IDE-BCDD-0001 102-2000008-001 Rev. 1 CLINICAL TRIAL REPORT
STATISTICAL ANALYSIS REPORT

A Prospective, Randomized, Blinded Study of 3-Dimensional Electrical Impedance Tomography Imaging (MEM device) . Modality as an Adjunct to Diagnostic Mammography Radiologic Exam in Detecting and Isolating Breast Cancer to Determine the Necessity of Fine Needle Aspiration, Core Biopsy or Open Biopsy

<u>Sponsor Information:</u>

Technology Commercialization International, Inc.,
One Park Square, Suite 210
6501 Americas Parkway NE
Albuquerque, NM 87110
Tel: (505) 247-4100, Fax: (505) 888-5244, E-mail: tci@tcimed.com
Contact Person: J. David Meister, PE

<u>Device Name:</u>

3D EIT Cancer Camera™, MEM -Breast Cancer Detection Device (BCDD™) Model 2 02
(U.S. Patent No. 6,167,300)

<u>Intended Use:</u>

This 3-D Electrical Impedance Tomography (3D EIT) medical imaging device is designed to detect and isolate breast carcinoma by measuring cell conductivity and displaying tissue conductivity in 3-dimensional tomographic images. It is intended that this imaging modality be used as an adjunct to diagnostic mammography radiologic exam in detecting and isolating breast cancer to further determine the necessity of fine needle aspiration, core biopsy or open biopsy.

> CAUTION- The 3-D Electrical Impedance Tomography (3D EIT) medical imaging device is an investigational device, limited by United States Law to investigational use.

TABLE OF CONTENTS

Section **Page**

Statistical Analysis of Clinical Trial Results

A clinical trial of a 3-dimensional electrical impedance tomography (3D EIT) medical device for breast cancer detection began in October 2001 and was completed in March 2002. The clinical trial was performed in accordance with trial protocol approved by the Western Institutional Review Board (WIRB) in Olympia, Washington. This formal clinical trial involved 40 female subjects who had completed a diagnostic radiological examination and the patient was scheduled for fine needle aspiration (FNA), core biopsy or open biopsy as a result of the diagnostic mammography usually followed by ultrasound. Most diagnostic radiology reports indicated BIRADS 4.

The testing sequence for the clinical trial is described as follows:

1. Prior to FNA or Biopsy, each patient was examined using 3D EIT, a 3-dimensional electrical impedance tomography imaging modality. The 3D EIT examination was performed before any invasive procedure was conducted on the breast.

2. After 3D EIT examination a fine needle aspiration (FNA) or biopsy was performed and the tissue samples sent to Pathology. The pathologic histology report formed the basis for statistical comparison of both 3D EIT and Diagnostic Radiology in this statistical analysis.

3. Statistical analysis of the results of the clinical trials compares 3D EIT and diagnostic radiology to histology results of biopsy.

A diagram of this testing sequence is given in Figure 1.

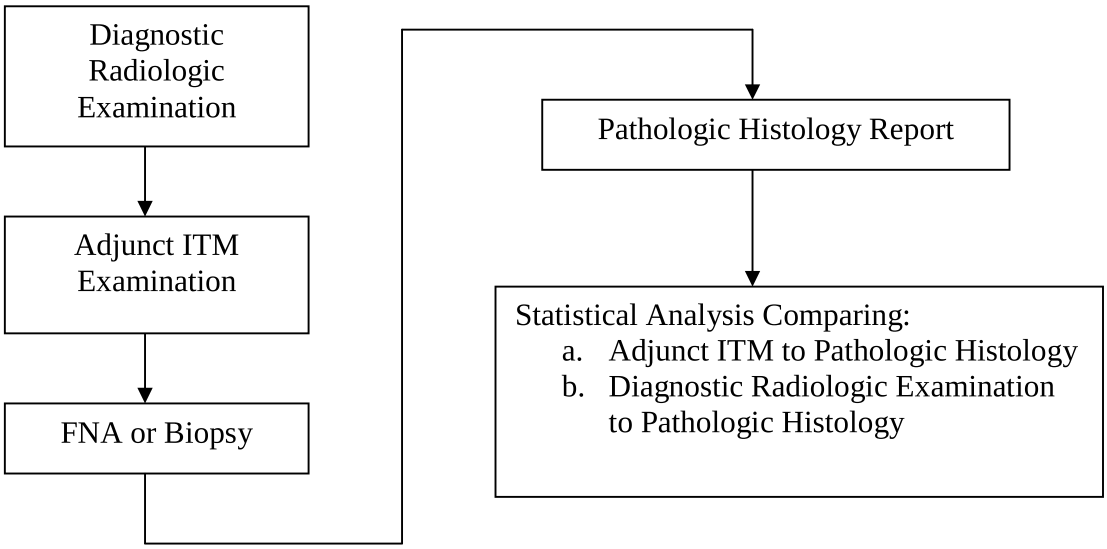


Diagram of the Testing Sequence for the Detection of Breast Cancer

Test results for all patients were compared to pathologic histology, which allows the 3D EIT and radiology diagnoses to be classified as follows.

	Pathologic Histology	
Diagnosis	**Malignant**	**Benign**
Malignant	True positive	False positive
Benign	False negative	True negative

Comparison with pathologic histology allows the diagnoses of 3D EIT and radiology to be classified as either "correct" or "not correct" and these classifications form the basis of the statistical analyses.

Statistical Methodology

Statistical analyses of the test results are based on the McNemar test after first applying a blocking technique employed in the Cochran test. Both of these nonparametric test procedures are presented in Conover (1999). Details of the Cochran Test are presented first and these are followed by details of the McNemar test.

The Cochran Test

Data for the Cochran Test

3D EIT and radiology were applied independently to each of 40 female patients as outlined in the above test sequence. The diagnosis of each modality was recorded as either "correct" or "not correct" when compared to results of fine needle aspiration or biopsy. The respective results were coded as "1" or "0" as shown in Table 1.

This tabular summary corresponds to the format given in Conover for the Cochran test. Conover also presents the following assumptions for the Cochran test.

Assumptions for the Cochran Test

1. The patients were randomly selected from the population of all possible patients selected for biopsy
2. The outcomes of the two technologies (3D EIT and diagnostic radiology) may be dichotomized as either "0" or "1" for each patient

Test Statistic for the Cochran Test

The simplified form of the test statistic T_1 for the case of two treatments (3D EIT and diagnostic radiology) may be written as follows:

$$T_1 = \frac{(C_1 - C_2)^2}{\sum_{i=1}^{r} R_i (2 - R_i)} \tag{1}$$

where C_1 and C_2 are the column totals in Table 1, R_i is the total of the i^{th} row in that table, and r is the number of patients.

Null Distribution of the Test Statistic for the Cochran Test

The null distribution of T_1 can be approximated with a chi-square distribution having 1 degree of freedom. The decision rule is to reject H_0 if T_1 exceeds the upper 1-α quantile of a chi-square distribution with 1 degree of freedom.

Hypotheses for the Cochran Test

The null hypothesis can be stated as:

H_0: P(Correct diagnosis for 3D EIT) = P(Correct diagnosis for diagnostic radiology) (2)

The alternative hypothesis can be either two-tailed:

H_1: P(Correct diagnosis for 3D EIT) $\neq$ P(Correct diagnosis for diagnostic radiology) (3)

or one-tailed:

H_1: P(Correct diagnosis for 3D EIT) > P(Correct diagnosis for diagnostic radiology) (4)

Note that the null hypothesis in Equation 2 states that 3D EIT and diagnostic radiology are equally effective in their respective agreement with FNA and biopsy results. It does not speak individually to false positives or false negatives, but rather it speaks to the totality of the level of agreement of the two technologies with the "truth" as determined by FNA or biopsy.

The McNemar test was used to analyze the test results based on the dichotomization of the test results for both 3D EIT and diagnostic radiology as shown in Table 1. Details of the McNemar test are now presented.

The McNemar Test

Data for the McNemar Test

The data for the McNemar test consist of observations of two sets of independent bivariate random variables (X_i, Y_i), i = 1, 2, …, *r* that result from breast cancer examinations on *r* women as described in the test protocol. The two sets of bivariate random variables are defined as follows for 3D EIT and for diagnostic radiology.

$X_i = 0$ if 3D EIT is not correct $Y_i = 0$ if diagnostic radiology is not correct

$X_i = 1$ if 3D EIT is correct $Y_i = 1$ if diagnostic radiology is correct

Table 1. Statistical Summary of the Test Results Utilized by the Cochran Test for the Clinical Trial Consisting of 40 Female Subjects

		Coded Results: 1 = correct 0 = not correct		**Row Totals**
Patient Notebook No.	**Pathologic Histology**	**Diagnostic 3D EIT Examination**	**Diagnostic Radiology Examination**	$\mathbf{R_i}$
1	Benign	1	0	1
2	Benign	0	1	1
3	Benign	1	0	1
4	Benign	0	0	0
5	Malignant	1	1	2
6	Malignant	1	1	2
7	Malignant	1	1	2
8	Malignant	1	1	2
9	Malignant	1	1	2
10	Benign	1	0	1
11	Benign	0	0	0
12	Benign	1	0	1
13	Benign	1	0	1
14	Benign	1	0	1
15	Malignant	1	1	2
16	Benign	0	0	0
17	Malignant	1	1	2
18	Benign	1	0	1
19	Benign	1	0	1
20	Benign	1	0	1
21	Benign	0	0	0

		Coded Results: 1 = correct 0 = not correct		**Row Totals**
Patient Notebook No.	**Pathologic Histology**	**Diagnostic 3D EIT Examination**	**Diagnostic Radiology Examination**	R_i
22	Benign	0	0	0
23	Benign	1	0	1
24	Malignant	1	1	2
25	Malignant	1	1	2
26	Benign	1	0	1
27	Benign	1	0	1
28	Malignant	1	1	2
29	Malignant	1	1	2
30	Benign	0	0	0
31	Benign	1	0	1
32	Benign	0	0	0
33	Benign	1	0	1
34	Benign	0	0	0
35	Malignant	1	0	1
36	Malignant	0	1	1
37	Benign	1	1	2
38	Benign	1	0	1
39	Benign	1	1	2
40	Malignant	1	1	2
	Totals:	$C_1 = 30$	$C_2 = 16$	$C_1+C_2 = 46$

Thus, there are four possible values for (X_i, Y_i): (0, 0), (0, 1), (1, 0), and (1, 1). These outcomes can be summarized in a 2×2 contingency table as shown in Table 2.

Table 2. Contingency Table Used to Classify Test Results

		Diagnostic Radiology Results	
		$Y_i = 1$ (correct)	**$Y_i = 0$ (not correct)**
Diagnostic 3D EIT Results	**$X_i = 1$ (correct)**	*a* = number of pairs where $X_i = 1$ and $Y_i = 1$	*b* = number of pairs where $X_i = 1$ and $Y_i = 0$
	$X_i = 0$ (not correct)	*c* = number of pairs where $X_i = 0$ and $Y_i = 1$	*d* = number of pairs where $X_i = 0$ and $Y_i = 0$

Hypotheses for the McNemar Test

The null hypothesis to be tested is the same as stated in Equation 2.

Test Statistic for the McNemar Test

Conover gives the following form for the McNemar test statistic that is used to test H_0:

$$T_2 = \frac{(b-c)^2}{b+c} \tag{5}$$

where *b* is defined in the contingency table in Table 2 as the number of pairs where $X_i = 1$ and $Y_i = 0$ and *c* is defined as the number of pairs where $X_i = 0$ and $Y_i = 1$.

It is worth noting that T_2 does not depend on either *a* or *d*. These cases represent "ties" (either $X_i = 1$ and $Y_i = 1$ or $X_i = 0$ and $Y_i = 0$). Ties are not included in the test statistic because they provide no useful information on the comparison of 3D EIT and diagnostic radiology.

The useful information for the comparison of interest as stated in H_0 is provided by *b* and *c* since these are the cases where the 3D EIT and radiology diagnoses disagree. If H_0 is true, these disagreements should occur randomly and *b* and *c* should be approximately equal. If H_0 is not true, then the probability of a correct diagnosis is not the same for 3D EIT and diagnostic radiology and *b* and *c* can be expected to differ considerably. The test statistic T_2 is designed to be sensitive to differences in *b* and *c* with $T_2 = 0$ when $b = c$ and $T_2 > 0$ when $b \neq c$. Hence, H_0 will be rejected only for large values of T_2 as explained in the next paragraph.

Decision Rule for the McNemar Test

The decision rule is to reject H_0 if T_2 exceeds the upper $1-\alpha$ quantile of the chi-square distribution with 1 degree of freedom.

Equivalence of the Cochran and McNemar Tests for Two Treatments

Conover (p. 256) shows the equivalence of the Cochran and McNemar tests when there are two treatments. That development is as follows:

The column total C_1 in Table 1 is the number of 1s (correct responses) for 3D EIT, which is equivalent to $a + b$ in the 2×2 contingency table summary for the McNemar test. Likewise, C_2 is the number of 1s (correct responses) for diagnostic radiology, which is equivalent to $a + c$. Thus, the numerator of the test statistic in Equation 1 can be written as:

$$(C_1 - C_2)^2 = (a + b - a - c)^2 = (b - c)^2 \tag{6}$$

The term $R_i(2\text{-}R_i)$ in the denominator of the test statistic in Equation 1 takes on the following possible values.

C_i	C_j	R_i	$R_i(2\text{-}R_i)$
1	1	2	0
0	0	0	0
1	0	1	1
0	1	1	1

The last two rows of this summary represent cases where there is a difference between 3D EIT and diagnostic radiology. Hence, the denominator in Equation 1 is merely the total number of patients for which there is a difference between the diagnoses of 3D EIT and diagnostic radiology. This total is represented by $b + c$ in the 2×2 contingency table.

Thus, the test statistic in Equation 1 can be expressed as

$$T_1 = \frac{(C_1 - C_2)^2}{\sum_{i=1}^{r} R_i(2 - R_i)} = \frac{(b-c)^2}{b+c} \tag{7}$$

Note that the right-hand side of Equation 7 is the same as T_2 in Equation 5.

Statistical Analysis of Clinical Trial Results

The totals in Table 1 produce the following test statistic for the Cochran test using Equation 1.

$$T_1 = \frac{(30-16)^2}{18} = \frac{196}{18} = 10.89 \tag{8}$$

The 2 × 2 contingency table summary similar to the one in Table 2 for the McNemar test appears in Table 3.

Table 3. Contingency Table Summary of Clinical Trial Test Results

		Diagnostic Radiology Results	
		$Y_i = 1$ (correct)	**$Y_i = 0$ (not correct)**
Diagnostic 3D EIT Results	**$X_i = 1$ (correct)**	$a = 14$	$b = 16$
	$X_i = 0$ (not correct)	$c = 2$	$d = 8$

Substituting the values in Table 3 into Equation 5 gives

$$T_2 = \frac{(30-16)^2}{16+2} = \frac{196}{18} = 10.89 \tag{9}$$

which is the same as the result for the Cochran test using Equation 8.

H_0 is rejected with a level of significance of $\alpha = 0.05$ since $T_2 > 3.841$, which is the upper 0.05 quantile of the chi-square distribution with 1 degree of freedom. The p-value is defined (see Iman, 1994) as the probability of getting the observed value or a value more extreme in the direction of the alternative hypothesis, assuming the null hypothesis is true. The p-value for T_2 is approximated by the chi-square distribution with 1 degree of freedom as $P(T_2 \geq 10.89) = 0.00097$. This small value provides strong evidence for rejecting H_0 in favor of the *two-tailed* alternative hypothesis in Equation 3.

Exact Distribution of the McNemar Test Statistic

Conover gives the following alternative form of the McNemar test statistic:

$$T_3 = b \tag{10}$$

T_3 is a very useful form of the test statistic as it can easily be used to perform either one- or two-tailed tests as identified in Equations 4 and 3, respectively. The exact null distribution of the test statistic T_3 is given in Conover as a binomial distribution with parameters n and $p = 0.5$, where $n = b + c$.

As explained previously, b and c are approximately equal when H_0 is true. If H_0 is not true, b and c will differ considerably. The probabilities associated with observed differences in b and c are quantified by the following cumulative binomial distribution with $p = 0.5$.

$$P(T_3 \leq t) = \sum_{x=0}^{t} \binom{n}{x} (0.5)^x (1-0.5)^{n-x} \tag{11}$$

Table 4 gives the cumulative binomial probabilities from Equation 11 for $n = b + c = 16 + 2 = 18$ (see Table 3) and $p = 0.5$. This table is used to determine rejection regions for either a one- or two-tailed test of the null hypothesis. Table 4 indicates that H_0 should be rejected when $T_3 \geq 13$ for a one-tailed test with $\alpha \approx 0.05$. The exact level of significance associated with this rejection region is found as:

$$\alpha = P(T_3 \geq 13) = 1 - P(T_3 \leq 12) = 1 - 0.95187 = 0.04813$$

Note that this level of significance is slightly lower than the desired level of 0.05. However, this is as close as possible 0.05 for the binomial distribution given in Table 4. To explain further, suppose the rejection region was defined as $T_3 \geq 12$. The corresponding exact level of significance is given as:

$$\alpha = P(T_3 \geq 12) = 1 - P(T_3 \leq 11) = 1 - 0.88106 = 0.11894$$

This level of significance is much larger than the desired value of 0.05 and is not as close to 0.05 as the previous one. Since these two rejection regions have associated levels of significance above and below the desired value 0.05, they are the only two reasonable choices. Of these two choices, the rejection region ($T_3 \geq 13$) with an associated level of significance closest to the desired 0.05 is selected.

Since $T_3 = 16$, H_0 is rejected and in favor of the one-tailed alternative in Equation 4. This indicates that the probability of a correct diagnosis with 3D EIT is significantly higher that the probability of a correct diagnosis for radiology. The p-value is found from Table 4 as follows:

$$\text{p-value} = P(T_3 \geq 16 \mid n=18, p=0.5) = 1 - P(T_3 \leq 15 \mid n=18, p=0.5) = 1 - 0.99934 = 0.00066$$

The smallness of this value provides strong support for the one-tailed alternative hypothesis given in Equation 4:

P(Correct diagnosis for 3D EIT) > P(Correct diagnosis for radiology)

In fact, H_0 would have been rejected for a choice of $\alpha < 0.001$.

Table 4. Cumulative Binomial Distribution for n = 18 and p = 0.5

t	$P(T_3 \leq t)$	
0	0.00000	
1	0.00007	
2	0.00066	
3	0.00377	
4	0.01544	
5	0.04813	
6	0.11894	**Do not**
7	0.24034	**Reject H_0**
8	0.40726	
9	0.59274	
10	0.75966	
11	0.88106	
12	0.95187	
13	0.98456	
14	0.99623	
15	0.99934	**Reject H_0**
16	0.99993	
17	1.00000	
18	1.00000	

The rejection of H_0 for the one-tailed test implies that 3D EIT has a higher level of agreement with either FNA or biopsy than does diagnostic radiology. This decision does not single out either false positives or false negatives, but rather addresses the totality of correct decisions by 3D EIT and diagnostic radiology.

Confidence Intervals for the Percent of Correct Diagnoses

The percent of correct diagnoses for patients in this clinical trial can be obtained from Table 3 for 3D EIT as 30/40 = 75% while diagnostic radiology has a correct percentage of only 16/40 = 40%. Approximate 95% confidence intervals (see Iman, 1994) for these proportions can be found as:

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad (12)$$

where n is the sample size, $\hat{p}$ is the point estimate of the percent of correct diagnoses and $\hat{q} = 1 - \hat{p}$.

Equation 12 gives the approximate 95% confidence interval for the percent of correct diagnoses for 3D EIT as 61.6% to 88.4% for $\hat{p} = 0.75$, $\hat{q} = 0.25$ and $n = 40$. The corresponding approximate 95% confidence interval for diagnostic radiology is 24.8% to 55.2% for $\hat{p} = 0.40$, $\hat{q} = 0.60$ and $n = 40$. These intervals do not overlap as shown in the display in Figure 2, which is consistent with the results of the test of H_0.

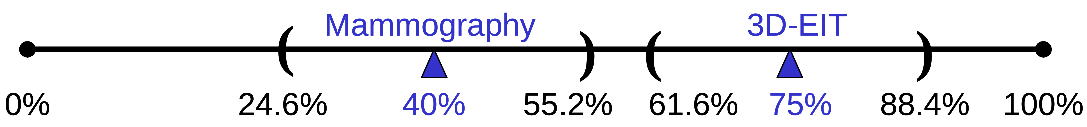


Figure 2. 95% Confidence Intervals for the Percent of Correct Diagnoses for Diagnostic Radiology Examination and 3D EIT

Sensitivity and Specificity

3D EIT diagnostic examination and diagnostic radiology examination are two imaging modalities used to detect and diagnose breast cancer. Thibodeau (1988) provides definitions of sensitivity and specificity that can be adapted to breast cancer diagnoses as follows. If all females with breast cancer have a diagnosis of malignancy, then the modality is *sensitive* to the presence of breast cancer. If all females without breast cancer have a diagnosis of benign, then the modality is *specific* to the absence of breast cancer. Moreover, Thibodeau indicates that if a test is both sensitive and specific, its results are clearly interpretable.

Sensitivity and specificity can be defined as conditional probabilities based on a 2×2 contingency table that compares the results for 3D EIT diagnostic examination with the biopsy results. A second table compares diagnostic radiology examination with the biopsy results. Table 5 gives the summary comparison of 3D EIT diagnostic results with pathologic histology and Table 6 gives the corresponding summary for comparing diagnostic radiology results with pathologic histology.

The conditional probability that defines sensitivity for 3D EIT is obtained directly from Table 5 as follows:

Table 5. Comparison of 3D EIT Diagnostic Results to Pathologic Histology

		Pathologic Histology		
		Malignant	**Benign**	**Totals**
3D EIT Diagnosis	**Malignant**	13	9	22
	Benign	1	17	18
	Totals	14	26	40

Table 6. Comparison of Diagnostic Radiology Results to Pathologic Histology

		Pathologic Histology		
		Malignant	**Benign**	**Totals**
Radiology Diagnosis	**Malignant**	13	23	36
	Benign	1	3	4
	Totals	14	26	40

Sensitivity = Pr(3D EIT diagnosis is malignant | Biopsy is malignant) = 13/14 = 0.926

Likewise, the conditional probability that defines sensitivity for diagnostic radiologic examination is obtained directly from Table 6 as follows:

Sensitivity = Pr(Radiology diagnosis is malignant | Biopsy is malignant) = 13/14 = 0.926

Note that the sensitivity measure is the same for both 3D EIT and diagnostic radiology.

The conditional probability that defines specificity for 3D EIT is obtained directly from Table 5 as follows:

Specificity = Pr(3D EIT diagnosis is benign | Biopsy is benign) = 17/26 = 0.654

Similarly, the conditional probability that defines specificity for diagnostic radiology is obtained directly from Table 6 as follows:

Specificity = Pr(Radiology diagnosis is benign | Biopsy is benign) = 3/26 = 0.115

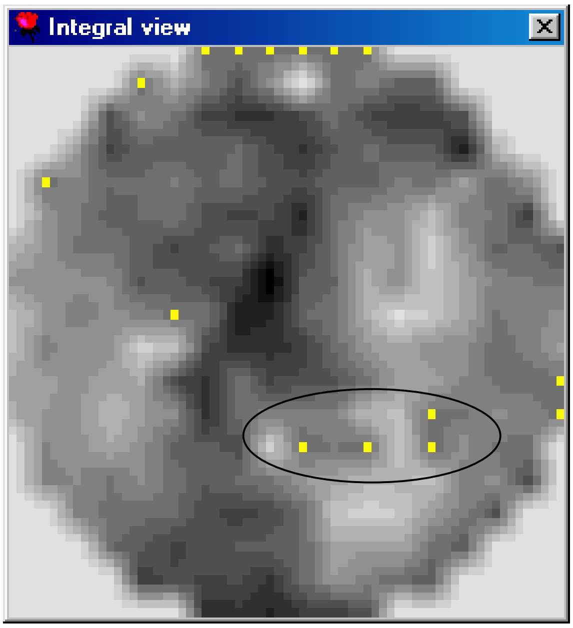
The specificity for diagnostic radiology examination is approximately 1/6 that of 3D EIT. Moreover, based on Thibodeau's previous observation, 3D EIT technology is both sensitive and specific, and its results are clearly interpretable. The same cannot be said of diagnostic radiology.

Note that the sensitivity and specificity results are consistent with the comments in the next section about false positives and false negatives. Moreover, as explained in the next section, the lone false negative for 3D EIT may easily have been avoided at test time had proper electrical contact been made in the area of the lesion during the 3D EIT test. Eliminating this lone false negative would increase the sensitivity for 3D EIT to 14/14 = 1.000 and the specificity would remain unchanged.

False Positives and False Negatives

The results summarized in Table 1 show that 3D EIT diagnostic examination had nine false positives (pathology is benign, but diagnosis is malignant) while diagnostic radiological examination had 23 false positives. In addition, diagnostic radiological examination (Patient Notebook No. 35) and 3D EIT diagnostic examination (Patient Notebook No. 36) each had one false negative (pathologic histology indicates malignancy, but diagnosis is benign). These data indicate that 3D EIT has a significantly lower rate of false positives without an increase in false negatives.

An examination of the data images for Patient Notebook 36 showed that complete electrical contact was not made in the area of the lesion during the 3D EIT diagnostic examination.



Repositioning the 3D EIT to establish complete electrical contact in all likelihood would have resulted in a diagnosis of malignancy and the single false negative would not have occurred.

This would change the value of c in Table 3 from 2 to 1 and the value of n to 16 + 1 = 17. The test statistic $T_3 = 16$ would not change and the p-value obtained from a binomial distribution with $n = 17$ and $p = 0.5$ would be further reduced to 0.00001.

Power and Sample Size Considerations for a One-Tailed Test

Power is the probability that H_0 is rejected given that H_0 is *false*. In the case of a one-tailed test, this is expressed as:

$$\text{Power} = P(\text{Reject } H_0 \mid H_0 \text{ is not true}) = P(\text{Reject } H_0 \mid n, p > 0.5) \qquad (13)$$

Since n = 18 in the analysis of the clinical trial results, Equation 13 would be rewritten as

$$\text{Power} = P(\text{Reject } H_0 \mid n = 18, p > 0.5) = P(T_3 \geq 13 \mid n = 18, p > 0.5) \qquad (14)$$

Equation 14 can be used to find the power for any value of $p > 0.5$. As an illustration, consider the case with $n = 18$ and $p = 0.6$. The cumulative null and alternative distributions are given side-by side in Table 7 to illustrate the power calculation for $p = 0.6$.

Table 7. Cumulative Binomial Distributions for the Null Case with n = 18 and p = 0.5 and the Alternative Case with p = 0.6

	$p = 0.5$		$p = 0.6$
t	$P(T_3 \leq t)$		$P(T_3 \leq t)$
0	0.00000		0.00000
1	0.00007		0.00000
2	0.00066		0.00003
3	0.00377		0.00021
4	0.01544		0.00128
5	0.04813		0.00575
6	0.11894	**Do not**	0.02028
7	0.24034	**Reject H_0**	0.05765
8	0.40726		0.13471
9	0.59274		0.26316
10	0.75966		0.43656
11	0.88106		0.62572
12	0.95187		0.79124
13	0.98456		0.90583
14	0.99623		0.96722
15	0.99934		0.99177
16	0.99993	**Reject H_0**	0.99868
17	1.00000		0.99990
18	1.00000		1.00000

The power is found from Table 7 as:

$$\text{Power} = P(\text{Reject } H_0 \mid n = 18, p=0.6) = P(T_3 \geq 13 \mid n = 18, p=0.6) = 0.20876 \qquad (15)$$

Other power calculations can be made using different values of n and p.

Figure 3 shows power curves for a one-tailed test with $\alpha = 0.05$ and n = 10, 20, 30, 40, and 50. Note that power increases as the value of p becomes larger than the hypothesized value 0.5 for the test statistic T_3. These curves are obviously useful in determining the power for any alternative value of p and a sample size n. They can also be used to determine the sample size needed to detect a deviation of size δ or larger from the hypothesized value of p = 0.5 as is now explained in detail.

Let p^* designate the alternative value of p where $p^* > 0.5$ and let $\delta = p^* - 0.5$. That is, δ is the amount by which the alternative value of p exceeds the hypothesized value of $p = 0.5$. Consider the horizontal dashed line appearing at 0.8 on the vertical axis in Figure 3. This dashed line intersects each of the five power curves in Figure 3. A vertical line can be dropped to the horizontal axis at each of these intersection points to identify p^* for each value of n. For example, a vertical line dropped to the horizontal axis for the power curve corresponding to n = 10 intersects the horizontal axis at approximately 0.84. It follows that $\delta = 0.84 - 0.5 = 0.34$. This means that a sample of size $n = 10$ will detect deviations greater than or equal to 0.34 with a power of 0.8 when $\alpha = 0.05$.

Similar calculations on the remaining power curves in Figure 3 show that $\delta = 0.255$ for $n = 20$; $\delta = 0.218$ for $n = 30$; $\delta = 0.198$ for $n = 40$; and $\delta = 0.185$ for $n = 50$. Clearly, δ decreases as n increases, which simply means that larger sample sizes are required to detect smaller deviations.

A plot of δ versus n is given in Figure 4. This curve can be used to determine the sample size required to detect a difference of at least δ for a one-tailed test having power = 0.8 with $\alpha \approx 0.05$.

Power Curves

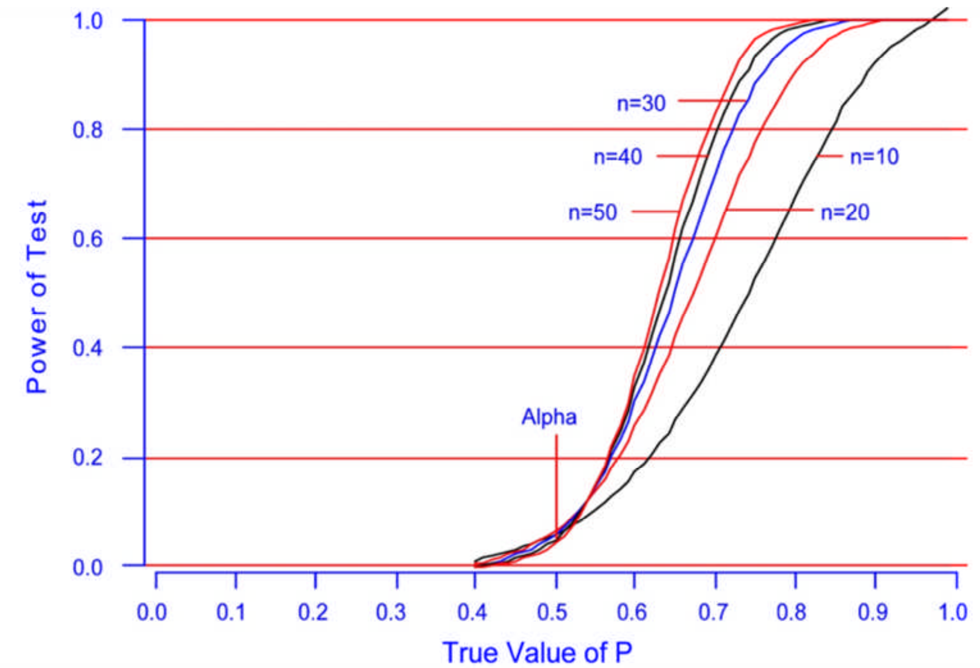


Figure 3. Power Curves for a One-Tailed Test of the Null Hypothesis

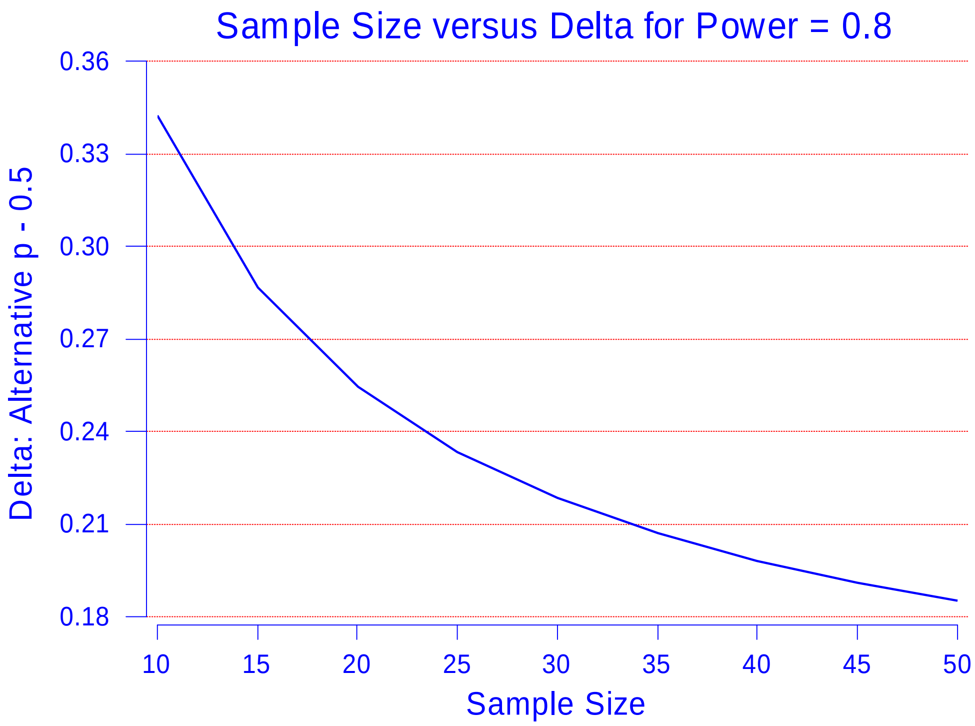


Figure 4. Sample Size Requirements for a Given Delta to be Detected with Power = 0.8 with a One-Tailed Test

References

1. Conover, W. J. (1999). **Practical Nonparametric Statistics**, 3ed, Wiley, NY.
2. Iman, R. L. (1994). **A Data-Based Approach to Statistics**, Duxbury, Belmont, CA.
3. Thibodeau, L. A. (1998). Sensitivity and Specificity, **Encyclopedia of Statistical Sciences**, Vol 8, 370-372, Wiley, NY.

ATTACHMENT A

Bayesian Analysis

Statistical analyses can be roughly thought of as either classical or Bayesian. Classical testing procedures can be characterized as follows:

1. They rely on underlying distributional assumptions
2. The test statistic is referenced to some distribution to determine a p-value
3. Results are frequently expressed as a confidence interval
4. They have no way of incorporating information on the known performance of a particular methodology

Nonparametric tests such as the Cochran and McNemar tests used in the preceding analyses ease some of the restrictions about distributional assumptions while still providing reliable results. However, they do not address items 3 and 4. On the other hand, Bayesian analyses are probabilistically based rather than relying on confidence intervals, which lack a probabilistic interpretation.

With regard to item 4, the preceding analyses based on the Cochran and McNemar tests were based entirely on the results of the clinical trials. Any prior information regarding the known performance of a particular modality is ignored. This information is valuable and can easily be incorporated into a Bayesian analysis, but classical test procedures have no way to utilize such information.

Bayesian analyses are based on Bayes' Theorem, which is now explained.

Bayes' Theorem

Bayes' theorem can be formally stated as follows:

$$p(\theta \mid x) = \frac{p(\theta)\, s(x \mid \theta)}{\int p(\theta)\, s(x \mid \theta) d\theta} \qquad (1A)$$

where the region of integration is taken over all possible values of θ (the success rate of breast cancer diagnoses). Equation 1A uses three probability distributions, which are defined as follows:

Probability Distribution	Common Name	How Determined
p(θ)	Prior distribution	Frequently based on expert knowledge (subjective)
s(x | θ)	Conditional probability model	Objective information (i.e. clinical trial results)
p(θ | x)	Posterior distribution	A combination of subjective and objective information

Expressed in words, Bayes' theorem produces a posterior distribution for θ by reweighting the prior distribution (or subjective information) with the objective information (or current clinical trial results). The steps needed to combine the prior information with results of the clinical trials in a Bayesian analysis are now explained.

Step 1. Determining the Prior Distribution p(θ)

Bayes' theorem takes advantage of expert knowledge about success rates for procedures used in the diagnosis of breast cancer. For example, suppose experts on mammography agree that it provides a correct diagnosis between 15% and 55% of the time. The data in Table 1 are consistent with rates in this range as they show 23 false positives and one false negative out of 40 patients, which means the diagnoses were correct 16/40 = 40% of the time. The first step in a Bayesian analysis is to express this expert information in the form of a *prior* distribution that can be used in Bayes' Theorem. The Bayesian approach is not very sensitive to these assumptions.

The range of 15% to 55% is modeled very closely with a beta distribution having parameters p = 10 and q = 15. A graph of the cumulative distribution function for this beta prior appears in Figure 1A. The density function for the beta prior with parameters p and q is as follows:

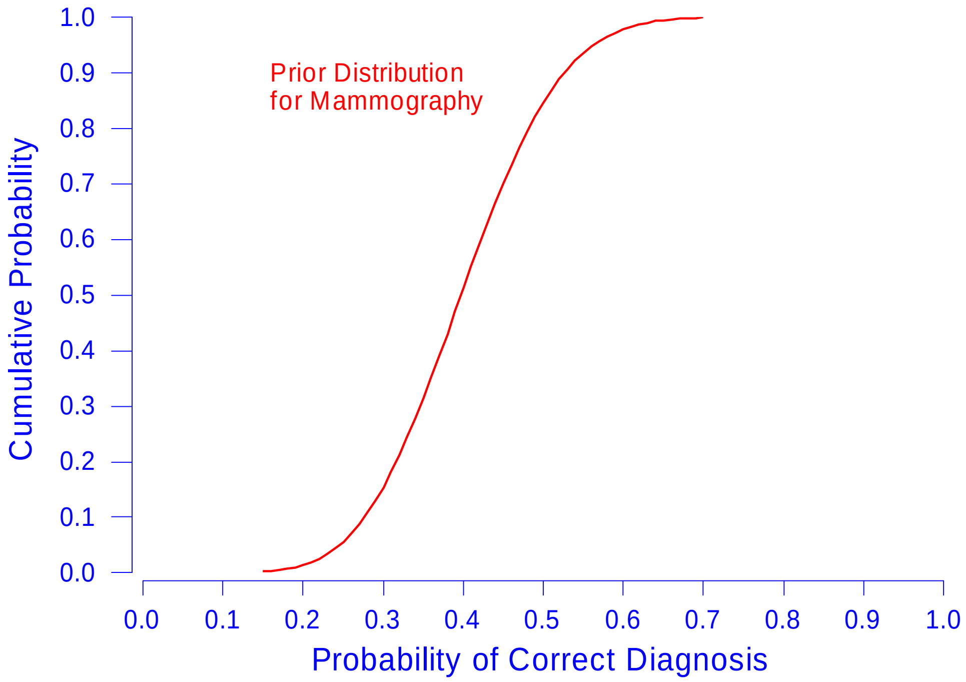


Figure 1A. Cumulative Distribution Function for the Beta Prior for Diagnostic Radiology, which has Parameters p = 10 and q = 15

$$p(\theta \mid p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)}\theta^{p-1}(1-\theta)^{q-1} \qquad 0 \le \theta \le 1 \quad p,q > 0 \tag{2A}$$

The expected value and variance of the beta distribution are:

$$E(\theta) = \frac{p}{p+q} \tag{3A}$$

and

$$V(\theta) = \frac{pq}{(p+q+1)(p+q)^2} \tag{4A}$$

Step 2. The Conditional Probability Model s(x | θ)

The clinical trial data consist of *x* successes in *n* independent trials with a fixed rate of success, θ. These assumptions correspond to those for a binomial random variable whose probability distribution is given as follows:

$$s(x \mid \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \qquad x = 0, 1, 2, \ldots, n \qquad 0 \le \theta \le 1 \tag{5A}$$

Step 3. The Posterior Distribution p(θ | *x*)

The posterior distribution p(θ | *x*) is obtained by substituting Equations 2A and 5A into Equation 1A. After some simplification, the density of the posterior distribution can be expressed as follows:

$$p(\theta | p, q, n, x) = \frac{\Gamma(p+q+n)}{\Gamma(p+x)\Gamma(q+n-x)} \theta^{p+x-1} (1-\theta)^{q+n-x-1} \qquad 0 \le \theta \le 1 \tag{6A}$$

$$p, q > 0$$

$$x = 0, 1, 2, \ldots, n$$

This is also a beta density and has parameters *p* + *x* and *q* + *n* - *x*. Thus, combining a beta prior with a binomial likelihood function produces another beta distribution. This is known as the reproductive property of the beta-binomial. The mean and variance for the posterior distribution are found from Equations 3A and 4A as follows:

$$E(\theta) = \frac{p+x}{(p+x)+(q+n-x)} = \frac{p+x}{p+q+n} \tag{7A}$$

and

$$V(\theta) = \frac{(p+x)(q+n-x)}{(p+q+n+1)(p+q+n)^2} \tag{8A}$$

Step 4. Application to Clinical Trial Results

The parameters of the prior distribution were given in Step 1 as $p = 10$ and $q = 15$. The expected value and variance for the prior distribution for diagnostic radiology are found from Equations 3A and 4A as 0.40 and 0.0092.

The clinical trials show that diagnostic radiology had $x = 16$ successes in $n = 40$ independent trials. From Equation 6A, the beta posterior distribution for diagnostic radiology has parameters $p + x = 10 + 16 = 26$ and $q + n – x = 15 + 40 - 16 = 39$. The expected value and variance for the posterior distribution for diagnostic radiology are found from Equations 7A and 8A as 0.40 and 0.0036. This agrees exactly with the observed mean 16/40 = 0.40 for diagnostic radiology. The variance is reduced by a factor of 3 from 0.0092 to 0.0036. Thus, the incorporation of the diagnostic radiology clinical trial data through Bayes' theorem has very little impact on the prior distribution. This is an indication that the prior distribution for θ is reasonable for diagnostic radiology.

The Bayesian analysis for 3D EIT diagnostic results uses the same prior as diagnostic radiology results since the null hypothesis in Equation 2 states that 3D EIT and diagnostic radiology have the same probability of making a correct decision. The clinical trials for 3D EIT showed $x = 30$ successes in $n = 40$ independent trials. From Equation 6A, the beta posterior distribution for 3D EIT has parameters $p + x = 10 + 30 = 40$ and $q + n – x = 15 + 40 - 30 = 25$. The expected value and variance for the posterior distribution for 3D EIT are found from Equations 7A and 8A as 0.615 and 0.0036. This compares to an observed mean of 30/40 = 0.75 for 3D EIT. Thus, the incorporation of the clinical trial data for 3D EIT through Bayes' theorem has weighted the mean of the prior distribution (0.40) *upward* toward the observed mean for 3D EIT.

An easy way to grasp the results of the Bayesian analyses is to compare the distribution function of the prior distribution given in Figure 1A to the posterior distribution functions for diagnostic radiology and 3D EIT. These graphs appear in Figure 2A. This figure shows that the posterior distribution for diagnostic radiology is quite close to its prior with the posterior having less spread (i.e. smaller variance). On the other hand, the posterior distribution for 3D EIT has shifted well to the right of the prior distribution, which is indicative of 3D EIT having a higher percentage of correct diagnoses.

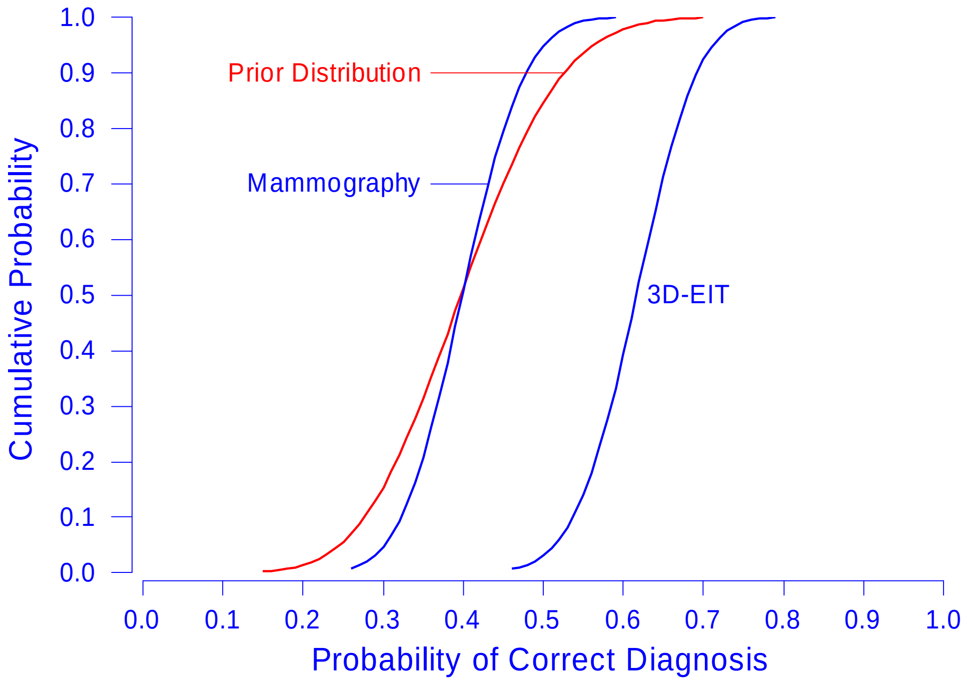


Figure 2A. Percent of Correct Diagnosis for Diagnostic Radiology and 3D EIT based on a Bayesian Analyses of Clinical Trial Data Using a Beta Prior with Parameters p = 10 and q = 15

Selected quantiles from the distribution functions in Figure 2A are summarized in Table 1A. Note that the median (0.50 quantile) for diagnostic radiology is 0.399 in Table 1A while the median for 3D EIT is 0.617. Moreover, the 0.95 quantile for diagnostic radiology (0.501) is less than the 0.05 quantile for 3D EIT. In other words, the probability that diagnostic radiology has a correct diagnosis *more than half the time* (0.501) is 0.05. On the other hand, the probability that 3D EIT has a correct diagnosis *more than half the time* is greater than 0.95.

Table 1A. Estimated Quantiles from the Bayesian Analyses

	Probability of Correct Diagnosis		
Quantile	**Prior**	**Diagnostic Radiology**	**3D EIT**
0.01	0.194	0.266	0.472
0.05	0.246	0.303	0.515
0.10	0.277	0.323	0.538
0.25	0.332	0.358	0.575
0.50	0.397	0.399	0.617
0.75	0.465	0.440	0.657
0.90	0.526	0.478	0.692
0.95	0.563	0.501	0.712
0.99	0.630	0.544	0.748

Robustness of the Prior Distribution

As mentioned earlier, the Bayesian approach is reasonably robust with respect to the choice of the prior distribution. For example, a non-informative prior that assumes the percent of correct diagnoses is uniformly distributed between 0 and 1 (a very gross assumption) can be described as a beta density with $p = 1$ and $q = 1$. The mean of this non-informative prior is 0.5. With this prior, the posterior mean for diagnostic radiology is 0.4048 compared to 0.40 for the previous analysis. Likewise, the posterior mean for 3D EIT is 0.738, which is close to the observed value 0.75.