theranos	LDT Validation Report	Theranos TIBC Assay	Rev:
		CL-RPT-14070	1
Description	Validation Report for Modified Siemens Assay of Total Iron Binding Capacity (TIBC) in Lithium Heparin Plasma		
Originator: Curtis Schneider		Date: 10/15/2013	

Validation of Modified Siemens Total Iron Binding (TIBC) Capacity Assay

Author(s):

Signature:	Date: 1/2/14
Name: Paul Patel, Ph.D.	Title: Team Lead, General Chemistry

Reviewer(s):

Signature:	Date:
Name:	Title:
Signature:	Date: 1/3/2014
Name: Daniel Young, Ph.D.	Title: Vice President

Approver(s):

Signature:	Date: 1/2/14
Name: Adam Rosendorff, M.D.	Title: Laboratory Director

9/19/15

Sunil S. Dhawan M.D.

Total Iron Binding Capacity Plasma Assay

Overview

Ingested iron is absorbed primarily from the intestinal tract and is temporarily stored in the mucosal cells as Fe(III)-ferritin. Ferritin provides a soluble protein shell to encapsulate a complex of insoluble ferric hydroxide-ferric phosphate. On demand, iron is released into the blood and transported as Fe(III)-transferrin.

Transferrin is the primary plasma iron transport protein, which binds iron strongly at physiological pH. Transferrin is generally only 25% to 30% saturated with iron. The additional amount of iron that can be bound is the unsaturated iron-binding capacity. The total ion binding capacity (TIBC) can be indirectly determined using the sum of serum iron (Total Iron) and UIBC. Knowing the molecular weight of the transferrin and that each molecule of transferring can bind 2 atoms of iron, TIBC and transferrin concentration is interconvertible.

Percent saturation (100 x serum iron/TIBC) is usually normal or increased in persons who are iron deficient, pregnant, or are taking oral contraceptive medications. Person with chronic inflammatory processes, hemochromatosis or malignancies generally display low transferrin.

I. Method Principle

The Siemens Total Iron Binding Capacity (TIBC) method uses two reagents in a sequential process that is monitored spectrophotometrically.

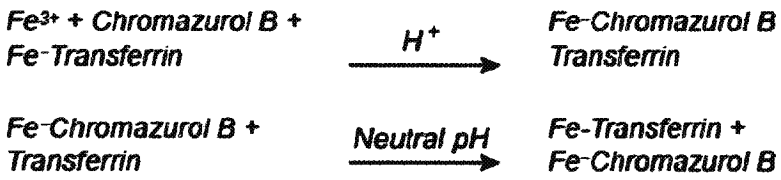
Step 1

a. The system adds reagent R1, an acidic buffer containing an iron-binding dye (Chromoazurol B) and ferric chloride to the plasma sample.
b. The low pH of R1 releases iron from transferrin
c. The iron forms a colored complex with the dye at the end of this first step. The colored complex represents both the plasma iron and excess iron already present in R1.

Step 2

a. The system then adds Reagent R2, a neutral buffer.
b. The pH shifts, resulting in a large increase in affinity of transferrin for iron
c. The plasma transferrin rapidly binds the iron by abstracting it from the dye-iron complex.
d. The observed decrease in absorbance of the color dye-iron complex is directly proportional to the total iron-binding capacity of the plasma sample.

Reaction Equation



II. Definitions and Abbreviations

The following definitions and abbreviations are used in this document and related documents and attachments:

a. **Accuracy:** Accuracy is defined by CLSI as the closeness of agreement between a test result and an accepted reference value. Trueness, a related CLSI term, is the closeness of agreement between the average of a number of replicate measured quantity values and a reference quantity value.

b. **Analyte:** Component represented in the name of a measurable quantity. The closely related term measurand is defined as the particular quantity subject to measurement.

c. **Analytical sensitivity:** There are several alternative uses of this term. Most commonly, and for the purposes of this Validation Plan, it is used interchangeably with limit of detection. It is also used to describe the ability of an analytical method to assess small variations of the concentration of an analyte, such as the slope of the calibration curve (IUPAC).

d. **Analytical specificity:** Ability of a test or procedure to correctly identify or quantify an entity, including in the presence of interfering substance(s) or phenomena.

e. **Calibration:** Set of operations that establish, under specified conditions, the relationship between values of quantities indicated by a measuring instrument or measuring system, or values represented by a material measure or a reference material, and the corresponding values realized by standards. Under CLIA, calibration refers to the process of testing and adjusting an instrument, kit, or test system, to provide a known relationship between the measurement response and the value of the substance being measured by the test procedure (42 CFR 493.1217).

f. **Calibrator:** A substance, material, or article intended to be used to establish the measurement relationships of a diagnostic medical device.

g. **CLIA:** Clinical Laboratory Improvement Amendments of 1988. Congressional legislation that defined and requires specific quality assurance practices in clinical laboratories.

h. **CLSI:** Clinical and Laboratory Standards Institute.

i. **Coefficient of Variation:** The ratio of the standard deviation to the average, often multiplied by 100 and expressed as a percentage, abbreviated as %CV.

j. **Colorimetry:** A technique used to determine the concentration of colored compound(s) in solution.

k. **Interfering substance:** A substance or quantity thereof that is not the measurand but that affects the result of the measurement.

l. **IUPAC:** International Union of Pure and Applied Chemistry

m. **LDT:** Laboratory –developed Test.

n. **Linearity:** Linearity is the ability of a quantitative analytical method to provide results that are directly proportional to the concentrations of an analyte in test samples, within a given measuring interval. It is an important parameter to confirm when evaluating an analytical method because it verifies correct interpolation of results between points.

o. **LMR:** Lower end of the measuring range is the lowest level at which defined conditions, including all stated characteristic of the method, are met.

p. **LoB:** Limit of Blank is the highest value in a series of results on a sample that contains no analyte.

q. **LoD:** Limit of Detection is the lowest amount of analyte in a sample that can be detected with stated probability, although perhaps not quantified as an exact value.

r. **LoQ:** When used without a prefix, the Limit of Quantitation is the lowest actual concentration at which an analyte is reliably detected and at which uncertainty of the test result is less than or equal to the goal set by the manufacturer or laboratory. The term may also be used with prefixes L for lower (LLOQ) and U for upper (ULOQ), respectively. Note: LoB < LoD ≤ LoQ.

s. **Matrix:** All components of a material system, except the analyte. A specimen matrix is the biological milieu in which an analyte exists (e.g., plasma, serum, urine, or other body fluids).

t. **Measuring Interval (reportable range; analytical measurement range or AMR):** A measuring interval consists of all numeric values between the lower and upper numeric values for which a method can produce quantitative results suitable for clinical use. Where applicable, a linearity study is frequently used to establish or verify the measuring interval that can be reported for a measurement method. Alternatively, the lower limit of the measuring interval may be assigned as the LoQ (LLOQ).

u. **Precision:** Precision is the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. It is usually expressed numerically in terms of standard deviation (SD) or percent Coefficient of Variation (%CV).

v. **Reference interval:** The interval between and including two reference limits. It is common practice to define a reference limit so a stated fraction of the reference values is less than or equal, or greater than or equal, to the respective upper or lower limit.

w. **SOP:** Standard Operating Procedure.

x. **Spectrophotometry:** The quantitative measurement of the transmission (or reflection) properties of a material as a function of wavelength.

y. **Testing System:** The entirety of the testing process, including instrument, sample, reagents, supplies, and procedures. Personnel are sometimes included in the definition.

III. Pre-clinical Validation

a. Analytical Measurement Range

i. Limits of Blank, Detection and Quantitation

The limits of blank, detection, and quantitation were determined to be 38.7 µg/dL, 51.8 µg/dL and 58.7 µg/dL respectively.

Limit of blank

CLSI guideline EP17-A section 4.3.1

Level	Number of samples	N	Mean	SD
Blank	1	20	23.0	9.6

Alpha	5%
Parametric LoB	38.7

Limit of detection

CLSI guideline EP17-A section 4.3.2

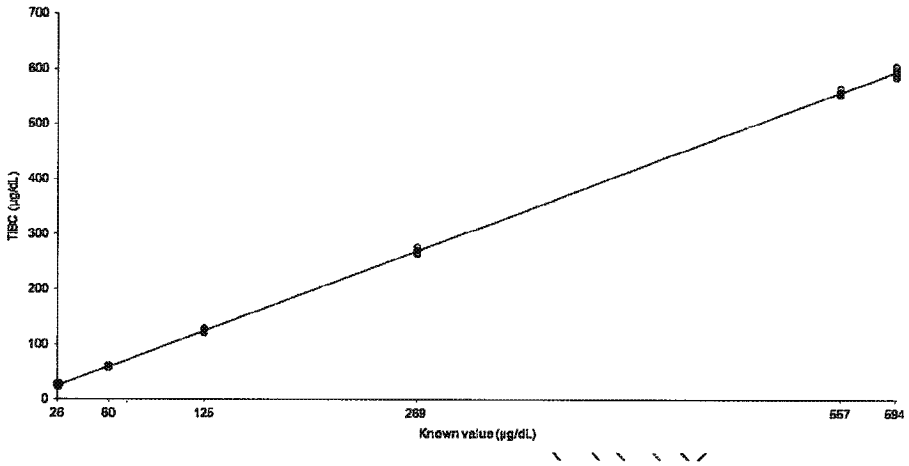
Level	Number of samples	N	Pooled SD
Low	1	20	7.9

Beta	5%
Parametric LoD	51.8

The lower limit of quantitation has been established at 58.7 µg/dL (13.5% CV and 97.8% recovery).

ii. Linearity

The Analytical Measurement Range (AMR) including linear measurement interval has been determined for TIBC in plasma. This method is linear from 26.0 – 594.0 µg/dL within the 10% allowable non-linearity in this interval.



A linear relationship fits the data better than a nonlinear relationship over the measuring interval.

Level	Mean	Linear fit	Nonlinear fit	Nonlinearity	Allowable nonlinearity
1	26.0	26.2	-	-	2.6
2	60.0	60.2	-	-	6.0
3	125.4	125.2	-	-	12.5
4	269.8	269.3	-	-	26.9
5	557.4	557.5	-	-	55.7
6	594.4	594.5	-	-	59.4

A linear relationship fits the data better than a nonlinear relationship over the measuring interval.

Nonlinearity is less than allowable nonlinearity: 10% upto 600µg/dL then 10%.
Performance requirement verified over the measuring interval.

b. Analytical Specificity

The analytical specificity for this assay was determined by testing the effect of hemoglobin (100 mg/dL), bilirubin (10 mg/dL) and triglycerides (200 mg/dL) on plasma samples spiked with the interferents and then compared with un-spiked controls. TIBC interference testing was performed at 502 µg/dL. Non-interference was defined as the mean result from testing of spiked samples within 10% of the mean of the un-spiked samples. Recoveries were within 99.0% to 101.0% (see table below).

Table 1. Interference Testing For Total Iron Binding Capacity.

Analyte (µg/dL)	% Recovery		
	Interferent		
	Bilirubin (10 mg/dL)	Hemoglobin (100 mg/dL)	Triglycerides (400 mg/dL)
TIBC	99	100.0	101.0

*N/A Not Applicable

No significant interference was observed.

c. Precision

Level = Level 1

Number of observations	80
Number of runs	40
Number of days	20
Runs per day	2
Replicates per run	2
Mean	248.2

	SD	95% CI	CV	Allowable Total SD
Repeatability	0.0	0.0 to 0.0	0.0%	-
Between-run	5.1		2.1%	-
Between-day	2.1		0.8%	-
Within-laboratory	5.6	4.5 to 7.2	2.2%	49.6

Imprecision is less than allowable total imprecision: 20% upto 500µg/dL then 20%.

Level = Level 2

Number of observations	80
Number of runs	40
Number of days	20
Runs per day	2
Replicates per run	2
Mean	324.9

	SD	95% CI	CV	Allowable Total SD
Repeatability	0.0	0.0 to 0.0	0.0%	-
Between-run	8.0		2.5%	-
Between-day	3.3		1.0%	-
Within-laboratory	8.7	7.1 to 11.2	2.7%	65.0

Imprecision is less than allowable total imprecision: 20% upto 500µg/dL then 20%.

Level = Level 3

Number of observations	80
Number of runs	40
Number of days	20
Runs per day	2
Replicates per run	2
Mean	400.9

	SD	95% CI	CV	Allowable Total SD
Repeatability	0.0	0.0 to 0.0	0.0%	-
Between-run	4.4		1.1%	-
Between-day	4.0		1.0%	-
Within-laboratory	5.9	4.8 to 7.9	1.5%	80.2

Imprecision is less than allowable total imprecision: 20% upto 500µg/dL then 20%.

The percent CV reported as zeros in the above precision summary are most likely a consequence of rounding the values in StatisPro.

IV. Clinical Validation

a. Method Comparison with Predicate (Accuracy/Comparability)

To test the accuracy of the assay on the Theranos System, forty four (44) unique patient samples were screened on the predicate method (Siemens, Advia) and on the Theranos method. One (1) sample was excluded as an outlier (mean absolute difference greater than 4). Using the predicate method fifteen (15) values were within the reference range (250 - 450 µg/dL) and twenty eight (28) were above the reference range. Based on the results of the data examination, either a simple linear regression or alternative procedures were used to estimate expected (average) bias and the confidence interval of expected bias at the desired medical decision level(s) as per CLSI guidance EP09-A2. StatisPro was used for bias calculations. These estimates were compared with internal criteria to judge the acceptability of the Theranos method. Each sample was run in duplicate on the predicate, and the average used for comparison to the Theranos method. Some samples were stored before analysis on both methods. If the confidence interval for the predicted bias includes the defined

acceptable bias or if the acceptable bias is greater than the higher limit of the confidence interval of the predicted bias, then the data do not show that the bias of the Theranos method is different from the acceptable bias or there is a high probability (97%) that the predicated bias is acceptable, respectively. The acceptable bias at each medical decision level was determined based on the total allowable error (TEa) minus the measured precision at the level closest to that decision level. Total allowable error (TEa) was taken from American Proficiency Institute (API) peer proficiency testing criteria or CLIA proficiency testing criteria for acceptable analytical performance, as printed in the Federal Register February 28, 1992;57(40):7002-186, when available. The TEa for Total Iron Bing Capacity is 20%. The table below shows the allowable bias and precision at 3 levels (values shown in parentheses) and the corresponding closest medical decision limits.

Table 2. Allowable Bias and Precision at the Medical Decision Levels

Medical Decision Levels (µg/dL)	**240 (248)**	**322 (325)**	**500 (401)**
Precision (%)	**0.8**	**1.0**	**1.0**
Allowable Bias (%)	**19.2**	**19.0**	**19.0**

Method comparison was performed with un-spiked (normal) and spiked samples. A comparison plot between Siemens-Advia (predicate) and the Theranos method shows spiked samples having a better agreement between the two methods as demonstrated by the distribution of data along the identity line. The un-spiked (normal) samples show a distinctive positive bias with the Theranos method suggesting that the normal samples respond differently between the two methods. In order to estimate the magnitude of the systematic error twenty new normal samples were tested by the two methods. See graph below showing a method comparison with normal samples (figure 1.).

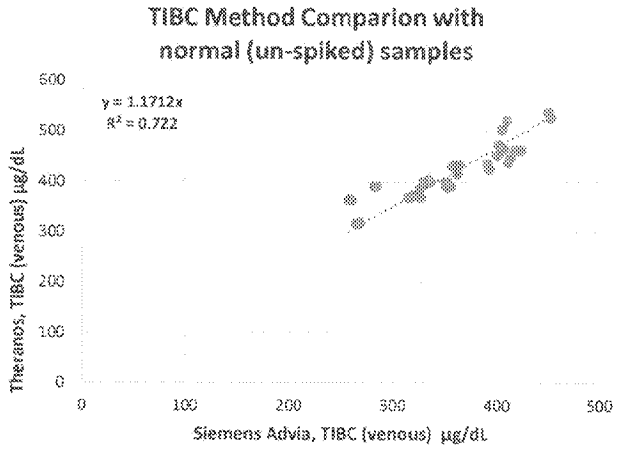


Figure 1. Graph showing method comparison of normal samples: Theranos method versus Predicate method (Siemens-Advia). A total of N= 19 normal samples were tested (1 sample was excluded as an outlier).

The range of values obtained for normal samples with the predicate and Theranos test methods were 257 - 452 µg/dL and 316 - 536 µg/dL respectively. Using the method comparison regression coefficient from the plot (figure 1.) a correction (slope: 1.17) was applied to assign new Theranos values to all normal samples (≤ 452 µg/dL by predicate method). A method comparison plot of predicate versus Theranos with the re-assigned normal (corrected) values and spike values is shown in the figure 2 below.

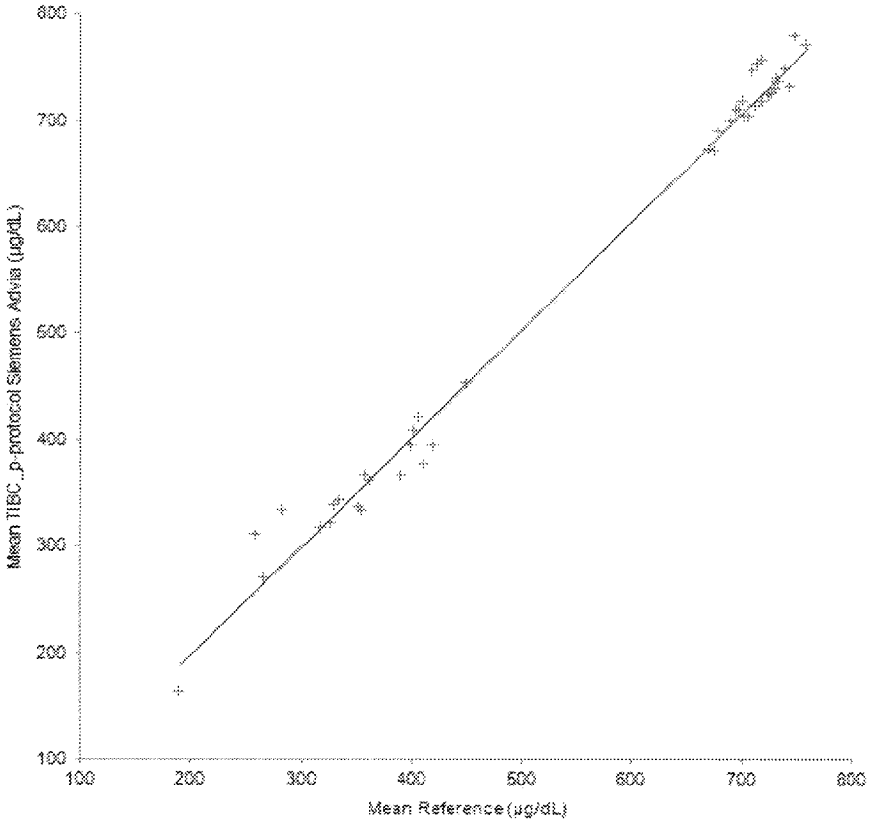


Figure 2. Graph showing Theranos method versus Predicate Method (Siemens Advia).

Simple linear regression was used to establish a slope, intercept and an r^2. The slope, intercept and clinical correlation were determined to be 1.02, -3.95 and 0.99 respectively.

Comparability

CLSI guideline EP09-A2-IR section 7

Level ID	Value	Difference	SE	95% CI	Allowable difference
	240.000000	-0.0397042	5.41373664	-0.9729672 to 10.89355	48.0000000
	322.000000	1.3027818	4.40910327	-7.6015829 to 10.207146	64.4000000
	500.000000	4.2169590	2.93580638	-1.7120223 to 10.145940	100.0000000

Difference is less than allowable bias: 20% upto 500µg/dl. then 20%.

The difference between the two methods is not greater than the allowable difference. The performance requirement is not verified.

b. Matrix Comparison

A method comparison study was performed to evaluate the effect of Lithium Heparin as an anti-coagulant in primary collection tubes on TIBC recovery using the predicate Siemens Advia method. Serum samples were used as reference. Matched serum and plasma samples were drawn and tested from each healthy volunteers and plasma recoveries were compared to the corresponding serum recoveries for each donor. Matched serum and plasma spiked samples were also included to fully cover the measuring range. A total of fifty three samples were tested.

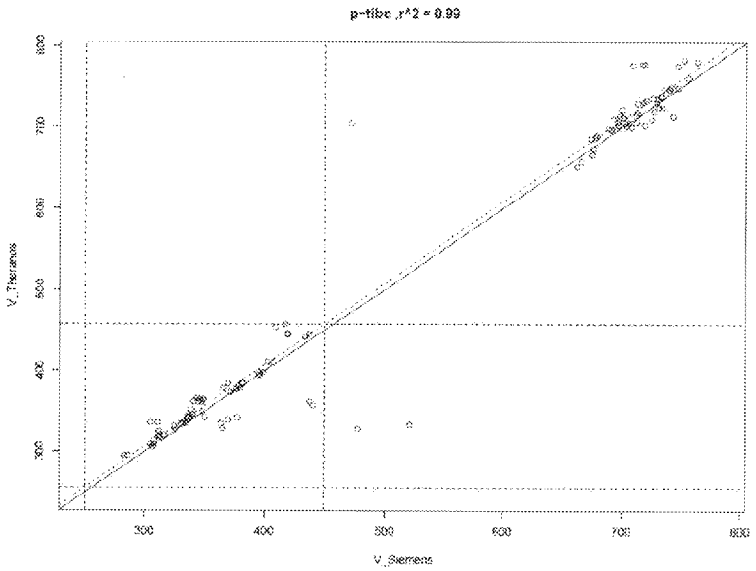


Figure 3. Graph showing method comparison of matched serum versus plasma samples with the Predicate Method (Siemens-Advia).

Simple linear regression was used to establish a slope, intercept and an r^2. The slope, intercept and clinical correlation were determined to be 1.00, 4.50 and 0.99 respectively.

c. Transference and Verification of Reference Interval (Venous)

Reference ranges were modified by applying the regression equation to the lower and upper reference limits of existing reference interval to generate a new reference range. New reference ranges were verified with venous samples using twenty (20) new normal subjects. For a reference range to pass verification, 95% of values should fall within the upper and lower reference limits and 5% or fewer values fall outside of the upper and lower reference limits (CLSI-C28A3). For venous verification 20 (100%) values fell within the new reference range and 0 (0%) values fell outside the new reference range. See graphs (figure 4.) below for venous samples verification.

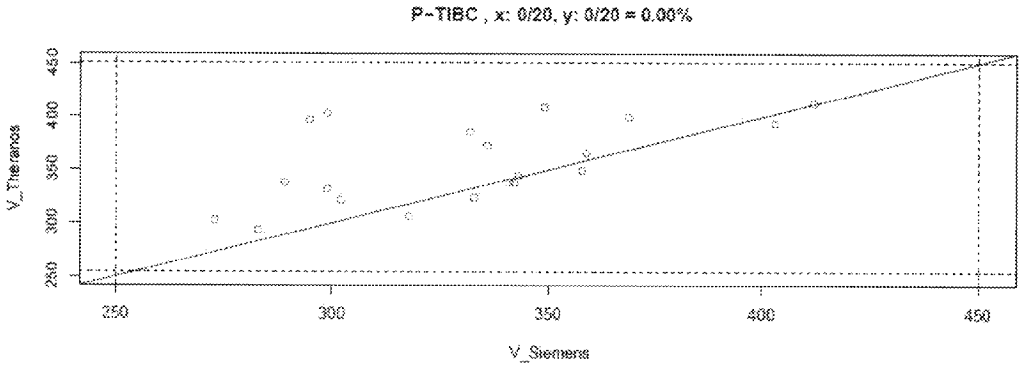


Figure 4. Graph showing venous sample reference range verification.

d. Verification of Reference Interval with Finger Stick Samples

New reference ranges were also verified with venous matched finger sticks (Lithium heparin) from a total of twenty (20) new normal subjects. The finger stick samples were collected in a Theranos blood collection device (BCD) configured with two separate Lithium heparin vessels. For finger stick verification 20 values (100 %) fell within the new reference range and no value (0%) fell outside the new reference range. See graphs (figure 5.) below for finger stick samples verification.

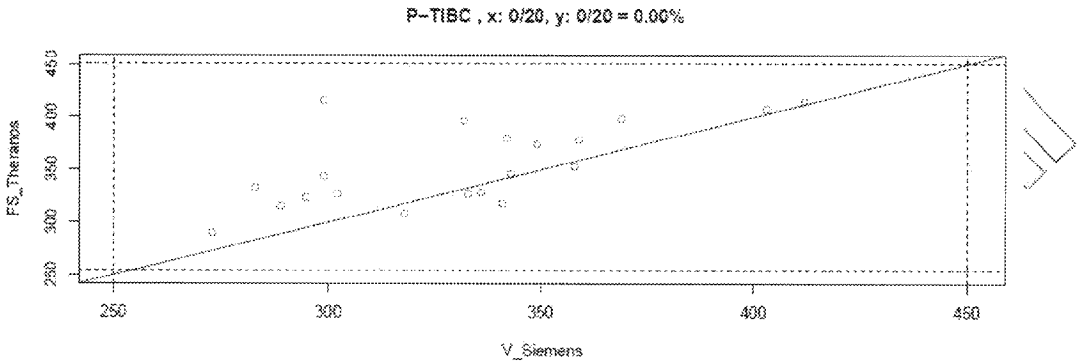


Figure 5. Graph showing Finger stick sample reference range verification.

To determine the corrected TIBC (Theranos) concentration the following correction will be required for samples ≤ 536 µg/dL: measured Theranos value/slope (1.17). The new reference range for finger stick total iron binding capacity was determined to be 251.0 – 455.0 µg/dL.

VI. Stability

a. Reagents

On-board Reagent Stability

System	Stability with Reagent Container Inserts*
ADVIA 1200	7 days
ADVIA 1650/1800	7 days
ADVIA 2400	7 days

For all systems, unopened reagents are stable until the expiration date printed on the product label when stored at 2°C - 8°C. Do not freeze the reagents.

For additional details, refer to the Methods Introduction section of the system-specific Operator's Guide.

b. Sample

Plasma samples for total iron binding capacity analysis are stable for 2 weeks at 2-8 °C, or at least 90 days at -20 °C.

c. Calibrators

Siemens Special Chemistry Calibrators should be stored at 2-8 °C, protected from light, and are stable until the expiration date on the vial label. Opened reconstituted calibrators are stable for 7 days, except for acid phosphatase, which is stable for at least 2 days.