SUMMARY OF SAFETY AND EFFECTIVENESS DATA

I. GENERAL INFORMATION

Device Generic Name:	Mammogram Image Analysis System
Device Trade Name:	MammoReader
Applicant's Name and Address:	Intelligent Systems Software, Inc. 311 Park Place Blvd. Suite 240 Clearwater, FL 33759
Date of Panel Recommendation:	Not applicable, see section XII
Premarket Approval Application (PMA):	P010038
Date of Good Manufacturing Practice Inspection:	December 13 and 14, 2001
Date of Notice of Approval to Applicant:	January 15, 2002

II. INDICATIONS FOR USE

The MammoReader is a computer system intended to identify and mark regions of interest on standard mammographic views to bring them to the attention of the radiologist after the initial reading has been completed. Thus, the system assists the radiologist in minimizing observational oversights by identifying areas on the original mammogram that may warrant a second review.

III. CONTRAINDICATIONS

There are no contraindications for the use of this device.

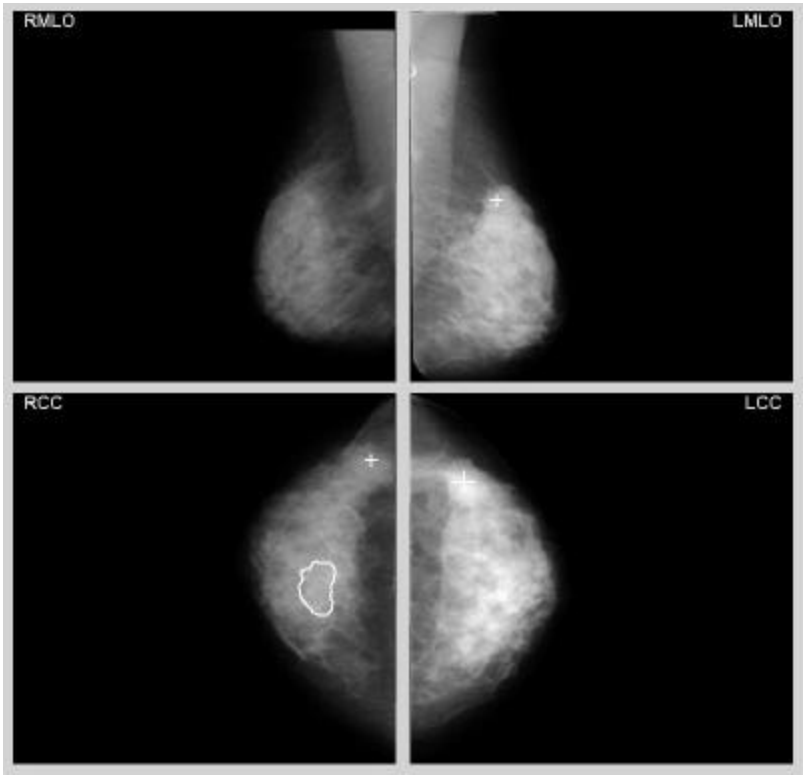
IV. WARNINGS AND PRECAUTIONS

Warnings and Precautions for use of the device are stated in the attached product labeling. (Attachment A)

V. DEVICE DESCRIPTION

The system consists of proprietary software, general-purpose computing equipment and an x-ray film digitizer. Mammograms are digitized using the x-ray film digitizer and the digital images are processed by image analysis software that automatically identifies suspicious regions in each image. The image analysis software has been designed to detect primary signs of breast cancer in mammogram images including microcalcification clusters, well-defined and ill-defined masses, spiculated lesions, architectural distortions and asymmetric densities. The results of the image analysis software are presented to the radiologist on an electronic display in the form of low-resolution images with markings indicating the locations of the suspicious regions in each image.

Hardware requirements include an x-ray film digitizer, a computer processor, disk storage, barcode readers, a barcode printer, and electronic displays. The device is delivered as an integrated system with specific installation configurations. The following figure shows an example case in which suspicious regions have been marked by the MammoReader system. The crosshairs mark potential lesion locations while closed outlines denote the extent of potential calcification clusters.



Example of a 4-view mammogram case, and the marks generated by the MammoReader. The crosshairs correspond to mass detections. The outlined region corresponds to a calcification cluster detection.

VI. ALTERNATIVE PRACTICES AND PROCEDURES

The field of mammography contains many standard practices and procedures that are well defined under the Mammography Quality Standards Act (MQSA) for maximizing the accuracy of reading screening mammograms. Although not required by MQSA, some clinics use “double reading” (a second human interpreter reviewing the same films) to increase the accuracy of screening mammography. The MammoReader or any other commercially available system approved for this intended use can be used as an alternative to the double reading of mammograms by two radiologists.

VII. MARKETING HISTORY

The MammoReader has never been marketed anywhere.

VIII. POTENTIAL ADVERSE EFFECTS OF DEVICE ON HEALTH

There are no known direct adverse effects on health or safety related to the use of the MammoReader. Indirect risks are that the device may fail to detect and mark some actionable region and/or may mark regions that are not actionable. Warnings are included with the device explaining these possibilities. Directions regarding the proper use of the markings generated by the system are also included.

IX. NON-CLINICAL STUDIES

The following sections summarize testing that was performed, the results of that testing, and a description of how the software detects cancers, so the radiologist may better understand how to use the MammoReader system properly.

Benchmark Testing

Benchmark testing is aimed at measuring the expected performance of the Image Analysis Module of the MammoReader system. Performance is defined in terms of the percentage of cancers that the system can detect and the average number of marks displayed for each image. A large database of digitized mammograms has been used to benchmark the true positive and false positive detection rates for the MammoReader. Each case in the database contains information about the location and type of each cancer as identified by an experienced radiologist. All cancers were verified by biopsy. All negative exams were verified by normal follow-up.

To estimate the sensitivity of the MammoReader system, the Image Analysis Module was run on 465 cancer cases. There were 440 cases with a single cancer, 23 cases with two cancers, and 2 cases with three cancers. In 169 cases, clustered calcifications were the only sign of cancer. The remaining 296 cases had other types of malignancy present (a mass, architectural distortion, or asymmetry). The MammoReader detected at least one cancer region in 89.3% (±1.4%) of the cancer cases. For cases that had malignant calcification clusters as the only sign of cancer, the MammoReader detected at least one cluster 91.0% (±2.2%) of the time. The detection rate for the remaining cancer cases that contained malignant masses was 87.4% (± 1.9%).

To estimate the false positive rate of the MammoReader system, the Image Analysis Module was run on 265 negative cases. There was an average of 0.83 marks per image. There were an average of 0.12 calcification marks per image and 0.71 mass marks per image.

Benchmark testing of the MammoReader shows that on average 1.53 true positive prompts are generated per case among cancer cases and 3.32 false positive prompts are generated for non-cancer cases. Since the system limits the number of detections per case, the MammoReader generates an average of 2.32 false positive prompts in cancer cases. This means that even among cancer cases the fraction of marked regions that are cancerous is estimated to be only 40%.

Repeatability Testing

Some variability in the outcome of the MammoReader is expected from one run to the next regardless of whether films are digitized multiple times on one system or on different systems. The source of this variability is the scanner. The software produces the same output for the same digital image. To measure this variability and its potential effect on estimates of system sensitivity, a repeatability study was designed and executed. Sixty (60) cases were used, each having a single cancer detected by the MammoReader in at least one view during a baseline run. These cases were selected from all consecutive screening cancer cases with negative prior exams collected at one clinical site. Fifty-seven (57) of the cases had a cancer identified by a radiologist in both views. The remaining three cases had a cancer identified in one view.

Excluding the baseline run, the 60 cases were digitized two times on each of three different scanners. The analysis results of the MammoReader were recorded and reviewed to determine if the lesions were detected for each run. A probability measure of the repeatability of the outcome of the system was calculated based on the number of runs in which the MammoReader results matched the baseline results in terms of either detecting or not detecting a particular cancer region. The estimated probability (± standard error) of repeating the same outcome in the same view for a particular lesion (or image level repeatability) is 93.3% (± 2.3%). The probability of repeating the outcome of either detecting the lesion in at least one view or not detecting the lesion in either view (or case level repeatability) is 97.5% (± 2.0%).

Of the 60 cases used in the repeatability study, 19 have calcification clusters visible in both views and 41 have other signs of cancer. For the calcifications cases, the probability of repeating the same outcome on the image and case level is 95.5% (± 3.4%) and 99.0% (± 2.3%) respectively. For cases with other signs of cancer, the image level repeatability is 92.2% (± 3.0%), and the case level repeatability is 96.8% (± 2.7%).

Compared to the baseline run, an average of 5.6% of the instances of cancer was missed in the 6 test runs that were detected in the baseline run. An almost identical average of 5.7% of the instances of cancer was detected in the 6 test runs that were missed in the baseline run. Statistical analysis shows no significant difference in the overall sensitivity between all possible pairs of runs.

X. CLINICAL TESTING

Definitions

Term	Definition
Screening mammogram	A routine mammogram obtained for a woman who did not exhibit physical symptoms of breast cancer, and who was not undergoing further evaluation for a specific finding, at the time of the exam.
Diagnostic mammogram	A mammogram obtained for a patient experiencing physical symptoms consistent with breast cancer, or to evaluate a specific finding.
Current exam	A diagnostic or screening mammogram that led to the diagnosis of breast cancer, or a diagnostic mammogram that was obtained in the process of diagnosing breast cancer in a symptomatic woman.
Prior exam	The most recent screening mammogram with a negative assessment obtained prior to a current exam.
BI-RADS™	Breast Imaging Reporting and Data System.
BI-RADS™ Assessment	A value of 0 to 5 that indicates the overall assessment of a mammogram as follows: 0 – other procedures required 1 – negative 2 – negative exam; benign findings 3 – probably benign, short interval follow-up suggested 4 – suspicious abnormality, biopsy should be considered 5 – highly suggestive of malignancy
Actionable lesion	A lesion that a radiologist would consider suspicious enough to warrant further evaluation including additional imaging or biopsy for the patient.
Recall rate	The percentage of cases in which the radiologist finds actionable lesions.
Unaided review	A conventional review of a screening mammogram performed by a trained radiologist for the purpose of identifying signs of breast cancer.
Aided review	A review of a screening mammogram that involves using the results of the MammoReader before making a final assessment for a case.

Sensitivity Study Results Summary

Three radiologists independently reviewed each of 327 current and prior mammography exams of confirmed cancer cases. The cases were consecutive cancers for which prior mammograms could be located -- diagnosed at one of 6 different clinical sites. The sites included a private radiology practice, a university-affiliated hospital, a large not-for-profit hospital, two breast health centers and a medical center affiliated with a large Health Maintenance Organization. The current and prior mammography exams were performed between 1995 and 2000. The prior exam was obtained up to 31 months prior to the current exam, and the average patient age was 65 with a range of 39 to 87 years. All cancer cases with available prior screening exams assessed as negative at the time of screening were used excluding:

1. Cases with missing films.
2. Cases in which the lesion was noted in the prior report.
3. Cases with implants.

Two sets of three experienced radiologists participated in this study. All six radiologists were MQSA certified, had an average of 17 (range of 12 to 24) years of experience in mammography, and read an average of 6,328 (range of 2,078 to 16,000) mammograms in the year prior to the study.

Each radiologist had access to the films of both current and prior exams as well as their reports. Each of the three radiologists was asked to assess independently whether a lesion described in the current report was visible in the prior exam. If judged visible in the prior, the radiologist was asked to assess subjectively whether the identified lesion would warrant patient recall in a standard screening environment if the attention of the radiologist is drawn to that lesion.

The MammoReader was run on all 327 current and prior exams. A lesion was considered detected by the MammoReader if the system correctly marks the location of the lesion in at least one view as determined by a radiologist.

Table 1 shows the number of cancer cases in which the radiologist located a lesion in the current exam, and the percent of these cases in which the lesion was detected by the MammoReader. The table also shows a breakdown of these percentages for diagnostic and screening current exams, as well as detection rates for cases in which a calcifications cluster was identified versus other types of lesions.

Table 1: Detection results on current exams in the Sensitivity Study (there is overlap between cases with calcification clusters and cases with other types of lesions).

	Total Cases	Lesion Located in Current Exam	Lesion Detected by Mammo-Reader	% Detected of Total Located	95% C.I. Lower Bound	95% C.I. Upper Bound
All current exams	327	322	268	83.23	78.69	87.14
Diagnostic current exams	96	94	79	84.04	75.05	90.78
Screening current exams	231	228	189	82.89	77.37	87.54
Screening currents with calcification clusters		72	66	91.67	82.74	96.88
Screening currents with other types of lesions		163	129	79.14	72.09	85.10

The MammoReader is intended for use as an aid to help reduce the chances of missing a lesion due to oversight in the course of normal screening. The prior screening exams originally read as negative were targeted in the study as cases that could potentially have benefited from the MammoReader. Because of the way the system is intended to be used in standard practice, the system benefit in detecting a cancer that would

otherwise be overlooked will be realized only if the radiologist agrees that a particular region marked by the system is suspicious. Hence, the detection of a lesion in the prior exam originally read as negative is meaningful only for lesions that a radiologist is most likely to deem actionable if his or her attention is somehow focused on the location of that lesion.

Table 2 shows the number of prior cases in which the same lesion in the prior exam was deemed actionable by all three radiologists, by two out of three or by one out of three. Each case is counted only once in one of the three categories (i.e. if a case has one lesion deemed actionable by all three radiologists and a second lesion deemed actionable by two out of three, the case is counted only once in the 3/3 category and so on). The table also shows the number of cases in which the same lesion was detected correctly by the MammoReader system.

The last row of Table 2 shows weighted sums of the cases in each category. The cases in each category are weighted by the fraction of radiologists who deemed the cases to be actionable. Cases in the top category are weighted by 3/3 or 1. Cases in the next category are weighted by 2/3, and cases in the last category are weighted by 1/3.

Table 2: Prior cases with lesions judged actionable by 3, 2, or 1 of 3 radiologists. The weighted sum of the total cases actionable is the result of multiplying the number in each category by the fraction of radiologists who deemed the cases actionable within that category then adding the resulting weighted numbers.

	Number of Cases	% Deemed Actionable of 327 Prior Exams	95% C.I. Lower Bound	95% C.I. Upper Bound	Number Detected by Mammo-Reader	% Deemed Actionable and Detected of 327 Prior Exams	95% C.I. Lower Bound	95% C.I. Upper Bound
Deemed actionable by 3 of 3	53	16.21	12.38	20.66	39	11.93	8.62	15.94
Deemed actionable by 2 of 3	56	17.13	13.20	21.65	41	12.54	9.15	16.62
Deemed actionable by 1 of 3	56	17.13	13.20	21.65	26	7.95	5.26	11.43
Weighted Sum	109	33.33	28.24	38.73	75	22.94	18.49	27.88

A larger percentage of the lesions found actionable by all three radiologists were detected by the MammoReader (39/53 or 74%) than the lesions found actionable by only one of the three radiologists (26/56 or 46%). This indicates that the system is more likely to detect lesions that are detected by and/or appear suspicious to more of the radiologists.

When more radiologists agree that a particular lesion is actionable, it is less likely that the lesion would be missed during routine screening if the attention of the radiologist had somehow been focused on that lesion. The weighted sum of 109 provides an estimate of the total number of cases, out of the 327 prior exams that are most likely to be deemed actionable if the MammoReader points out the location of the lesion to an interpreting radiologist during routine screening. The weighted sum of 75 provides an estimate of the number of cases with actionable lesions actually detected by the MammoReader.

This study shows that about 33% (109/327) of the cancer cases collected in the study show actionable signs of the cancer in the prior exams that were originally read as negative. The MammoReader detected the actionable lesion in 69% (75/109) of these cases, or 23% (75/327) of all prior exams.

The time intervals between the current and prior exams on which lesions were deemed actionable, and were detected by the MammoReader, ranged from 4 to 27 months with an average of 14 months. This study showed that 23%, or approximately one in four to five of those women diagnosed with breast cancer and who

had had an earlier screening mammogram within the prior 4 to 27 months, could have had their cancers discovered earlier by an average of 14, and a range of 4 to 27, months.

Specificity Study Results Summary

Ten radiologists affiliated with three different clinical sites participated in this study. The radiologists were MQSA certified, had an average of 10 (range of 1 to 20) years of mammography experience, and read an average of 2,818 (range of 849 to 7,013) mammograms in the year prior to the study. The sites included a university affiliated research center, a not-for-profit large hospital, and a private radiology group practice.

Three hundred consecutive routine-screening exams obtained between July 19 and July 29, 2000 were collected from one of the sites. The cases were not pre-selected in any way and hence, each case may or may not have been read as negative at the time of screening. Cases were only excluded if any of the films were missing or were not available. The average patient age was 51 (range of 34 to 82).

All 300 cases were run through the MammoReader. The system generated 205 calcification clusters and 894 mass marks. Since these cases are routine screening exams, we assume that virtually all of these marks are false detections. This results in a false detection rate of 0.17 clusters and 0.75 masses per image, or an overall false detection rate of 0.92 marks per image. Note that these results are for a different data set than the normal cases used in the benchmark testing described in Section 7.1.

The cases were randomized such that each radiologist read a different set of 150 cases aided and the remaining 150 cases unaided. Further, each case was read by a different randomly selected group of five radiologists aided and by the other five radiologists unaided. The study resulted in 150 aided and 150 unaided readings per radiologist, or a total of 1,500 aided and 1,500 unaided readings across all radiologists.

Each radiologist had access to the screening exam and its prior exam if available. The cases were read in the same sequence within a session and the radiologists were instructed to review each case in the conventional manner before reviewing the MammoReader results. The radiologists did not know if a case was to be read aided or unaided until they finished their conventional review and requested the MammoReader results. Radiologist assessment was provided in the form of an overall BI-RADS assessment for each case.

The recall rate, defined as the percentage of cases assigned a BI-RADS assessment of 0, 4 or 5 out of the total cases reviewed, was calculated for the 1,500 aided and 1,500 unaided readings across all radiologists. Aided and unaided recall rates were also calculated for each radiologist individually. Table 3 shows the average recall rates for each radiologist and the average recall rates across all aided versus unaided readings. The average recall rates (and exact 95% confidence intervals) are 18.1% (16.1%, 20.1%) for the aided readers, and 15.1% (13.3%, 17.0%) for the unaided readers. The difference between aided and unaided recall rates of 3% (95% confidence interval of 1% to 5%) is statistically significant (P = 0.03).

Table 3: Recall rates for aided and unaided readings including BI-RADS 0, 4 or 5 as recall (N/A in this table indicates that the recall rate is not available because radiologist had less than one year experience).

RAD	Years Exp.	Mamm. Read per Year	Aided Recall Count	Unaided Recall Count	**Aided Recall Rate (%)**	Lower Bound for C.I.	Upper Bound for C.I.	**Unaided Recall Rate (%)**	Lower Bound for C.I.	Upper Bound for C.I.	Reported Recall Rate (%)
1	12	2,415	24	21	**16.00**	10.53	22.86	**14.00**	8.88	20.60	12.17
2	5	849	45	29	**30.00**	22.80	38.01	**19.33**	13.35	26.57	17.50
3	4	3,000	32	20	**21.33**	15.07	28.76	**13.33**	8.34	19.84	15.00
4	11	7,013	42	37	**28.00**	20.98	35.91	**24.67**	18.00	32.36	16.35
5	14	2,078	23	27	**15.33**	9.98	22.11	**18.00**	12.21	25.10	8.25
6	1	1,000	45	30	**30.00**	22.80	38.01	**20.00**	13.92	27.30	N/A

7	15	3,886	10	12	**6.67**	3.24	11.92	**8.00**	4.20	13.56	7.80
8	20	3,076	24	19	**16.00**	10.53	22.86	**12.67**	7.80	19.07	14.80
9	4	2,876	13	18	**8.67**	4.70	14.36	**12.00**	7.27	18.30	13.70
10	14	1,983	13	13	**8.67**	4.70	14.36	**8.67**	4.70	14.36	10.70
Total			271	226	**18.07**	16.15	20.11	**15.07**	13.29	16.98	

When reading cases aided, the radiologists were asked to review the cases in the conventional manner before reviewing the outcome of the MammoReader, and subsequently making their final assessments. The radiologists had the option to make free notes on each case. Four of the 10 radiologists made specific notes indicating that the MammoReader caused them to recall a case that they would not have recalled otherwise. The radiologists indicated this in 5, 1, 1, and 2 out of the 150 cases they each read unaided and then aided.

Based on these notes, it is estimated that an average of 2 to 3 (between 1% and 2%) of the 150 cases read aided by each radiologist were given a recall assessment as a result of reviewing the outcome of the MammoReader and re-examining the films. This along with the measured aided and unaided recall rates indicate that use of the MammoReader can result in a small increase in the recall rate. This increase is acceptable considering the improvement in earlier detection.

XI. CONCLUSIONS DRAWN FROM CLINICAL STUDIES

The clinical data demonstrated that use of the MammoReader would have helped the mammographer detect a certain percentage of cancers on screening mammography that had been missed, with an acceptable increase in callback rate.

An unblinded retrospective study of 327 cancer cases showed that 23% (95% CI, 18-28%) of women diagnosed with breast cancer, who had had prior screening mammograms, could have had their cancers discovered earlier, by an average of 14 months, with use of the MammoReader. The study was not designed to measure the percentage of additional cancers that would be detected in a screening population, which includes both women who have and who have not had prior screening mammograms. The percentage derived from the study was based on the number of cancers marked by the system and judged to be actionable by mammographers in mammograms obtained up to 31 months prior to diagnosis. This percentage may be overestimated because the study radiologists were asked to render unblinded judgments as to the actionability of a lesion identified in retrospect.

In a second clinical study, done independently of the first, ISSI found that use of the MammoReader increased the callback (false positive) rate from 15.1% to 18.1%, a relative increase of 20% (95% CI, 6.6%-32.9%). The callback rate is necessarily increased, because the MammoReader is intended only to alert a mammographer to <u>additional</u> areas on the mammogram that she or he may have failed to notice. In particular, it is specifically intended <u>not</u> to change a mammographer's decision to work up a lesion (i.e., to call the patient back) based on the initial (unaided) review of the mammogram.

XII. PANEL RECOMMENDATION

In accordance with the provisions of section 515(c)(2) of the act as amended by the Safe Medical Devices Act of 1990, this PMA was not referred to the Radiological Devices Panel, an FDA advisory committee, for review and recommendation because the information in the PMA substantially duplicates information previously reviewed by this panel.

XIII. CDRH DECISION

The sponsor's manufacturing and control facilities were inspected on December 13 & 14, 2001, and they were found to be in compliance with Good Manufacturing Practice Regulations.

Based on the review of the information submitted the PMA (which includes all modules and amendments), the device has been found to be reasonably safe and effective for its intended use when used in accordance with the instructions for use. CDRH worked with ISSI and refined the labeling so that it accurately described the capabilities of the device as demonstrated by the clinical trials that were conducted.

FDA issued an approval order on January 15, 2002

XIV. APPROVAL SPECIFICATIONS

Directions for use: See attached labeling.

Hazards to Health from Use of the device: See Indications, Contraindications, Warnings, and Precautions in the attached labeling.

Postapproval Requirements and Restrictions: See approval order.