therefore, the expert team members were asked to write down, on a 3-point scale, how certain they are about their judgments of the relative performance of the alternatives. This 3-point scale was converted into a corresponding deviation from the initial judgments of each team member. When an expert indicated that he/she is 'very certain' about a judgment, the corresponding deviation is zero. The indication 'moderately certain', corresponds to a deviation of 2 points and the indication 'uncertain' corresponds to a deviation of 4 points. A deviation of 2 points on the fundamental scale for pairwise comparisons (table 1) reflects one level in definition. The deviations can be negative (a loss in the performance of PAM) and positive (a gain in the performance of PAM). This way, three scenarios, most negative, average, and most positive, resulted that reflect the overall performance of PAM. These scenarios can be used to identify margins for the predicted performance of PAM. Uncertainty about an important factor will lead to a larger deviation in the overall performance of PAM than uncertainty about a less important factor. Therefore, next to improving the performance of PAM with respect to this factor, it is also important to reduce uncertainty. When more (fundamental) research is done and more certainty exists the margins in the overall performance scenario become smaller.

RESULTS

Two expert teams were invited, for two separate AHP sessions. In the first session, the performance of PAM II in comparison with X-ray mammography and ultrasound was assessed (alternatives a-c, bottom box figure 3). In the second session, the performance of PAM II in comparison with MRI was assessed (alternatives d-e, bottom box figure 3). This separate assessment was carried out because other performance data and patient groups exist for both sets of alternatives. In appendix 3, a program of the two sessions is provided. The results of both sessions will be presented separately.

Overall results first session: X-ray mammography & ultrasound vs. PAM II

The overall results of the first session, in which the three alternatives X-ray mammography & ultrasound, X-ray mammography & PAM II and PAM II were compared, are presented in table 3 and figure 4. The numbers between brackets in the column headings reflect the relative importance weights of the criteria. The numbers in the row headings reflect the overall relative preference for the alternatives. The other numbers in table 3 reflect the relative preference for the alternatives with respect to the criteria in the corresponding columns. From figure 4b it becomes clear that the most preferred diagnostic alternative for imaging breast cancer is PAM II (relative weight .442, table 3). The combined use of mammography and PAM II is the least preferred alternative, presumably because of the least preferred performance on all criteria besides effectiveness (table 3). Effectiveness however is the most important criterion. The most important subcriterion is sensitivity (2.1, see figure 4a), which means that a change in the performance of the alternatives with respect to sensitivity, is most likely to change the overall preference. The combined use of X-ray mammography and ultrasound is expected to show the highest sensitivity (.41, table 3), the use of solely photoacoustics is expected to show the lowest sensitivity (.25, table 3). Safety/risks is also an important criterion, but because all alternatives are relatively save, the performance of the alternatives with respect to this criterion will not be decisive for the overall preference. The expert team indicates that the radiation dose for x-ray mammography is not a very important safety factor for a diagnostic application of the alternatives. For a use of the alternatives in screening, it is expected that radiation dose will be more important. Chemical exposure is rated more important than physical exposure, because the expert team is relatively more familiar with the low physical risks, and expects that the time span of the possible injury caused by chemical exposure will be larger.

Criteria	Costs (.097)				Effectiveness (.551)		Patient comfort (.085)			Safety/risks (.267)		
	1.1 (.42)	**1.2** (.32)	**1.3** (.16)	**1.4** (.10)	**2.1** (.87)	**2.2** (.13)	**3.1** (.34)	**3.2** (.11)	**3.3** (.55)	**4.1** (.14)	**4.2** (.36)	**4.3** (.51)
Alternatives												
Mammography & Ultrasound .303	.15	.10	.19	.23	.41	.43	.11	.39	.79	.11	.33	.12
Mammography & PAM II .256	.06	.14	.05	.11	.34	.33	.08	.09	.10	.08	.33	.08
PAM II .442	.79	.76	.76	.66	.25	.24	.81	.52	.11	.81	.33	.80

Table 3 Results of Mammography, US & PAM comparison

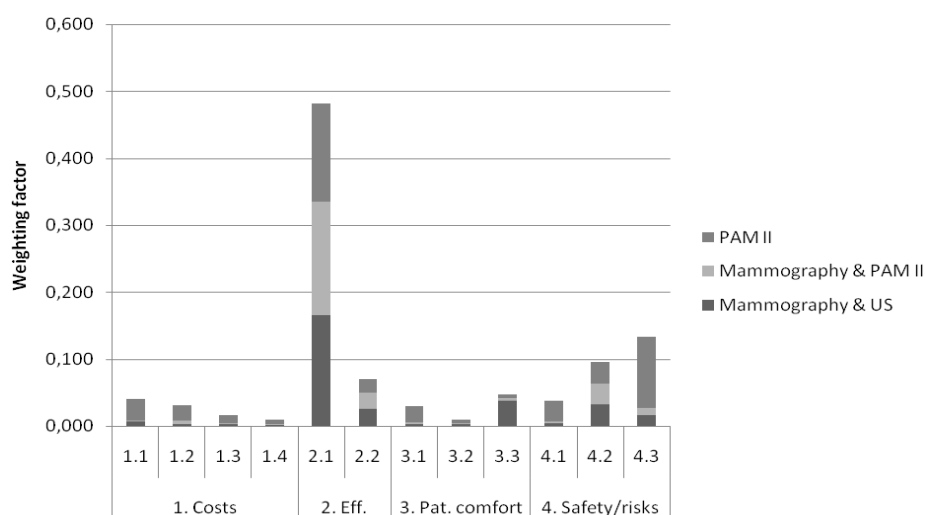***Figure 4a*** *Importances of the criteria, overall results first session. 1.1 = Scan time, 1.2 = Manpower, 1.3 = Price, 1.4 = Peripheral equipment, 2.1 = Sensitivity, 2.2 = Specificity, 3.1 = Body contact, 3.2 = Environmental factors, 3.3 = Time between scan and results, 4.1 = Physical exposure, 4.2 = Chemical exposure, 4.3 = Bodily burden*
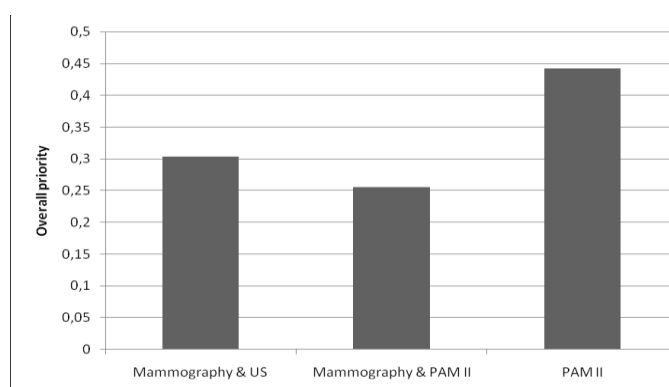

Figure 4b *Preferences for the alternatives, overall results first session*

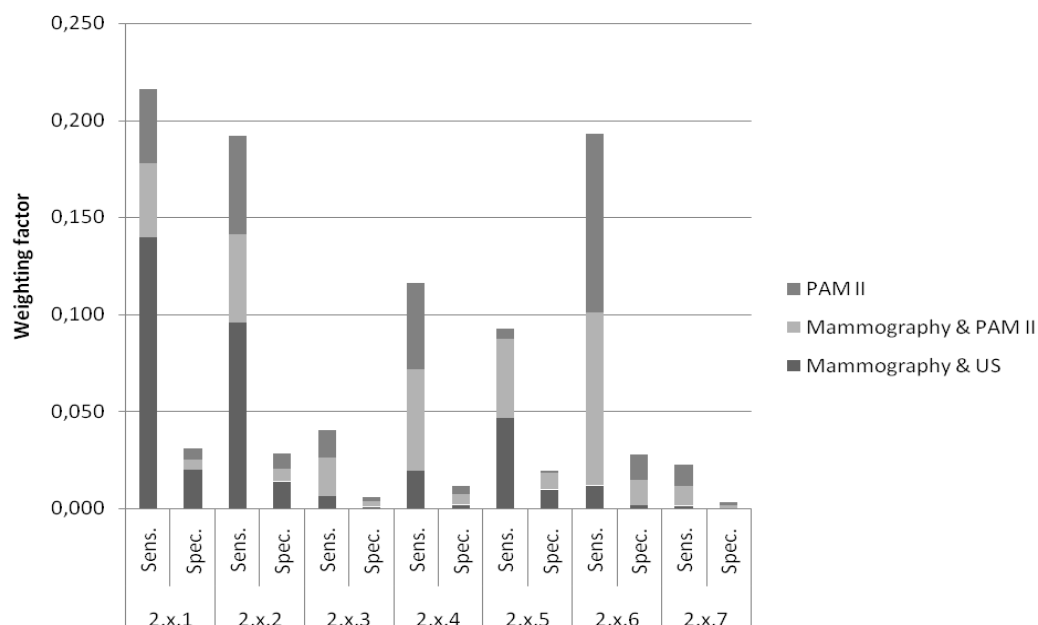| Table 4 Results of Mammography, US & PAM comparison, effectiveness subcriteria														
Criteria **Alternatives**	**Sensitivity** **(.874)**							**Specificity** **(.126)**						
	2.x.1 (.25)	2.x.2 (.22)	2.x.3 (.05)	2.x.4 (.13)	2.x.5 (.10)	2.x.6 (.22)	2.x.7 (.03)	2.x.1 (.25)	2.x.2 (.22)	2.x.3 (.04)	2.x.4 (.09)	2.x.5 (.15)	2.x.6 (.22)	2.x.7 (.03)
Mammography & Ultrasound .347	.65	.50	.16	.17	.50	.06	.06	.65	.50	.16	.17	.50	.06	.06
Mammography & PAM II .351	.18	.24	.50	.45	.44	.46	.46	.18	.24	.50	.45	.44	.46	.46
PAM II .302	.17	.26	.34	.38	.06	.48	.48	.17	.26	.34	.38	.06	.48	.48

Figure 5a *Importances of the effectiveness subcriteria, first session. Sens. = sensitivity, Spec. = specificity. 2.x.1 = mass margins, 2.x.2 = mass shape, 2.x.3 = mass size, 2.x.4 = location mass, 2.x.5 = Ca ++, 2.x.6 = vascularization, 2.x.7 = oxygen saturation.*
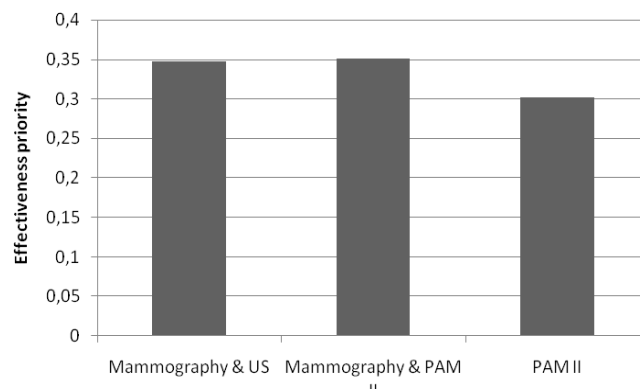

Figure 5b *Preferences for the alternatives with respect to effectiveness, first session.*

Effectiveness subcriteria

The effectiveness subcriteria have been more extensively assessed. The results of this effectiveness subcriteria analysis are presented in table 4 and figure 5. The effectiveness subcriteria have been assessed only by the radiologist and technical experts. Due to time restrictions, the performance of the alternatives with respect to the sensitivity and specificity subcriteria has only been assessed once. It is assumed that the performance of the alternatives will be equal with respect to the same sensitivity and specificity subcriteria. From figure 5b and table 4, it becomes clear that the performance of all alternatives with respect to effectiveness is almost equal, with PAM II being the least preferable option.

Sensitivity is a lot more important than specificity. The most important subcriteria that determine the sensitivity are the presentation of mass margins, mass shape and vascularization in the images. For the mass margins it is important to be able to discriminate between sharp defined margins and unsharp, diffuse margins. PAM II mainly displays the blood vessels that are situated in or around the tumor. The expert team is not sure whether blood vessels grow at the exact margin, therefore PAM II is the least preferred alternative for displaying mass margins. With respect to mass shape it is important to be able to discriminate between regular and irregular structures. The expert team indicates that ultrasound is better at displaying structures within the mass, therefore the combined use of mammography and ultrasound is the most preferred alternative with respect to this criterion. The presence of vascularization is highly indicative for malignancy. Vascularization is best visualized by PAM II (whether or not accompanied by x-ray mammography). Ultrasound makes use of Color Doppler to measure flow, but this feature is only used to search for large vessels that are characteristic for benign fibroadenomas. Also the location of the mass and the presence of Ca ++ (in a specific shape) are important for the sensitivity of the alternatives. Ca ++ is only visible by mammography, therefore both alternatives that make use of mammography have the highest performance on this aspect. With respect to the location of the mass it is important that the full breast is imaged (preferably in 3D) and that the fasci interface between subcutaneous fat and water can be imaged with high quality. At this interface, most lobuli are situated, in which tumor growth may occur. The alternatives that make use of PAM II are most preferred with respect to the visualization of the location of the mass. The size of the mass and the oxygen saturation are relatively unimportant for the sensitivity of the alternatives. The exact size of the mass is, even pathologically, difficult to determine, and determination of the size is highly operator dependent. The physical contrast mechanisms on which the different imaging techniques are based all result in a different representation of the tumor, and therefore of the tumor size. Only the minimum size that can be detected is of importance, not the exact size. It is doubted by some of the experts if PAM II will be able to measure oxygen saturation levels, some of the experts do not provide judgments with respect to this criterion. With respect to specificity, the order of importances of the subcriteria is equal to the order of importances of the subcriteria with respect to sensitivity.

Overall results second session: MRI vs. PAM II

The overall results of the second session, in which the alternatives MRI and PAM II were compared, are presented in table 5 and figure 6.From figure 6b, it becomes clear that PAM II is preferred over MRI for imaging breast cancer. The priorities for the different alternatives are however not far from each other (.57 vs. .43). The most important criterion is sensitivity (2.1, see figure 6a), which means that a change in the performance of the alternatives with respect to sensitivity, is most likely to change the overall preference. The performance of MRI with respect to sensitivity (.52, table 5) is expected to be a bit better than the performance of PAM II (.48, table 5). On all criteria besides effectiveness, PAM II is preferred over, or equally preferred to, MRI. Safety/risks is the second most important criterion. Both physical and chemical exposure are important subcriteria that determine the performance with respect to safety/risks. Chemical exposure is rated a bit more important than physical exposure, because the expert team is relatively more familiar with the physical risks, and expects that the time span of the possible injury caused by chemical exposure, due to the injection of Gd contrast agent in MRI, will be larger.

Effectiveness subcriteria

The results of the effectiveness subcriteria analysis of the second session are presented in table 6 and figure 7. From figure 7b and table 6 becomes clear that the performance of both alternatives with respect to effectiveness is almost equal, with PAM II being the least preferable option.

Table 5 Results of MRI & PAM comparison												
Criteria	Costs (.113)				Effectiveness (.611)		Patient comfort (.047)			Safety/risks (.228)		
	1.1	**1.2**	**1.3**	**1.4**	**2.1**	**2.2**	**3.1**	**3.2**	**3.3**	**4.1**	**4.2**	**4.3**
Alternatives	(.26)	(.23)	(.34)	(.17)	(.80)	(.20)	(.17)	(.26)	(.57)	(.32)	(.56)	(.12)
MRI .43	.50	.38	.36	.14	.52	.54	.32	.18	.50	.38	.15	.20
PAM II .57	.50	.62	.64	.86	.48	.46	.68	.82	.50	.62	.85	.80

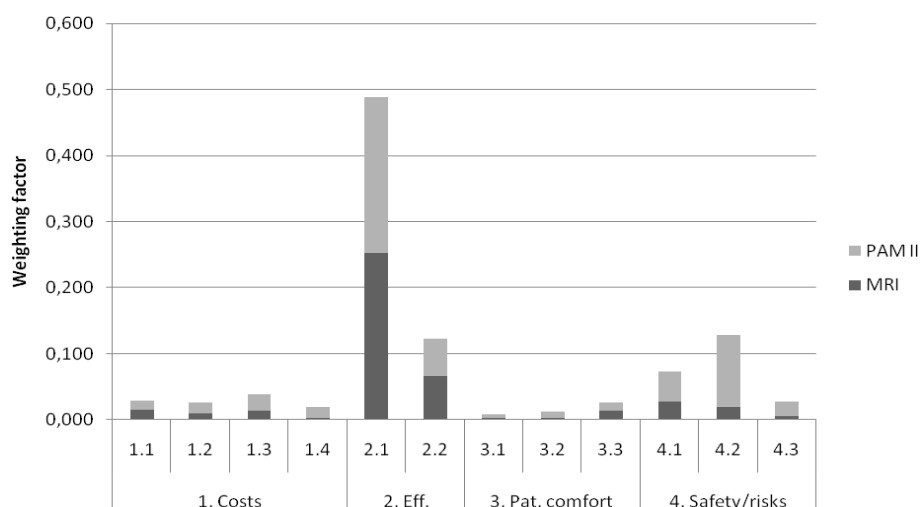*Figure 6a* Importances of the criteria, overall results second session. 1.1 = Scan time, 1.2 = Manpower, 1.3 = Price, 1.4 = Peripheral equipment, 2.1 = Sensitivity, 2.2 = Specificity, 3.1 = Body contact, 3.2 = Environmental factors, 3.3 = Time between scan and results, 4.1 = Physical exposure, 4.2 = Chemical exposure, 4.3 = Bodily burden
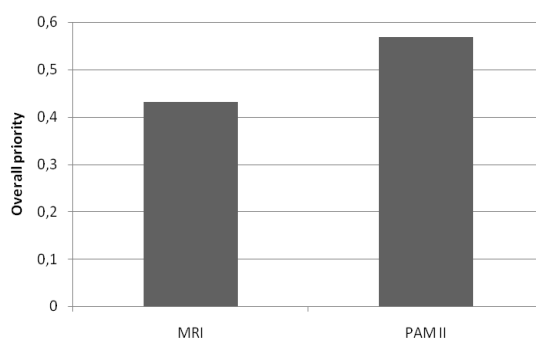
Figure 6b Preferences for the alternatives, end results second session

Criteria	Sensitivity (.799)							Specificity (.201)						
Alternatives	2.x.1 (.14)	2.x.2 (.14)	2.x.3 (.03)	2.x.4 (.01)	2.x.5 (.06)	2.x.6 (.37)	2.x.7 (.25)	2.x.1 (.40)	2.x.2 (.12)	2.x.3 (.02)	2.x.4 (.02)	2.x.5 (.06)	2.x.6 (.18)	2.x.7 (.20)
MRI .521	.67	.66	.80	.66	.50	.57	.14	.67	.66	.80	.66	.50	.57	.14
PAM II .479	.33	.34	.20	.34	.50	.43	.86	.33	.34	.20	.34	.50	.43	.86

Table 6 MRI & PAM comparison, effectiveness subcriteria

Figure 7a *Importances of the effectiveness subcriteria, second session. Sens. = sensitivity, Spec. = specificity. 2.x.1 = mass margins, 2.x.2 = mass shape, 2.x.3 = mass size, 2.x.4 = location mass, 2.x.5 = Ca ++, 2.x.6 = vascularization, 2.x.7 = oxygen saturation.*

Figure 7b *Preferences for the alternatives with respect to effectiveness, second session.*

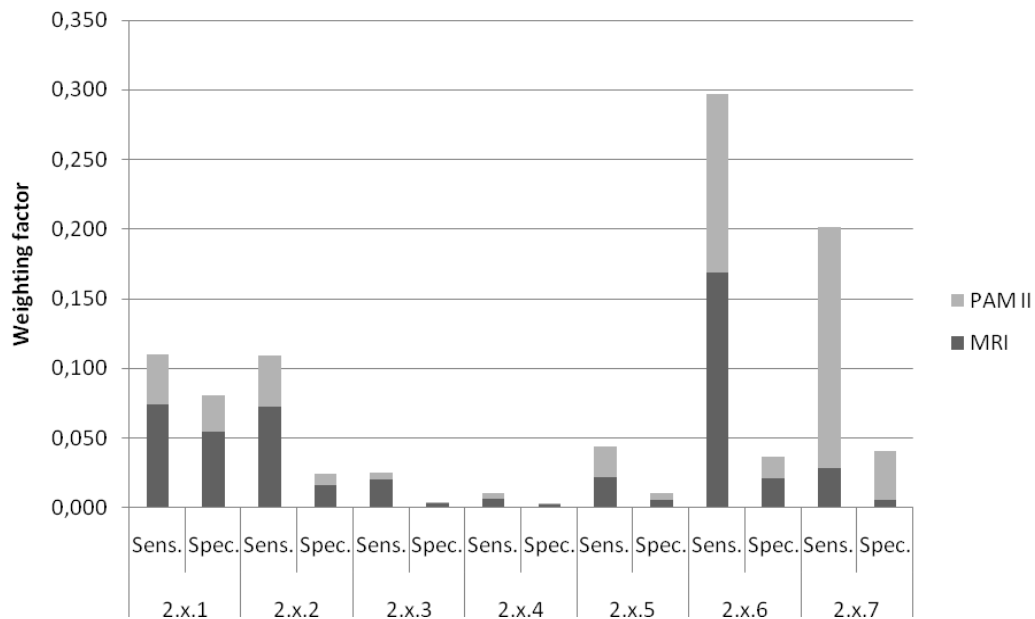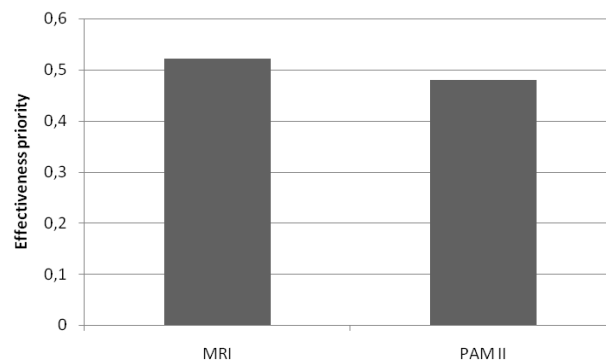Sensitivity is a lot more important than specificity. The most important subcriteria that determine the specificity are vascularization and oxygen saturation. Vascularization can be visualized both by MRI and photoacoustics, the detail resolution of MRI is however higher than that of PAM II, which results in a slight preference for MRI with respect to imaging vascularization. Providing information about the oxygen saturation of the blood surrounding the tumor is only possible by PAM II, which results in a higher performance of PAM II with respect to this criterion. The presentation of mass margins and mass shape are also important subcriteria of sensitivity. MRI uses contrast agent to visualize the lesions, this contrast agent may leak into the surroundings of the lesion (few cells thickness), which makes it harder to visualize the exact margins. MRI is however the most preferred alternative for both these criteria, because of the higher contrast and resolution that can be obtained with MRI. The visibility of the margins and shape depends among others on the product of these factors. Mass size, the location of the mass and Ca++ are relatively unimportant criteria for the sensitivity of the alternatives. With respect to specificity the visualization of mass margins becomes more important, and the visualization of vascularization is less important. When the margins are solid and clearly defined, this is the most important indicator of a benign lesion.

Scenario analysis

The tables in appendix 6 present the results of the analysis of three scenarios of the first and second session. In figure 8a, the effectiveness preferences for the three scenarios for the first session are presented. In figure 8b, the resulting overall preferences for the three scenarios are presented.

Figure 8a Preferences for the alternatives, effectiveness results first session. Neg. = negative scenario, av. = average scenario, pos. = positive scenario.

Figure 8b Preferences for the alternatives, results first session. Neg. = negative scenario, av. = average scenario, pos. = positive scenario.
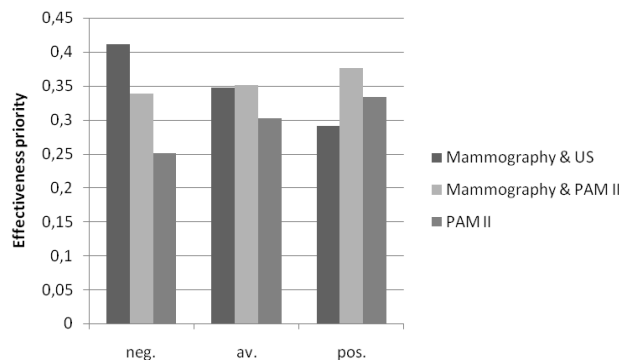

Figure 9a Preferences for the alternatives, effectiveness results second session. Neg. = negative scenario, av. = average scenario, pos. = positive scenario
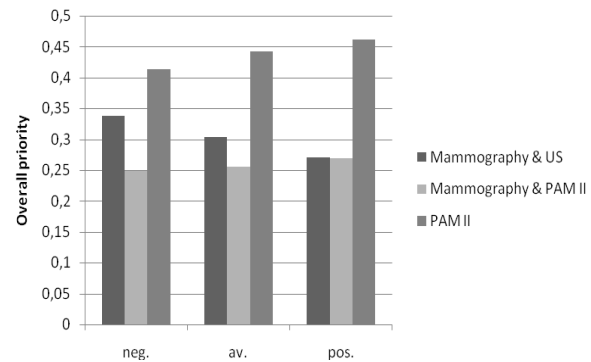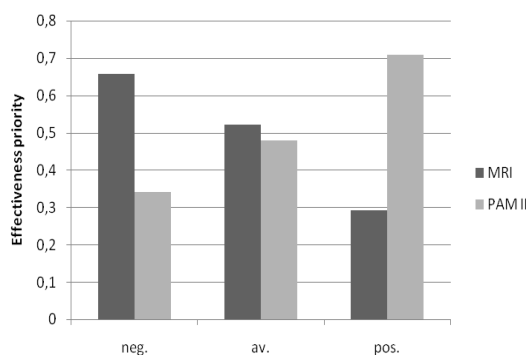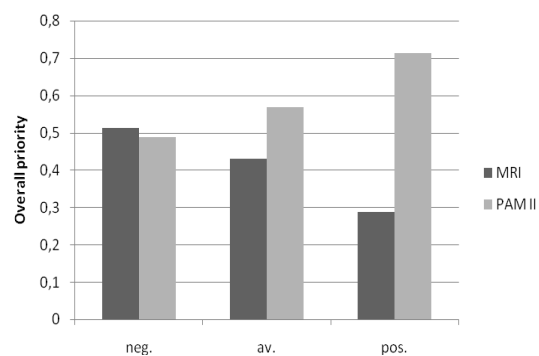

Figure 9b Preferences for the alternatives, results second session. Neg. = negative scenario, av. = average scenario, pos. = positive scenario.

With increasing positivity, PAM II becomes a more preferred option, its performance with respect to effectiveness remains however worse than that of the combined use of PAM II and X-ray mammography. Also in the overall results, the preference for PAM II increases with increasing positivity. PAM II is in all scenario's, also in the negative scenario, the most preferred alternative. The combined use of x-ray mammography and PAM II is only in the positive scenario not the least preferred alternative.

In figure 9a, the effectiveness preferences for the three scenarios for the second session are presented. In figure 9b, the resulting overall preferences for the three scenarios for the second session are presented. With increasing positivity, PAM II becomes a more preferred option, however only in the positive scenario the performance of PAM II with respect to effectiveness is better than the performance of MRI. In the overall results, PAM II is the most preferred alternative in the average and positive scenario, with increasing priority values for increasing positivity.

Certainty vs. Group variance
Figure 12 presents a scatter plot of the group average of the certainty about each judgment that was made by the members of the expert team (in the second part of the sessions) versus the group variance of these judgments.
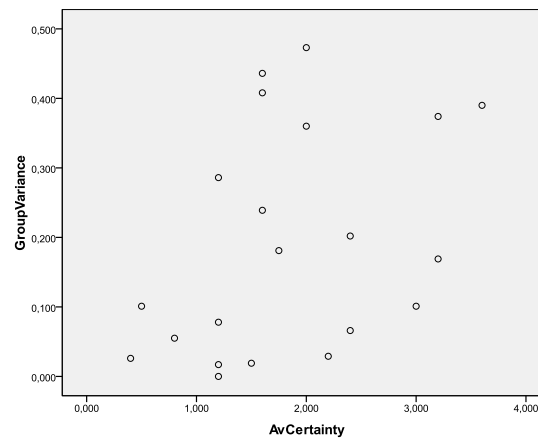

Figure 10 Scatter plot of the certainty group average vs. the group variance

The Spearman correlation was computed; a significant correlation coefficient of .445 (p = .043) exists for this data set. This represents a moderate linear relation between the group average of the certainty about each judgment and the group variance of these judgments.

Subgroup analysis
To check if there are discrepancies between the judgments of the technology creators, and the other stakeholders, a subgroup analysis has been carried out. Subgroups have been identified based on the interest they have concerning PAM. The composition of the different subgroups is presented in table 7.
In figures 11 and 12, graphs are presented with the results of the different subgroups. Figure 11 presents the results of the first session in which X-ray mammography, ultrasound and PAM II were compared, figure 12 presents the results of the second session in which MRI and PAM II were compared. (In the first session, no manager was available, therefore the 'B' subgroup is lacking.) In these results, the performance of the alternatives with respect to sensitivity and specificity is based on a rough estimation of the expert team, as opposed to the results of the more extensive assessment by means of sensitivity and specificity subcriteria, presented in figures 5 and 7.
From figure 11a becomes clear that the technology creators focus more on effectiveness and costs than the technology users. For the technology users, patient comfort and safety/risks are more important than for the technology creators. The relative order of the criteria is equal for both subgroups. The technology user judgments show that the performance of PAM II is best with respect to patient comfort and safety/risks. The relatively lower performance in effectiveness does not prevent PAM II from resulting as the number one alternative for the technology users. The larger importance of effectiveness, and the relatively lower performance of PAM II in effectiveness, results in a lower position of PAM II in the overall preferences of the technology creators. The combined use of mammography and PAM II is preferred the most by them.

Table 7 Subgroups		
Selection criteria	**Subgroup**	**Members**
Interest	Technology buyer (B)	Manager
	Technology users (U)	Nurse practitioner/radiology assistant
		Radiologist
	Technology creators (C)	Physicist – Technical design of PAM
		Laser physicist - Technical design PAM principles
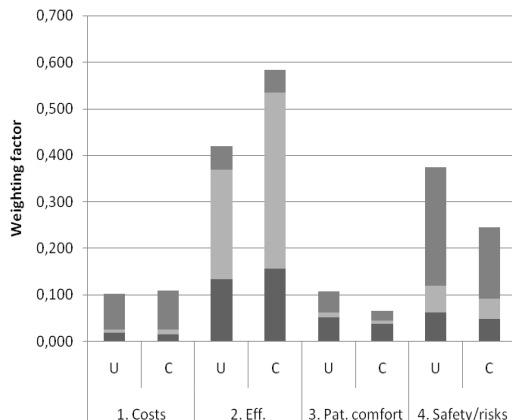
Figure 11a *Importances of the criteria, first session. U = Technology user subgroup, C = Technology creator subgroup.*
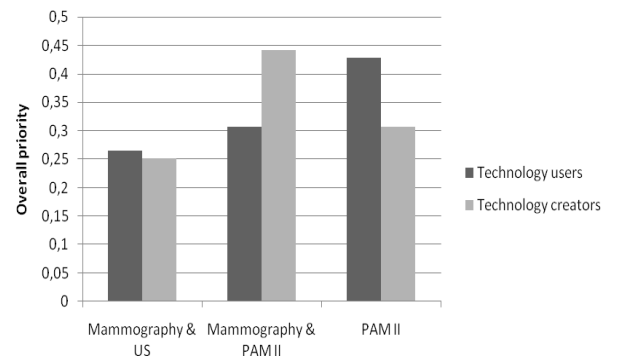
Figure 11b *Preferences for the alternatives, first session.*
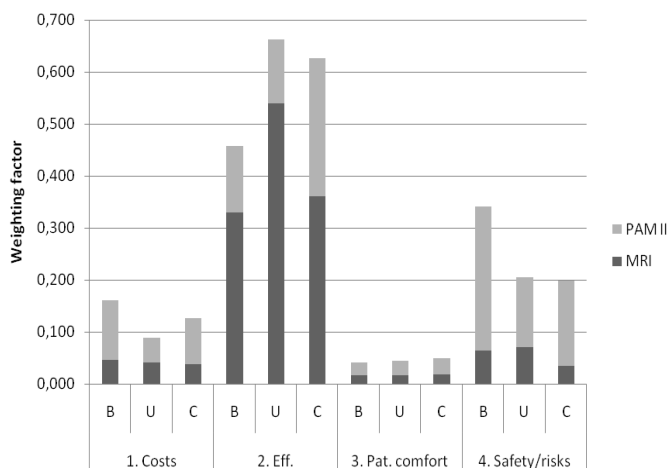
Figure 12a *Importances of the criteria, second session. B = Technology buyer subgroup, U = Technology user subgroup, C = Technology creator subgroup.*
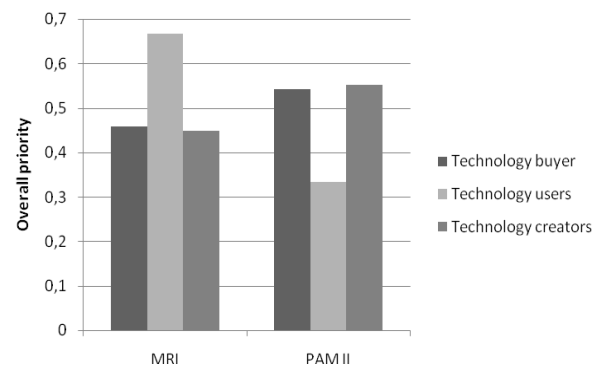
Figure 12b *Preferences for the alternatives, second session.*

Figure 12a shows that for the technology buyer, costs and safety are more important than for the other subgroups. For the technology creators, costs are more important than for the technology users, and effectiveness is a bit less important than for the technology users. For the technology buyer, effectiveness is also less important than for the other subgroups. The relative order of the criteria is equal for all subgroups.

The technology buyer and technology creators subgroup show similar results for the overall judgment of the performance of the alternatives (figure xb). PAM II is slightly better, this is the result of the higher performance on costs and safety criteria. Furthermore, the technology creators subgroup expects a higher relative performance of PAM II on effectiveness than the other subgroups. The technology users expect a relatively low performance of PAM II on effectiveness, and because this criteria is a lot more important than the other criteria on which PAM II performs best, in the overall results PAM II is the least preferred alternative.

House of Quality
The most important technological improvements for PAM II follow from the most important clinical criteria. In table 8, the most important criteria are presented that resulted from the AHP analysis, together with their overall relative weighting factors. These criteria are used as input in the House of Quality analysis. Also the criterion 'time between scan and results' has been added, because PAM II is the least preferred option with respect to this criterion.

Table 8 Most important criteria resulting from the AHP analysis			
First session: X-ray, US & PAM II		**Second session: MRI & PAM II**	
Criterion	**Relative weight**	**Criterion**	**Relative weight**
Bodily burden	.133	Vascularization	.182
Mass margins	.119	Oxygen saturation	.123
Mass shape	.106	Mass margins	.067
Vascularization	.106	Mass shape	.067
Time between scan and results	.047	Time between scan and results	.027

The designer of PAM II has filled in a template House of Quality that was retrieved from the QFD Online website. [QFD online] The resulting House of Quality is presented in appendix 7.
In table 9, the resulting technical criteria, together with their relative weights, target values and accomplishment difficulty are presented.

Table 9 Results House of Quality analysis			
Technical criterion	**Target value**	**Difficulty** **(0 = easy, 10 = extremely difficult)**	**Relative weight**
Reconstruction algorithm quality		5	22.1
Sensitivity detector	0.5 Pa	8	20.4
Detector bandwidth	150 %	9	16.7
Number of wavelengths	5	6	15.8
No. of parallel detector elements	640	7	8.5
Max. power	MPE	5	7.9
Design		8	7.8
Repetition rate laser	100 Hz	8	0.7

From table 9 can be concluded that the quality of the reconstruction algorithm, the sensitivity of the detector, the bandwidth of the detector and the number of wavelengths are the most important technical aspects to improve. The reconstruction algorithm are the mathematics that produce a reconstruction of the imaged object out of different projections. Important input for the reconstruction algorithm are statistical knowledge of the data acquisition process and geometry of the data imaging system. The sensitivity of the detector determines how much of the original signal is caught (compared to background noise) and can be used to create an image. The detector bandwidth determines the ability of the detector to detect the signals of absorbers over a range of sizes. Furthermore, the number of wavelengths determines the penetration depth of the light and the absorption contrast at the tumor site. [Manohar *et al.* 2008] The visualization of the most important criteria resulting from the AHP analysis depends for a large part on the resolution of the imaging system. The resolution is mainly determined by the measurement time, the detector sensitivity and the quality of the reconstruction algorithm. [Manohar *et al.* 2008]

DISCUSSION AND CONCLUSION

The overall results indicate that, based on the performance on costs, effectiveness, patient comfort and safety/risks criteria assessed, PAM II will be the most preferred alternative. The results of the different alternatives are however close to each other.

Most important for the performance of the different diagnostic imaging techniques is effectiveness; the performance on sensitivity aspects mainly determines how the different techniques are appreciated by the experts. Sensitivity is a complex criterion, there has been discussion about the exact definition within the context used for this research. A subdivision into clinical imaging subcriteria has been made, in order to clarify the importance of and performance on benign and malignant tumor aspects that radiologists assess in the breast images. The presentation of vascularization is one of the most important indicators of malignancy. Furthermore, results of the judgments made by the expert team show that the presentation of mass margins and mass shape are also important preconditions for providing a high-sensitivity diagnosis. The added value of information about the level of oxygen saturation remains unclear. In the assessment of MRI versus PAM II this subcriterion scores high in importance, in the assessment of mammography, ultrasound and PAM II this subcriterion is the least important. This criterion is unique for PAM II, because the other diagnostic imaging techniques are not able to provide any information about the level of oxygen saturation. When the radiologists should agree that oxygen saturation offers important information about the malignancy of a mass, it is expected that the need for information about this unique feature will boost the development of PAM II. A unique feature of mammography is its ability to provide information about Ca ++, which may be important for the sensitivity of the diagnosis. Due to the possession of this unique feature, mammography remains a popular imaging technique for breast cancer. If the developers of PAM II decide to focus on further development of oxygen saturation level determination, they should provide the radiologists with clear information about how this information is presented in the image, and radiologists should agree on how to use this information to underpin their diagnosis. Besides effectiveness, safety/risks is also an important criterion. The diagnostic imaging techniques used in the hospital are all relatively safe, and judgments were subjective to knowledge bias and risk perception of the different experts. This resulted for example in the technology users showing a higher importance of physical exposure (radiation risk) than the engineers, who all are physicists. Chemical exposure resulted as the most important safety/risks subcriterion. Also bodily burden obtained a high score, but from the discussion that followed can be concluded that the experts were not that concerned about the risk of bodily burden, more about the discomfort. Costs and patient comfort resulted as the least important criteria for a diagnostic breast imaging device.

There is some uncertainty about the way PAM II will perform with respect to the clinical imaging subcriteria. This uncertainty is converted into three different scenario's, that reflect the relative performance of PAM II in the most negative situation, an average situation and in the most positive situation. In order to achieve this, the level of uncertainty has been related to a deviation from the initial judgments of the team members. Only in the most negative scenario, MRI will be a more preferred alternative than PAM II. All other scenario's result in PAM II being the most preferred option. The largest uncertainty exists about the performance of PAM II with respect to oxygen saturation (results not presented). Furthermore, also a relatively large uncertainty exists about the performance of PAM II with respect to mass margins and mass shape (results not presented). These criteria are all relatively important, it is therefore advised to carry out more research and explain more about how these factors will become visible in a photoacoustic image. It has been examined if the average level of uncertainty in the expert group is related to the group variance.

To check if there are discrepancies between the judgments of the technology creators, and the other stakeholders, a subgroup analysis has been carried out. The relative order of importance of the criteria is equal for all expert team members. The technology creators are more focused on the effectiveness, whereas the technology users show a higher importance of patient comfort and safety/risks. Therefore it is advised for the creators to emphasize the good performance of PAM II with respect to these criteria; according to the users, PAM II already shows the best relative performance. The technology buyer logically shows a higher importance of costs than the technology creators, but also safety/risks is a more important criterion for the technology buyer than for the other subgroups. This could be explained by the fact that in the hospital in which the manager is working, the current focus is on improving safety of medical devices. The higher importance reflects this management focus. The technology creators can make use of this management focus by emphasizing the good performance of PAM II with respect to safety (and costs); according to the buyer, PAM II already shows the best relative performance. It is remarkable that the technology creators show a lower preference for PAM

II, compared to mammography and ultrasound, than the technology users do. For the technology users, PAM II is the best option, the technology creators however prefer PAM II only in addition to X-ray mammography. This result may have been caused by the low agreement between the technology creators about the relative performance of PAM II with respect to sensitivity. One of the creators has a very modest attitude towards PAM II, and assigned a low relative weight to PAM II. This will be further explained below, in the discussion of the methods used. The technology users show a higher preference for MRI compared to PAM II, whereas both the technology buyer and the technology creators prefer PAM II over MRI. This is mainly caused by the judgment of the users with respect to effectiveness; according to the users MRI performs a lot better than PAM II. The creators however also show a more favorable judgment of MRI with respect to performance on effectiveness (they know the limits of PAM II), therefore it is assumed that they will focus the improvement of PAM II such that they can approach the effectiveness of MRI.

The most important criteria that have resulted from the AHP analysis are used in a House of Quality analysis. The results of the House of Quality analysis show that investing in the development of a good reconstruction algorithm, a high sensitivity detector, a large detector bandwidth, and a multi-wavelength light source will be most rewarding. A large detector bandwidth is expensive, but is expected to improve the performance of PAM II on the most important criteria for a diagnostic breast imaging device.

The AHP analysis has been carried out in two different sessions: in the first session, X-ray mammography, ultrasound and PAM II were compared, in the second session, MRI and PAM II were compared. The overall results show that PAM II is preferred over all other alternatives. Based on the results presented in this research, the best position for PAM II within the current diagnostic track would be as a substitute for the combined use of X-ray mammography and ultrasound, at the start of the diagnostic track. From the scenario analysis follows that, for all scenarios that result from the comparison of X-ray mammography and ultrasound versus PAM II, PAM II is the most preferred alternative, even in the negative scenario. However, this is not true for the comparison of MRI versus PAM II, in which MRI is preferred over PAM II in the negative scenario. Also, body contact was judged to be an important criterion in the first session, and PAM II shows a better performance with respect to this criterion because the breast of the patient is not forced between two plates as during X-ray mammography. In the comparison with MRI, body contact was less important. Another important criterion in the first session was vascularization, PAM II has a considerably better performance than ultrasound with respect to this criterion. Furthermore, the subgroup analysis shows that PAM II is the most preferred alternative by the user subgroup when compared to X-ray mammography and ultrasound, but when compared to MRI, the users prefer MRI. The users are thought to be important stakeholders in the investment decision in hospitals, therefore it is advised to position PAM as substitute for X-ray mammography and ultrasound. A disadvantage of positioning PAM II at this point is the importance of Ca ++, and the low performance of PAM II with respect to imaging this factor. Therefore it is thought that X-ray mammography will still be used as an adjunct to PAM II. Time will show if the information provided by PAM II will be sufficiently for omitting Ca ++ information. An aspect of ultrasound examination that was not assessed (because this is another part of the diagnostic track), is the possibility it offers for guided biopsies. In practice, this happens during the first examination, when needed. This is something that is not possible with PAM II. Therefore it is assumed that in case biopsies are needed, ultrasound will still be used as imaging method, however as an adjunct and not as primary imaging method.

Discussion of methods used

In this research, AHP analysis was used to answer the research questions. With the aid of Team Expert Choice, an expert panel assessed sets of pairwise comparisons. Through discussion, knowledge was shared across the disciplines and disagreements were tried to be overcome. The quality of this research depends for a large part on the criteria, subcriteria and alternatives chosen. The criteria used in this analysis were identified mainly on the basis of literature and interviews with different experts. Costs was identified as a potentially important criterion, and was further subdivided into other criteria. The assessment of these costs subcriteria however appeared not to be useful. The subcriteria were not fully independent, and the performance on the subcriteria depends for a large part on the setting in which the device will be used. MRI can for instance also be used for other applications, and the investment of an expensive MRI depends for a large part on the expectations of the occupancy rate due to these other applications. Also, the size and specialization of the hospital in which the devices will be used may be of influence. Next to this, scan time may for example not be important until the device is fully booked. Furthermore, the definition of costs is a bit vague, one could for example ask what the follow-up costs are of a wrong (and right) diagnosis; this way costs and effectiveness are also interrelated. AHP

was in this case not suited for cost-effectiveness analysis, and assessment of the costs subcriteria did not provide a clear view of the most important cost components. Cost-optimization can be calculated, and should not be determined based on this type of analysis. Effectiveness was subdivided into sensitivity and specificity. During the analysis, it appeared that the expert team had troubles with the provided definitions of these subcriteria. During the first session, these definitions were discussed and adapted to unambiguous definitions upon which all experts agreed. During the rest of the analysis, the new definitions were used. Sensitivity and specificity were further subdivided into clinical imaging subcriteria. It is expected that the assessment of the performance of the alternatives with respect to these subcriteria provides more genuine results than the first rough estimation of the performance with respect to sensitivity and specificity, that was assessed in the subgroup analysis. This depends however on the quality with which the clinical criteria that are chosen represent the total diagnostic input that is needed from the imaging devices. MRI was identified as one of the diagnostic alternatives. During the second session it appeared however that MRI mainly serves a screening-type goal: identifying new lesions, as opposed to the diagnostic-type goal: discriminating between a benign and malignant lesion. This screening- type goal was not represented in the diagnostic track that was provided by Medisch Spectrum Twente, and was therefore not used. The diagnostic-type goal is assessed in this research, which should be kept in mind when interpreting the results. During the second session this caused some discussion, and sometimes screening-type arguments were used to underpin a judgment (e.g. 'most tumors originate in this area, so this area should be imaged with good quality').

The medical, social, industrial or technical backgrounds of the panel members influence the factors that are incorporated into the assessment, as well as the corresponding weighting factors. Therefore it is essential that a representative group of experts is represented in the expert team. In the first session, a manager was lacking, which could have caused an importance bias away from costs. Looking at the results of the second session however shows that the relative importance of costs is only slightly higher than in the first session. It is therefore not expected that costs are underappreciated in the first session, and that the presence of a manager in the first session would have resulted in a change in the overall outcome. Two of the experts were very dominantly present in the discussion about the judgments. Some of the other experts reacted on this dominance by heavily defending the opposite judgments, and some of the experts changed their opinion according to the one of the dominant team member. The reaction depended mainly on the (hierarchical) regard the other team members had with respect to the dominant member. The best group composition for AHP would be one in which all stakeholders have the same authority in their field, and all stakeholders are equally distributed over the fields. In this analysis, relatively many engineers were represented, each with its own specialism. Their knowledge about PAM II and the other technologies assessed reached beyond the information that was provided to prepare before the AHP analysis, which sometimes caused some discussion in which sudden new information became available that changed the opinions of the stakeholders. On the one hand, this is positive, because all stakeholders have learned from each other through the discussions that took place. On the other hand, this is negative, because sometimes these discussions took quite some time and were not needed to provide the judgment that had to be assessed. It may be good to take more time before the AHP for sharing knowledge, e.g. by presentations of the different stakeholders, or by providing more extensive preparation literature.

At certain times during the assessment, the expert team expressed the uncertainty about judgments they provided. Also the data that was provided in the performance matrix was not trusted by all experts. Only with respect to the sensitivity subcriteria, the uncertainty in the performance of the alternatives was assessed. A scenario analysis provided insight in the margins of the relative performance of the new technology due to uncertainty. For a complete overview, this approach may also be useful for the other criteria. It has been examined if the average level of uncertainty in the expert group is related to the group variance. A significant moderate correlation of .445 (p = .043) was found between these factors. This moderate correlation reflects that the variation of judgments within a group is not merely the result of uncertainty, but also the result of real disagreement about the performance of an alternative. Despite this disagreement, it was in most cases through discussion possible to reach a certain level of consensus, but sometimes the disagreement kept existing. From the correlation it can be concluded that it is not sufficient to conduct the uncertainty just from the group variance, researchers should ask explicitly about this information and use this for example in a scenario analysis as was done in this research. Part of the moderate correlation can also be explained by the personality of the team members. A comparison of the average certainty and variance of the average certainty for all team members shows that one of the members is 'quite sure' about almost all judgments (cautious personality), and

one of the team members is 'very sure' about almost all judgments (self-confident personality). The other team members show more variance in their judgments.

To check if there are discrepancies between the judgments of the technology creators and the other stakeholders, a subgroup analysis has been carried out . It should be noted that the subgroups used in the subgroup analysis were small (1-2 persons), and sometimes the deviation in judgments between team members was large. Therefore, results of the subgroup analysis are not statistically significant. In order to determine if differences are significant, larger subgroups are needed. However, a proper use of the AHP method does not allow for larger expert teams than about 10 persons. The largest disagreement between the technology users (first session) existed about the relative importances of the patient comfort subcriteria: the nurse practitioner thinks the time between the scan and the moment the patient receives the results is most important (.694, data not presented), whereas the radiologist indicates that body contact is the most important criterion (.713, data not presented). The largest disagreement between the technology creators was about the relative performance of PAM II with respect to specificity (.594 vs. .091, first session). This is expected to be due to a discussion about the definitions of sensitivity and specificity in the first session, which may have caused some confusion. The expert team indicated that it was hard to provide a rough estimation of the performance of the alternatives with respect to these criteria, because of the complicated definition. During the more extensive assessment of these criteria, the agreement between the subgroup members was higher.

The overall inconsistency of the judgments was low. Some sets of judgments resulted in higher inconsistencies because of extreme preferences for one of the alternatives. At first sight, for some of the judgments this seemed logical (e.g. the combined option is always more expensive than the single technique), but for other judgments this was not the case (e.g. it is not always true that the combined use of techniques produces a higher sensitivity than the use of a single technique). This could have resulted in some performance shifts. Due to time restrictions and inexperience not all of these judgments have been reassessed.

The overall AHP hierarchy that was assessed was extensive, and resulted in a lot of pairwise comparisons. Each pairwise comparison was discussed, which may have demanded a lot of energy from the experts. At the end of the sessions, fatigue may have caused some distraction. This resulted for some participants in providing extreme judgments, and in discussions about unimportant aspects. It was hard to keep the team focused until the last judgment.

It can be concluded that AHP is a suitable method to assess technologies that are in an early state of development, there are however some points that need attention. In this analysis, AHP appeared not to be useful for the assessment of costs subcriteria. For a new technology that is going to be used in a hospital, a lot depends on the current setting in which similar examinations take place. This setting can differ between hospitals, which made the AHP assessment difficult. It is therefore advised to define the setting in which the new technology will be used on beforehand. Another drawback is the little information that is available about the new technology, and the difficult judgment about the quality of this information. During early stage development, there is not much published about a new technology, and in the information provided for AHP assessment is often one-sided originating from the developer. Not all team members may have the same basic level of understanding about the technology , which in this case sometimes led to unnecessary discussions about irrelevant details. So, knowledge sharing before the AHP assessment is important. The last point of attention is the motivation of the team members. Through the whole assessment, serious attention and input are needed from the experts. Dominance of one or more team members can lead to biased results. Therefore, distraction because of too extensive hierarchical AHP structures or by one of the members steering the discussion into the wrong direction should be prevented. To be able to do this, the discussion leader of the AHP should also be well informed about the exact content of the issues being addressed. When all of these precautions are taken into account, AHP is able to provide a useful prediction of the possible future success of a new technology. In this case the developers of PAM II know now that their device has a realistic chance of success, which may be a good motivation for continuing their work. During the AHP analysis, useful interdisciplinary discussions took place that provided the developers of PAM II with new insights. The developers of PAM II have indicated some technical aspects of the device that might need improvement, the results of the House of Quality analysis presented in this research can be used for priority setting of the implementation of these technical improvements. Furthermore, from this research can be concluded that in case PAM II will be further developed for diagnostic use, it can best be positioned at the start of the diagnostic track, as a substitute for the combined use of X-ray mammography and ultrasound. In the further development

of PAM II, more specific issues and questions that arise in this part of the diagnostic track should be examined, in order to make PAM II optimally suitable for the diagnostic demand.

RECOMMENDATIONS

This AHP analysis has provided useful insights into the potential success of PAM II. The most important aspects that need to be improved in order to improve its position with respect to the alternative imaging techniques are improving the reconstruction algorithm quality, increasing the sensitivity of the detector, increasing the detector bandwidth and increasing the number of wavelengths. More research should be carried out with respect to specific issues and questions that exist at the start of the diagnostic track, in order to make PAM II optimally suitable for the diagnostic demand. Furthermore, it is advised to execute more research to be able to determine the sensitivity and specificity of the technique. This will reduce the uncertainty that exists at the users. The spread in performance of PAM II between the different scenarios assessed is large and will diminish with increasing certainty, which makes a better assessment of the relative performance of PAM II possible. The medical experts were not familiar with the oxygen saturation parameter, and could therefore not indicate if this criterion will be valuable. It is advised to execute more research about the exact way in which this parameter can support diagnosis. The radiologists were curious about the visualization of this parameter in the image, it is therefore advised to let them judge several concepts, and develop the visualization in cooperation with them.

With respect to the use of the AHP analysis method, the following can be recommended:
- The assessment of costs subcriteria should be provided by an other method, AHP was not suitable in this case.
- The goal, setting and definitions of the criteria should remain clear for all participants during the assessment. In this research, the definitions of the criteria were projected on a second screen. This proved to be helpful, but was not at all times sufficient. The discussion leader of the AHP session should check at all times if it is clear what is asked from the expert.
- Before assessment of the judgments, a session should be organized which is focused on knowledge-sharing, on order to guarantee a same basic understanding of all issues addressed by all team members.
- The AHP hierarchy used in this research was quite extensive, which caused some loss of focus of the team members at the end of the sessions. It is advised to keep the hierarchy as short as possible, or to arrange more sessions in which parts of the hierarchy are assessed on different days.

The AHP method proved to be suitable for answering the research questions. With this method the possibilities of PAM II can be further explored. It will be useful to investigate the likely success of the device in screening, because a lot of people can then be reached, and the current screening methods do not provide satisfactory results with respect to effectiveness. The expert team has indicated that they expect that other criteria will be more important for screening than for diagnosis. For example, the radiation risk will be much more important, and patient comfort will also be more important because a high quality screening program demands a good cooperation of the patients. Next to a change of the goal, also the performance of PAM II with respect to other alternatives could be investigated. Nuclear medicine becomes more popular within cancer diagnosis, and also new applications of MRI are being developed, such as MR elastography, that may become serious competitors of PAM II.

The results of this analysis promise a potentially successful future for PAM II. I hope the results of this research are a good motivation for the developers of PAM II to continue their research.
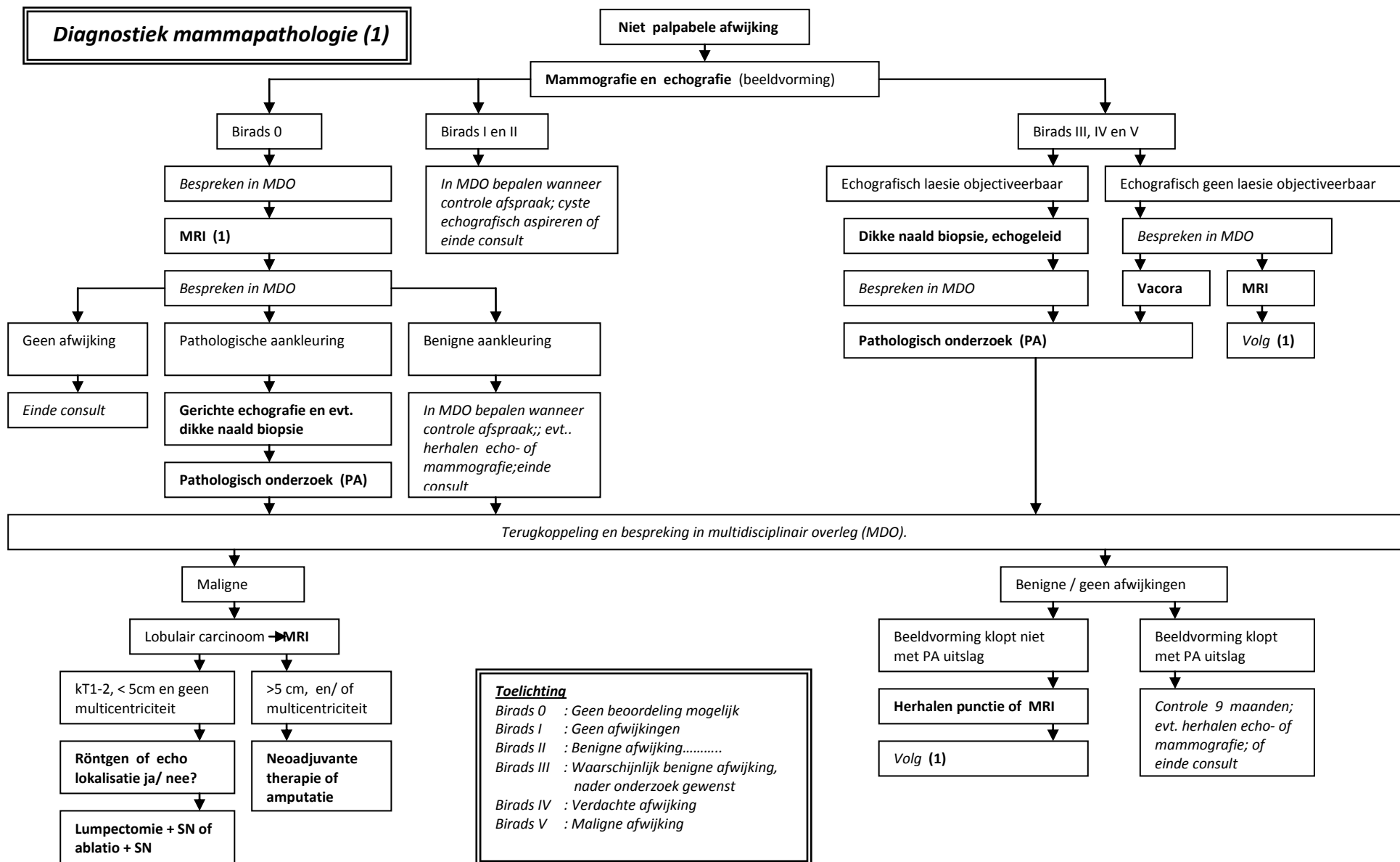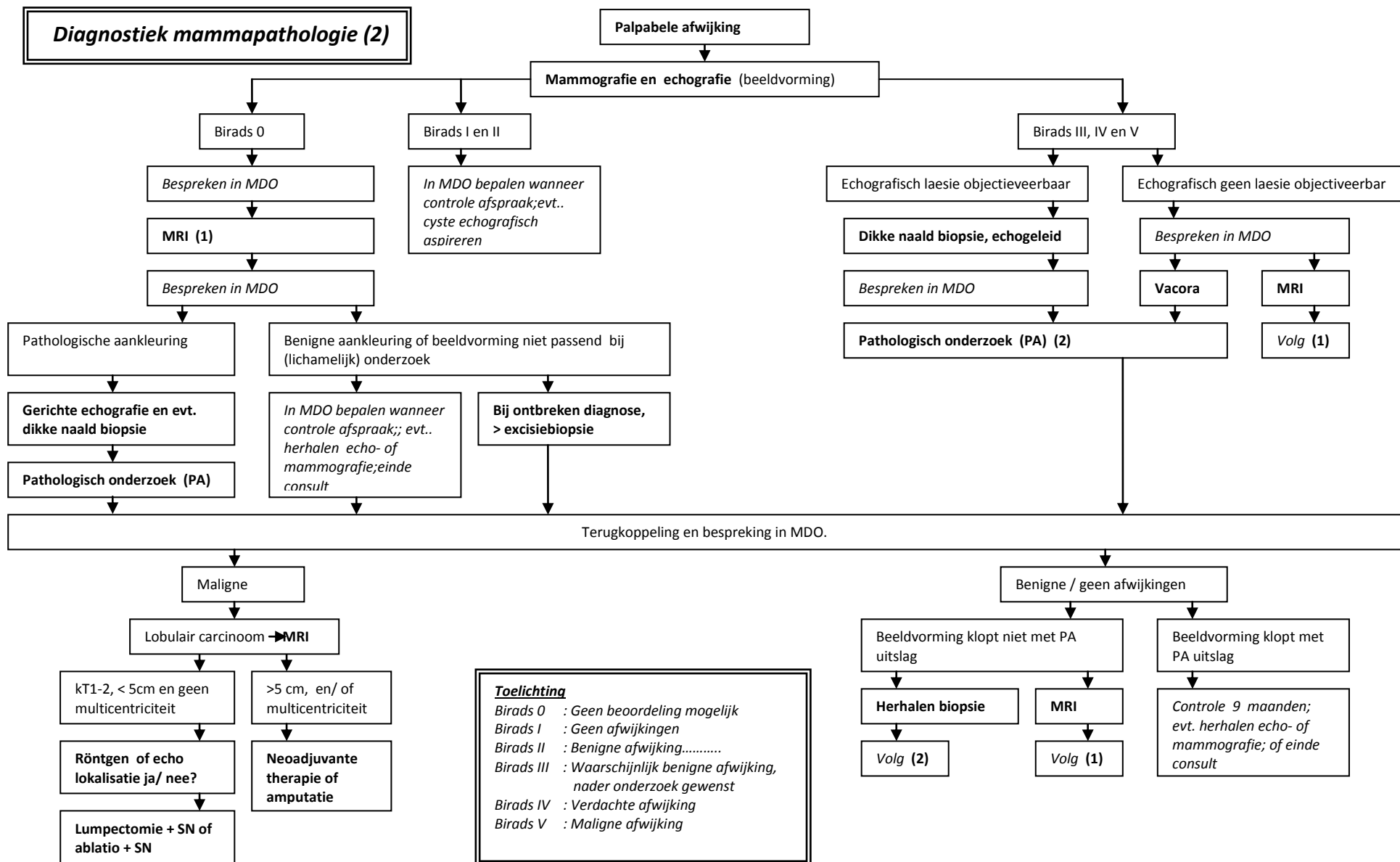

ACKNOWLEDGEMENTS

REFERENCES

Ali S., Coombes R.C. (2002). "Endocrine-responsive breast cancer and strategies for combating resistance". In: *Nature Reviews Cancer.* Vol. 2, nr. 2, p. 101-112.

Beeldvormende technieken binnen de kankerbestrijding, Vizier op de toekomst (2005). De werkgroep 'Beeldvormende Technieken binnen de Oncologie' van de Signaleringscommissie Kanker (SCK) van KWF Kankerbestrijding, KWF Kankerbestrijding, The Netherlands.

Berg W.A., Birdwell R.L., Gombos E.C., Wang S., Parkinson B.T., Raza S., Green G.E., Kennedy A., Kettler M.D. (2008). *Diagnostic Imaging: Breast.* Salt Lake City, Utah: Amirsys Inc.

Bevolkingsonderzoek Borstkanker Nederland. 28 November 2007. Bevolkingsonderzoek Borstkanker in Nederland, The Netherlands. 25 May 2009.
<http://www.bevolkingsonderzoekborstkanker.nl/>

BI-RADS® Atlas (2003). Breast Imaging Reporting and Data System Atlas, Reston. American College of Radiology (ACR). Excerpted text online available. 24 June 2009.
<http://www.acr.org/SecondaryMainMenuCategories/quality_safety/BIRADSAtlas.aspx>

Borstkanker. 22 Aug 2008. KWF Kankerbestrijding, The Netherlands. 7 May 2009.
<http://www.kwfkankerbestrijding.nl/>

Cijfers Centrum voor Mammacare 2008 (2008). Centrum voor Mammacare (CvM). Internal document Medisch Spectrum Twente.

Cijfers, kerncijfers: Meest voorkomende kankersoorten. 1 Feb 2009. Vereniging van Integrale Kankercentra, The Netherlands. 7 May 2009.
<http://www.ikcnet.nl/>

Decker T., Hungerman D., Böcker W. (2009). "Prognostische und prädiktive Faktoren invasiver Mammakarzinome: Update 2009". In: *Der Pathologe.* Vol. 30, nr. 1, p. 49-55.

Douma K.F.L., Hummel J.M., Bueno-de-Mesquita J.M., van Harten W.H. (2007). "Methodology of Constructive Technology Assessment in Health Care". In: *Intl J. of Technology Assessment in Health Care.* Vol. 23, nr. 2, p. 162-168.

Hauser J.R., Clausing D. (1988). "The House of Quality". In: *Harvard Business Review.* Vol. 66, nr. 3, p. 63-73.

Hummel J.M., van Rossum W., Verkerke G.J., Rakhorst G. (2000). "Assessing Medical Technologies in Development, A new paradigm of Medical Technology Assessment". In: *Intl J. of Technology Assessment in Health Care.* Vol. 16, nr. 4, p. 1214-1219.

Hummel J.M. (2001). *Supporting medical technology development with the analytic hierarchy process.* Groningen: Rijksuniversiteit Groningen, faculty of Medical Sciences.

Koertshuis R. (2008). *Stroomschema's diagnostiek CvM.* Internal document Medisch Spectrum Twente.

Kohn L.T., Corrigan J.M., Donaldson M.S., eds. (2000). *To Err Is Human: Building a Safer Health System.* Institute of Medicine, Committee on Quality of Health Care in America. National Academy Press, Washington, D.C.

Liberatore M.J., Nydick R.L. (2007). "The analytic hierarchy process in medical and health care decision making: A literature review". In: *European Journal of Operational Research.* Vol. 189, p. 194-207.

Mammacarcinoom: Landelijke richtlijn met regionale toevoegingen (2008). Versie 1.1. Nationaal Borstkankeroverleg Nederland (NABON), Vereniging van Integrale Kankercentra (VIKC), Kwaliteitsinstituut voor de Gezondheidszorg CBO, The Nederlands.

Manohar S., Vaartjes S.E., van Hespen J.C.G., Klaase J.M., van den Engh F.M., Steenbergen W., van Leeuwen T.G. (2007). "Initial results of in vivo non-invasive cancer imaging in the human breast using near-infrared photoacoustics". In: *Optics express.* Vol. 15, nr. 19, p. 12277-12285.

Manohar S., Kharin A., de Vos C., van Hespen J.C.G., Steenbergen W., van Leeuwen T.G. (2008). *The Twente Photoacoustic Mammoscope: Towards clinical trials.* Biophysical Engineering Group, Faculty of Science and Technology, University of Twente, The Netherlands.

NHS Executive (1996). *Promoting clinical effectiveness : a framework for action in and through the NHS.* NHS Executive, United Kingdom.

NVKF. Het vakgebied Klinische Fysica 8 May 2009. Nederlandse Vereniging voor Klinische Fysica, The Netherlands. 11 June 2009.
<http://www.nvkf.nl/>

Polyak K. (2008). "Is BreastTumor Progression Really Linear?". In: *Clinical Cancer Research.* Vol. 14, nr. 2, p. 339-341.

QFD Online. Free House of Quality Templates for Excel. 2007. Quality Function Deployment online. 27 November 2009.
<http://www.qfdonline.com/>

Saaty T.L. (1980). *The Analytic Hierarchy Process: Planning, Piority Setting, Resource Allocation.* New York: McGraw Hill.

Sculpher M., Drummond M., Buxton M. (1997). "The iterative use of economic evaluation as part of the process of Health Technology Assessment". In: *Journal of Health Services Research and Policy.* Vol. 2, nr. 1, p. 26-30.

Siebert M., Clauss L.C., Carlisle M., Casteels B., de Jong P., Kreuzer M., Sanghera S., Stokoe G., Trueman P., Lang A.W. (2002). "Health Technology Assessment for Medical Devices in Europe, What must be considered". In: *Intl J. of Technology Assessment in Health Care.* Vol. 18, nr. 3, p. 733-740.

Sluijs J.P. van der, Janssen P.H.M., Petersen A.C., Kloprogge P., Risbey P.S., Tuinstra W., Ravetz J.R. (2004). *RIVM/MNPGuidance for Uncertainty Assessment and Communication: Tool Catalogue for Uncertainty Assessment.* Netherlands Environmental Assessment Agency, National Institute for Public Health and the Environment (RIVM), The Netherlands.

Teo N.B., Shoker B.S., Jarvis C., Martin L., Sloane J.P., Holcombe C. (2003). "Angiogenesis and invasive recurrence in ductal carcinoma in situ of the breast". In: *European Journal of Cancer.* Vol. 39, nr. 1, p. 38-44.

Vaartjes S. (2006). *First clinical trials of the Twente Photoacoustic Mammoscope and instrumentation for optical characterization of breast phantoms.* Master thesis Biomedical Engineering. University of Twente. The Netherlands.

Vallejo-Torres L., Steuten L.M.G., Buxton M.J., Girling, A.J., Lilford R.J., Young T. (2008). "Integrating Health Economics Modelling in the Product Development Cycle of Medical Devices: A Bayesian Approach". In: *Intl J. of Technology Assessment in Health Care.* Vol. 24, nr. 4, p. 459-464.

Visser O. and Van Noord K.J., eds. (2005). *Feiten + Fabels over kanker in Nederland.* Vereniging van Integrale Kankercentra, Utrecht, The Netherlands.

Wallner P.E., Konski A. (2008). "The Impact of Technology on Health Care Cost and Policy Development". In: *Elsevier, Seminars in Radiation Oncology*. P. 194-200.

Xu R.X., Young D.C., Mao J.J., Povoski S.P. (2007). "A prospective pilot clinical trial evaluating the utility of a dynamic near-infrared imaging device for characterizing suspicious breast lesions". In: *Breast Cancer Research.* Vol. 9, nr. 6, p. R88.

APPENDIX 1:

FLOWCHARTS DIAGNOSTIC TRACK MAMMACARE DEPARTMENT MST

Diagnostiek mammapathologie (1)

Niet palpabele afwijking

Mammografie en echografie (beeldvorming)

Birads 0

Birads I en II

Birads III, IV en V

Bespreken in MDO

In MDO bepalen wanneer controle afspraak; cyste echografisch aspireren of einde consult

Echografisch laesie objectiveerbaar

Echografisch geen laesie objectiveerbaar

MRI (1)

Dikke naald biopsie, echogeleid

Bespreken in MDO

Bespreken in MDO

Bespreken in MDO

Vacora

MRI

Geen afwijking

Pathologische aankleuring

Benigne aankleuring

Pathologisch onderzoek (PA)

Volg (1)

Einde consult

Gerichte echografie en evt. dikke naald biopsie

In MDO bepalen wanneer controle afspraak;; evt.. herhalen echo- of mammografie;einde consult

Pathologisch onderzoek (PA)

Terugkoppeling en bespreking in multidisciplinair overleg (MDO).

Maligne

Benigne / geen afwijkingen

Lobulair carcinoom ➜ MRI

Beeldvorming klopt niet met PA uitslag

Beeldvorming klopt met PA uitslag

kT1-2, < 5cm en geen multicentriciteit

>5 cm, en/ of multicentriciteit

Herhalen punctie of MRI

Controle 9 maanden; evt. herhalen echo- of mammografie; of einde consult

Röntgen of echo lokalisatie ja/ nee?

Neoadjuvante therapie of amputatie

Volg (1)

Lumpectomie + SN of ablatio + SN

Toelichting
Birads 0 : Geen beoordeling mogelijk
Birads I : Geen afwijkingen
Birads II : Benigne afwijking………..
Birads III : Waarschijnlijk benigne afwijking, nader onderzoek gewenst
Birads IV : Verdachte afwijking
Birads V : Maligne afwijking

Diagnostiek mammapathologie (2)

Palpabele afwijking

Mammografie en echografie (beeldvorming)

Birads 0

Birads I en II

Birads III, IV en V

Bespreken in MDO

In MDO bepalen wanneer controle afspraak;evt.. cyste echografisch aspireren

Echografisch laesie objectieveerbaar

Echografisch geen laesie objectiveerbar

MRI (1)

Dikke naald biopsie, echogeleid

Bespreken in MDO

Vacora

MRI

Bespreken in MDO

Bespreken in MDO

Pathologisch onderzoek (PA) (2)

Volg **(1)**

Pathologische aankleuring

Benigne aankleuring of beeldvorming niet passend bij (lichamelijk) onderzoek

Gerichte echografie en evt. dikke naald biopsie

In MDO bepalen wanneer controle afspraak;; evt.. herhalen echo- of mammografie;einde consult

Bij ontbreken diagnose, > excisiebiopsie

Pathologisch onderzoek (PA)

Terugkoppeling en bespreking in MDO.

Maligne

Benigne / geen afwijkingen

Lobulair carcinoom → **MRI**

Beeldvorming klopt niet met PA uitslag

Beeldvorming klopt met PA uitslag

kT1-2, < 5cm en geen multicenticiteit

>5 cm, en/ of multicenticiteit

Herhalen biopsie

MRI

Controle 9 maanden; evt. herhalen echo- of mammografie; of einde consult

Röntgen of echo lokalisatie ja/ nee?

Neoadjuvante therapie of amputatie

Volg **(2)**

Volg **(1)**

Lumpectomie + SN of ablatio + SN

Toelichting
Birads 0 : Geen beoordeling mogelijk
Birads I : Geen afwijkingen
Birads II : Benigne afwijking………..
Birads III : Waarschijnlijk benigne afwijking, nader onderzoek gewenst
Birads IV : Verdachte afwijking
Birads V : Maligne afwijking

APPENDIX 2:

BACKGROUND INFORMATION: Health Technology Assessment (HTA)

Health Technology Assessment

During the last decade, Health Technology Assessment (HTA) has been increasingly used as an evaluation approach to enable decisions on coverage and reimbursement of new technologies. [Siebert *et al.* 2002, Douma *et al.* 2007] Mainly under the influence of policy pressure, HTA's generally are composed of clinical efficacy and cost-effectiveness analysis (CEA) studies, with every country adapting the HTA to its own needs. [Douma *et al.* 2007]

HTA is the collective name given to a number of activities applying systematic methods of scientific inquiry to the evaluation and use of new or existing healthcare technologies. The overall objective of HTA is to provide robust and objective information for decision making in healthcare at different levels. [Siebert *et al.* 2002]

A major drawback of HTA is that it presumes a "ceteris paribus" (static) situation of technology development, whereas it has become evident that environment and technology are often dynamic and mutually influencing each other. To improve the effectiveness of new technologies, influencing changes is sometimes needed rather than studying changes. Commitment from clinicians and other stakeholders is therefore needed, and all relevant aspects of technology and environmental interaction should be covered. [Douma *et al.* 2007] To achieve this, Constructive Technology Assessment (CTA), a further development of HTA, has to be carried out.

During the development of medical technology, CTA, or economic evaluation, can be carried out in different stages, according to the different phases of technology development. The earliest stage of economic evaluation is stage 1 assessment, which should be undertaken once the basic science of the technology has been investigated. In stage 1 economic evaluation, a new technology is characterized against a baseline and aims to describe the likely economic characteristics of the innovation. The input for this analysis is early data on patients, if available, and independent clinical judgment. The effectiveness can also be assessed by e.g. using outcomes as perceived by patients. [Sculpher *et al.* 1997]

At this early stage, CTA may especially be important for medical device developers. The rapidly increasing range and expense of new medical devices means that there is increasing pressure for them to articulate the superior value of their products. Early assessment will not only help companies reduce their failure rates, but also help ensure that patients and other users of medical devices gain access to the most beneficial technologies. A CTA is a promising tool to support this in three ways [Vallejo-Torres *et al.*, 2008]:

- By allowing the estimation of potential cost-effectiveness to be part of the investment decision process and to avoid investing in a technology that could never be cost-effective
- By supporting companies to prioritize between several competing possibly cost-effective concepts or prototypes
- By identifying from early stages of development those parameters that have the largest impact on the likely cost-effectiveness of the product to direct scarce research resources

However, despite these benefits, there are only a limited number of papers published on methods of CTA applied at an early developmental stage in the development of health care technology. [Douma *et al.*, 2007]

In order to find an answer to the research questions presented in this paper, a CTA of the PAM was carried out. Taking into account the limited amount of data available regarding photoacoustic mammography, the analysis should be based on the available evidence concerning the current technology that the new device aims to substitute or will compete with, and expert opinion and/or assumptions regarding the likely impact on cost and effectiveness of the new device. [Vallejo-Torres *et al.*, 2008]

The current technologies that the PAM aims to substitute or has to compete with, are determined first. Information is gathered on the nature of these technologies and their organizational settings by means of literature research and documentation analysis.

Second, experts are identified who will provide the input needed to determine the factors that will result in a successful design of the PAM. According to Wallner (2008), the following stakeholders are often involved in technology assessments [Wallner *et al.* 2008]:

- Technology end users (patients and families)
- Physicians
- Technology Creators
- Patient Advocacy groups
- Government Agencies
- Payers

In addition to this list, also Medical Physicists and Medical Engineers are thought to be important stakeholders, because they have an important role in assessing the safety and quality of new devices used in their hospitals, and a large share in the hospital's capital budgeting. [NVKF]

From each of the stakeholder groups, one or two opinion leaders or experts in mammography are asked to participate in this research.

APPENDIX 3:

PROGRAM AHP SESSIONS

PROGRAM OF THE AHP

X-RAY MAMMOGRAPHY, ULTRASOUND AND PAM II
Morning program, location: Haaksbergerzaal 3, MST

9.00 - 9.15	**ARRIVAL AND COFFEE;**
9.15 - 10.00	**INTRODUCTION AND BACKGROUNDS;**

9.15 - 9.30   Introduction about photoacoustic mammography,
by S. Manohar, PhD;

9.30 - 9.40   Introduction about Team Expert Choice;

9:40 - 10.00   Explanation of the evaluation structure;

10.00 - 11.30   **DISCUSSION ABOUT THE EVALUATION OF THE DIAGNOSTIC ALTERNATIVES, SUPPORTED BY TEAM EXPERT CHOICE (PART I);**

10.00 - 10.20   Discussion about the importances of costs, effectiveness, patient comfort and safety/risks;
All participants

10.20 - 11.00   Discussion about the importances of costs, effectiveness, patient comfort and safety/risks subcriteria;
All participants

11.00 - 11.30   Discussion about the pursued quality of the diagnostic alternatives, with respect to costs, effectiveness, patient comfort and safety/risks subcriteria;
All participants

11.30 - 11.40   **COFFEE;**

11.40 - 12.55   **DISCUSSION ABOUT THE EVALUATION OF THE DIAGNOSTIC ALTERNATIVES, SUPPORTED BY TEAM EXPERT CHOICE (PART II);**

11.40 - 12.20   Discussion about the importances of the 'sensitivity & specificity' subcriteria;
Medical specialists, technical specialists

12.20 - 12.55   Discussion about the pursued quality of the diagnostic alternatives, with respect to the 'sensitivity & specificity' subcriteria;
Medical specialists, technical specialists

12.55 - 13.00   **PRESENTATION OF FINAL RESULTS MORNING PROGRAM;**

13.00   **END MORNING PROGRAM**

Lunch, location: Restaurant MST, Haaksbergerstraat

13.00 – 14.00   **LUNCH;**

MRI AND PAM II
Afternoon program, location: Ariënszaal 2, MST

14.00 - 14.15   **ARRIVAL AND COFFEE;**

14.15 - 15.00   **INTRODUCTION AND BACKGROUNDS;**

       14.15 - 14.30   Introduction about photoacoustic mammography,
                                by S. Manohar, PhD;

       14.30 - 14.40   Introduction about Team Expert Choice;

       14:40 - 15.00   Explanation of the evaluation structure;

15.00 - 16.20   **DISCUSSION ABOUT THE EVALUATION OF THE DIAGNOSTIC ALTERNATIVES, SUPPORTED BY TEAM EXPERT CHOICE (PART I);**

       15.00 - 15.20   Discussion about the importances of costs, effectiveness, patient comfort and safety/risks;
                                 All participants

       15.20 - 16.00   Discussion about the importances of costs, effectiveness, patient comfort and safety/risks subcriteria;
                                 All participants

       16.00 - 16.20   Discussion about the pursued quality of the diagnostic alternatives, with respect to costs, effectiveness, patient comfort and safety/risks subcriteria;
                                 All participants

16.20 - 16.30   **COFFEE;**

16.30 - 17.40   **DISCUSSION ABOUT THE EVALUATION OF THE DIAGNOSTIC ALTERNATIVES, SUPPORTED BY TEAM EXPERT CHOICE (PART II);**

       16.30 - 17.15   Discussion about the importances of the 'sensitivity & specificity' subcriteria;
                                 Medical specialists, technical specialists

       17.15 - 17.40   Discussion about the pursued quality of the diagnostic alternatives, with respect to the 'sensitivity & specificity' subcriteria;
                                 Medical specialists, technical specialists

17.40 - 17.45   **PRESENTATION OF FINAL RESULTS AFTERNOON PROGRAM;**

17.45   **END**

APPENDIX 4:

DEFINITIONS CRITERIA

The following definitions of the criteria were used in the AHP analysis:

1. Costs
   The costs are defined as the costs for the organization were the diagnostic examinations take place. In most cases, this will be the hospital.

   1.1. Scan time
   The scan time is defined as the time the whole imaging procedure takes, starting the moment the preparations for the first patient start and ending the moment the preparations for the next patient can start. In X-ray mammography, it is common practice to acquire images of both breasts (bi-lateral) for comparison. If indicated by the medical specialists, this may also be the case for PAM and MRI images. This should be kept in mind.
   Reading of the images is also taken into account in this criterion, in case reading takes place on-site (during the examination). [Berg *et al.* 2008]
   By keeping the scan time as short as possible, the highest amount of patients can be imaged a day. Then, the costs will be recouped sooner.

   1.2. Manpower
   The manpower needed, is the amount of personnel that is needed to acquire the images in a diagnostic mammography setting. This will include all personnel needed to control the device and to support the patients. Also, the expert level of this personnel has to be taken into account. It is assumed that personnel with a higher expert-level will induce higher costs.

   1.3. Price (+ maintenance & disposables)
   The price is the price for all the equipment that is needed in a single setting of each alternative, including the price of the maintenance that is needed to assure the quality of the equipment. The depreciation period for the different devices should also be taken into account. Also, the price of disposable materials (x no. of patients) has to be taken into account when evaluating this criterion.

   1.4. Peripheral equipment (ICT + environment)
   In order to fully benefit from all necessary options, or for optimal safety assurance of the screening mammography alternatives, specific peripheral equipment, ICT facilities and/or environmental facilities may be needed. The costs for these additional items have to be taken into account in this criterion.

2. Effectiveness
   Clinical effectiveness has been defined by the NHS as: "The extent to which specific clinical interventions, when deployed in the field for a particular patient or population, do what they are intended to do - i.e. maintain and improve health and secure the greatest possible health gain from the available resources". [NHS Executive 1996] In this case, the accuracy of the tests is of great importance, which is reflected by the sensitivity and specificity of the techniques.

   2.1. Sensitivity*
   Sensitivity indicates the sensitivity of a diagnostic test or technique. One can calculate the sensitivity by dividing the amount of true positives by the sum of the true positives and the false negatives. This is expressed in a percentage which can vary between 0 and 100%. A high sensitivity e.g. indicates that there is only a small amount of tumors which is missed by the test. The sensitivity of the test is mainly determined by the amount of false negatives. [SCK 2005]
   In this case, when evaluating the use of diagnostic devices, a high sensitivity indicates that there is only a small amount of malignant lesions missed by the test (or classified as benign). The question to be asked in the evaluation of this criterion is how well the subcriteria are able to determine if a lesion is malignant/are indicators of a malignant lesion.
   During the AHP, a shorter definition has been used, namely: Avoiding false negatives.

   2.2. Specificity*
   Specificity indicates the specificity of a diagnostic test or technique. One can calculate the specificity by dividing the amount of true negatives by the sum of the true negatives and the false positives. This is expressed in a percentage which can vary between 0 and 100%. A low specificity e.g. indicates that many benign lesions are classified as malignant. The specificity of the test is mainly determined by the amount of false positives. [SCK 2005]

In this case, when evaluating the use of diagnostic devices, a high specificity indicates that there is only a small amount of benign lesions classified as malignant. The question to be asked in the evaluation of this criterion is how well the subcriteria are able to determine if a lesion is benign/are indicators of a benign lesion.

During the AHP, a shorter definition has been used, namely: Avoiding false positives.

2.x.1 Mass margins

The margins of a mass have different appearances in images, that may be indicators of malignancy. Different appearances are for example: surrounding, lobular, obscured, turbid or spicular. [Berg *et al.* 2008, BI-RADS Atlas 2003, Polyak 2008]

2.x.2 Mass shape

The shape of a mass can also be an indicator of malignancy. Different appearances of masses are for example: round/oval, or lobular. Also the shapes within a mass can be important for diagnosis. [Berg *et al.* 2008, BI-RADS Atlas 2003, Polyak 2008]

2.x.3 Mass size

To examine if a lesion has grown with respect to previous images, it may be important to be able to determine the exact size of a mass. [Decker *et al.* 2009]

2.x.4 Location mass

The location of a mass/lesion can be important for diagnosis. Full breast imaging may be an important option, but also zooming in on a specific area and displaying this area with high quality. [Berg *et al.* 2008]

2.x.5 Ca ++

Ca ++, or microcalcifications, can, depending on the way they present, be an important indicator of the presence or absence of malignancy. [Berg *et al.* 2008]

2.x.6 Vascularization

When a tumor grows, small blood vessels grow around it (angiogenesis) for nutrition supply and waste removal. A number of studies have shown that the degree of vascularity within an invasive breast carcinoma may be of prognostic value. Several other studies have also shown that various premalignant lesions of the breast can induce angiogenesis in animal experimental systems and in the human breast. [Teo *et al.* 2003]

2.x.7 Oxygen saturation

Oxygen saturation is thought to be indicative of the speed with which the tumor is growing: malignant tissues may have lower oxygen saturation due to imbalanced oxygen supply and uptake and increased blood volume due to angiogenesis. [Xu *et al.* 2007]

3. Patient comfort

Patient comfort is defined as the way the patient both mentally and physically experiences the total clinical intervention. This can be important for the cooperation of the patient.

3.1. Body contact

This criterion comprises all aspects that relate with the direct contact the body has with the device (physical ergonomical aspects). The posture of the patient for a certain time span, the amount of compression of the breasts, the temperature of contact media, and protective measures that should be arranged, such as wearing safety goggles or removing piercings, all are examples this criterion.

3.2. Environmental factors

The environment of the patient is different for the different imaging alternatives. Examples of factors influencing this criterion are the people who are present in the same room as the patient, the way communication with the patient is arranged, sound the device produces, and the direct space around the patient.

3.3. Time between scan and results

The period between the execution of the scan and when the patient receives the result is reflected by this criterion. Images that require a lot of post-processing, or need a multi-expert assessment, can delay the diagnostic process.

4. Safety/risks

In its report 'To Err is Human' [Kohn *et al.* 2000], the IOM defines patient safety as "freedom from accidental injury". In this case, the use of each diagnostic alternative may result in specific (additional) risks for the patient, which possibly result in accidental injury. For evaluation of this criterion, both the chance of occurring and the severity of the injury should be taken into account.

4.1. Physical exposure

Physical exposure means exposure to radiation (X-ray, laser light), (ultrasonic or radiofrequent) waves, or high magnetic field strengths/gradients, that can be accompanied by biological damage.

4.2. Chemical exposure

Chemical exposure means exposure of the body to chemical substances that are need during the procedure, such as contrast agent, gel or fluid.

4.3. Bodily burden

Bodily burden means the amount of breast compression, but also injection of contrast agents. This may result in (severe) pain or inflammation/injury.

* to further elicit these characteristics of a diagnostic test (device) table A4.1 and the formula's below can be used.

Table A4.1 describes the characteristics of a test in relation to real disease (gold standard). The false-positive is represented by "b" and a false negative result is represented by "c". The formulas to calculate all the test characteristics are also showed.

Table A4.1 Test results			
		Disease	
		yes	no
Test	positive	a	b
	negative	c	d

Formulas:

Specificity = d / b+d

Accuracy = a+d / a+b+c+d

Prevalence = a+c / a+b+c+d

Sensitivity = a / a+c

Predictive Value:
    positive = a / a+b
    negative = d / c+d

Considering sensitivity and specificity you can choose what test is necessary or helpful, but when results are available, the most important information is predictive value. Results of a test can be positive or negative. In case the test is positive or abnormal, it is necessary to know some important information about the disease, and then calculate the positive predictive value, which expresses how many times the positive result of the test really represents disease. This value is better achieved with specific tests and depends on prevalence, sensitivity and specificity according to the equation:

$$\text{Positive predictive value} = \frac{(\text{Sensitivity x Prevalence})}{(\text{Sensitivity x Prevalence}) + ((1 - \text{Specificity}) \times (1 - \text{Prevalence}))}$$

APPENDIX 5:

PERFORMANCE MATRIX DIAGNOSTIC BREAST IMAGING TECHNIQUES

Performance Matrix

Op de volgende pagina's is een performance matrix weergegeven, waarin per alternatief voor elk criterium dat tijdens de AHP-analyse geëvalueerd gaat worden kort enkele informatie staat vermeld. Deze informatie is bedoeld als achtergrond informatie, waarop u uw oordeel zou kunnen baseren. De discussie zal dan ook niet gaan over de inhoud van deze matrix, waar nodig kan deze geraadpleegd worden.

Deze performance matrix behandelt alleen de criteria die geëvalueerd gaan worden. Voor het laatste onderdeel van de discussie, waarbij alleen de medici en technici aanwezig zijn, kan de technische performance van elk alternatief ook waardevolle achtergrondinformatie zijn. Momenteel wordt nog gewerkt aan een technische performance matrix, deze zal tijdens de discussie worden uitgedeeld.

Indicatie, sensitiviteit en specificiteit
Om alle informatie in een context te plaatsen is hieronder voor elk bestaand alternatief de klinische indicatie binnen de mammadiagnostiek weergegeven, evenals de sensitiviteit en specificiteit. Voor fotoakoestische mammografie is nog geen klinische indicatie bekend, wel is een grove schatting gemaakt van de sensitiviteit en specificiteit die met deze techniek behaald zou kunnen worden.

- Röntgenmammografie

Indicatie: Röntgenmammografie is vaak de primaire onderzoeksmethode voor symptomatische vrouwen met een leeftijd > 30 jaar.
Sensitiviteit: ± 85.5%, de sensitiviteit wordt hoger naarmate de borst uit meer vetweefsel bestaat.
Specificiteit: ± 87.7%

- Echografie

Indicatie: Echografie heeft de voorkeur als onderzoeksmethode bij vrouwen met een leeftijd < 30 jaar, zwangere vrouwen en vrouwen die borstvoeding geven. Daarnaast wordt het veel toegepast in combinatie met röntgenmammografie om een verdachte massa's verder te onderzoeken, en wordt het veel toegepast om puncties te begeleiden.
Sensitiviteit: Wanneer echografie wordt gebruikt als aanvulling op röntgenmammografie, stijgt de sensitiviteit met ongeveer 6.5 – 14%
Specificiteit: Wanneer echografie wordt gebruikt als aanvulling op röntgenmammografie, stijgt de specificiteit naar > 90%

- MRI

Indicatie: MRI wordt in de mammadiagnostiek toegepast bij een selecte patiëntengroep; hieronder valt bijvoorbeeld de diagnose van de postoperatieve borst of diagnostiek wanneer okselklieren positief zijn en op basis van röntgenmammografie en echografie geen diagnose gesteld kan worden.
Sensitiviteit: 86 - 100%, hoger dan röntgenmammografie, vooral voor vrouwen met dens borstweefsel.
Specificiteit: 20 – 90%, hangt sterk af van patiënt type.

- Fotoakoestische mammografie

Sensitiviteit: De eerste meetresultaten resulteerden in vergelijkbare conclusies die met een gecombineerde diagnose van röntgenmammografie en echografie waren gesteld. De correlatie met de pathologisch bepaalde tumorgrootte was hoog. Dit suggereert een vergelijkbare sensitiviteit met die van röntgenmammografie + echografie. Daarnaast werden bij een patiënt ringvormige patronen gezien, een maligne indicatie, terwijl in de mammo- en echografische afbeeldingen benigne eigenschappen werden herkend. Pathologie bevestigde maligniteit, wat erop kan duiden dat de sensitiviteit van

fotoakoestische mammografie iets hoger is dan van röntgenmammografie + echografie, ongeveer gelijk aan die van MRI omdat met behulp van deze techniek ook de ringvormige patronen zichtbaar worden.

Specificiteit: Fotoakoestische mammografie is op hetzelfde principe als MRI gebaseerd: angiogenese. Doordat ook benigne laesies soms vascularisatie laten zien en daardoor aankleuren is de specificiteit van MRI laag. Het wordt verwacht dat fotoakoestische mammografie hier ook last van heeft. De toevoeging van een extra optie aan fotoakoestiek: speed of sound imaging, kan deze specificiteit verhogen. Bij speed of sound imaging worden afbeeldingen gecreëerd die dezelfde informatie bevatten als echografie afbeeldingen. Met echografie kan een hoge specificiteit worden behaald (in aanvulling op rontgenmammografie), dit zou ook met fotoakoestiek kunnen.

Categorie	Subcriteria	Röntgenmammografie	Echografie	MRI	Fotoakoestische mammografie (PAM II)
1. Kosten	1.1 Scan tijd	20 min, beide borsten	15 min, enkele borst + oksel	30 min, beide borsten	15-20 min, enkele borst
	1.2 Werknemers	1 laborante (positionering en beoordelen beeldkwaliteit)	1 radioloog, 1 (vrouwelijke) laborante	2 laborantes (ivm dubbel check contrastvloeistof)	2 laborantes (klinisch en technisch)
	1.3 Prijs	Aanschaf: € 500.000 Onderhoud: 8% van de aanschafprijs per jaar Max levensduur: 10 jaar (functioneel)	Aanschaf: € 250.000 – € 400.000, afhankelijk van merk & opties Onderhoud: 10% van de aanschafprijs per jaar Max levensduur: 5 jaar (functioneel) Kan ook voor andere doeleinden gebruikt worden.	Aanschaf: € 1-2 miljoen Onderhoud: 10% van de aanschafprijs per jaar Max levensduur: 15-20 jaar (afschrijving op software & spoelen) Kan ook voor andere doeleinden gebruikt worden. (± 15% van de tijd voor mammo-onderzoek, groeit)	Aanschaf: € 400.000 Onderhoud: € 10.000 per jaar Max levensduur: 5-7 jaar (technisch)
	1.4 Randvoorwaarden	Afscherming ruimte: € 10.000 - € 50.000 Ruimte wordt alleen voor deze apparatuur benut.	Speciale transducer (bij aanschafprijs). Geen spec. eisen aan ruimte. Ruimte kan ook voor andere doeleinden gebruikt worden.	Bouw en afscherming ruimte: € 100.000 - € 500.000 Ruimte wordt alleen voor deze apparatuur benut. Spoelen + software: € 50.000 Bijvullen helium: € 5000 per jaar	Laserveiligheid afscherming (o.a. interlock): €10.000 Ruimte kan ook voor andere doeleinden gebruikt worden.
2.x. Sensitiviteit/ specificiteit	2.x.1 Randen massa	*Afhankelijk van technische eigenschappen*	*Afhankelijk van technische eigenschappen*	*Afhankelijk van technische eigenschappen*	Alleen bloedvaten zichtbaar, met speed-of-sound imaging ook anatomische informatie.
	2.x.2 Vorm massa	*Afhankelijk van technische eigenschappen*	*Afhankelijk van technische eigenschappen*	*Afhankelijk van technische eigenschappen*	Alleen bloedvaten zichtbaar, met speed-of-sound imaging ook anatomische informatie.
	2.x.3 Grootte massa	*Afhankelijk van technische eigenschappen*	Niet geschikt voor zeer grote tumoren.	Veel gebruikt om preoperatief tumorgrootte te bepalen	Alleen bloedvaten zichtbaar, met speed-of-sound imaging ook anatomische informatie.

Categorie	Subcriteria	Röntgenmammografie	Echografie	MRI	Fotoakoestische mammografie (PAM II)
	2.x.4 Locatie massa	2D projectie techniek. Zowel full breast als inzoomen mogelijk.	2D techniek, exacte lokalisatie sterk afhankelijk van vaardigheid en kennis radioloog. Lokale afbeelding. Beperkte diepteresolutie.	3D techniek. Zowel full breast als inzoomen mogelijk.	3D tomografische techniek. Full breast (radius = 40 mm). Beperkte diepteresolutie (afhankelijk van reconstructie algoritme).
	2.x.5 Ca++	In vet weefsel zichtbaar als witte puntjes. In dens weefsel moeilijk zichtbaar. Vergroting mogelijk voor nadere inspectie.	Meestal te klein voor detectie, soms zichtbaar als echorijke structuur in massa.	-	-
	2.x.6 Vascularisatie	-	Flow real time zichtbaar met Color Doppler	Zichtbaar met behulp van Gd contrastmiddel, individuele vaten zichtbaar.	Bloed wordt zichtbaar door de absorptie van licht. Individuele vaten niet zichtbaar.
	2.x.7 Zuurstofsaturatie	-	-	Eventueel mogelijk met spectroscopie.	PAM II kan informatie geven over de zuurstofsaturatie (%) van het bloed rond de tumor.
3. Patiënt comfort	3.1 Lichamelijk contact	De patiënt staat rechtop, haar borst wordt ingeklemd tussen twee platen gedurende enkele seconden.	De patiënt ligt op haar rug met haar hand achter haar hoofd. Een transducer wordt met lichte druk over haar borst bewogen, tussen de transducer en de borst zit (koude) contact gel.	De patiënt ligt op haar buik met haar borsten in een speciale spoel. Piercings en sieraden moeten van het lichaam verwijderd worden. De patiënt moet tijdens een sequentie (tot 8 min) zo stil mogelijk blijven liggen. Voor het scannen wordt een infuus met contrastmiddel aangesloten.	De patiënt ligt op haar buik met haar borst in een cup met water op lichaamstemperatuur. De borst wordt stil op de plaats gehouden door een magnetisch zuigmechanisme rond de tepel. De patiënt moet tijdens het onderzoek (tot 20 min) zo stil mogelijk blijven liggen

Categorie	Subcriteria	Röntgenmammografie	Echografie	MRI	Fotoakoestische mammografie (PAM II)
	3.2 Omgeving	Het mammogram wordt gemaakt door een laborante die de borst van de patiënt positioneert. Bij het maken van de foto staat ze achter een scherm. Directe communicatie is mogelijk.	Het echografisch onderzoek wordt uitgevoerd door een radioloog die tegenover de patiënt zit. Verder is in de ruimte een (vrouwelijke) laborante aanwezig. De radioloog legt uit wat hij op het scherm ziet.	De patiënt wordt op een onderzoekstafel een tunnel ingeschoven met een smalle opening (60-70 cm). Tijdens het scannen zijn geen andere personen in de ruimte, communicatie vindt plaats via de intercom. Het scannen produceert veel lawaai.	De laborante is tijdens de procedure in dezelfde ruimte als de patiënt aanwezig, directe communicatie is mogelijk. De patiënt draagt oogbescherming tijdens het scannen. De PAM maakt een luid klikkend geluid met een frequentie van ongeveer 10 Hz.
	3.3 Tijd tussen scan en uitslag	Direct uitslag mogelijk, vaak ook later dezelfde dag.	Dezelfde dag.	2-3 dagen.	2-3 dagen.
4. Veiligheid/ risico's	4.1 Fysische blootstelling	Blootstelling aan ioniserende straling. Kans op mammografie geïnduceerd carcinoom is afhankelijk van leeftijd patiënt (hoe ouder hoe lager). Patiënten < 25 jaar: 1 geïnduceerde tumor per 35.000 mammogrammen. Patiënten 25-30 jaar: 1 geïnduceerde tumor per 70.000 mammogrammen. Patiënten 50-65 jaar: 1 geïnduceerde tumor per 125.000 mammogrammen.	-	De magnetische velden kunnen storingen teweeg brengen in IED's (pacemakers) en hebben een sterke aantrekkingskracht op ferromagnetische materialen in de omgeving. Metalen implantaten kunnen opwarmen en mogelijk bewegen. Geleidende lussen kunnen brandwonden veroorzaken. Daarnaast kan schade ontstaan door opwarming onder invloed van RF energie en dient het gehoor beschermd te worden.	Wanneer laserveiligheids-maatregelen worden getroffen geen risico's. De vermogensdichtheid ligt onder de MPE, en wordt via een feedbackmechanisme gecontroleerd.

Categorie	Subcriteria	Röntgenmammografie	Echografie	MRI	Fotoakoestische mammografie (PAM II)
	4.2 Chemische blootstelling	Naast infectiepreventie maatregelen geen extra risico's.	Infectiepreventie maatregelen dienen te worden getroffen. De contact gel kan in enkele patiënten een allergische reactie teweeg brengen.	Het contrastmiddel, gadolineum, kan een toxische reactie in de nieren teweeg brengen. Daarnaast is ook een allergische reactie bij enkele patiënten mogelijk.	Naast infectiepreventie maatregelen geen extra risico's.
	4.3 Lichamelijke belasting	De druk die op de borst wordt uitgeoefend is ongeveer 150 N (± 15 kg)	-	De patiënt moet tijdens het onderzoek op haar buik liggen, wat een probleem kan vormen bij astmatische/obese patiënten.	De patiënt moet tijdens het onderzoek op haar buik liggen, wat een probleem kan vormen bij astmatische/obese patiënten.

APPENDIX 6:

RESULTS OF SCENARIO ANALYSIS

	Criteria	Sensitivity (.874)							Specificity (.126)						
		2.x.1 (.25)	**2.x.2** (.22)	**2.x.3** (.05)	**2.x.4** (.13)	**2.x.5** (.10)	**2.x.6** (.22)	**2.x.7** (.03)	**2.x.1** (.25)	**2.x.2** (.22)	**2.x.3** (.04)	**2.x.4** (.09)	**2.x.5** (.15)	**2.x.6** (.22)	**2.x.7** (.03)
Scenarios	**Alternatives**														
Negative scenario	Mammography & Ultrasound .411	.75	.64	.27	.33	.53	.07	.09	.75	.64	.27	.32	.53	.07	.09
	Mammography & PAM II .339	.13	.17	.47	.46	.40	.47	.42	.13	.17	.47	.46	.40	.47	.42
	PAM II .250	.12	.19	.26	.21	.07	.46	.49	.12	.19	.26	.21	.07	.46	.49
Average scenario	Mammography & Ultrasound .347	.65	.50	.16	.17	.50	.06	.06	.65	.50	.16	.17	.50	.06	.06
	Mammography & PAM II .351	.18	.24	.50	.45	.44	.46	.46	.18	.24	.50	.45	.44	.46	.46
	PAM II .302	.17	.26	.34	.38	.06	.48	.48	.17	.26	.34	.38	.06	.48	.48
Positive scenario	Mammography & Ultrasound .291	.54	.36	.11	.10	.44	.06	.06	.54	.36	.11	.10	.44	.06	.06
	Mammography & PAM II .376	.25	.30	.50	.44	.51	.46	.46	.25	.31	.51	.44	.51	.46	.46
	PAM II .333	.21	.34	.39	.46	.05	.48	.48	.21	.34	.39	.46	.05	.48	.48

Table A6.1 Results of Mammography, US & PAM comparison, effectiveness subcriteria

		Sensitivity (.799)							Specificity (.201)						
Criteria															
		2.x.1 (.14)	**2.x.2** (.14)	**2.x.3** (.03)	**2.x.4** (.01)	**2.x.5** (.06)	**2.x.6** (.37)	**2.x.7** (.25)	**2.x.1** (.40)	**2.x.2** (.12)	**2.x.3** (.02)	**2.x.4** (.02)	**2.x.5** (.06)	**2.x.6** (.18)	**2.x.7** (.20)
Scenarios	**Alternatives**														
Negative scenario	MRI .658	.85	.85	.88	.78	.66	.80	.18	.85	.85	.88	.78	.66	.80	.18
	PAM II .342	.15	.15	.12	.22	.34	.20	.82	.15	.15	.12	.22	.34	.20	.82
Average scenario	MRI .521	.67	.66	.80	.66	.50	.57	.14	.67	.66	.80	.66	.50	.57	.14
	PAM II .479	.33	.34	.20	.34	.50	.43	.86	.33	.34	.20	.34	.50	.43	.86
Positive scenario	MRI .291	.31	.35	.50	.43	.34	.31	.12	.31	.36	.50	.43	.34	.31	.12
	PAM II .709	.69	.65	.50	.57	.66	.69	.88	.69	.65	.50	.57	.66	.69	.88

Table A6.2 Results of MRI & PAM comparison, effectiveness subcriteria

APPENDIX 7:

HOUSE OF QUALITY

Title:	Improvement aspects of PAM II
Author:	M. Hilgerink
Date:	24-11-2009
Notes:	Roof: increase in one characteristic can/has to result in increase / decrease of other characteristic.
	This will indicate all possible consequences of a change of a characteristic

Legend

Symbol	Meaning	Value
⊖	Strong Relationship	9
O	Moderate Relationship	3
▲	Weak Relationship	1
++	Strong Positive Correlation	
+	Positive Correlation	
−	Negative Correlation	
▼	Strong Negative Correlation	
▼	Objective Is To Minimize	
▲	Objective Is To Maximize	
X	Objective Is To Hit Target	

Column #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Direction of Improvement: Minimize (▼), Maximize (▲), or Target (x)	▲	▲	X	▲	▲	▲	▲	▲							
Quality Characteristics	Number of wavelengths	Repetition rate laser	Max. power (MPE)	Detector bandwidth	Sensitivity detector	Detector area (no. of parallel detector elements)	Reconstruction algorithm quality	Design							

Row #	Max Relationship Value in Row	Relative Weight	Weight / Importance	Demanded Quality	1	2	3	4	5	6	7	8
1	9	23,4	0,1	Vascularization	⊖		O	⊖	⊖	▲	⊖	
2	9	15,1	0,1	Mass margins	O		▲	⊖	⊖	▲	⊖	
3	9	14,0	0,1	Mass shape	O		⊖	⊖	⊖	▲	⊖	
4	9	20,0	0,1	Oxygen saturation	⊖		▲	O	⊖	▲	⊖	
5	9	21,6	0,1	Bodily burden	▲	▲	▲			⊖		⊖
6	9	6,0	0,0	Time between scan and results	▲					▲	⊖	⊖
7												
8												
9												
10												

	1	2	3	4	5	6	7	8
Target or Limit Value	5	100 Hz		150%	0,5 Pa	640		
Difficulty (0=Easy to Accomplish, 10=Extremely Difficult)	6	8	5	9	8	7	5	8
Max Relationship Value in Column	9	1	9	9	9	9	9	9
Weight / Importance	504,7	21,6	253,0	532,1	651,8	272,6	705,8	248,2
Relative Weight	15,8	0,7	7,9	16,7	20,4	8,5	22,1	7,8