510(k) SUBSTANTIAL EQUIVALENCE DETERMINATION
DECISION SUMMARY
ASSAY ONLY

A. 510(k) Number:
K093957

B. Purpose for Submission:
New device

C. Measurand:
Human epididymis protein 4 (HE4)

D. Type of Test:
Quantitative, Automated chemiluminescence immunoassay on the ARCHITECT *i* Systems

E. Applicant:
Fujirebio Diagnostics, Inc.

F. Proprietary and Established Names:
ARCHITECT HE4 assay
ARCHITECT HE4 Calibrator kit
ARCHITECT HE4 Control kit

G. Regulatory Information:

1. Regulation section:
   21 CFR § 866.6010, Tumor-associated Antigen Immunological Test System
   21 CFR § 862.1150, Calibrator
   21 CFR § 862.1660, Quality Control Material (Assayed and Unassayed)
2. Classification:
   Class II ARCHITECT HE4 assay
   Class II calibrator
   Class I Quality control material
3. Product code:
   OIU - epithelial ovarian tumor associated antigen (HE4) Test
   JIT - Calibrator, Secondary
   JJX – Single (Specified) Analyte Controls (Assayed and Unassayed)
4. Panel:
   Immunology 82 (HE4)
   Chemistry 75, (Calibrator and Quality Control)

H. Intended Use:

1. Intended use(s):
   The ARCHITECT HE4 assay is a chemiluminescent microparticle immunoassay (CMIA) for the quantitative determination of HE4 antigen in human serum. The assay is to be used as an aid in monitoring recurrence or progressive disease in patients with epithelial ovarian cancer. Serial testing for patient HE4 assay values should be used in conjunction with other clinical methods used for monitoring ovarian cancer.

   ARCHITECT HE4 Calibrator Kit Intended Use
   The ARCHITECT HE4 Calibrators are for the calibration of the ARCHITECT i System when used for the quantitative determination of HE4 antigen in human serum.

ARCHITECT HE4 Control Kit Intended Use
The ARCHITECT HE4 Controls are used for the verification of the accuracy and precision of the ARCHITECT i System when used for the quantitative determination of HE4 antigen in human serum.

2. Indication(s) for use:
   Same as Intended Use
3. Special conditions for use statement(s):
   Prescription Use only
4. Special instrument requirements:
   Abbott ARCHITECT *i* System analyzers

I. Device Description:

The ARCHITECT HE4 assay is a two-step immunoassay for the quantitative determination of HE4 antigen in human serum using Chemiluminescent Microparticle Immunoassay (CMIA) technology. The kit is composed of the following:

1. ARCHITECT HE4 Reagent Kit (2P54)
   MICROPARTICLES 1 Bottle (6.6 mL) Anti-HE4 (mouse, monoclonal) coated microparticles in Phosphate Buffered Saline (PBS) buffer with protein (bovine) stabilizers and detergent. Minimum concentration: 0.1% solids. Preservative: ProClin 300.
   CONJUGATE 1 Bottle (5.9 mL) Anti-HE4 (mouse, monoclonal) acridinium-labeled conjugate in PBS buffer with protein (bovine) stabilizers and detergent. Minimum concentration: 50 ng/mL. Preservative: ProClin 300.
2. ARCHITECT HE4 Calibrator Kit (2P54-01)
   CALIBRATORS 6 Bottles (4.0 mL each) of ARCHITECT HE4 Calibrators. Calibrators A through F are prepared in PBS buffer with a protein (bovine) stabilizer. Preservative: ProClin 30.
   The calibrators are at the following concentrations:

Calibrator	Concentration (pmol/L)
CAL A	0
CAL B	30
CAL C	100
CAL D	250
CAL E	750
CAL F	1500

3. ARCHITECT HE4 Control Kit (2P54-10)
   CONTROLS 3 Bottles (8.0 mL each) of ARCHITECT HE4 Controls. The Low, Medium, and High Controls are prepared in PBS buffer with a protein (bovine) stabilizer. Preservative: ProClin 30.
   The calibrators are at the following concentrations:

Control	Concentration Target (pmol/L)	Concentration Range (pmol/L)
Low Control	50	35.0 – 65.0
Medium Control	175	122.5 – 227.5
High Control	700	490.0 – 910.0

4. Other Reagents:
   MULTI-ASSAY MANUAL DILUENT 1 Bottle (100 mL) ARCHITECT *i* Multi-Assay Manual Diluent containing phosphate buffered saline solution. Preservative: antimicrobial agent.
   PRE-TRIGGER SOLUTION Pre-Trigger Solution containing 1.32 % (w/v) hydrogen peroxide.
   TRIGGER SOLUTION Trigger Solution containing 0.35N sodium hydroxide.
   WASH BUFFER Wash Buffer containing phosphate buffered saline solution. Preservative: Antimicrobial Agents.

J. Substantial Equivalence Information:

1. Predicate device name(s):
   Fujirebio HE4 EIA Kit
2. Predicate 510(k) number(s):
   k072939
3. Comparison with predicate:

Similarities		
Item	Device	Predicate
	ARCHITECT HE4	Fujirebio HE4 EIA
Intended Use	Quantitative determination of HE4 antigen in human serum. The assay is to be used as an aid in monitoring recurrence or progressive disease in patients with epithelial ovarian cancer. Serial testing for patient HE4 assay values should be used in conjunction with other clinical methods used for monitoring ovarian cancer.	Same
Type of Specimen	Human Serum	Same
Antigen Detected	HE4	Same
Capture Antibody	Mouse monoclonal (2H5)	Same
Detection Antibody	Mouse monoclonal (3D8)	Same

Differences		
Item	Device	Predicate
Instrument System	ARCHITECT *i* System	Manual
Principle of Operation	Chemiluminscent Microparticle Immunoassay (CMIA)	Manual Enzymatic Immunoassay(EIA)
Reportable assay range	20 – 1500 pmol/L	15 – 900 pmol/L
Calibrators	6 Levels (0 – 1500 pmol/L)	6 Levels (0 – 900 pM)
Controls	3 Levels (50, 175 and 700 pmol/L) Supplied as separate kit	2 Levels (50 and 400pM) Supplied with Kit
Interpretation of Results	Calibrator Curve	Standard Curve

K. Standard/Guidance Document Referenced (if applicable):

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline - Second Edition (C28-A2)
Interference Testing in Clinical Chemistry; Approved Guideline - Second Edition (EP7-A2)
Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline-Second Edition (EP5-A2)

Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline (EP9-A2)
Stability Testing of In Vitro Diagnostic Reagents (EN13640)
Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline (EP6-A)
Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline (EP17-A)
Guidance Document for the Submission of Tumor Associated Antigen Premarket Notification [510(k)] to FDA
Guidance on Informed Consent for In Vitro Diagnostic Device Studies Using Leftover Human Specimens that are Not Individually Identifiable - Guidance for Sponsors, Institutional Review Boards, Clinical Investigators and FDA Staff

L. Test Principle:

The ARCHITECT HE4 assay is a two-step immunoassay for the quantitative determination of HE4 antigen in human serum using chemiluminescent microparticle immunoassay (CMIA) technology with flexible assay protocols, referred to as Chemiflex. In the first step, sample and anti-HE4 coated paramagnetic microparticles are combined. HE4 antigen present in the sample binds to the anti-HE4 coated microparticles. After washing, acridinium-labeled anti-HE4 conjugate is added. Following another wash cycle, pre-trigger and trigger solutions are added to the reaction mixture. The resulting chemiluminescent reaction is measured as relative light units (RLUs). A direct relationship exists between the amount of HE4 antigen in the sample and the RLUs detected by the ARCHITECT *i* System optics.
The ARCHITECT HE4 Calibrators are for the calibration of the ARCHITECT *i* System when used for the quantitative determination of HE4 antigen in human serum. The ARCHITECT HE4 Controls are used for the verification of the accuracy and precision of the ARCHITECT *i* System when used for the quantitative determination of HE4 antigen in human serum.

M. Performance Characteristics (if/when applicable):

1. Analytical performance:
   a. Precision/Reproducibility:

   The precision of the ARCHITECT HE4 assay was assessed at three (3) different sites. These studies were modeled after CLSI guideline EP5-A2 ("Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline – Second Edition, 2004"). At each site, 2 lots of reagent kit, calibrator kit, and control kit were utilized. At one site and 1 lot, the ARCHITECT *i*2000 instrument system was used while the other sites used the ARCHITECT *i*2000$_{SR}$ instrument system. Only the ARCHITECT *i*2000$_{SR}$ System data will be used to communicate performance claims in the proposed ARCHITECT HE4 package insert.
   Six samples were tested in 2 replicates using 2 lots of reagents, at 2 separate runs per day, for 20 days (not necessarily consecutive) on 1 ARCHITECT *i*2000$_{SR}$ System or 1 ARCHITECT *i*2000 System. A single calibration curve was used throughout the study. The kit controls were tested and evaluated for each run to determine assay validity. The 6 samples were the following:
   Panel 1: pooled human serum - HE4 concentration is approximately 40 pmol/L

Panel 2: pooled human serum spiked with native HE4 antigen - HE4 concentration is approximately 200 pmol/L
Panel 3: pooled human serum spiked with recombinant HE4 Antigen (kit Calibrator/Control antigen) - HE4 concentration is approximately 1000 pmol/L
Low Control: recombinant HE4 Antigen spiked into PBS buffer with a protein (bovine) stabilizer - HE4 concentration is approximately 50 pmol/L
Medium Control: recombinant HE4 Antigen spiked into PBS buffer with a protein (bovine) stabilizer - HE4 concentration is approximately 175 pmol/L
High Control: recombinant HE4 Antigen spiked into PBS buffer with a protein (bovine) stabilizer - HE4 concentration is approximately 700 pmol/L
Human pool 1 was a specimen containing HE4 concentration levels that could be deemed as natural. However, it is difficult to obtain samples at higher HE4 concentration levels due to the infrequent prevalence of ovarian cancer in young healthy women. So the use of spiked samples is sufficient.
Within-run, between-run, between-day, and total imprecision, expressed as standard deviation and %coefficient of variation (%CV) for both lots and for each lot separately, was calculated. The following table summarizes data for both combined lots:

Sample	Mean	Between lot		Between day		Between Run Within-day		Within Run		Total		Upper 95% Conf. Limit	
		SD	% CV	SD	% CV	SD	%CV	SD	%CV	SD	%CV	SD	%CV
Control L	49.5	0.203	0.4%	0.675	1.4%	0.543	1.1%	1.4	2.8%	1.658	3.4%	1.862	3.8%
Control M	171.3	4.345	2.5%	2.148	1.3%	3.897	2.3%	3.589	2.1%	7.181	4.2%	8.064	4.7%
Control H	667.8	27.377	4.1%	12.881	1.9%	10.9168	1.6%	16.519	2.5%	36.159	5.4%	40.607	6.1%
Panel 1	38.7	0.415	1.1%	0.494	1.3%	0.747	1.9%	1.144	3.0%	1.511	3.9%	1.697	4.4%
Panel 2	186.3	4.798	2.6%	1.377	0.7%	3.089	1.7%	6.104	3.3%	8.469	4.5%	9.511	5.1%
Panel 3	1097.5	23.522	2.1%	21.386	1.9%	11.114	1.0%	33.582	3.1%	47.559	4.3%	53.409	4.9%

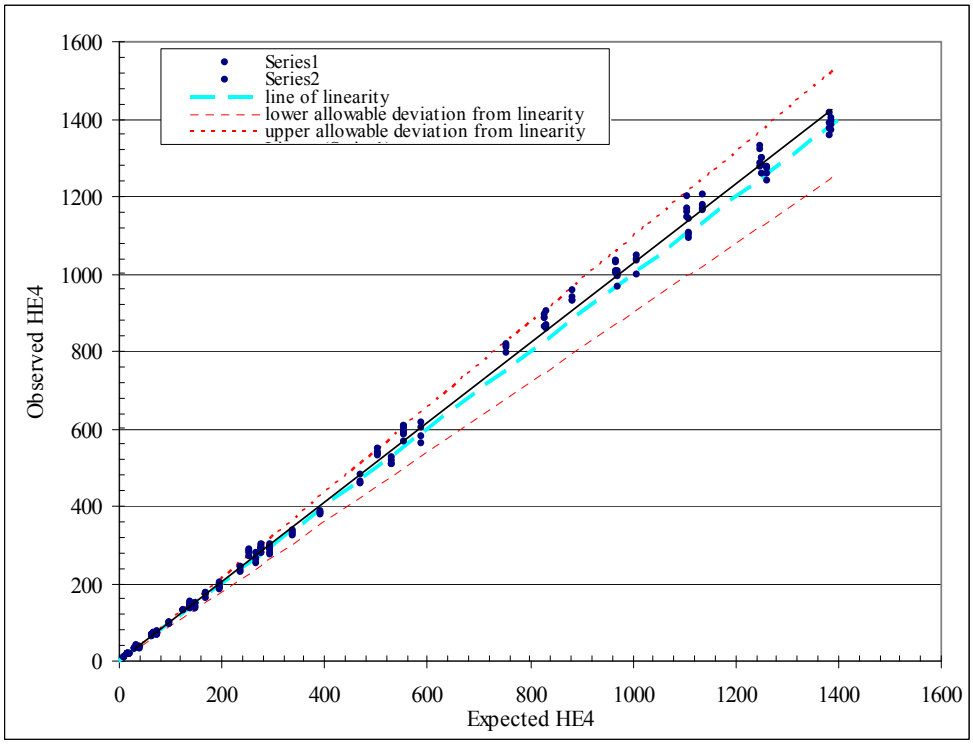
The upper 95% confidence limit for total imprecision for all samples is less than or equal to 6.1%CV. The ARCHITECT HE4 assay meets the predetermined acceptance criteria of ≤ 10% total CV.

b. Linearity/assay reportable range:

The assay measures concentrations between 20 and 1500 pM. To assess the linearity of the assay in the reportable range, CLSI guidance EP6-A, entitled "Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline", was used to design the linearity experiment. Four concentration levels ranging from 200 to 1500 pmol/L were prepared from a panel of ovarian cancer patients prior to dilution. A fifth panel was prepared from normal female serum supplemented with native HE4 (derived from cell culture). A series of 10 dilutions was prepared from these 5 starting pools. The un-diluted and diluted series were tested in 4 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ System. A single calibration curve was used throughout the study. The assay kit controls were tested and evaluated for each run to determine assay validity. The expected concentration value at each dilution

was calculated for each sample using the dilution factor and the initial concentration of the undiluted sample pool. The mean observed HE4 concentration of 4 replicates at each dilution was calculated from the assay results. A polynomial regression analysis of observed and expected HE4 concentration was performed for first-, second-, and third-order polynomials using a weighted regression method. If the non-linear coefficients of the second or third order polynomial are statistically significant then the deviation from linearity was investigated. For 2 of the 5 specimen pools, significant non-linearity was observed. Deviation from linearity was investigated. The assay met the bias allowance for deviation from linearity defined as ± 10% for concentration levels above 20 pmol/L and ± 2 pmol/L for samples at or below 20 pmol/L. The assay meets the requirements of the CLSI Guideline EP6A for linearity from 20 to 1500 pmol/L.

Analysis of the sponsor's data was performed without using weighted regression analysis. For all specimen pools with expected HE4 concentrations ranging from 20 pmol/L to 1400 pmol/L, the second and third order polynomials were significant. The deviation from linearity, expressed as a percentage deviation of the non-linear model, for all samples ranged from -2.9% to 7.3% deviation from linearity. The mean percentage deviation from linearity compared with the expected HE4 concentration was -0.043%. The deviation from linearity met the sponsor's specification of ± 10% for HE4 concentrations above 20 pmol/L. The following graph illustrates this analysis.



The individual replicates for the dilution series of the 5 specimen pools is indicated in the graph. The line of linearity and the observed best fit line of a linear model are also shown. The upper and lower allowable deviation from linearity are plotted based on the sponsor's specification for deviation from

linearity. The graph shows that the individual datapoints of the dilution series of all 5 specimen pools, though not statistically linear, are sufficiently close to the line of linearity and the sponsor's specification for deviation from linearity. The data is acceptable and indicates sufficient linearity in the assay range of 20 to 1500 pmol/L.

High Sample Dilution Study

To evaluate the recommended dilution procedure of 1:10 for out of range samples (1500 – 15,000 pmol/L) as described in the package insert for the assay, recombinant HE4 antigen was added to 6 serum samples to create out-of-range conditions. The final HE4 concentrations before automatic dilution were approximately 2100, 4300, 6500, 8600, 10,800, and 14,000 pmol/L. The expected value for each sample was determined from the concentration and volume of HE4 antigen added and the endogenous concentration and volume of serum sample. Samples were tested undiluted and diluted using the auto-dilution function (1:10 dilution) on the ARCHITECT *i*2000$_{SR}$ instrument in 2 replicates. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. The mean concentration value was calculated for each sample. The percent recovery of observed HE4 concentration relative to the expected concentration was calculated. The mean recoveries for the six (6) individual samples ranged from 91 to 106%. The mean percent recovery for samples was 99%. The assay meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples.

There is a slight proportional bias in observed HE4 and Expected HE4 for the samples tested. %recovery increased with increasing expected HE4 concentrations, though the mean %recovery is within the acceptance criteria. From the data it is possible to calculate the %bias between observed and expected HE4 concentration. The %recoveries and calculated %bias are as follows:

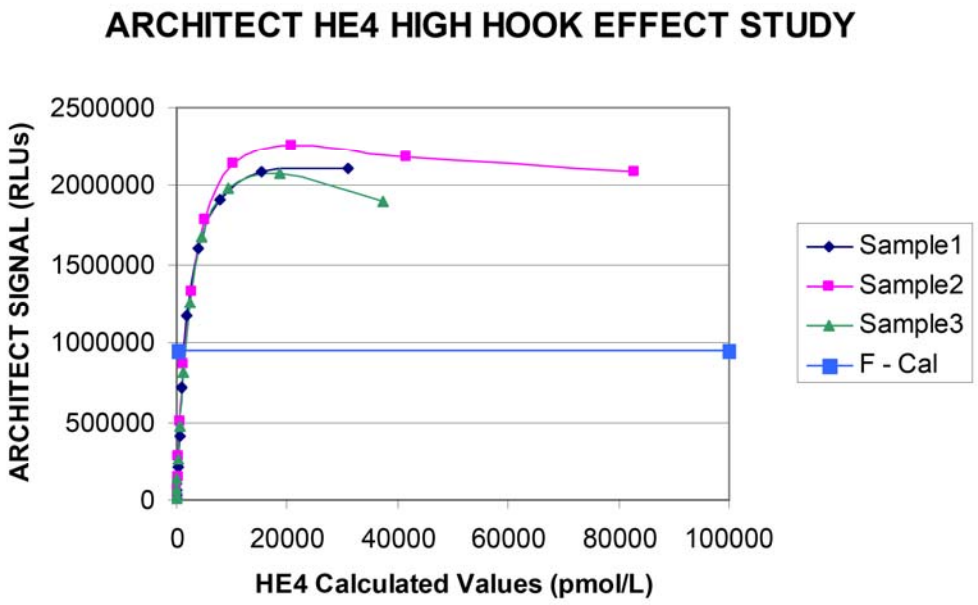
Sample	Expected Value (pmol/L)	Measured Value (n=2) (pmol/L)	Measured Value x 10 (Dil Factor)	Percent Recovery (%)	% bias Obs – Exp
1	2211.6	201.6	2015.5	91	-8.9%
2	4345.6	421.3	4213	97	-3.1%
3	6550.2	644.9	6449	98	-1.5%
4	8705.5	882	8820	101	1.3%
5	10851.6	1147.2	11472	106	5.7%
6	13953.9	1432.6	14326	103	2.7%

High Dose hook effect

To determine the effect of extreme levels of analyte in the assay, one pooled serum sample (un-spiked) and two normal sera were supplemented with recombinant HE4 antigen. These samples were then serially diluted in Wash Buffer such that one or more of the dilutions fell within the reportable range of the assay but above the B calibrator (30 pmol/L). All samples were run in 2

replicates using the assay on the ARCHITECT *i*2000$_{SR}$ System. A single calibration curve was used throughout the study. The kit controls were tested and evaluated for each run to determine assay validity. The instrument response values (i.e., Relative Light Units) of the neat and diluted samples were plotted against their respective calculated HE4 value. For each sample and its corresponding dilutions, the calculated HE4 value was determined by multiplying the mean value of the samples by their respective dilution factor when the HE4 concentration was within the range of the assay. The values of sample dilutions within the standard curve but above the B calibrator were used to calculate the neat sample concentration.

The results are graphically summarized by the sponsor as follows:



The data in the figure shows that increasing levels of analyte increases the response of the assay to a plateau level of approximately 2,000,000 RLU's. The RLU values of the samples did not diminish with increasing antigen concentration as would be expected with high dose hook (prozone) effect but maintained a response level in excess of the RLU value of the F Calibrator.

c. *Traceability, Stability, Expected values (controls, calibrators, or methods):*

Kit Calibrators

The kit contains 6 bottles (4.0 mL each) of ARCHITECT HE4 Calibrators. Calibrators A through F are prepared in PBS buffer with a protein (bovine) stabilizer. Preservative: ProClin 300. The calibrators are at the following concentrations:

Calibrator	Concentration (pmol/L)
CAL A	0
CAL B	30

CAL C	100
CAL D	250
CAL E	750
CAL F	1500

There is currently no known internationally recognized consensus reference method or reference material for standardization. Calibrator value is related to a sponsor maintained reference preparation. The calibrators are manufactured volumetrically and are referenced to this standard. HE4 assay values are expressed as pmol/L.

The material used to standardize the assay is recombinant Ig-HE4 fusion protein consisting of a Human Fc antibody fragment and Human Epididymis protein (HE4). The material was produced in a stably transfected Chinese Hamster Ovary (CHO) cell line. The cell line was grown in serum free medium. The cell culture supernatant was purified by Protein A affinity chromatography. The concentration of the resulting material was determined by MALDI-TOF Mass Spectrometry This material is the primary source material from which secondary and bulk kit calibrators are prepared. Bulk testing is performed by measuring the Relative Light Unit (RLU) of each calibrator as compared to the RLU of the reference calibrators when tested on the ARCHITECT System. When the bulk calibrators are within specifications, they are aseptically filtered and filled into their final containers. Quality Control testing of the filled calibrators as bottled components is performed followed by finished goods testing of the packed Calibrator Kit using packaged Reagent Kits. The calibrator curve and controls must meet predetermined acceptance criteria.

Kit controls

The control kit contains 3 Bottles (8.0 mL each) of ARCHITECT HE4 Controls. The Low, Medium, and High Controls are prepared in PBS buffer with a protein (bovine) stabilizer. Preservative: ProClin 300. The controls are at the following concentrations:

Control	Target Concentration (pmol/L)	Range
Low	50	35-65
Medium	175	122.5 – 227.5
High	700	490 – 910

The Controls are prepared by adding HE4 Antigen to a diluent to achieve the desired concentrations. After passing Quality Control testing the bulk material is then used to prepare the Low, Medium and High Controls. Each control is prepared by diluting the bulk material with the diluent. After mixing and equilibration, the controls are tested by measuring the RLU of each control as compared to the RLU of the reference controls. Adjustments are made if necessary. When the bulk controls are within specifications, they are aseptically filtered and filled into their final containers. Quality Control testing of the filled

controls as bottled components is performed followed by finished goods testing of the packed Control Kit using packaged Reagent Kits, packaged Calibrator Kits. Controls must meet predetermined acceptance criteria.

ARCHITECT HE4 Calibrator and Control Stability Studies

A study was conducted to evaluate the stability of the ARCHITECT HE4 Calibrator and Control Kits when all bottled components are manually opened and closed. The study was performed to support the package insert claim instructing the end user to tightly close the caps on the bottles after each use and return to 2-8°C storage. The stability studies for the ARCHITECT HE4 Calibrator and Control material are performed in compliance to the CEN standard, 13640 – Stability Testing of In Vitro Diagnostic Reagents. Three (3) lots of kit Calibrators and Controls were set up for the study. Sufficient kit Calibrators and Controls are stored at 2-8°C and used to evaluate three (3) panels, described below, at monthly time points. All of the bottles were opened and closed at time 0. One calibrator lot is stored inverted. The panels (Panel 1, Panel 2, and Panel 3) consist of 3 HE4 concentrations spanning the calibration range of the Assay. The Panels are pooled serum samples containing either recombinant or native HE4 antigen. The Panels are stored at ≤ -60°C. At time zero and each subsequent time point the following run was or will be performed:

Run 1 – A calibration curve using calibrators in duplicate with 1 replicate of HE4 controls and 2 replicates of samples from panels.
Run 2 – HE4 controls in single replicates and panels in 2 replicates.

At each time point, all runs are required to meet predetermined acceptance criteria for the calibration curve, control recovery, and panel recovery. Intended storage data currently supports 9 months dating for kit Calibrators and Controls. This study is ongoing and will be supplemented with additional timepoints.

Sample stability - fresh serum samples

A study was conducted to determine the stability of HE4 antigen in fresh serum samples stored at 2-8°C. Serum samples from 10 healthy volunteers were collected. The samples were stored continuously at 2-8°C. Each sample was tested in 2 replicates using the ARCHITECT HE4 assay on the ARCHITECT *i*2000$_{SR}$ System and tested at Day 0 (within 8 hours of draw), Day 2, Day 3, Day 4, and Day 5. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 2 replicates of each sample tested at Day 2, Day 3, Day 4, and Day 5 was compared to the corresponding mean concentration at Day 0 (control).

The mean of daily percent recoveries across the 10 samples was within 100 ± 10% for the study period. As a result, the ARCHITECT HE4 assay meets the predetermined acceptance criteria of 100 ± 10% of Day 0 value for the daily mean and supports the package insert sample storage claim of up to 4 days at 2-8°C.

Sample Stability – Frozen samples

A study was conducted to determine the stability of HE4 antigen in frozen serum samples. Serum samples from 11 healthy volunteers were collected. Each sample was aliquoted into a minimum of 4 tubes and frozen at ≤ -10°C. One (1) fresh/unfrozen aliquot was retained as the control (Day 0). The Day 0 control was tested 3 in replicates using the ARCHITECT HE4 assay on 1 ARCHITECT *i*2000$_{SR}$ System within 8 hours of draw. The remaining aliquots were stored continuously at ≤ -10°C and tested in 3 replicates using the ARCHITECT HE4 assay on the same ARCHITECT *i*2000$_{SR}$ System at Day 4 and Day 5. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 3 replicates of each sample tested at Day 4 and Day 5 was compared to the corresponding mean concentration at Day 0 (control).

All samples recovered within 100 +/-15% of Day 0 on Day 4, 5, 6, and 7. Testing supports the package insert recommendations.

Sample Stability - Room Temperature Sample

A study was conducted to characterize the stability of HE4 antigen in fresh serum samples at room temperature and to support the ARCHITECT HE4 package insert claim for room temperature storage. Serum samples from 10 healthy volunteers were collected. Each sample was tested in 2 replicates using the ARCHITECT HE4 assay on the ARCHITECT *i*2000$_{SR}$ System. The samples were stored continuously at 25 ± 2°C and run at Time 0 (within 8 hours of draw), 24 ± 2 hours and 48 ± 2 hours. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 2 replicates of each sample tested at Time 0 (control) was compared to the corresponding mean concentration at 24 hours and 48 hours.

The range of percent recoveries for each individual sample was between 97% - 113% for 24 hours and 97% - 114% for 48 hours. The percent recovery for the mean of all samples was 105% for 24 hours and 106% for 48 hours. As a result, the ARCHITECT HE4 assay meets the predetermined acceptance criteria of 100 ± 10% of Time 0 value and supports the package insert sample storage claim of up to 24 hours at room temperature.

Sample Stability - Freeze / Thaw

A study was conducted to determine the effects of multiple freeze/thaw cycles on serum samples analyzed with the ARCHITECT HE4 assay. Twelve (12) fresh (less than 36 hours from the time of the draw) serum samples were spiked with native HE4 antigen to values between 125 and 1100 pmol/L and used for this study. Each sample was aliquoted into 12 tubes. One aliquot was retained as the unfrozen control and the remaining aliquots were subjected to repeated freeze/thaw cycles. The samples were frozen at ≤ -20°C for a minimum of 30 minutes followed by thawing them at room temperature. At the end of each cycle one vial from each set was removed and stored at 2-8°C until the assay was ready

to run. A total of ten (10) freeze/thaw cycles were completed for each serum sample. For each aliquot, the control and the test sample sets were tested in 2 replicates using the ARCHITECT HE4 assay on the ARCHITECT *i*2000$_{SR}$ System. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The recovery of the cycle was determined by comparing the mean test sample value tested against the unfrozen control sample mean value.

The average recovery per cycle for the 12 samples ranged from 93% to 128%. The percent recoveries for individual results were within +/-15% of the unfrozen control through cycle 6. The data supports a conclusion that samples can be frozen and thawed six (6) times for the ARCHITECT HE4 assay without compromising the assay.

d. Detection limit:

To characterize the Limit of Blank (LoB), Limit of Detection (LoD) and Limit of Quantitation (LoQ) of the ARCHITECT HE4 assay, an experimental protocol was utilized modeled after CLSI guideline EP17-A ("Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline – 2004").

A Low Level Sample Set was prepared by diluting 2 normal human serum samples in kit calibrator A (0 pmol/L HE4) to obtain HE4 concentrations as shown in the table below.

Panel	Endogenous HE4 Levels (pmol/L)	Dilution in Calibrator A	Expected Value (pmol/L)
1 2	27.1	1:20 1:40	1.4 0.7
3 4	20.6	1:10 1:60	2.1 0.3

Calibrator A and Low Level Sample Set were tested in 5 replicates, once per day for three (3) days on 2 ARCHITECT *i*2000$_{SR}$ Systems with 3 kit reagent lots (N= 5x3x2x3 = 90). Each of the 3 reagent lots were run on both ARCHITECT *i*2000$_{SR}$ Systems. One lot of kit calibrators and controls was used per instrument system. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity.

To determine the Limit of Blank (LoB), the mean and standard deviation (SD) were calculated for each of the runs of the Calibrator A. The total mean and average SD for the Calibrator A is -0.10 ± 0.09 pmol/L. The LoB was 0.05 pmol/L based on equation: LoB = Mean + 1.645*SD$_{A\ Cal}$.

To determine the Limit of Detection (LoD), the individual total means and SDs, and the average SD were calculated for each of the runs of the Low Level Sample Set. The results are summarized in the following table:

Panel	Endogenous HE4 Levels (pmol/L)	Dilution in Calibrator A	Expected Value (pmol/L)	Observed value (pmol/l)	estimated SD
4	27.1	1:20	0.3	0.120	0.07
2		1:40	0.7	0.44	0.07
1	20.6	1:20	1.4	1.22	0.07
3		1:10	2.1	2.060	0.10

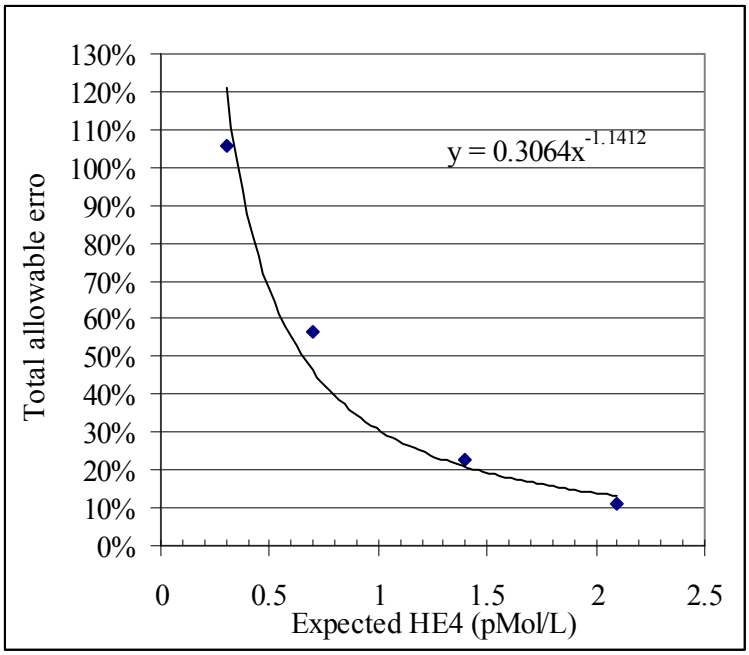
The LoD was calculated to be 0.18 pmol/L based on the following formula: LoD = LoB + 1.645 $_{SDSample\ Set}$= 0.05 + (1.645 * 0.08).

To determine the Limit of Quantitation (LoQ), the mean of the differences between the expected values and measured values was calculated for each replicate of each member of the Low Level Sample Set to determine the total error. The average individual biases and SDs of each member of the Low Level Sample Set were calculated by run. The average bias and average SD for all concentrations in the Low Level Sample set was -0.15 +/- 0.08 pmol/L. The Total Error is 0.31 pmol/L (28% of the mean expected value) based on the following equation: Total Error = | Bias | + 2*SD_S = |- 0.15| + (2 * 0.08). The LoQ claimed was 0.18 pmol/L.

Since the %CV and %bias are a function of the expected concentration, the calculated LoQ should not base on the average total error for the 4 samples. The observed and expected HE4 concentrations with the total biases and %CV of variation should be individually calculated for the 4 panel set. The following table summarizes the %CV of variation for each individual expected HE4 concentration, the %bias from expected HE4 concentration, and the calculated total allowable error for each expected HE4 concentration:

Panel	Expected Value	Observed value	estimated SD	%CV	%bias	allowable total error
4	0.3	0.120	0.0700	23.3%	-60.0%	105.7%
2	0.7	0.44	0.0700	10.0%	-37.1%	56.7%
1	1.4	1.22	0.0700	5.0%	-12.9%	22.7%
3	2.1	2.060	0.1000	4.8%	-1.9%	11.2%

A plot of the allowable total error vs. the expected HE4 value shows the relationship.



Extrapolating from the smoothed model of the total allowable error for the chosen total error of 30%, the HE4 concentration is 1.0 pmol/L. At this HE4 concentration the observed bias of the observed HE4 concentration from the expected concentration was -17.6% and the %CV of the observed HE4 was 7%. Therefore, the LoQ estimate is 1 pmol/L with 30% total allowable error. This value is less than the acceptance criteria (< 20 pmol/L).

The LoD (0.2 pmol/L), and LoQ (1.0 pmol/L) meet the sponsor predetermined acceptance criteria of LoD < 15 pmol/L and LoQ < 20 pmol/L. These data support the performance claims of the assay.

e. *Analytical specificity:*

Lipid interference

To evaluate the potential interference from lipids in the assay, CLSI guideline EP7-A ("Interference Testing in Clinical Chemistry, Approved Guideline") was used to design the interference experiments. Six (6) human serum samples were used in this study. The endogenous HE4 concentrations of samples 3 through 6 were augmented by the addition of HE4 antigen (recombinant) prepared from a stock solution. Samples 1 and 2 were not spiked with additional HE4 antigen. Each HE4 serum sample was split into 2 aliquots. One aliquot was spiked with a lipids solution to achieve 3000 mg/dL lipid. One aliquot was spiked with an equal volume of phosphate-buffered solution (PBS) for use as a control sample. Serum samples were not diluted more than 20%. The control and the lipid supplemented samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 2 replicates of each sample with the added lipids was compared to the mean value of the corresponding control sample without lipids to calculate the percent recovery. The average recovery was 98%

ranging from 94% to 104% for samples with a 3000 mg/dL lipid concentration. Less than 10% average interference was observed in the assay with specimens containing an elevated level of lipids. Testing meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples. These data support the performance claims of the assay.

Bilirubin interference

Interference from bilirubin was assessed in a manner similar to lipid interference. Twelve (12) human serum samples with HE4 values distributed across the assay range were supplemented with bilirubin as a potential interfering substance. Each HE4 serum sample was split into 2 aliquots. One aliquot was spiked with a bilirubin solution (unconjugated bilirubin in 10 mM NaOH) to achieve 20 mg/dL bilirubin. One aliquot was spiked with an equal volume of 10 mM NaOH for use as a control sample. Serum samples were not diluted more than 5%. The control and the bilirubin supplemented samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. Percent recoveries were calculated as in lipid interference evaluation. The average recovery was 98% ranging from 92% to 103% for samples with a 20 mg/dL bilirubin concentration. Less than 15% individual interference and less than 10% average interference was observed in the assay with specimens containing an elevated level of bilirubin. Testing meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples. These data support the performance claims of the assay.

Hemoglobin interference

Interference from hemoglobin was assessed in a manner similar to lipid interference. Six (6) human serum samples with HE4 values distributed across the assay range were supplemented with hemoglobin as a potential interfering substance. Each HE4 serum sample was split into two (2) aliquots. One aliquot was spiked with a hemoglobin solution to achieve 500 mg/dL hemoglobin. One aliquot was spiked with an equal volume of PBS for use as a control sample. Serum samples were not diluted more than 5%. The control and the hemoglobin supplemented samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. Percent recoveries were calculated as in lipid interference evaluation. The average recovery was 108% ranging from 104% to 113% for samples with a 500 mg/dL hemoglobin concentration. Less than 15% individual interference and less than 10% average interference was observed in the assay with specimens containing an elevated level of bilirubin. Testing meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples. These data support the performance claims of the assay.

Low and High Protein interference

Interference from high protein concentrations was assessed in a manner similar to lipid interference. Six (6) serum samples were split into four (4) equal aliquots, one (1) aliquot was used as the control and one (1) aliquot was used to measure the protein concentration of the samples. Based on the measured protein concentration for each of the six (6) serum samples, one (1) aliquot was diluted with PBS to obtain 3 g/dL total protein, and one (1) aliquot was spiked with human serum albumin protein to obtain 12 g/dL final protein concentration. HE4 was also supplemented to achieve different HE4 concentrations. Serum samples were not diluted more than 5% by addition of HE4. The control and the protein supplemented samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. Percent recoveries were calculated as in lipid interference evaluation. The average recovery was 103% ranging from 101% to 107% for samples with a 3 g/dL protein concentration. The average recovery was 103% ranging from 101% to 108% for samples with a 12 g/dL protein concentration. Less than 15% individual interference and less than 10% average interference was observed in the assay with specimens containing an elevated level or lowered level of protein. Testing meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples. These data support the performance claims of the assay.

Human anti-mouse antibody interference

To evaluate the potential interference from Human Anti-Mouse Antibodies (HAMA) in the assay, CLSI guideline, EP7-A was used to design the interference experiments. Preparation of Test Samples using six (6) HAMA positive human serum samples was as follows:

1. Each HAMA serum sample was split into three (3) equal aliquots.
2. One (1) aliquot was spiked with approximately 90 pmol/L of recombinant HE4 antigen (spike 1) prepared from a stock solution. No sample was diluted more than 5% through spiking of the HE4 antigen.
3. One (1) aliquot was spiked with approximately 450 pmol/L of recombinant HE4 antigen (spike 2) prepared from a stock solution. No sample was diluted more than 5% through spiking of the HE4 antigen.
4. One (1) aliquot was the unaltered control used to determine the endogenous HE4 concentration in each HAMA sample.

The Expected HE4 Concentrations of HAMA samples spiked with HE4 Antigen were calculated as follows:

$[(V0 \times C0) + (VS \times CS)] / (V0 + VS)$ Where:

V0 = Volume of unspiked sample
C0 = Concentration of unspiked sample (control)
VS = Volume of spiking solution
CS = Concentration of spiking solution

The control and the HAMA samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to

determine assay validity. The percent recovery was calculated from the mean Observed HE4 Concentration obtained for each HAMA sample compared to the Expected HE4 Concentration of that sample.
The range of recoveries for the individual samples (HAMA values ranging from 45.4 to 154.6 ng/mL) was from 91% to 108%, the average recovery was 102%. The range of recoveries for each HAMA level tested was from 95 to 106%. Less than 15% interference for each individual sample tested and less than 10% average interference for each HAMA level tested was observed in the ARCHITECT HE4 assay. As a result, the testing meets the sponsor's predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recoveries of each HAMA level. These data support the performance claims of the assay.
Rheumatoid arthritis interference
To evaluate the potential interference from rheumatoid factor (RF) antibodies in the assay, a design similar to the HAMA interference was performed. Preparation of Test Samples using six (6) HAMA positive human serum samples was as follows:

1. Each RF serum sample was split into three (3) equal aliquots.
2. One (1) aliquot was spiked with approximately 90 (spike 1) pmol/L of HE4 antigen (recombinant) prepared from a stock solution. No sample was diluted more than 5% through spiking of the HE4 antigen.
3. One (1) aliquot was spiked with approximately 450 (spike 2) pmol/L of HE4 antigen (recombinant) prepared from a stock solution. No sample was diluted more than 5% through spiking of the HE4 antigen.
4. One (1) aliquot was the unaltered control used to determine the endogenous HE4 concentration in each RF sample.

The Expected HE4 Concentrations of RF samples spiked with HE4 Antigen were calculated similarly to the calculation for HAMA interference. The control and the RF samples were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ analyzer. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The percent recovery was calculated from the mean Observed HE4 Concentration obtained from each RF sample supplemented with HE4 antigen compared to the Expected HE4 Concentration of that sample.

The range of recoveries for the individual samples (RF values ranging from 21 to 445 IU/mL) was from 98% to 111%; the average recovery was 103%. The range of recoveries for each RF level tested was from 100 to 108%. Less than 15% interference for each individual sample tested and less than 10% average interference for each RF level tested was observed in the assay. As a result, the testing meets the sponsor's predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recoveries of each RF level. These data support the performance claims of the assay.

Interference from various chemotherapeutic agents

To evaluate the potential interference from several chemotherapeutic agents in the assay, CLSI guideline EP7-A was used to design the interference experiments. Two (2) levels of human serum supplemented with native HE4 antigen were spiked with each chemotherapeutic agent listed below as a potential interfering substance. The various chemotherapeutic substances were prepared for spiking using a particular solvent. Control samples were spiked with an equal volume of each respective solvent without the agent. The control samples and the samples supplemented with chemotherapeutic substances were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ System analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 2 replicates of each sample with the added therapeutics was compared to the mean value of the corresponding control sample with solvent only to calculate the percent recovery.

	HE4 level 1	HE4 level 2	
Chemotherapeutic agent	Percent Recovery (%)	Percent Recovery (%)	Mean Percent Recovery (%)
Carboplatin 500 ug/ml	105	99	102
Cisplatin 165 ug/ml	98	97	98
Clotrimazole 0.3 ug/ml	103	99	101
Cyclophosphamide 500 ug/ml	106	101	104
Dexamethasone 10 ug/ml	105	97	101
Doxorubicin 1.2 ug/ml	105	100	103
Leuvcovorin 2.7 ug/ml	105	102	104
Melphalan 2.8 ug/ml	104	98	101
Methotrexate 45 ug/ml	100	96	98
Paclitaxel 3.5 ng/ml	102	101	102

The range of the recoveries for all samples tested was 96% to 106%, the average recovery for each chemotherapeutic agent tested was 98% - 104%. Less than 15% interference for each individual sample tested and less than 10% average interference for each therapeutic tested was observed in the assay. As a result, testing meets the sponsor's predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of each therapeutic tested. These data support the performance claims of the assay.

Cross-reactivity from other tumor markers

To evaluate the potential interference from other known tumor markers (such as CA125, CA15-3, CA19-9, CEA, and AFP) in the assay, semi-purified tumor marker solutions were prepared from in-house cultures or obtained from an outside source (CEA and AFP). 100 times, 50 times and 10 times the normal level of the respective tumor marker were added to kit Calibrator A. The target concentrations are as follows:

Cross Reactant	100X Concentration	50X Concentration	10X Concentration
CA 125	3500 U/mL	1750 U/mL	350 U/mL
CA 15-3	3500 U/mL	1750 U/mL	350 U/mL
CA 19-9	3500 U/mL	1750 U/mL	350 U/mL
CEA	1000 µg/L	500 µg/L	100 µg/L
AFP	400 µg/L	200 µg/L	40 µg/L

The samples supplemented with the various tumor markers but lacking added HE4 were tested in 2 replicates using the assay on the ARCHITECT *i*2000$_{SR}$ system analyzer. A single calibration curve was used throughout the study. The kit Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from the 2 replicates of each sample with the added tumor marker antigen was calculated. The mean results of the replicates are as follows:

Cross Reactant marker	Observed HE4 (pmol/L) at 100X	Observed HE4 (pmol/L) at 50X	Observed HE4 (pmol/L) at 10X
CA 125	0	0	0
CA 15-3	11.6	5.6	1.3
CA 19-9	0.1	0.1	0.1
CEA	0	0.2	0
AFP	0.2	0.1	0

All individual samples supplemented with the various tumor markers resulted in an HE4 result of < 15 pmol/L (labeling claim for the Limit of Detection). As a result, these data support the performance claims of the assay.

Serum tube type

A study was conducted to evaluate whether commonly used serum collection tubes interfere with the ARCHITECT HE4 assay. Serum samples from 10 healthy volunteers were collected in the following blood collection tube types (BD vacutainer tubes) and tested within eight hours:

Closure	Tube type	Descriptor name
With Conventional Stopper	Red/Gray Top – SST Red Top – Serum (Control)	"SST" "Red Top"
With Hemogard™ Closure	Gold Top – SST Red Top – Serum	"Gold" "Red H"

Each blood collection tube type for each donor sample was split into 2 aliquots. One aliquot for each of two donor subjects was spiked with HE4 antigen (recombinant) giving 5 different increasingly higher HE4 concentrations. The specimens were not diluted more than 10%. One (1) aliquot was left un-spiked (neat). The spiked and the non-spiked serum samples for each tube type were tested in 2 replicates using

the assay on the ARCHITECT *i*2000$_{SR}$ System. A single calibration curve was used throughout the study. The ARCHITECT HE4 Controls were tested and evaluated for each run to determine assay validity. The mean concentration obtained from 2 replicates of each sample tested was compared to the corresponding spike or neat mean concentration for the conventional serum red top tube sample (control).

The range of percent recoveries for each individual sample was between 97% - 108% for non-spiked samples and 95% - 113% for spiked samples. The percent recovery for the mean of all samples was 105% for Red/Gray Top (SST), 101% for Red Top Hemogard™ (Red H) and 102% for Gold Top (Gold). The assay meets the predetermined acceptance criteria of 100 ± 15% for individual sample recoveries and 100 ± 10% for the mean recovery of all samples. Serum collected in serum separator tubes with a conventional or Hemogard closure will not likely cause interference in the assay.

f. Assay cut-off:

There is no assay cut-off for monitoring the progression of epithelial ovarian cancer using this marker.

2. Comparison studies:

a. Method comparison with predicate device:

A comparison of HE4 test results for the Architect HE4 assay and the manual HE4 immunoassay was performed with the following objectives:

1. Provide data for method comparison for inclusion in the package insert.
2. Determine the correlation and agreement of the ARCHITECT HE4 Assay with the HE4 EIA.
3. Verification of the performance of the quality control (QC) procedures as stated in the proposed ARCHITECT HE4 Package Insert.
4. Verification of user testing in single replicates

Two sites provided serum specimens for the study. For each site, the minimum number and patient groups were: 60 specimens from apparently healthy females, 60 specimens from women with benign gynecologic disease, and 60 specimens from women with epithelial ovarian cancer. Specimens were obtained from 4 commercial serum vendors. Samples were collected under an Institutional Review Board (IRB)-approved protocol or considered remnant samples and the IRB concluded that informed consent was not needed. The following inclusion/exclusion criteria were utilized for specimen collection:

Inclusion criteria:

1) Samples must be collected in red top tubes or serum separator tubes (SST)
2) Sample Information
   a) Age or Date of Birth
   b) Gender
   c) Collection Date
   d) Menopausal status
   e) Ethnicity

Exclusion criteria

3) Samples collected in tubes other than red top tubes or SST

4) Stored at 2-8°C for more than 3 days or at room temperature for more than 24 hours
5) No sample information
6) >5 Freeze/Thaw Cycles
7) Obvious microbial contamination

To determine the statistical relationship between the ARCHITECT HE4 Assay and the manual HE4 immunoassay, linear regression, Deming regression, and Passing-Bablok regression analyses were performed at each of the 2 laboratory sites. One site was the manufacturer's site and 1 external site. A minimum of 180 serum samples were tested at each site. Both Pearson and Spearman correlation coefficients were determined for all comparisons. Ninety-five percent (95%) confidence intervals on the regression parameters were calculated. Outliers were evaluated per CLSI EP9-A2.3. Bias plots were generated to assess the agreement between the two assays as per CLSI EP9-A2.3.

The acceptance criteria for samples with HE4 EIA values from 20 to 900 pM are as follows:

a) Passing-Bablok slope: 1.00 ± 10% (i.e., 0.90 to 1.10)
b) Correlation Coefficient ≥ 0.90

For this study, the sponsor chose to limit one comparison to the common assay range for both assays. The ARCHITECT HE4 assay range is 20 - 1,500 pmol/L and the HE4 EIA assay range is 15 – 900 pM. The common range for the comparison was 20 to 900 pmol/L.

The ARCHITECT HE4 assay was performed following the instructions for use specified in the proposed package insert. Sample testing was performed on 1 lot of reagents using 1 ARCHITECT *i*2000$_{SR}$ System per site. A single calibration curve was used throughout the study at each site. The kit controls were tested and evaluated for each run to determine assay validity. Sample dilutions for ARCHITECT HE4 values above the range of the assay were diluted using the on-board dilution protocol at a dilution of 1:10.

The manual HE4 immunoassay was run according to the package insert. The samples were tested in 2 replicates at each site. The mean of the 2 replicates was used for the data analyses. Sample dilutions for manual HE4 immunoassay above the range of the assay were diluted manually at a dilution of 1:10 with the exception of 1 sample (EOC74) at the external site which was diluted at dilution of 1:5.

At the sponsor's testing site, the following samples were included for analysis:

65 samples from apparently healthy females
65 samples from women with benign gynecologic diseases
63 samples from women with epithelial ovarian cancer
193 samples total

Disease Category	Frequency
Healthy Females	65

Premenopausal	26
Postmenopausal	39
Benign Gynecologic Diseases	65
Endometriosis	41
Fibroids	24
Epithelial Ovarian Cancer	70
Total	200

Of the 70 subjects with epithelial ovarian cancer (EOC), 7 specimens were excluded from analysis due to HE4 values above 900 pmol/L.

At the external testing site, the following samples were included for analysis:

65 samples from apparently healthy females
65 samples from women with benign gynecologic diseases
67 samples from women with epithelial ovarian cancer
197 samples total

Disease Category	**Frequency**
Healthy Females	65
Premenopausal	5
Postmenopausal	60
Benign Gynecologic Diseases	65
Endometriosis	10
Fibrocystic breast	2
Fibroids	36
HPV	3
Ovarian cysts	2
Polycystic ovaries	12
Epithelial Ovarian Cancer	70
Total	200

Of the 70 subjects with epithelial ovarian cancer (EOC), 2 specimens were excluded from analysis due to HE4 values above 900 pmol/L and 1 specimen was excluded due to HE4 value below 20 pmol/L.

The patients providing specimens from each site are modestly different. At the external site most normal females were postmenopausal (92%) while at the sponsor's site approximately 50% were postmenopausal women. At the external site, approximately one-third of women with benign gynecologic disease had diseases other than endometriosis and fibroid disease while at the sponsor's site all women with benign disease had endometriosis and fibroid disease.

The sponsor compares assay results at each site as well as for all subjects combined. Analysis will concentrate only on combined data from each site.

The racial composition of all study subjects was the following:

Ethnicity	n	Proportion
African American	8	2%
Caucasian	386	97%
Hispanic	5	1%
Unknown	1	0%

The median age for three disease groups of subjects tested are as follows:

Age (Yrs) by Disease	n	Median	95% CI	
benign	130	46.0	43.0	49.0
cancer	140	56.0	52.0	59.0
normal	130	63.0	60.0	65.0

All benign disease groups were pooled into a single category.

Of all women, 64% were post-menopausal and 36% were pre-menopausal. The median HE4 values by menopausal status are as follows:

HE4 values by Menopausal Status		n	Median	95% CI
ARCHITECT HE4	Post-Menopausal	255	60.40	53.40 to 66.00
	Pre-Menopausal	145	42.30	38.80 to 45.10
manual HE4	Post-Menopausal	255	60.48	56.20 to 65.84
	Pre-Menopausal	145	45.86	42.42 to 50.05

Note that the HE4 values for each assay are essentially identical in post-menopausal women. The ARCHITECT HE4 values are lower than the manual HE4 values in pre-menopausal women. It is not clear if this difference is clinically significant.
The median HE4 values by disease group are as follows:

		N	Median	95% CI
ARCHITECT HE4 by Disease	benign	130	40.25	38.20 to 43.70
	cancer	140	105.60	77.50 to 140.20
	normal	130	49.90	44.40 to 54.10
manual HE4 by Disease	benign	130	44.63	42.42 to 47.21
	cancer	140	112.05	84.65 to 131.19
	normal	130	51.90	47.60 to 54.72

Among cancer and normal subjects, the median HE4 values for each assay (Architect and manual immunoassay) are equivalent. For subjects with benign diseases, the median ARCHITECT HE4 value is lower than the HE4 value using the manual immunoassay. It is not clear if this difference is clinically significant.

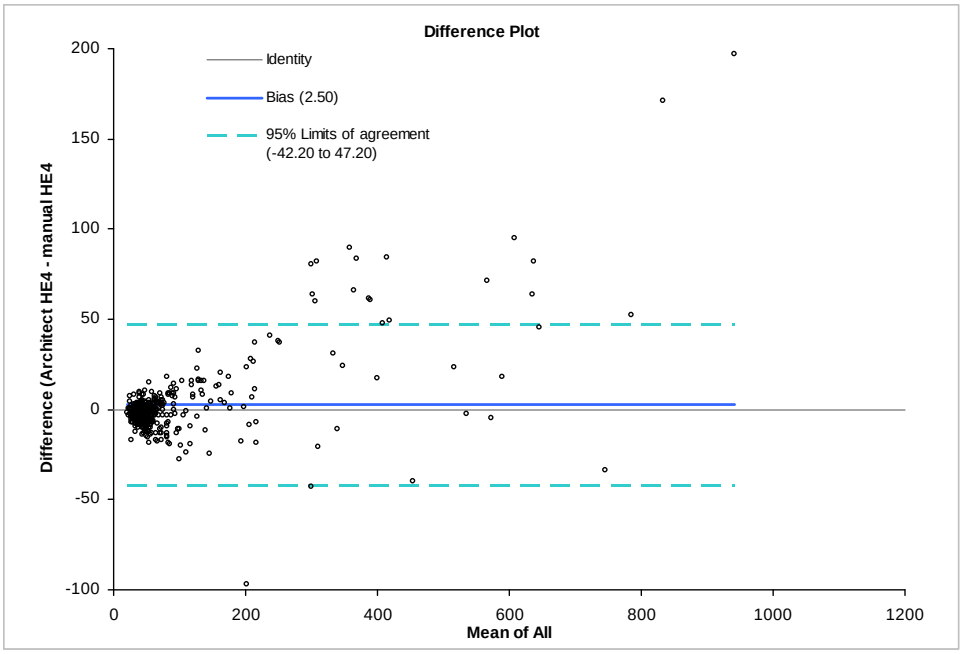
The sponsor analyzed for both combined sites different measures of equivalent test results for the Architect HE4 assay and the manual HE4 immunoassay. To assess the bias between assays, the sponsor chose 3 measures of assessment

(when selecting the manual HE4 immunoassay concentration range of 20-900 pM). The 3 measures used for this evaluation were:

1. Actual concentration difference (bias) between the 2 assays (defined as bias = ARCHITECT HE4 result – HE4 EIA result),
2. Percent difference (% bias) between the 2 assays (defined as %bias = [(ARCHITECT HE4 result – HE4 EIA result) / HE4 EIA result] x 100%,
3. Ratio bias between the 2 assays (defined as ratio bias = (ARCHITECT HE4 value) / (HE4 EIA value).

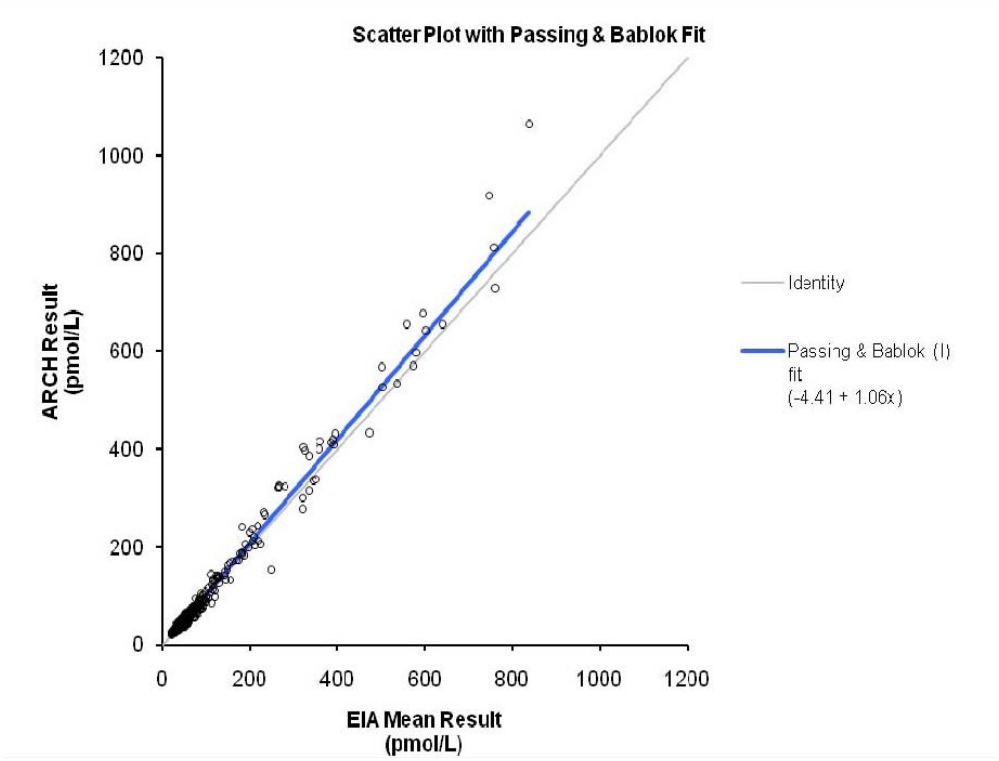
For the concentration range of 20 to 900 pmol/L (manual HE4 assay), the mean difference (bias) in values was determined to be 2.2 pmol/L (95% confidence interval of the difference 0.1 to 4.3) for 390 subjects. The mean percentage bias relative to the manual assay result was -2.3% (95% confidence interval from -3.5% to -1.1%). The ARCHITECT HE4 assay has, on average, a higher value than the manual HE4 assay.

A graphical plot of differences in assay result (Bland-Altman plot) is as follows:



Based on this plot for HE4 values above 300 pmol/L, the ARCHITECT HE4 assay has a proportion of differences greater than 50 pmol/L. ARCHITECT HE4 results are more frequently higher than the manual HE4 immunoassay above 300 pmol/L. This difference in HE4 concentration above 300 pmol/L is likely because of HE4 values from ovarian cancer subjects.

The sponsor shows a figure illustrating the relationship between the ARCHITECT HE4 assay values and the HE4 EIA values for the 20-900 pM concentration range as determined by a Passing-Bablok regression analysis. This figure is as follows:



The slope of the best fit line for 390 subjects with manual HE4 immunoassay values between 20 and 900 pmol/L was 1.06 (95% confidence interval 1.03 to 1.08). The intercept of the best fit line was -4.4 pmol/L (95% confidence interval -5.5 to –3.0). The Spearman rank correlation coefficient was 0.97. The slope result met the sponsor's acceptance criteria of a slope between 0.9 and 1.10.

Based on the relationship in the Passing-Bablok regression analysis, the %bias between assay methods is calculated as follows:

manual HE4	predicted Architect HE4	%bias
58.6	57.3	-2.2%
100	102.3	2.3%
125	129.4	3.5%
150	156.6	4.4%
175	183.7	5.0%
200	210.8	5.4%
250	265.1	6.1%
300	319.4	6.5%
500	536.6	7.3%

Note that for values below 150 pmol/L, the upper limit of the normal range, the %bias between assays is below 5%. Above 300 pmol/L, the %bias between assays

is 6.5% or slightly above. At HE4 concentrations above 300 pmol/L, subjects in the comparison had ovarian cancer. Therefore, it is likely that at 300 pmol/L or higher, a subject would have ovarian cancer and likely would be undergoing serial surveillance monitoring for cancer recurrence or would be receiving therapy. Therefore, the difference is likely not clinically significant.

For the concentration range of >20 pmol/L (manual HE4 assay; range of values 20 to ~8000 pmol/L), the mean difference (bias) in values was determined by the sponsor to be -4.1 pmol/L (95% confidence interval of the difference -10.1 to 1.8) for 399 subjects. The mean percentage bias relative to the manual assay result was -2.7% (95% confidence interval from -3.9% to -1.4%). The ARCHITECT HE4 assay has, on average, a lower value than the manual HE4 assay when the manual assay is > 20 pmol/L. A figure is shown illustrating the relationship between the ARCHITECT HE4 assay values and the HE4 EIA values for HE4 values > 20 pmol/L in the manual assay (n = 399) as determined by a Passing-Bablok regression analysis. The slope of the best fit line with manual HE4 immunoassay values > 20 pmol/L was 1.04 (95% confidence interval 1.02 to 1.07). The intercept of the best fit line was -3.7 pmol/L (95% confidence interval -5.1 to -2.4). The Spearman rank correlation coefficient was 0.97. The slope result met the sponsor's acceptance criteria of a slope between 0.9 and 1.10.

The sponsor also sought to show the equivalence of replicates of the same serum sample when tested in the Architect HE4 assay. This was assessed by comparing replicate 1 vs. replicate 2 from the same serum sample on 195 samples less than 1500 pmol/L. The sponsor shows a figure showing the bias between the 1st replicate value and the 2nd replicate value. The overall average bias was 0.3% (95% Limits of Agreement: -9.0% to 9.5%). The sponsor further indicated the relationship between the 1st replicate value and the 2nd replicate value for the 20 to 1500 pmol/L concentration range as determined by a Passing-Bablok Regression Analysis. The slope of the best fit line was 1.01 (95% confidence interval 1.00 to 1.02). The intercept of the line was -0.45 pmol/L (95% confidence interval -1.05 to 0.18). The Spearman rank correlation coefficient was 0.99. The Passing-Bablok slope and intercept indicate close agreement between replicates since the slope is equivalent with 1.0 and the intercept is equivalent with 0. Thus the results support testing of HE4 serum samples in single replicates.

For all the method comparison studies, the results met the sponsor's acceptance criteria and are acceptable.

b. Matrix comparison:
No comparison was made with other specimen testing matrices since only human serum is the testing matrix.

3. Clinical studies:
   a. Clinical Sensitivity:
   A retrospective clinical study was performed utilizing remnant serial serum samples from 76 subjects diagnosed with Epithelial Ovarian Cancer (EOC) after the completion of chemotherapy. The serial serum samples (a minimum of 3 serial draws) were taken from pre- and postmenopausal women 18 years of age or older with a diagnosis of EOC who were undergoing serial surveillance

monitoring of cancer progression during clinical follow-up. Subject samples could have been drawn either during therapy, or following the completion of therapy for the treatment of EOC. A clinician assessment was made at each clinical follow-up at which a sample draw from each subject was taken. At each clinical follow-up time, a clinical assessment of each subject at that follow-up time categorized the disease status of subjects into 1 of the following subgroups:

1. No Evidence of Disease – NED: A complete lack of clinical evidence of disease.
2. Stable Disease: Clinical evidence that the disease has not changed since last assessment.
3. Progressive Disease: Clinical evidence of growth in the primary tumor or the appearance of new tumors since the last assessment.
4. Responding Disease: Clinical evidence that there is a shrinking of the primary tumor and no evidence of new tumors

HE4 serum levels for all serum samples were measured at the sponsor's testing site in single replicates using several kit lots. The assay was performed following the appropriate quality control procedures and manufacturer's instructions for use specified in the proposed package insert. Sample testing was performed on 1 ARCHITECT *i*2000$_{SR}$. A single calibration curve was used throughout the study.

For purposes of calculating clinical sensitivity and other performance parameters, the clinical disease status was condensed into 2 categories: progressive disease and non-progressive disease. A percentage change in successive HE4 values (14%) was used to categorize the HE4 readings into those that were and were not significantly elevated. The percent change greater than or equal to 14% categorized an HE4 reading into a positive elevation; while a percent change less than 14% categorized an HE4 reading into a negative elevation. The two category HE4 test results were cross-tabulated against progression or non-progression in disease.

The clinical sensitivity of the 14% elevation in HE4 values (from the preceding serum sample) was 53.5% (exact binomial 95% confidence interval 43.2% to 63.6%). Of 99 clinical events categorized as progressive disease, a percent elevation at least 14% was detected by the assay in 53 events across all subjects. Further details of the study are below.

*b. Clinical specificity**:***

For the same clinical study summarized in the clinical sensitivity section, the clinical specificity of the 14% elevation in HE4 values (from the preceding serum sample) was 78.5% (exact binomial 95% confidence interval 73.7% to 82.9%). Of 331 clinical events categorized as non-progressive disease, a percent elevation less than 14% was detected by the assay in 260 events across all subjects. Further details of the study are below.

c. Other clinical supportive data:

The specific objectives of the study included:

- Performance evaluation of ARCHITECT HE4 assay as an aid for monitoring recurrence or progressive disease for patients diagnosed with EOC in a representative patient cohort.
- Performance verification of the quality control procedures as stated in the proposed package insert

Study design

A retrospective clinical study was performed utilizing remnant serial serum samples from 76 subjects diagnosed with Epithelial Ovarian Cancer (EOC) after the completion of chemotherapy. The serial serum samples (a minimum of 3 serial draws) were taken from pre- and postmenopausal women 18 years of age or older with a diagnosis of EOC who were undergoing serial surveillance monitoring of cancer progression during clinical follow-up. Subject samples could have been drawn either during therapy or following the completion of therapy for the treatment of EOC. A clinician assessment was made at each clinical follow-up at which a sample draw from each subject was taken. At each clinical follow-up time, a clinical assessment of each subject at that follow-up time categorized the disease status of subjects into 1 of the following subgroups:

1. No Evidence of Disease – NED: A complete lack of clinical evidence of disease.
2. Stable Disease: Clinical evidence that the disease has not changed since last assessment.
3. Progressive Disease: Clinical evidence of growth in the primary tumor or the appearance of new tumors since the last assessment.
4. Responding Disease: Clinical evidence that there is a shrinking of the primary tumor and no evidence of new tumors

HE4 serum levels for all serum samples were measured at the sponsor's testing site in single replicates using several kit lots. The assay was performed following the appropriate quality control procedures and manufacturer's instructions for use specified in the proposed package insert. Sample testing was performed on 1 ARCHITECT *i*2000$_{SR}$. A single calibration curve was used throughout the study.

Study population

No specific inclusion/exclusion criteria are listed in the clinical study summary. However, subjects were pre- and postmenopausal women 18 years of age or older with a diagnosis of EOC. The following inclusion/exclusion criteria are listed in a pre-IDE clinical study outline (I090501) for this study submitted for review:

Inclusion Criteria

- Female, age ≥ 18 years
- Diagnosed with EOC
- Currently receiving or completed therapy for treatment of EOC
- A minimum of 3 serial draws available
- Appropriate clinical data
- Minimum 0.4 mL volume of serum available

- Normal appearance of sample

Exclusion Criteria

- No diagnosis of EOC
- Did not undergo therapy for the treatment of EOC
- Less than 3 serial draws available
- Insufficient sample volume
- Multiple freeze-thaw cycles
- Icteric, lipemic, hemolytic, substantial particulates

These criteria are acceptable and represent the sponsor's description of the study population. All subjects providing remnant serial serum samples were from a tertiary cancer center through a commercial vendor. Informed consent was noted as obtained utilizing a general Institutional Review Board (IRB) approved consent form allowing for the subjects' samples and clinical data to be used for future research purposes. No samples were specifically drawn for this study.
The clinical data collected for each subject included the following:

- patient ID
- date of birth
- race/ethnicity
- menopausal status
- date of EOC diagnosis
- initial diagnosis method and surgical information (if available)

For each subject sample draw, the following information was collected:

- chemotherapeutic treatment information (onset date, end date, regimen)
- imaging information (date, type, findings)
- physical exam date and findings
- sample draw date
- Clinical Disease Status
- additional comments

Sample Size

The sample size was calculated to be approximately 89 subjects with a minimum of 3 draws per subject. The statistical analysis of the sample size was discussed in the clinical study outline (I090501) previously submitted for review.

Sponsor Statistical Analysis Plan

The object of the trial is to demonstrate that the performance of the ARCHITECT HE4 assay is clinically equivalent to that of the already-cleared HE4 EIA. This will be achieved by verifying the diagnostic performance of the ARCHITECT HE4 assay in differentiating between visits in which patients' disease status is progressing vs. visits in which the disease is not progressing. The study will be considered successful if the overall concordance is statistically significantly higher than 62.8% using a test at the one-sided 2.5 % significance level. Specifically, at each visit, the attending clinician will classify the patient's status

into the 4 categories listed above. These four categories will be reduced to two categories–one class indicating disease progression, and a second class of the above remaining 3 categories indicating lack of disease progression.

The ARCHITECT HE4 will also be categorized into 2 classes. It is anticipated that the total random variability of the assay will be a coefficient of variation of no more than 10%. If this is the case, the total random variability of the difference between two assays is no more than 14% corresponding to a difference that was statistically significantly different from zero. Following this categorization, if the assay corresponding to a visit is more than 14% higher than the assay at the immediately preceding visit, the HE4 will be scored as statistically significantly elevated; otherwise it will be scored as not statistically significantly higher. A 2x2 cross-tabulation of all follow-up visits will be made.

Diagnostic performance of the ARCHITECT HE4 will be assessed by its positive and negative concordance, and the overall concordance, with clinical progression. ARCHITECT will be considered clinically equivalent in performance to the EIA if its overall concordance is at least 62.8%, a figure clinically compatible with that established in the HE4 EIA submission.

Each patient will provide a profile of binary scores for HE4 elevation and clinical progression. These profiles will be analyzed using generalized estimating equation model. The null hypothesis to be tested will be rephrased in terms of the log odds ratio. The target of 62.8% concordance corresponds to a log odds ratio of 0.87, and so the analysis will be in terms of a hypothesis test of the log odds ratio.

Writing θ for the log odds ratio for the association between the ARCHITECT HE4 assay and clinical status, this leads to the statistical hypothesis

H_0: $\theta \leq 0.87$ which, for success, will have to be rejected in favor of

H_a: $\theta > 0.87$

Testing will be one-sided at the 2.5% significance level. The 2x2 cross-tabulation will be used to calculate estimates of the positive, negative and overall concordance. Standard errors and confidence intervals for these quantities will be calculated from the generalized linear model outputs.

Subject disposition

A total of 525 samples for 80 subjects were initially obtained. Fourteen serum samples from 4 subjects were excluded from analysis because the diagnosis was something other than EOC after further review of the subject's medical chart. Five samples from 5 subjects were excluded because the disease status was not determined at the time of serum sampling. A total of 506 samples from 76 subjects were used for the analyses.

Subject Demographics

The mean age of subjects was 54 years old (95% confidence interval from 50.3 to 56.9). The majority of subjects were Caucasian (84.2%); 5.3% were African-American. Only 2 (2.6%) subjects were premenopausal subjects; the remaining

were post-menopausal women. The following table represents the distribution of clinical stages:

Category	Number	Percent of total
stage I	8	11%
stage II	7	9%
stage III	43	57%
stage IV	5	7%
Low Malignant Potential	1	1%
incomplete stage/unstaged	9	12%
undetermined/Not available	3	4%
sum	76	

There was a total of 506 serum sample draws for the 76 subjects. The mean number of sample draws per subject was 6.7 ranging from 3 to 31 draws. The length of time (in days) over which the subjects were monitored ranged from 36 to 2614 days (median of 368 days). The median interval (in days) between successive visits was 83 ranging from 0 to 1120.

Statistical analysis

When categorizing the percentage change in successive HE4 readings into those that were/were not significantly elevated and cross-tabulated against progression/non-progression in disease, the following table results:

	disease status		
Elevation in HE4	progression	no progression	total
Elevated (≥ 14%)	53	71	124
Not elevated (< 14%)	46	260	306
total	99	331	430
prevalence (π) ± se	23.0%	from 19.1%	to 27.3%

The following table summarizes the performance parameters:

sensitivity ± se	53.5%	Exact binomial 95% confidence interval	from 43.2%	to 63.6%
specificity ± se	78.5%	Exact binomial 95% confidence interval	from 73.7%	to 82.9%
PPV ± se	42.7%	95% Conf. Int. based on positive likelihood ratio	from 36.2%	to 49.6%
NPV ± se	85.0%	95% Conf. Int. based on negative likelihood ratio	from 82.0%	to 87.6%
Total agreement	72.8%		68.3%	76.9%

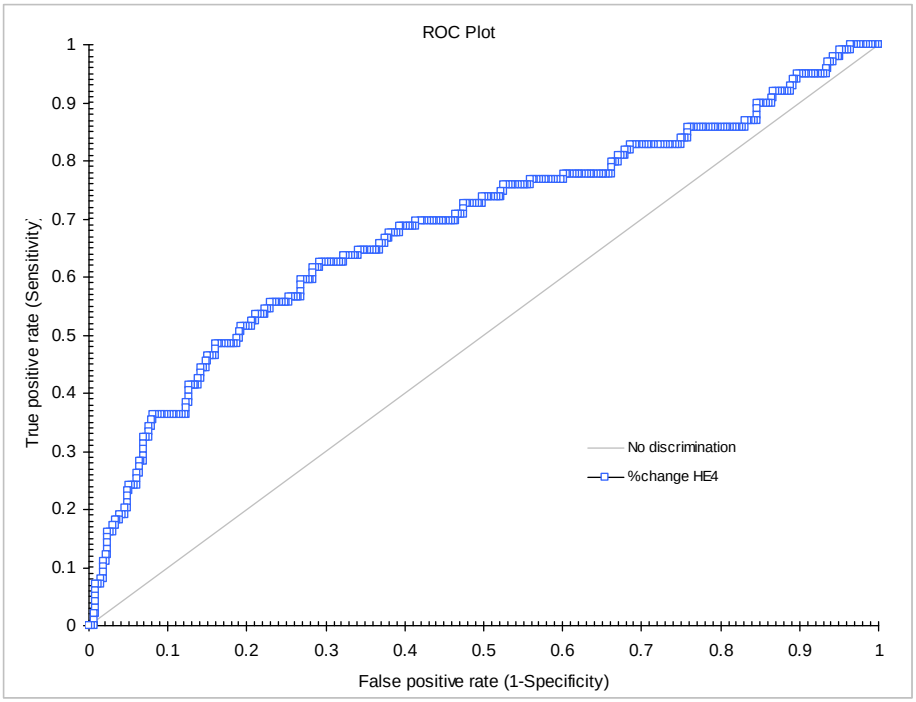
The sponsor calculates the sensitivity and specificity (and their standard errors by jackknife re-sampling) as: Sensitivity 53.4%, standard error 4.6%; Specificity

78.8%, standard error 1.9%, from which the False positive rate is 21.2%, standard error 1.9%.

The difference in sensitivity and the false positive proportion was 32.2% (95% confidence interval of the difference 22.4% to 42.0%). The difference in sensitivity and specificity is significantly greater than zero. The difference in specificity and 1-specificity in this analysis are equivalent. Therefore, the assay is informative with respect to progression/non-progression status.

The sponsor examines the total concordance of progression/non-progression status vs. %change in HE4 values to test the stated null hypothesis. The sponsor frames the hypothesis by converting the hypothesis into the logarithm of the odds ratio for the 2 x 2 table. A generalized linear model with a logistic link and autoregressive error structure was created. The 95% confidence interval for the log odds associated with elevated HE4 was determined to be 0.68 to 1.31 (mean 0.991). A test of the null hypothesis of the logarithm of the odds ratio was significant and the sponsor concludes that the overall percent concordance being less than or equal to 62.8% was rejected in favor of the hypothesis that the percent concordance exceeds 62.8%. Simple examination of the lower 95% confidence interval of the estimated overall agreement in the table above indicates that the total agreement exceeds 62.8% (lower limit 68.3% with mean concordance 72.8%). Thus the null hypothesis can be rejected in favor of the alternative hypothesis of total agreement greater than 62.8%.

The sponsor shows a receiver-operator characteristics (ROC) curve analysis of the percent change in HE4. The sponsor found that the area under the ROC curve was 0.685 with standard error 0.033; this area is statistically significantly better than the non-association area of 0.5. Analysis of the area and 95% confidence interval was very similar (area = 0.684; 95% confidence interval 0.618 to 0.750). The ROC curve is as follows:



The sponsor summarizes the analysis and suggests a series of %change in HE4 cutoff values with associated sensitivity and specificity performance values. The following table indicates analysis at different cutoff values for %change in HE4:

Percent Change in HE4	Sensitivity (%)	Specificity (%)	Lower CI for Specificity	Upper CI for Specificity
0	69 / 99 = 69.7%	58%	52%	63%
5	62 / 99 = 62.6%	69%	63%	74%
10	56 / 99 = 56.6%	75%	70%	79%
14	53 / 99 = 53.5%	78%	73%	83%
20	47 / 99 = 47.5%	84%	80%	88%
25	39 / 99 = 39.4%	87%	83%	91%
50	28 / 99 = 28.3%	94%	90%	96%
75	18 / 99 = 18.2%	97%	94%	98%
100	16 / 99 = 16.2%	98%	95%	99%

Note that for increasing %change in HE4 cutoff values, the sensitivity decreases and the specificity increases. There is currently no accepted %change in HE4 value that clinicians have decided is optimal for use. The labeling will present a similar table of differing sensitivity and specificity values at different %change in HE4 values so clinicians can choose a value that reflects their own preferences in sensitivity or specificity.

The following table shows the median (and 1st and 3rd quartile values) %change in HE4 for each clinical disease state.

Disease state	N	Min	1st quartile	Median	95% CI		3rd quartile	max
NED	183	-97.3%	-11.7%	-1.4%	-5.5%	1.4%	11.3%	3771%
Progression	99	-68.9%	-5.4%	16.8%	7.9%	24.6%	53.2%	2608%
Responding	61	-93.9%	-42.7%	-15.6%	-30.3%	-9.4%	-3.4%	71.6%
Stable	87	-81.7%	-15.0%	2.0%	-6.1%	5.5%	18.5%	8808%

The sponsor makes the following conclusions on the median values:

- The NED and Stable subjects have very similar summary statistics,
- Over 75% of visits in which the subject was responding to treatment, showed a decrease in HE4,
- Nearly 75% of the subjects showing disease progression had an increase in HE4.

It should also be noted that subjects with both NED and stable disease had a median %change in HE4 equivalent with 0% change (since the 95% confidence interval of the median includes 0%). Subjects with no evidence of disease (NED) had a %change in HE4 between -11.7% in the first quartile and 11.3% in the third quartile. This indicates that for 75% of subjects with NED, the %change in HE4 is less than ± 14% and this change in HE4 will not be informative. Subjects with responding disease had a median %change that was significantly below 0% change (95% confidence interval does not include 0). Subjects with progression had a median %change significantly above 0% change (confidence interval does not include 0). This indicates that the median %change for subjects with progression, responsive disease, and stable disease categories are related to the disease state of the subject. This difference in median %change can be generalized to a 3 x 3 table of subject counts for all visits where the HE4 concentration is above 140 pmol/L as follows:

HE4 > 140 pM	Disease status			
	responsive	stable	Progression	total
< -14%	10	9	6	25
-14% ≤ x ≤ 14%	6	11	15	32
> 14%	1	18	38	57
total	17	38	59	114

The association of a percent change in HE4 in the 3 assay categories shown with disease status is significant ($p < 0.001$, Chi square test). The following indicates the sensitivity and specificity of the %change in HE4 for each disease status:

Se responding =	58.8%	± 0.119	Sensitivity – (1-Specificity) =	43.4%	From 18.9%	To 67.8%
Sp responding =	84.5%	± 0.037				
Se stable =	28.9%	± 0.074	Sensitivity – (1-Specificity) =	1.3%	From -16.3%	To 18.9%
Sp stable =	72.4%	± 0.051				
Se progression =	64.4%	± 0.062	Sensitivity – (1-Specificity) =	29.9%	From 12.3%	To 47.4%
Sp progression =	65.5%	± 0.064				

The informativeness of a decreasing HE4 (< - 14% while HE4 concentration remains above 140 pM) for responding disease status is significant since the difference in sensitivity minus 1-specificity is greater than 0 (95% confidence interval of the difference from 18.9% to 67.8%). The informativeness of an increasing HE4 (>14% while HE4 concentration remains above 140 pM) for progressive disease is significant since the difference in sensitivity minus 1-specificity is greater than 0 (95% confidence interval of the difference from 12.3% to 47.4%).

Since 75% of subjects with NED had a %change in HE4 between -14% and 14%, this percent change in HE4 is not informative. However, a different type of cutoff may be informative in subjects with NED vs subjects without NED (i.e. active disease; progression, stable disease, or responding disease statuses). The median HE4 concentration of subjects with no evidence of disease was 49.7 pmol/L. This median concentration is significantly lower than the median HE4 concentrations of subjects with progressive, stable, or responding disease. Of 183 events where the disease state was categorized as NED, 129 NED events (70.5%) had HE4 concentrations less than the overall median HE4 concentration of 71.4 pmol/L for all 4 disease states. The median HE4 concentration and the 95% confidence interval of the median for the 4 disease states are as follows:

Clinical Disease Status	No. of events	Median HE4 (pmol/L)	Lower 95% CI	Upper 95% CI
NED	183	49.7	45.2	53.9
Progression	99	208.8	139.5	398.2
Responding	61	75.1	63.6	104.2
Stable	87	89.2	68.4	177.4

The sponsor also notes that comparison of the NED subjects with the active subjects shows that the HE4 level itself differs substantially between the 4 groups, and so may be used within the cancer subject group to differentiate NED from active cancer subjects. A ROC analysis by the sponsor of the Active and NED groups was carried out and showed that the maximum of the sum of sensitivity and specificity occurred if an HE4 level of 140 pmol/L was used to differentiate between the 2 groups.

Since subjects with NED may have low HE4 concentrations, a cutoff of 140 pmol/L HE4 (the upper limit of normal HE4 values) was selected to categorize subjects with and without elevation in HE4. Comparison of these 2 categories with disease status NED or not NED results in the following 2 x 2 table:

	NED	not NED	
HE4 ≤ 140	179	133	312
HE4 > 140	4	114	118
	183	247	430

The sensitivity of HE4 values less than 140 pmol/L was 97.8%. This indicates that 98% of subjects with no evidence of disease will have HE4 concentrations

less than 140 pmol/L. The specificity of HE4 values at 140 pmol/L is 46.2%. For subjects without NED, 46.2% will have HE4 values greater than 140 pmol/L. The informativeness of HE4 concentrations for subjects with or without NED is statistically significant. Note that of 247 events where the disease state was categorized as not NED, 133 events (53.8%) had HE4 concentration less than 140 pmol/L. This reflects the proportion of events with a "false positive" result (i.e. HE4 ≤ 140 pmol/L). The proportion of NED events when the HE4 concentration is less than or equal to 140 pmol/L was 57.4% (179 of 312).

A table of the performance for distinguishing between NED and active disease for some other HE4 cutoffs is as follows:

HE4 cutoff	Sensitivity	Specificity
50	48%	88%
75	74%	66%
100	86%	57%
125	95%	52%
140	98%	49%

The implication from the table is that clinicians may choose a lower HE4 concentration. As a result, the sensitivity will decrease while the specificity will increase. Since there is no clinically accepted HE4 concentration for subjects with no evidence of disease, the clinician may choose from this range of HE4 in order to choose the sensitivity or specificity important to the evaluation of subjects in their care.

The following table combines all subject visits for all disease states and HE4 cutoff values;

		NED	responding	stable	progression	total
HE4 ≤ 140		179	44	49	40	312
HE4 > 140	< -14%	0	10	9	6	25
	-14% ≤ x ≤ 14%	0	6	11	15	32
	> 14%	4	1	18	38	61
	total	183	61	87	99	430

Note from the table that at the majority of subject visits, the HE4 concentration was less than 140 pmol/L, though the 4 disease states were present. Of the 312 subject visits where the HE4 value was less than 140 pmol/L, 179 visits occurred where the clinical state of the subject was NED (57.4%; 95% confidence interval based on positive likelihood ratio 54.5% to 60.2%). Of 61 subject visits where the %change in HE4 was greater than 14% and the HE4 concentration remained above 140 pmol/L, 38 visits occurred where the clinical statue of the subject was progression (62.3%; 95% confidence interval based on the positive likelihood ratio from 50.9% to 72.5%). The association of disease states and changes in HE4 of elevation above or below 140 pmol/l was significantly more than random association. Thus in the above table the HE4 result (both as a %change or elevation above 140 pmol/L) is statistically associated with disease states.

The sponsor makes the following conclusions regarding the clinical study:

1. The primary endpoint of the trial was to demonstrate that clinical disease progression is associated with an increase in HE4 from the preceding visit of at least 14%. The study found that a total concordance probability of at least 62.8% demonstrate that an increase in HE4 of at least 14% is significantly associated with clinical cancer progression, the 95% confidence interval for the total concordance being 68.9% to 76.7%. Clinicians may wish to use other cutoffs to reflect their preferences in the tradeoff between sensitivity and specificity; a table of the impact of different cutoffs is given.
2. The HE4 concentration itself provides an indication of whether disease is still present compared to when no evidence of disease is clinically apparent. Nearly all (98%) of the subjects with no evidence of disease had HE4 levels below 140 pmol/L, as opposed to 51% of those with active disease manifestations. From this analysis, 54% of subjects without NED (133/247) had HE4 concentration below 140 pmol/l.

The clinical study supports the claimed indication for use of the HE4 assay.

4. Clinical cut-off**:**
There is currently no clinically accepted cut-off for use in monitoring cancer progression in epithelial ovarian cancer subjects with this assay. A cut-off for use in this situation could be a percentage change from a previously determined value and would be expected to correlate with the clinical state at the time of assay and clinical evaluation. The labeling contains various percentage changes in HE4 from a previous value as determined in the clinical study. The assay performance characteristics (sensitivity and specificity) at each cut-off are indicated for use in a serial surveillance monitoring situation where clinical outcome is categorized as cancer progression/non-progression.
5. Expected values/Reference range:
A total of 1626 samples were tested from which results were used to determine the distribution of the ARCHITECT HE4 concentration values in various benign and malignant conditions as well as in apparently healthy women. Apparently healthy women were used to establish the normal and reference ranges for the ARCHITECT HE4 assay. The study was performed on single-point human serum samples obtained from commercial vendors, specimen banks or leftover specimens from a previous sponsor study. These samples were collected under an Institutional Review Board (IRB) approved protocol or considered remnant samples and the IRB concluded that informed consent was not needed. The following information was collected when the specimen was obtained:
    - Age or Date of Birth
    - Gender
    - Collection Date
    - Menopausal status
    - Ethnicity

Samples were not collected when stored at 2-8°C for more than 3 days or at room temperature for more than 24 hours; when there was no sample information; when >5

Freeze/Thaw Cycles had been undergone; when obvious microbial contamination was present.
HE4 serum levels were measured in single replicates following the appropriate quality control procedures and manufacturer's instructions for use specified in the proposed package insert. Sample testing was performed on 1 lot of reagents using 4 ARCHITECT *i*2000$_{SR}$ Systems. One calibration curve on each instrument was used. The HE4 Controls were tested and evaluated for each run to determine assay validity. Sample dilutions for HE4 values above the range of the assay were diluted using the on-board instrument dilution protocol at a dilution of 1:10.
The following summarizes the samples tested:
400 apparently healthy females
519 women with benign gynecologic diseases
314 women with epithelial ovarian cancer
49 women with other benign diseases
50 women with endometrial cancer
50 women with bladder cancer
50 women with gastrointestinal cancer
50 women with breast cancer
50 women with lung cancer
44 women with Congestive Heart Failure
50 pregnant women
1626 samples total
A summary of the specific diagnosis for the "Benign Gynecological Diseases" and "Other Benign Diseases" are summarized in the tale below.

Benign Gynecological Conditions	Sample Size (n)	Other Benign Conditions	Sample Size (n)
Cystadenomas / Adenofibromas	119	Anemia	4
Ovarian Cysts (all types)	100	Elevated Progesterone	1
Sex Cord Stromal Tumors	20	Fibrocystic Changes	28
Germ Cell Tumors	28	Fibrocystic Tumors	16
Endometriosis /Endometriomas	96	Total	49
Normal	6		
Uterine Fibroids	70		
Other Conditions	80		
Total	519		

The "Other Benign Gynecological Diseases" category included a total of 80 subjects with the following conditions: abscess (3), adhesions (2), Brenner tumor (2), Candida (1), Chlamydia (9), cervical dysplasia (1), cystic mesothelioma (1), fibrocystic breast (2), HPV (6), hydrosalpinx (7), leiomyomas (10), myomas (6), pelvic inflammatory disease (1), polycystic ovaries (27), and salpingitis (2).
Apparently healthy women were ~95% Caucasian and 50% premenopausal with a mean age of 49 years (range 14 – 93 years). To determine the upper limit of the normal range for the HE4 assay, calculations of the 95th and 97.5th percentiles were

performed for the overall set of apparently healthy females in the study as well as by menopausal status. The pre- and postmenopausal HE4 normal value cutoffs were chosen based on the 95th and 97.5th percentile of the population tested. The sponsor recommends that each laboratory establish its own reference value for the population of interest. The summary values are as follows:

Cohort	Sample Size	95th Percentile	97.5th Percentile
Apparently Healthy Females	400	98.2 pmol/L	145.2 pmol/L
Premenopausal	210	65.3 pmol/L	90.6 pmol/L
Postmenopausal	190	125.9 pmol/L	163.0 pmol/L

The sponsor has chosen 70 pmol/L as the upper limit of normal for pre-menopausal women and 140 pmol/L for post menopausal women. The data are accurate and appropriate.

Assay values from serum specimens of subjects with various benign and malignant conditions, along those from all the apparently healthy females are presented to indicate the distribution of values in these conditions. A summary table of subjects with benign conditions is as follows:

	All Healthy Females	Pregnant Females	All Benign Gynecologic	All Non-Gynecologic Benign	Congestive Heart Failure (all post-menopausal)
n	400	50	519	49	44
Median Age (years)	50	27	47	79	78
% Caucasian	95%	50%	87%	100%	84%
% African American	2%	40%	7%	0%	7%
median HE4 (pmol/L)	39.7	35.1	52.9	88	98.5
2.5th Percentile	23.5	21.5	22.6	39.4	38.9
97.5th Percentile	145.2	58.0	173.3	675.9	345.1
%< 70 pmol/L	89%	98%	84%	33%	34%
% < 140 pmol/L	97%	100%	96%	82%	70%

A summary table of subjects with cancer conditions is as follows:

	ovarian cancer		endometrial cancer		breast cancer	
	pre-menopausal	post-menopausal	pre-menopausal	post-menopausal	pre-menopausal	post-menopausal
n	67	247	12	38	19	31
Median Age (years)	44	61	39	64	45	64
% Caucasian	99%	94%	75%	87%	84%	65%
% African American	1%	2%	0%	3%	16%	3%
median HE4 (pmol/L)	87	201.9	64.6	69.5	70.1	70.1
2.5th Percentile	27.7	31.3	32	40.6	44.3	44.3
97.5th Percentile	1039.7	5463.8	158.6	1279.9	3214.4	3214.4
% > 70 pmol/L	60%	76%	42%	50%	42%	71%
% > 140 pmol/L	34%	59%	17%	24%	16%	29%

	Colon cancer	Lung cancer	Bladder cancer
n	50	50	50
Median Age (years)	67	76	72
% Caucasian	94%	98%	100%
% African American	4%	0%	0%
median HE4 (pmol/L)	62.7	97.2	127.9
2.5th Percentile	38.2	43.5	39.4
97.5th Percentile	251	259.7	1062.1
% > 70 pmol/L	32%	78%	76%
% > 140 pmol/L	14%	34%	48%

The sponsor shows a more comprehensive table of the percent of subjects with several categories of HE4 concentrations for each cohort of subjects.

	Number of Subjects	0.0 – 70.0 pmol/L	70.1 – 140 pmol/L	140.1 – 500 pmol/L	500.1 – 1500 pmol/L	> 1500 pmol/L
APPARENTLY HEALTHY	400					
Females (Premenopausal)	210	95.7%	2.9%	1.4%	0%	0%
Females (Postmenopausal)	190	81.58%	13.68%	4.74%	0%	0%
BENIGN CONDITIONS	662					
Benign Gynecological Disease	519	83.82%	11.94%	3.47%	0.58%	0.19%
Premenopausal	306	91.18%	6.86%	1.31%	0.33%	0.33%
Postmenopausal	213	73.24%	19.25%	6.57%	0.94%	0%
Pregnancy	50	98.0%	2.0%	0%	0%	0%
Other (non-gyn) Benign Disease	49	32.65%	48.98%	12.24%	6.12%	0%
Congestive Heart Failure	44	34.09%	36.36%	27.27%	2.27%	0%
CANCER	564					
Ovarian Cancer	314	27.71%	18.79%	27.71%	16.88%	8.92%
Premenopausal	67	40.30%	25.37%	19.40%	13.43%	1.49%
Postmenopausal	247	24.29%	17.0%	29.96%	17.81%	10.93%
Endometrial Cancer	50	52.0%	26.0%	18.0%	2.0%	2.0%
Breast Cancer	50	38.0%	38.0%	12.0%	6.0%	6.0%
Gastrointestinal Cancer	50	68.0%	18.0%	14.0%	0%	0%
Lung Cancer	50	22.0%	44.0%	34.0%	0%	0%
Bladder Cancer	50	24.0%	28.0%	34.0%	10.0%	4.0%
TOTAL	1626					

The data are reasonable.

N. Proposed Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Part 809.10.

O. Conclusion:

The submitted information in this premarket notification is complete and supports a substantial equivalence decision.