theranos	LDT Validation Report	Theranos Total Bilirubin Assay	Rev:
		CL RPT-14054	1
Description	Validation Report for Modified Siemens Assay of Total Bilirubin in Lithium Heparin Plasma		
Originator: Curtis Schneider		Date: 09/24/2013	

Validation of Modified Siemens Total Bilirubin Assay

Author(s):

Signature:	Date: 11/7/13
Name: Paul Patel, Ph.D.	Title: Team Lead, General Chemistry

Reviewer(s):

Signature:	Date:
Name:	Title:
Signature: SURAJ SAKSENA	Date: 11/7/13
Name: Daniel Young, Ph.D.	Title: Vice President

Approver(s):

Signature:	Date: 11/7/13
Name: Adam Rosendorff, M.D.	Title: Laboratory Director

9/19/11

Sunil S. Dhawan M.D.

Total Bilirubin Plasma Assay

I. Overview

Bilirubin is one of the most commonly tested analytes to assess liver function. Approximately 85% of the total bilirubin produced is derived from the heme moiety of hemoglobin, while the remaining 15% is produced from the red blood cell precursors destroyed in the bone marrow and from the catabolism of other heme-containing proteins. After production in peripheral tissues, bilirubin is rapidly taken up by hepatocytes where it is conjugated with glucuronic acid to produce mono- and diglucuronide, which are excreted in the bile.

A number of inherited and acquired diseases affect one or more of the steps involved in the production, uptake, storage, metabolism, and excretion of bilirubin. Bilirubinemia is a frequent and direct result of these disturbances.

Jaundice can occur as a result of problems at each step in the metabolic pathway. Disorders may be classified as those due to increased bilirubin production (eg, hemolysis and ineffective erythropoiesis), decreased bilirubin excretion (eg, obstruction and hepatitis), and abnormal bilirubin metabolism (eg, hereditary and neonatal jaundice).

The most commonly occurring form of unconjugated hyperbilirubinemia is that seen in newborns and referred to as physiological jaundice. Elevated unconjugated bilirubin in the neonatal period may result in brain damage (kernicterus). Treatment options are phototherapy and, if severe, exchange transfusion.

The rare genetic disorders, Crigler-Najjar syndromes Type I and Type II, are caused by a low or absent activity of bilirubin UDP-glucuronyl-transferase. In Type I, the enzyme activity is totally absent, the excretion rate of bilirubin is greatly reduced and the serum concentration of unconjugated bilirubin is greatly increased. Patients with this disease may die in infancy owing to the development of kernicterus.

The increased production of bilirubin that accompanies the premature breakdown of erythrocytes and ineffective erythropoiesis results in hyperbilirubinemia in the absence of any liver abnormality. In hepatobiliary diseases of various causes, bilirubin uptake, storage, and excretion are impaired to varying degrees. Thus, both conjugated and unconjugated bilirubin is retained and a wide range of abnormal serum concentrations of each form of bilirubin may be observed. Both conjugated and unconjugated bilirubin are

increased in hepatitis and space-occupying lesions of the liver; and obstructive lesions such as carcinoma of the head of the pancreas, common bile duct, or ampulla of Vater.

II. Method Principle

The bilirubin is oxidized by vanadate at about pH 2.9 to produce biliverdin. In the presence of the detergent and the vanadate, both conjugated (direct) and unconjugated bilirubin are oxidized. This oxidation reaction causes the decrease in the optical density of the yellow color, which is specific to bilirubin. The decrease in optical density at 451/545 nm is proportional to the total bilirubin concentration in the sample. The concentration is measured as an endpoint reaction.

Reaction Equation

III. 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. Method accuracy is used in a different sense by the American Association of Pharmaceutical Scientists where it is expressed as percent relative error (%RE). 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.

IV. Pre-clinical Validation

a. Analytical Measurement Range

i. Limits of Blank, Detection and Quantitation

The limits of blank, detection, and quantitation determined were 0.00 mg/dL, 0.00 mg/dL, and 0.20 mg/dL (100% recovery), respectively.

Limit of blank

CLSI guideline EP17-A section 4.3.1

Level	Number of samples	N	Mean	SD
Blank	1	20	0.00	0.00

Alpha	5%
Parametric LoB	0.00

Limit of detection

CLSI guideline EP17-A section 4.3.2

Level	Number of samples	N	Pooled SD
Low	1	19	0.00

Beta	5%
Parametric LoD	0.00

Limit of quantitation

CLSI guideline EP17-A section 5.1

Level	Number of samples	N
Low	1	19

Bias	0.00
Pooled imprecision	0.00
95% total error	0.00
Allowable error	-
LoQ	0.00

Level	Sample	n	Assigned value	Mean	Median	SD	CV
Blank	1	20	0	0.00	0.00	0.00	-
Low	1	19	0.2	0.20	0.20	0.00	0.0%

ii. Linearity

The Analytical Measurement Range (AMR) including linear measurement interval has been determined by Siemens. Refer to the **Analytical Range** section of the manufacturer product information insert for additional details.

b. Analytical Specificity

The analytical specificity for this assay was determined by observing the effects of hemoglobin (100 mg/dL) and triglycerides (400 mg/dL) on the recovery of total bilirubin (0.5 mg/dL) in spiked plasma samples. No significant interference (NSI) was determined if the mean analyte concentration of an interferent-spiked sample reported within 10% of the mean analyte concentration of an un-spiked sample. Recoveries of total bilirubin in the presence of hemoglobin and triglycerides were each 100% (see table below).

Analyte:	% Recovery in the presence of each interferent:	
Total Bilirubin (mg/dL)	Hemoglobin (100 mg/dL)	Triglycerides (400 mg/dL)
0.5	100*	100*

* NSI observed at interferent level tested.

c. Precision

Level = L1

Number of observations	78
Number of runs	39
Number of runs excluded	1
Number of days	20
% of days with 1 run	5%
Runs per day	2
Replicates per run	2

CLSI guideline EP05-A2 section 10.4 recommends a minimum of 40 runs

Mean	0.879

	SD	95% CI	CV
Repeatability	0.028	0.023 to 0.036	3.2%
Between-run	0.000		0.0%
Between-day	0.136		15.5%
Within-laboratory	0.139	0.107 to 0.201	15.9%

Level = L2

Number of observations	80
Number of runs	40
Number of days	20
Runs per day	2
Replicates per run	2
Mean	2.184

	SD	95% CI	CV
Repeatability	0.030	0.024 to 0.038	1.4%
Between-run	0.000		0.0%
Between-day	0.103		4.7%
Within-laboratory	0.107	0.082 to 0.153	4.9%

Level = L3

Number of observations	**78**
Number of runs	**39**
Number of runs excluded	**1**
Number of days	**20**
% of days with 1 run	**5%**
Runs per day	**2**
Replicates per run	**2**

CLSI guideline EP05-A2 section 10.4 recommends a minimum of 40 runs.

Mean	**6.630**		
	SD	**95% CI**	**CV**
Repeatability	**0.107**	**0.088 to 0.137**	**1.6%**
Between-run	**0.000**		**0.0%**
Between-day	**0.064**		**1.0%**
Within-laboratory	**0.124**	**0.106 to 0.150**	**1.9%**

V. *Clinical Validation*

a. Method Comparison with Predicate (Accuracy/Comparability)

To test the accuracy of the assay on the Theranos System, 37 unique patient samples were screened on the predicate method (Siemens, Advia) and on the Theranos method. Using the predicate method twenty (27) values were within the reference range (0.3— 1.2 mg/dL), none were below the reference range, and ten (10) 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 Bilirubin is 20%. The table below shows the allowable bias and precision at 2 levels (values shown in parentheses) and the corresponding closest medical decision limits.

Table 1. Allowable Bias and Precision at the Medical Decision Levels

Medical Decision Levels (mg/dL)	**1.4 (0.9)**	**2.5 (2.2)**
Precision (%)	**15.5**	**4.7**
Allowable Bias (%)	**4.5**	**15.3**

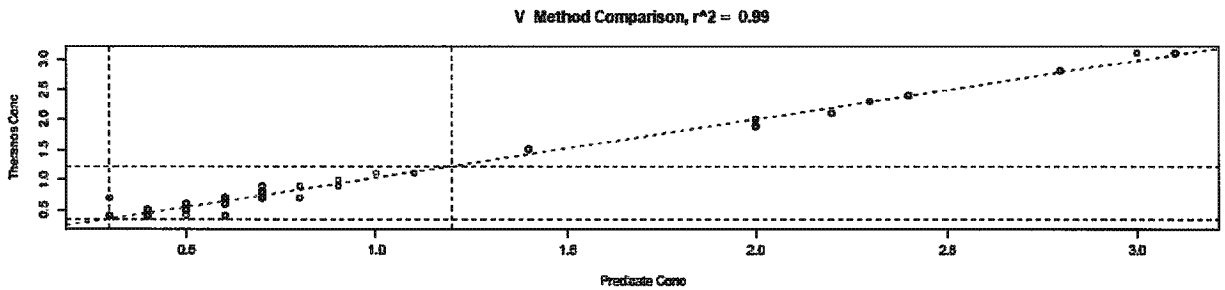

Figure 1. Graph showing Theranos method versus Predicate Method (Siemens Advia). Simple linear regression was used to establish a slope, intercept and an r2. The slope, intercept and clinical correlation were determined to be 0.971, 0.062 and 0.99 respectively.

Comparability

CLSI guideline EP09-A2-IR section 7

Level ID	Value	Difference	SE	95% CI	Allowable difference
	1.4	0.02	0.013	-0.01 to 0.05	0.18
	2.5	-0.01	0.025	-0.06 to 0.04	0.31
	20.0	-0.51	0.287	-1.09 to 0.08	2.51

Difference is less than allowable bias: 12.55%.

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

b. 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 eight (28) 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. For venous verification 27 (96.4%) values fell within the new reference range and 1 (3.6%) values fell outside the new reference range. See graphs below for venous samples verification.

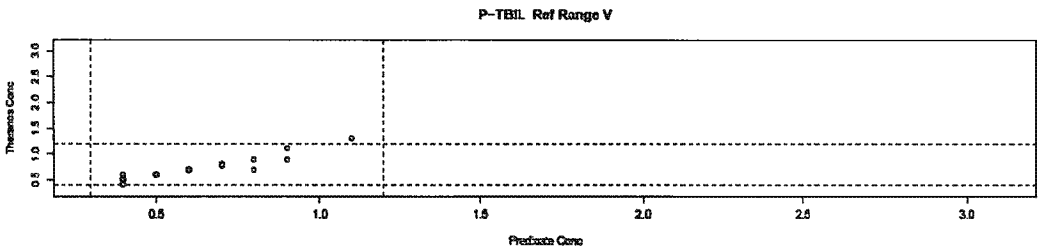

Figure 2. Graph showing venous sample reference range verification.

c. Verification of Reference Interval with Finger Stick Samples

New reference ranges were also verified with venous matched finger sticks from twenty six (26) new normal subjects. For finger stick verification 26 (100%) values fell within the new reference range and 0 (0%) values fell outside the new reference range. See graphs below for finger stick samples verification.

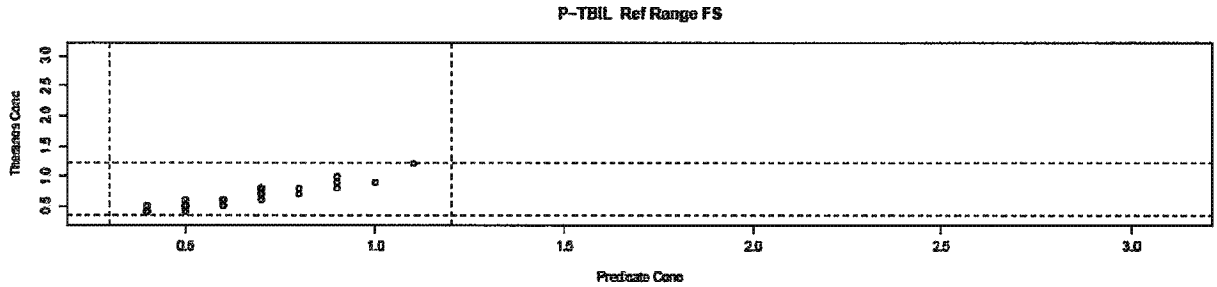

Figure 3. Graph showing Finger stick sample reference range verification.

Theranos.Analyte	Anti-coagulant	Existing Reference Range (mg/dL)		New Reference Range (mg/dL)	
		Reference Range (low)	Reference Range (High)	Transfered RR (low)	Transfered RR (high)
Total Bilirubin	Heparin	0.3	1.2	0.4	1.2

The new reference range for finger stick Total Bilirubin was determined to be 0.4 – 1.2 mg/dL.

VI. Stability

a. Reagents

On-board Reagent Stability

System	Stability
ADVIA 1200	60 days
ADVIA 1650/1800	60 days
ADVIA 2400	60 days

For all systems, unopened reagents are stable until the expiration date printed on the product label when stored at 2° - 35°C. Do not freeze reagents.

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

b. Sample

Plasma samples for Total Bilirubin analysis are stable for 2 weeks at 2-8 °C, or at least 90 days at -20 °C when stored protected from light.

c. Calibrators

Siemens Chemistry Calibrators should be stored at 2-8 °C, protected from light, and are stable until the expiration date on the vial label. Opened calibrators are stable for 48 hours, except for total and direct, which are stable for 8 hours.

REVISION HISTORY

Revision Level	Effective Date	Initiator	ECO Number
A	11/06/2013	A. Rosendorff	CL ECO-00117

Section Number	Description and Justification of Changes
All	Initial Release