EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR
Banyan Brain Trauma Indicator

DECISION MEMORANDUM

A. DEN Number:

DEN170045

B. Purpose for Submission:

De Novo request for evaluation of automatic class III designation of the Banyan Brain Trauma Indicatory (BTI)

C. Measurands:

Ubiquitin C-terminal hydrolase-L1 (UCH-L1) and glial fibrillary acidic protein (GFAP)

D. Type of Test:

Manual enzyme-linked immunosorbent assay, semi-quantitative

E. Applicant:

Banyan Biomarkers, Inc.

F. Proprietary and Established Names:

Banyan BTI

G. Regulatory Information:

1. Regulation section:

   21 CFR §866.5830

2. Classification:

   Class II (special controls)

3. Product code:

   QAT

4. Panel:

   Immunology (82)

H. Indications for use:

1. Indications for Use:

   The Banyan BTI is an in vitro diagnostic chemiluminescent enzyme-linked immunosorbent assay (ELISA). The assay provides a semi-quantitative measurement of the concentrations of ubiquitin C-terminal hydrolase-L1 (UCH-L1) and glial fibrillary acidic protein (GFAP) in human serum, and is used with the Synergy 2 Multi-mode Reader.

   The assay results obtained from serum collected within 12 hours of suspected head injury are used, along with other available clinical information, to aid in the evaluation of patients 18 years of age and older with suspected traumatic brain injury (Glasgow Coma Scale score 13-15). A negative assay result is associated with the absence of acute intracranial lesions visualized on a head CT (computed tomography) scan.

   The Banyan BTI is for prescription use only.

2. Special conditions for use statement(s):

   For prescription use only

   For in vitro diagnostic use

3. Special instrument requirements:

   BioTek Instruments, Inc., Synergy 2 Multi-mode Reader, model SL.

I. Device Description:

The Banyan BTI consists of two kits, one for the UCH-L1 assay components and one for GFAP assay components. Each kit is packaged individually in a box and consists of the following: 96-well microtiter strip plate, each well coated with mouse monoclonal UCH-L1 antibody or mouse monoclonal GFAP capture antibody (1 plate); UCH-L1 or GFAP calibrators (1 vial); UCH-L1 or GFAP calibrator diluent (1 vial, 4 mL); UCH-L1 or GFAP control 1 (1 vial); UCH-L1 or GFAP control 2 (1 vial); mouse monoclonal UCH-L1 or mouse monoclonal GFAP detection antibody (1 vial, 0.23 mL); UCH-L1 or GFAP detection antibody diluent (2 vials, 6.5 mL per vial for UCH-L1 or 1 vial, 14 mL for GFAP); ready-to-use assay diluent (2 vials, 5 mL per vial for UCH-L1 or 1 vial, 10 mL for GFAP), chemiluminescent substrate solution A (2 vials, 4.5 mL per vial) and solution B (2 vials, 4.5 mL per vial); a wash tablet and four adhesive plate seals. Components within the same kit are intended to be used together. In each kit, sufficient quantities of each component are provided to test samples from up to 30 patients. Each kit is stored at 2°C to 8°C until ready for use.

J. Standard/Guidance Document Referenced:

CLSI EP05-A3, Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline–Third Edition

CLSI EP6-A: Evaluation of Linearity of Quantitative Measurement Procedures; A Statistical Approach; Approved Guideline

CLSI EP07-A2: Interference Testing in Clinical Chemistry; Approved Guideline–Second Edition

CLSI EP09-A2: Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline

CSLI EP17-A2: Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline- Second Edition

CLSI EP25-A, Evaluation of Stability of In Vitro Diagnostic Reagents

CLSI C28-A3c: Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline- Third Edition

K. Test Principle:

The Banyan BTI is a semi-quantitative test that determines UCH-L1 and GFAP concentrations in human serum using two separate chemiluminescent sandwich immunoassays. Samples (clinical specimens, controls, or standards) are pipetted into wells of a microplate that are coated with a monoclonal antibody that can capture the target (UCH-L1 or GFAP) protein, thereby immobilizing the target to the well. After washing away unbound protein, a second UCH-L1 or GFAP mouse monoclonal antibody that has been conjugated to the enzyme horseradish peroxidase (HRP) is added to the well. After washing away unbound HRP-conjugated antibody, the chemiluminescent substrate is added to the well. The HRP enzyme catalyzes a specific reaction with the chemiluminescent substrate, which produces light that is detected at 300 nm to 700 nm with the Synergy 2 Multi-mode Reader, a 96-well plate-based luminometer. The amount of light generated is proportional to the amount of conjugated antibody in the well. The results from the wells containing standards are used to create a dose-response curve to quantify the amount of UCH-L1 or GFAP in the sample.

The reader GEN5 IVD software installed on a personal computer processes the data generated and determines the validity of results for each specimen tested. If the results for a specimen do not meet predetermined validity criteria, the result is invalid. For specimens with valid results, the concentration (average of duplicate results) of the analyte is reported along with the categorization of the semi-quantitative result relative to the cutoff value, if the concentration is between the pre-established lower and upper limits of quantitation. If the concentration is above the upper limit of quantitation, a concentration will not be reported, but the result will be reported as ‘Above’ the cutoff value. If the concentration is below the lower limit of quantitation, a concentration will not be reported, but the result will be reported as ‘Below’ the cutoff value.

Results of the UCH-L1 and GFAP assays are reported separately by the reader but must be combined to determine the Banyan BTI result (Positive, Negative, or Not Reportable) because the reader does not report the final result. The Banyan BTI result must be interpreted by the laboratory professional according to the table below:

Banyan UCH-L1 Kit Result (relative to cutoff)[A]	**Banyan GFAP Kit Result** (relative to cutoff)[B]	**Banyan BTI Result**
Below	Below	Negative
Below	Above	Positive
Above	Below	Positive
Above	Above	Positive
Invalid or No Result	Below	Not Reportable[C]
Invalid or No Result	Above	Positive
Below	Invalid or No Result	Not Reportable[C]
Above	Invalid or No Result	Positive
Invalid or No Result	Invalid or No Result	Not Reportable[C]

[A] Above = UCH-L1 concentration is equal to or above 327 pg/mL; Below = UCH-L1 concentration is below 327 pg/mL
[B] Above = GFAP concentration is equal to or above 22 pg/mL; Below = GFAP concentration is below 22 pg/mL
[C] Clinical samples with Invalid Results or No results that yield a Not Reportable Banyan BTI result may be retested once to obtain an interpretable Negative or Positive result

The test outcome for a specimen is considered 'Positive' when the measured serum concentration of either UCH-L1 and/or GFAP (in pg/mL) is above its clinical cutoff value. Conversely, the test outcome is considered 'Negative' when the measured serum concentrations of both UCH-L1 and GFAP are below their respective cutoff values.

Specimens that fail to meet the predetermined validity criteria yield 'Invalid Result' outcomes. Specimens with 'Invalid Result' outcomes for either protein that lead to a 'Not Reportable' qualitative result may be retested for that protein once to obtain a 'Negative' or 'Positive' result. The qualitative results for the retested samples are re-interpreted according to the above table.

A 'No Result' outcome occurs when a run is aborted. The reader aborts the run to prevent sensor damage from light generated from a sample that contains a very high concentration of target analyte. In this situation, the resulting report lists 'No Result' for all samples located in the microtiter plate row containing the sample that caused the abort, and for all samples in subsequent rows on the microtiter plate. When this occurs, the reader software generates an

error message that includes an error code and the number of the well that triggered the run abort. The assay result (relative to cutoff) for the specimen that triggered the run abort is considered Above, and the qualitative result for the applicable sample is then interpreted using the above table. All other specimens with 'No Result' outcomes that lead to a 'Not Reportable' qualitative result may be retested once to obtain a 'Negative' or 'Positive' result using a new assay kit. The qualitative results for the retested samples are re-interpreted according to the table above.

L. Performance Characteristics:

1. Analytical performance:

   All results met the manufacturer's pre-determined acceptance criteria.

   a. Precision/Reproducibility:

   Semi-quantitative precision: A study was conducted per CLSI guideline EP05-A3 to evaluate the within-laboratory precision of the UCHL-1 kit and the GFAP kit. Two separate panels, each consisting of five human sera samples with levels of UCH-L1 or GFAP that cover the measuring range of the respective kits, were tested at one site, using one instrument and one reagent lot, over the course of 20 days. The panel members were made of pooled human sera from healthy volunteers. Higher panel members were spiked with one or more positive clinical specimens (i.e., sera containing high level of endogenous UCH-L1 or GFAP) from subjects with a severe traumatic brain injury (TBI) or with the recombinant UCH-L1 or purified native GFAP protein as detailed in the tables below. Each panel member was tested each day with two runs per day and two replicate measurements per sample per run for a total of 80 replicates per sample. The results are summarized in the tables below for each kit.

UCH-L1 Kit										
Panel member	Mean (pg/mL)	N	Within-run		Between-run		Between-day		Within-laboratory	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV
1	196.9	80	8.4	4.3	0.0	0.0	2.8	1.4	8.9	4.5
2[A]	268.9	80	8.5	3.2	1.8	0.7	5.7	2.1	10.4	3.9
3[B]	390.6	80	9.1	2.3	5.1	1.3	6.8	1.7	12.5	3.2
4[C]	1163.7	80	32.5	2.8	0.0	0.0	0.0	0.0	32.5	2.8
5[C]	2118.5	80	46.2	2.2	27.0	1.3	3.6	0.2	53.7	2.5

[A] Pooled sera from normal donors and a patient with severe TBI
[B] Pooled sera from normal donors and patients with severe TBI
[C] Pooled sera from normal donors were spiked with recombinant UCH-L1 antigen

GFAP Kit										
Panel member	Mean (pg/mL)	N	Within-run		Between-run		Between-day		Within-laboratory	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV
1	6.8	79[C]	0.5	7.4	0.6	8.8	0.0	0.0	0.7	10.3
2	21.6	79[C]	0.6	2.8	1.1	5.1	0.0	0.0	1.3	6.0
3[A]	36.4	80	0.7	1.9	1.3	3.6	0.2	0.5	1.5	4.1
4[B]	138.5	80	5.0	3.6	13	0.9	2.4	1.7	5.6	4.0
5	252.7	80	5.2	2.1	5.7	2.3	4.4	1.7	8.9	3.5

[A] Pooled sera from normal donors and patients with severe TBI
[B] Pooled sera from normal donors and a patient with severe TBI were spiked with purified native GFAP antigen
[C] A total of two measurements were invalid and excluded from analysis

Qualitative precision: A total of 80 replicates of each panel member were performed to evaluate qualitative precision. The panel members were identical to those in the semi-quantitative precision study above; however, % correct call was calculated for each panel member based on the number of positive BTI results. Results are summarized in the tables below for each kit.

UCH-L1 Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	191.2	80	0/80	100
2[B]	282.2	80	0/80	100
3[C]	396.9	80	80/80	100
4[C]	1188.3	80	80/80	100
5[C]	2107.6	80	80/80	100

[A]Below cutoff; [B]Cut-off-20%; [C]Above cutoff

GFAP Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	9.3	80	0/80	100
2[B]	23.2	79[D]	28/79	65
3[C]	40.4	80	80/80	100
4[C]	146.5	80	80/80	100
5[C]	261.0	80	80/80	100
[A]Below cutoff; [B]Cut-off +20%; [C]Above cutoff: [D] Only three measurements instead of four on day 3				

Semi-quantitative internal reproducibility: A study was conducted per the CLSI guideline EP05-A3 to evaluate the effects of two major sources of variability: operator and reagent lot. The same two sample panels evaluated in the within-laboratory precision study were tested by three operators in four replicates per run and one run per day for five days with three reagent lots of each kit for a total of 180 replicates per panel member. Three Synergy 2 Multi-mode Readers were used randomly across operators throughout the study. The results are summarized in the tables below for each kit.

UCH-L1 Kit												
Panel member	Mean (pg/mL)	N	Within-run		Between-day		Between-operator		Between-lot		Within-laboratory	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV	SD	%CV
1	189.2	180	7.7	4.1	11.2	5.9	0.0	0.0	7.7	4.1	15.6	8.2
2[A]	289.3	180	10.0	3.5	13.2	4.6	0.0	0.0	15.1	5.2	22.4	7.7
3[B]	385.5	180	12.8	3.3	14.8	3.8	3.2	0.8	9.1	2.4	21.8	5.7
4[C]	1122.3	180	36.1	3.2	20.4	1.8	22.3	2.0	25.8	2.3	53.8	4.8
5[C]	2086.8	180	70.5	3.4	36.4	1.7	76.7	3.7	55.3	2.6	123.5	5.9
[A] Pooled sera from normal donors and a patient with severe TBI [B] Pooled sera from normal donors and patients with severe TBI [C] Pooled sera from normal donors were spiked with recombinant UCH-L1 antigen												

GFAP Kit												
Panel member	Mean (pg/mL)	N	Within-run		Between-day		Between-operator		Between-lot		Within-laboratory	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV	SD	%CV
1	8.2	180	0.6	7.3	0.8	9.8	0.0	0.0	0.8	9.8	1.3	15.9
2	25.3	180	1.0	4.0	1.3	5.1	0.3	1.2	1.9	7.5	2.5	9.9
3[A]	38.6	180	1.4	3.6	1.7	4.4	0.6	1.6	1.1	2.8	2.5	6.5
4[B]	141.6	180	5.5	3.9	3.7	2.8	0.0	0.0	2.1	1.5	6.9	4.9
5	247.4	180	6.2	2.5	6.8	2.7	3.5	1.4	5.4	2.2	11.3	4.6

[A] Pooled sera from normal donors and patients with severe TBI
[B] Pooled sera from normal donors and a patient with severe TBI were spiked with purified native GFAP antigen

Qualitative internal reproducibility: A total of 180 replicates of each panel member were tested to determine internal qualitative reproducibility. The panel members were identical to those in the semi-quantitative internal reproducibility study above; however, % correct call was calculated for each panel member based on the number of positive BTI results. Results are summarized in the tables below for each kit.

UCH-L1 Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	191.2	180	0/180	100
2[B]	282.2	180	9/180	95
3[C]	396.9	180	180/180	100
4[C]	1188.3	180	180/180	100
5[C]	2107.6	180	180/180	100

[A]Below cutoff; [B]Cut-off-20%; [C]Above cutoff

GFAP Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	9.3	180	0/180	100
2[B]	23.2	180	172/180	96
3[C]	40.4	180	180/180	100
4[C]	146.5	180	180/180	100
5[C]	261.0	180	180/180	100
[A]Below cutoff; [B]Cut-off +20%; [C]Above cutoff				

Semi-quantitative external reproducibility: To evaluate site-to-site reproducibility, two separate panels, each consisting of five samples with analyte levels spanning the measuring range of each kit, were tested in replicates of five at three U.S. sites. The panel members were pooled serum samples from healthy volunteers mixed with pooled serum samples containing high levels of endogenous UCH-L1 and GFAP from mTBI patients, except for UCH-L1 panel members 4 and 5, that consisted of a combination of endogenous and recombinant proteins to achieve the desired target UCH-L1concentrations. At each site, one operator performed one run on each of five non-consecutive days using one reagent lot of each kit and one Synergy 2 Multi-mode Reader. Combining all sites, a total of 75 replicates were obtained per panel member. The results are summarized in the tables below for each kit:

UCH-L1 Kit										
Panel member	Mean (pg/mL)	N	Within-run		Between-day		Between-site		Total	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV
1[A]	203.8	75	5.9	4.2	7.2	3.5	8.6	2.9	12.7	6.2
2[A]	305.1	75	7.9	4.5	10.6	3.5	13.7	2.6	19.1	6.2
3[A]	413.8	75	11.7	3.9	12.3	3	16.3	2.8	23.5	5.7
4[B]	1216.6	75	30.4	3.1	27.8	2.3	37.6	2.5	55.8	4.6
5[B]	2245.9	75	58.6	3.8	54.9	2.4	85.6	2.6	117.3	5.2
[A] Pooled sera from normal donors and patients with severe TBI [B] Pooled sera from normal donors and patients with severe TBI were spiked with recombinant UCH-L1 antigen										

GFAP Kit										
Panel member	Mean (pg/mL)	N	Within-run		Between-day		Between-site		Total	
			SD	%CV	SD	%CV	SD	%CV	SD	%CV
1[A]	8.1	73[B]	0.4	5.2	0.5	5.9	0.5	6.1	0.8	9.9
2[A]	25.4	75	1	3.8	0.8	3	0.1	0.6	1.3	4.9
3[A]	37.9	75	1.1	2.9	0.8	2.2	0.5	1.4	1.5	3.9
4[A]	146.8	74[B]	3.0	2.1	3.7	2.5	2.0	1.4	5.2	3.5
5[A]	258.7	75	8.0	3.1	1.1	0.4	4.6	1.8	9.3	3.6

[A] Pooled sera from normal donors and patients with severe TBI
[B] A total of three measurements were invalid and excluded from analysis

Qualitative external reproducibility: A total of 75 replicates of each panel member were performed to determine external qualitative reproducibility. The panel members were identical to those in the semi-quantitative external reproducibility study above; however, % correct call was calculated for each panel member based on the number of positive BTI results. Results are summarized in the tables below for each kit.

UCH-L1 Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	191.2	75	0/75	100
2[B]	282.2	75	6/75	92
3[C]	396.9	75	75/75	100
4[C]	1188.3	75	75/75	100
5[C]	2107.6	75	75/75	100

[A]Below cutoff; [B]Cut-off -20%; [C]Above cutoff

GFAP Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
1[A]	9.3	75	0/75	100
2[B]	23.2	75	75/75	100
3[C]	40.4	75	75/75	100

GFAP Kit				
Panel member	Mean (pg/mL)	Total replicates	Qualitative agreement	
			Number of positive BTI results	% Correct call
4[C]	146.5	75	75/75	100
5[C]	261.0	75	75/75	100
[A]Below cutoff; [B]Cut-off +20%; [C]Above cutoff				

b. *Linearity/assay reportable range:*

Linearity: Linearity and recovery characteristics of the UCH-L1 kit and GFAP kit were evaluated per the CLSI guideline EP6-A. A pool of human sera containing a high concentration of analyte was combined with a pool of human sera containing a low concentration of analyte to create a mid-pool that was then combined with the high and low pools to obtain additional dilutions for a total of 11 sample pools spanning across the measuring range of each assay. Each sample pool was tested in four replicates using one UCH-L1 kit lot and one GFAP kit lot. The observed values were evaluated against the calculated values and standard linear regression was performed. The results are summarized in the tables below for each kit:

UCH-L1 Kit				
Pool	Expected (pg/mL)	Measured mean (pg/mL)	%CV	%Recovery
1	73.3	73.3	2.9	100
2	139.3	138	5.3	99.1
3	244.9	244.9	3.3	100
4	304.3	300.7	2.1	98.8
5	363.6	361.6	4.1	99.4
6	601.2	592.0	2.6	98.5
7	865.2	861.4	1.5	99.6
8	1393.1	1286.0	3.6	92.3
9	2184.9	2115.2	2.5	96.8
10	2448.9	2336.0	3.7	95.4
11	2712.8	2712.8	3.5	100

Linearity was demonstrated throughout the measurable range of the UCH-L1 kit. The claimed linear range for the UCH-L1 kit is from 80 pg/mL (lowest calibrator) to 2560 pg/mL (highest calibrator). The regression equation for the linear range is $y=0.97x-4.7$, $R^2=0.99$. The 95% confidence interval (CI) of the slope was 0.95–1.00; the 95% CI of the intercept was −43.2–33.8.

GFAP Kit				
Pool	**Expected** (pg/mL)	**Measured mean** (pg/mL)	**%CV**	**%Recovery**
1	8.6	8.6	8.7	100
2	14.1	13.8	7.0	97.8
3	17.7	17.6	2.1	99.4
4	26.8	26.8	3.1	100.0
5	44.8	42.9	1.7	95.8
6	80.9	72.9	1.7	90.1
7	117.0	102.5	1.0	87.6
8	189.2	167.5	0.8	88.5
9	297.5	277.7	1.6	93.3
10	333.6	319.3	0.9	95.7
11	369.7	369.7	2.6	100.0

Linearity was demonstrated throughout the measuring range of the GFAP kit. The claimed linear range for the GFAP kit is from 10 pg/mL (lowest calibrator) to 320 pg/mL (highest calibrator). The regression equation for the linear range is y=0.92x +0.14, R^2=0.99. The 95% CI of the slope was 0.88–0.95; the 95% CI of the intercept was –3.88–4.17.

Hook effect: To show that a clinical specimen with very high concentrations of analyte would not cause false negative results ('Below' the cutoff) hook effect studies were performed. A sample with GFAP concentration of 100 ng/mL (100,000 pg/mL) was tested with the GFAP kit. A 'No Result' was observed due to the run abort which is interpreted as 'Above' the cutoff and is, therefore, not a negative result. For the UCH-L1 kit, a test sample with UCH-L1 concentration of 200.0 ng/mL (200,000 pg/mL) was tested. This sample also triggered the run to abort and generated a 'No Result' and not a false negative.

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

i) Traceability and value assignment: As there are no international reference standards for GFAP and UCH-L1, commercially available antigens for GFAP (HyTest GFAP antigen) and UCH-L1 (Origene-UCH-L1 antigen) were used to create Reference Standards for each kit. The UCH-L1 and GFAP Reference Standards were aliquoted and stored at ≤ -70°C. In addition, Quality Control (QC) Reference Panels consisting of three concentrations of analyte in the low, middle and high regions of the measuring range of each kit were manufactured from human serum pools,

aliquoted, and then stored at ≤ -70°C. The nominal antigen concentration of each QC Reference Panel member, UCH-L1 Reference Standard and GFAP Reference Standard was established using two development (pre-pilot) lots of each kit. New lots of UCH-L1 or GFAP antigen are value assigned against the appropriate Reference Standard by assaying dilutions of the antigen in a previously released kit lot and substituting the calibrators of that kit lot with dilutions of the Reference Standard to form a Standard Curve.

Calibrators and Controls*:* For both kits, the calibrators and controls are prepared from antigen that has been value assigned against the Reference Standard. The GFAP calibrators are prepared at six different levels with native protein purified from the human brain and assigned values ranging from 10–320 pg/mL. The UCH-L1 calibrators are prepared at six different levels from human UCH-L1 recombinant protein and assigned values ranging from 80–2560 pg/mL. The nominal values for the controls are assigned using the released calibrators. The calibrators and controls are spiked at the appropriate concentration in the Calibrator Diluent and lyophilized for storage at 2–8 ºC. Lot-to-lot consistency is achieved by adjusting the reconstitution volume of the calibrator so that the reported results for the QC Reference Panel members are within the predefined acceptable tolerance of the nominal assigned values.

ii) Assay kit stability:

Closed assay kit stability: An ongoing real-time stability study performed on three consecutively released lots of the UCH-L1 kit and three consecutively released lots of the GFAP kit supports a shelf-life claim of (b) (4) when stored at 2–8°C.

Open assay kit stability: An open-vial stability study was not conducted because both the UCH-L1 Kit and GFAP Kit are intended for single-use, and, therefore, no reagents are to be saved for testing or re-testing at a later time. The Banyan BTI Package Insert 'Materials Provided' states "*Do not reuse. Single-use only*"

iii) Sample stability and storage:

Nine serum samples spanning the measuring range of each kit (two close to the cutoff, two low, and five high values for UCH-L1 and one close to the cutoff, two low, and six high values for GFAP) were tested in five replicates to determine sample storage and freeze/thaw stability. The results support the following stability claims: If not tested immediately, serum separated from freshly collected venous blood can be stored at room temperature for up to 120 minutes. A minimum of 250µL of serum should also be stored at -80°C in the event of retesting due to a Not Reportable result. Serum samples should not be subjected to more than 5 freeze/thaw cycles. The Banyan BTI Package Insert 'Limitations of the Procedure' states that '*Inappropriate specimen collection, handling, preparation, storage, and transport may negatively impact assay performance*.'

d. Detection limit:

The lower limit of quantitation (LLoQ) and upper limit of quantitation (ULoQ) were established in accordance with the CLSI guideline EP17-A2. To establish the precision profile across the reportable range of each kit, two seven-member panels (one for each kit) spanning the UCH-L1 concentration range from 75 pg/mL to 3125.8 pg/mL and the GFAP concentration range from 7 pg/mL to 400 pg/mL were tested by three operators in four replicate measurements per run with one run per day for five days using three reagent lots. Three Synergy 2 Multi-mode Readers were used randomly across operators. A total of 60 measurements per panel member were obtained for each lot. The within-laboratory precision estimates of the three lots are summarized in the tables below for each kit:

UCH-L1 Kit												
Panel Member	**Lot 1**				**Lot 2**				**Lot 3**			
	Mean (pg/mL)	**N**	**Within-laboratory**		**Mean (pg/mL)**	**N**	**Within-laboratory**		**Mean (pg/mL)**	**N**	**Within-laboratory**	
			SD	%CV			SD	%CV			SD	%CV
1	84.6	60	6.5	7.7	67.5	60	5.4	8.0	66.0	60	6.5	9.8
2	190.9	60	13.0	6.8	196.5	60	12.2	6.2	180.2	60	16.0	8.9
3	305.0	60	17.4	5.7	288.8	60	15.4	5.3	273.9	60	17.1	6.2
4	381.8	60	20.0	5.2	396.9	60	19.4	4.9	377.9	60	20.4	5.4
5	1101.8	60	52.8	4.8	1152.6	60	38.8	3.4	1112.4	60	49.2	4.4
6	2089.0	60	109.6	5.2	2142.5	60	92.7	4.3	2029.8	60	126.5	6.2
7	2964.6	60	188.0	6.3	3003.1	60	175.6	5.8	2882.2	60	254.3	8.8

GFAP Kit												
Panel Member	**Lot 1**				**Lot 2**				**Lot 3**			
	Mean (pg/mL)	**N**	**Within-laboratory**		**Mean (pg/mL)**	**N**	**Within-laboratory**		**Mean (pg/mL)**	**N**	**Within-laboratory**	
			SD	%CV			SD	%CV			SD	%CV
1	6.2	60	0.5	8.1	8.0	60	0.8	10.0	6.0	60	0.4	6.7
2	7.8	60	0.8	10.3	9.2	60	1.2	13.0	7.6	60	1.0	13.2
3	27.3	60	1.6	5.9	23.4	60	1.9	8.1	25.3	60	1.6	6.3
4	39.5	60	2.2	5.6	38.9	60	2.9	7.5	37.3	60	1.6	4.3

GFAP Kit												
Panel Member	Lot 1				Lot 2				Lot 3			
	Mean (pg/mL)	N	Within-laboratory		Mean (pg/mL)	N	Within-laboratory		Mean (pg/mL)	N	Within-laboratory	
			SD	%CV			SD	%CV			SD	%CV
5	140.4	60	8.6	6.1	140.0	60	6.5	4.6	144.4	60	4.2	2.9
6	243.8	60	11.8	4.8	244.3	60	8.0	5.7	253.9	60	9.9	3.9
7	271.6	60	25.7	9.5	279.8	60	9.0	3.2	293.0	60	11.7	4.0

To estimate LLoQ and ULoQ of the UCH-L1 kit, precision profiles were generated using within-laboratory precision (as %CV) and mean measured concentrations of the panel members and fitted with a 3rd-order polynomial model. The %CVs at UCH-L1 concentrations of 79.0 pg/mL and 2561 pg/mL were calculated using each lot-specific 3rd-order polynomial equation. The resulting within-laboratory estimates (as %CV) for the three lots of UCH-L1 kit are summarized in the table below:

UCH-L1 Kit					
			%CV from precision profile		
LoQ	Concentration (pg/mL)	Specification	Lot 1	Lot 2	Lot 3
		%CV	%CV	%CV	%CV
LLoQ	79.0	15	7.8	8.1	10.2
ULoQ	2561.0	15	5.8	5.0	7.7

The %CV specification was met for all three lots and the LLoQ and ULoQ of the UCH-L1 kit were verified at 80.0 pg/mL and 2560.0 pg/mL, respectively.

To estimate the LLoQ and ULoQ of the GFAP kit, precision profiles were generated using within-laboratory precision (as SD) and mean measured concentrations of the panel members and fitted with a statistically significant linear regression model. The precision goal of 15%CV was converted to SD at GFAP concentrations of 9.0 pg/mL and 321.0 pg/mL and each lot-specific precision profile (as SD) was then solved for each concentration. The resulting within-laboratory estimates (as SD and %CV) for the three lots of the GFAP kit are summarized in the table below:

GFAP Kit									
				SD and %CV from precision profile					
LoQ	Concentration (pg/mL)	Specification		Lot 1		Lot 2		Lot 3	
		SD	%CV	SD	%CV	SD	%CV	SD	%CV
LLoQ	9.0	1.4	15	0.7	7.8	1	11.1	0.7	7.8
ULoQ	321.0	48.2	15	19.3	6	10.2	3.2	10.3	3.2

The SD and %CV specifications were met for all three lots and the LLoQ and ULoQ of the GFAP Kit were verified at 10.0 pg/mL and 320.0 pg/mL, respectively.

e. *Analytical specificity:*

i) *Endogenous interferences:*

Assay interference was assessed per the CLSI guideline EP07-A2 by testing pooled human serum spiked with recombinant UCH-L1 protein at a final concentration of 400 pg/mL and 1500 pg/mL or purified human native GFAP at a final concentration of 25 pg/mL and 70 pg/mL. Test samples were created by spiking with the potentially interfering substances listed in the table below at test concentrations recommended in CLSI EP07-A2. Control samples were spiked only with the appropriate solvent used to create the interfering substances panel. Test and control samples were analyzed in replicates of five, in one assay run, with one kit lot. The recovery was calculated by comparing measurements of the test and control samples.

Potential interfering substance	Test concentration
Endogenous	
Bilirubin (conjugated)	20 mg/dL
Triglycerides	3000 mg/dL
Hemoglobin	500 mg/dL
Rheumatoid factor	1000 IU
Human Albumin	12 g/dL
Human anti-mouse antibody (HAMA)	>160X[A]

Potential interfering substance	Test concentration
Exogenous	
Aspirin (Acetylsalicylic acid)	3.62 mmol/L
Acetaminophen	1324 µmol/L
Coumadin (Warfarin)	32.5 µmol/L
Ibuprofen	2425 µmol/L
Lopressor (Metoprolol)	18.7 µmol/L
Oxazepam	25 ng/mL
Plavix (Clopidogrel)	9 µg/mL
Cardene (Nicardipine)	400 ng/mL
(b) (4) Ethanol	5% (weight by volume)
Benzoylecgonine	37.5 ng/mL
d-Methamphetamine	125 ng/mL
EDDP[B]	125 ng/mL
Methadone	37.5 ng/mL
Methaqualone	37.5 ng/mL
Morphine	250 ng/mL
Phencyclidine	3.1 ng/mL
Propoxyphene	37.5 ng/mL
Secobarbital	25 ng/mL

[A] >160 times more activity than a known negative for ability to crosslink in a mouse system assay as characterized by (b) (4)

[B] 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine

For the UCH-L1 kit, hemoglobin, rheumatoid factor and HAMA showed a statistically significant difference from the control and % recovery exceeded the pre-specified acceptance limits. These substances were re-tested at lower doses. Results show no interference up to 62.5 mg/dL hemoglobin, 250 IU rheumatoid factor, and 40x HAMA. No interference was seen with the other interferents tested, as described in the table above.

For the GFAP kit, hemoglobin and HAMA showed a statistically significant difference from the control and % recovery exceeded the pre-specified acceptance

limits. Both substances were re-tested at lower doses. Results show no interference up to 62.5 mg/dL hemoglobin and 40x HAMA (i.e. forty times the activity of a known negative). No interference was seen with the other interferents tested, as described in the table above.

The Banyan BTI Package Insert 'Limitations of the Procedure' states that "*Levels of hemoglobin, rheumatoid factor, or HAMA exceeding the normal physiological concentration or activity in serum may have erroneously high Banyan UCH-L1 or Banyan GFAP results, potentially leading to a false positive Banyan BTI result.*'

ii) Cross-reactivity:

A panel comprised of proteins that have significant homology to either GFAP or UCH-L1 was evaluated for cross-reactivity. The GFAP panel of potential cross-reactants was spiked into the GFAP calibrator diluent and UCH-L3, the only protein identified with significant amino acid homology to UCH-L1 (~55%), was spiked into the UCH-L1 calibrator diluent at the concentrations listed in the table below:

Potential cross-reactant	**Test concentration[A]**	**N**	**Mean** (pg/mL)	**SD**	**%CV**
GFAP Kit					
Vimentin	354 ng/mL[1]	4	0	n/a	n/a
Desmin	127 ng/mL[2]	4	0	n/a	n/a
Peripherin	5 ng/mL[3]	4	0	n/a	n/a
Neurofilament light	68 pg/mL[4]	4	10.8	3.9	35.9%
Neurofilament medium	8.6 ng/mL[5]	4	0	n/a	n/a
Neurofilament heavy	77 ng/mL[6]	4	0	n/a	n/a
Keratin type II	10 ng/mL[7]	4	0	n/a	n/a
Internexin	77 ng/mL[B]	4	0	n/a	n/a
UCH-L1 Kit					
UCH-L3	354 ng/mL[C]	4	0	n/a	n/a

[A] Test concentrations of the identified cross-reactants (except for internexin and UCH-L3) are based on the highest concentration of each protein in the circulation as reported in: [1]Sun *et al.* (2010) *J Proteome Research* 9:1923; [2]Ma *et al*. (2009) *Mol Cell Proteomics* 8.8:1878; [3]Determined at MyBioSource, email communication; [4]Giottino *et al*. (2013) *PLOS One* 8: e75091; [5]Martinez-Morillo *et al*. (2015) *Clin Chem Lab Med* 53:1575; [6] Lu *et al.* (2015) *J Neurol Neurosurg Psychiatry* 86:565; [7]Sundstrom *et al.* (1990) *Int J Cancer* 46:604

[B] For internexin, the level of 77 ng/mL was chosen because this is the highest concentration of a neurofilament (i.e. neurofilament heavy) tested in the study

[C] For UCH-L3, the level of 354 ng/mL was chosen because this is the highest concentration of cross-reactant (Vimentin) tested with the GFAP kit

The spiked calibrator diluent samples were tested in replicates of four using one GFAP kit lot and one UCH-L1 kit lot. Results show that, in the absence of the assay-specific marker, there is significant (15.9%) cross-reactivity detected with the GFAP kit for neurofilament light at concentration of 68 pg/mL. No cross-reactivity in the UCH-L1 kit was observed when UCH-L3 was tested. As the serum concentration of neurofilament light in patients with neurodegenerative diseases such as Guillain-Barré syndrome, amyotrophic lateral sclerosis, Parkinson's disease, Alzheimer's disease, or Creutzfeldt-Jakob disease could exceed 68 pg/mL, the Banyan BTI Package Insert 'Limitations of the Procedure' states that "*Due to cross-reactivity of neurofilament light with the antibodies in the Banyan GFAP Kit, patients with neurodegenerative diseases such as Guillain-Barré syndrome, amyotrophic lateral sclerosis (ALS), Parkinson's disease, Alzheimer's disease, or Creutzfeldt-Jakob disease may have erroneously high Banyan GFAP results, potentially leading to a false-positive Banyan BTI result"*.

iii) Carryover/cross-contamination:

A study was performed to evaluate if cross-contamination and/or carryover occurs between samples in the plate wells during the assay procedure. Two test plates were run with an alternating pattern over all 30 available patient sample locations. The pattern consisted of two wells containing only the calibrator diluent (blank) followed by two wells containing the calibrator diluent spiked with GFAP purified from human brain lysate at a final concentration of 1 ng/mL (which is approximately 33 times the concentration of the highest assay calibrator). A control plate (consisting of all blank samples) was also tested. Testing was conducted using one GFAP kit lot and one Synergy 2 Multi-mode Reader. Results showed that carryover did not occur as none of the wells with blanks reported values higher than "<10 pg" (which is a value below the lowest calibrator). As UCH-L1 and GFAP measurements are performed in identical fashion, it is inferred that there is also no carryover with the UCH-L1 kit.

f. Assay cut-off:

The assay cutoffs were determined by analyzing a training dataset consisting of a completely independent population distinct from the subjects evaluated in the ALERT-TBI pivotal trial (see below). A total of 334 subjects (39.2% were female and 60.8% were male; mean age 48.3 years) with Glasgow Coma Scale (GCS) scores between 13–15 who had blood specimens collected within eight hours from the time of head injury were included in the training dataset. Of the 334 subjects, 102 had a positive adjudicated CT result, i.e., the subject had confirmed presence of an acute intracranial lesion per an independent Neuroimaging Review Committee. The K-fold cross-validation technique was used to aid in the selection of the optimal cutoff values for the two biomarkers. After performing 50 rounds of 10-fold cross-validations, the optimal cutoff values were selected as 327 pg/mL for the UCH-L1 kit and 22 pg/mL for the GFAP kit.

2. Comparison studies:

a. Method comparison with predicate device:

Not applicable

b. Matrix comparison:

Not applicable. Serum is the only claimed sample matrix. The Banyan BTI Package Insert 'Limitations of the Procedure' states that states "*Human serum specimens only are indicated for use. The use of plasma has not been validated and may adversely affect the test outcome*".

5. Clinical studies:

a. Clinical sensitivity and specificity:

The ability of Banyan BTI to assist physicians in determining the need for a computed tomography (CT) scan of the head in conjunction with other clinical information was evaluated in a prospective, multi-site pivotal trial (ALERT-TBI). This study enrolled subjects over the age of 18 who presented to a health care facility (HCF) or emergency department (ED) with suspected head injuries and GCS scores of 9–15 (mild-moderate TBI). Subjects had blood withdrawn within 12 hours of head injury and tested.

All subjects in the study received standard-of-care treatment which included a head CT scan performed within three hours of presenting to HCF or ED and within 12 hours of injury. All images were de-identified and transmitted from each site to the Core Imaging Laboratory. A Neuroimaging Review Committee, consisting of three neuroradiologists independent from the ALERT-TBI study, conducted an independent, blinded review of each CT scan without access to any other clinical or laboratory data except for age and

gender. The blinded review consisted of an assessment of image quality, scalp injuries, facial, cranial and skull based fractures, acute intracranial lesions, and incidental findings of potential clinical relevance. Two primary reviewers independently evaluated each subject CT scan to determine whether it was CT-positive, CT-negative, inconclusive, or unreadable as defined by the presence or absence of acute intracranial lesions, the presence of non-evaluable acute intracranial lesions, or the inability to fully assess the head CT scan, respectively. An acute intracranial lesion was defined as any trauma induced or related finding visualized upon head CT scan, and may have included acute epidural hematomas, acute subdural hematomas, indeterminate extra-axial lesions, cortical contusions, parenchymal hematomas, non-hemorrhagic contusions, ventricle compression, ventricular trapping, brain herniation, intraventricular hemorrhage, hydrocephalus, subarachnoid hemorrhage, petechial hemorrhage, global or focal brain edema and post traumatic ischemia. If the interpretations of the two primary reviewers did not agree, the scan was adjudicated by a third reviewer who was not a primary reviewer and who was blinded to the interpretations from the two primary reviewers. Disagreement was defined as any difference in result of 'CT-positive', 'CT-negative', 'Inconclusive', or 'Unreadable'.

A total of 2,011 subjects who fulfilled the study inclusion and exclusion criteria and gave informed consent were enrolled at 22 clinical sites in three countries. Two-thirds of the subjects (1,354/2,011 or 67.3%) were enrolled at U.S. sites. A total of 47 subjects (47/2011 or 2.3%) were excluded due to the following: unreadable head CT scan results (7/2011), inconclusive head CT scan result (3/2011), lost blood samples (2/2011), no head CT scan result (8/2011), no serum collected (16/2011), and withdrawal by the subject (11/2011). There were no withdrawals due to screen failure, adverse events, or deaths. Of the 1964 subjects remaining in the study, there were 17 subjects with a GCS of 9–12 (moderate TBI) which are excluded from the analyses of Banyan BTI clinical performance. Of the 1947 evaluable subjects, 1312 (67.4%) of the subjects were enrolled in the United States and 635 (32.6%) were enrolled in Germany and Hungary.

Subject enrollment demographics and CT scan results for all 1947 evaluable subjects are presented in the table below. Mean subject age was 48.9 years (range 18 to 98 years) and most subjects were either working (48.0%) or retired (25.3%). These demographic characteristics are similar between the CT scan-positive and CT scan-negative groups with the following exceptions: the percentage of white subjects is higher in the CT scan-positive group (81.7%) than in the CT scan-negative group (68.9%); the percentage of black or African American subjects is lower in the CT scan-positive group (13.3%) than in the CT scan-negative group (28.1%), and the percentage retired is higher in the CT scan-positive group (35.8%) than in the CT scan-negative group (24.6%). These differences were statistically significant (P=0.003 for white, P=0.0003 for black or African American, and P=0.009 for retired subjects).

	Head CT scan results		Total
	Positive	Negative	
N	120	1827	1947
Age[A], years			
Mean (SD)	58.8 (18.29)	48.3 (20.94)	48.9 (20.94)
Median	58.5	48.0	49.0
Range (min, max)	(20, 95)	(18, 98)	(18, 98)
Gender, N (%)			
Male	70 (58.3%)	1033 (56.5%)	1103 (56.7%)
Female	50 (41.7%)	794 (43.5%)	844 (43.3%)
Ethnicity, N (%)			
Hispanic or Latino	1 (0.8%)	89 (4.9%)	90 (4.6%)
Not Hispanic or Latino	118 (98.3%)	1737 (95.1%)	1855 (95.3%)
Not Reported	1 (0.8%)	1 (0.1%)	2 (0.1%)
Race[B], N (%)			
White	98 (81.7%)	1259 (68.9%)	1357 (69.7%)
Black or African American	16 (13.3%)	513 (28.1%)	529 (27.2%)
Asian	5 (4.2%)	24 (1.3%)	29 (1.5%)
Native Hawaiian/Pacific Islander	1 (0.8%)	2 (0.1%)	3 (0.2%)
American Indian or Alaska Native	1 (0.8%)	12 (0.7%)	13 (0.7%)
Unknown	1 (0.8%)	27 (1.5%)	28 (1.4%)

[A] Age was calculated relative to the date of informed consent
[B] Subjects could have indicated more than one race

Other baseline characteristics collected for all evaluable subjects include height, weight, body mass index (BMI), vital signs (systolic and diastolic blood pressure, respiratory rate, body temperature), and tobacco and alcohol use. These characteristics measured in all subjects were similar in the CT scan-positive and CT scan-negative groups, except for alcohol use, which was slightly lower but not statistically significant in the CT scan-positive group. Commonly reported medical history findings for all evaluable subjects were surgical and medical procedures, metabolism and nutrition

disorders, vascular disorders, and immune system disorders. Surgical and medical procedures occurred at higher frequency in the CT-negative group compared to the CT scan-positive group but the difference was not statistically significant. All other categories occurred at similar frequencies in both groups. Common neurological history reported findings were nervous system disorders and psychiatric disorders. Both categories occurred at similar frequencies in the CT scan-positive and CT scan-negative groups.

The following head injury characteristics were collected for all evaluable subjects and summarized in the table below. The mean time from head injury to blood draw was 3.5 hours. Most subjects had a GCS score of 15 (94/120 or 78.3% in CT scan-positive subjects and 1738/1827 or 95.1% in CT scan-negative subjects). The percentage of subjects with GCS scores of 13 and 14 were higher in the CT scan-positive subjects compared to the CT scan-negative subjects.

	Head CT scan results		**Total**
	Positive	**Negative**	
N	120	1827	1947
Time from head injury to examination (hours)[A]			
N	120	1825	1945[D]
Mean (SD)	1.85 (1.7)	1.66 (1.9)	1.67 (1.9)
Median	1.20	1.03	1.05
Range (min, max)	(0.3, 7.8)	(0.1, 33.4)	(0.1, 33.4)
Time from head injury to CT scan (hours)[A]			
N	120	1825	1945[D]
Mean (SD)	2.82 (2.1)	2.76 (2.1)	2.76 (2.1)
Median	2.06	2.15	2.13
Range (min, max)	(0.5, 10.9)	(0.2, 33.5)	(0.2, 33.5)
Time from head injury to blood draw (hours)[A]			
N	120	1824	1944[E]
Mean (SD)	3.75 (1.9)	3.49 (2.1)	3.50 (2.1)
Median	3.26	3.13	3.17
Range (min, max)	(0.3,9.3)	(0.3,35.3[F])	(0.3,35.3[F])
GCS score			
9	0 (0.0%)	0 (0.0%)	0 (0.0%)
10	0 (0.0%)	0 (0.0%)	0 (0.0%)
11	0 (0.0%))	0 (0.0%)	0 (0.0%)

	Head CT scan results		Total
	Positive	Positive	
GCS score			
12	0 (0.0%)	0 (0.0%)	0 (0.0%)
13	7 (5.8%)	15 (0.8%)	22 (1.1%)
14	19 (15.8%)	74 (4.1%)	93 (4.8%)
15	94 (78.3%)	1738 (95.1%)	1832 (94.1%)
Neurological assessment Number (%) of subjects experiencing:			
Loss of Consciousness	82 (68.3%)	737 (40.3%)	819 (42.1%)
Confusion	44 (36.7%)	316 (17.3%)	360 (18.5%)
Vomiting	14 (11.7%)	134 (7.3%)	148 (7.6%)
Vomiting Two or More Episodes	10 (8.3%)	64 (3.5%)	74 (3.8%)
Post traumatic Amnesia (PTA)	81 (67.5%)	559 (30.6%)	640 (32.9%)
Retrograde PTA≥ 30 min.	22 (18.3%)	68 (3.7%)	90 (4.6%)
Persistent Anterograde PTA	42 (35.0%)	178 (9.7%)	220 (11.3%)
Seizures	2 (1.7%)	11 (0.6%)	13 (0.7%)
Subjects with Drug or Alcohol Intoxication at Time of Presentation to Facility	30 (25.0%)	380 (20.8%)	410 (21.1%)
Dangerous Mechanism of Injury[B]	27 (22.5%)	376 (20.6%)	403 (20.7%)
Physical Evidence[C]			
Visible Trauma Above the Clavicle	101 (84.2%)	1125 (61.6%)	1226 (63.0%)
Suspected Open or Depressed Skull Fracture	14 (11.7%)	46 (2.5%)	60 (3.1%)
Signs of Basal Skull Fracture	10 (8.3%)	27 (1.5%)	37 (1.9%)
Presence of Neurosurgical Lesions	5 (4.2%)	0 (0.0%)	5 (0.3%)

[A] Time since head injury calculated relative to time that subject was first examined by medical personnel at facility
[B] Dangerous mechanism of injury was pedestrian struck by a motor vehicle, an occupant ejected from a motor vehicle, or a fall from an elevation of 3 or more feet or 5 stairs
[C] Prior to head CT
[D] Two subjects did not have time from head injury to examination and CT scan recorded
[E] Three subjects did not have time from head injury to blood draw recorded
[F] Two subjects had blood draws taken more than 12 h (18 h and 25.3 h) from head injury

The most common head CT findings in the 120 subjects with CT-positive scans were scalp injury (96.7%), subarachnoid hemorrhage (59.2%), the presence of incidental findings (58.3%), and acute subdural hematoma (47.5%). Other frequently reported findings included cranial fractures (26.7%), parenchymal hematoma (20.0%), facial fractures (17.5%), skull based fractures (15.0%), and indeterminate extra-axial lesions (15.0%). All other findings occurred in less than 10% of CT-positive subjects.

Venous whole blood samples were collected into BD Vacutainer SST Venous Blood Collection tubes under routine methods within three hours of presentation to the HCF or ED, but not longer than 12 hours from injury. Serum samples processed by clinical site personnel and shipped to a third-party biorepository where they remained stored frozen at -80°C until testing after completion of the ALERT-TBI. Testing of the serum specimens was conducted at three clinical testing sites blinded to the subject's diagnosis and clinical status. Specimens requiring retest were retested by the same testing site after completion of the initial round of testing.

The test results were used to differentiate the intended patient population into two groups:

1. patients with a very low probability of having acute intracranial lesions for whom the physician may recommend not undergoing neuroimaging via head CT scan, thus avoiding exposure to unnecessary radiation, and
2. all other suspected head injury patients for whom the physician may recommend undergoing neuroimaging via head CT scan as currently occurs under standard of care.

To estimate clinical performance characteristics, the Banyan BTI result was compared to the consensus head CT scan result for each patient. The performance estimates are summarized in the 2x2 table below. Of the 1947 evaluable subjects, 120 had positive CT scan results. Of the 120 subjects with positive CT scan results, 117 had a positive Banyan BTI result (sensitivity = 97.5%). The remaining three CT scan positive subjects had negative results from the Banyan BTI test. The rate of false negative (FN) results was 2.5% (3/120). None of the five subjects identified with a lesion requiring surgical intervention had a FN result suggesting that Banyan BTI correctly classified all these five CT-positive subjects as assay positive. Of the 1827 subjects with negative CT scan results, 666 had a negative Banyan BTI result (specificity = 36.4%). The rate of False Positive (FP) results was 63.2% (1161/1827). Overall, there were 669 subjects with negative Banyan BTI results. Of these, 666 had negative CT scan results. The Negative Predictive Value (NPV) of the assay was 99.6% (666/669). The potential benefit of the assay would be a reduction in unnecessary CT scans by approximately one third (36.5% or 666 of 1827 subjects had true negative assay results). The Positive Predictive Value (PPV) of the assay was 9.2%. The Likelihood Ratio Negative (LRN) of the assay was 0.07 (one-sided, lower, exact 95% confidence limit: 0.170). The Likelihood Ratio positive (LRP) of the assay was 1.53 (one-sided, lower, 95% confidence limit:1.468). The results showed that the Banyan BTI is characterized by high sensitivity and high

NPV, which supports clinical utility as an aid in the evaluation of the need for a CT scan in subjects presenting with a GCS score of 13 to 15 and a negative Banyan BTI result.

		Head CT scan result		Total
		Positive	Negative	
Banyan BTI result	**Positive**	117	1161	1278
	Negative	3	666	669
	Total	120	1827	1947

Sensitivity = 117/120 (97.5%); 95% CI[A]: 92.9%–99.5%

Specificity = 666/1827 (36.5%); 95% CI[A]: 34.2%–38.7%

Negative predictive value (NPV) = 666/669 (99.6%); 95% CI[A]: 98.7%–99.9%

Positive predictive value (PPV) =117/1278 (9.2%); 95% CI[A]: 7.6%–10.9%

[A] Two-sided, exact 95% binomial confidence interval (CI) using the Clopper-Pearson method

Analyses of assay performance by gender and time from injury relative to blood draw are shown in the table below. There was little variation in NPV and PPV between males and females and with increasing time from injury. These data indicate that gender differences and differences between head injury characteristics did not translate into statistically significant differences in assay performance.

	Sensitivity N (%) (95% CI)	**Specificity** N (%) (95% CI)	**NPV** N (%) (95% CI)	**PPV** N (%) (95%CI)
All subjects N=1947	117/120 (97.5%) (92.9–99.5)	666/1827 (36.5%) (34.2–38.7)	666/669 (99.6%) (98.7–99.9)	117/1278 (9.2%) (7.6–10.9)
Gender				
Male N=1103 (56.7%)	69/70 (98.6%) (92.3–100)	367/1033 (35.5%) (32.6–38.5)	367/368 (99.7%) (98.5–100)	69/735 (9.4%) (7.4–11.7)
Female N=844 (43.3%)	48/50 (96.0%) (86.3–99.5)	299/794 (37.7%) (34.3–41.1)	299/301 (99.3%) (97.6–99.9)	48/543 (8.8%) (6.6–11.5)

	Sensitivity N (%) (95% CI)	Specificity N (%) (95% CI)	NPV N (%) (95% CI)	PPV N (%) (95%CI)
Time from injury to blood draw				
0–4 hours N=1469 (75.4%)	85/86 (98.8%) (93.7–100)	493/1383 (35.6%) (33.1–38.2)	493/494 (99.8%) (98.9–100	85/975 (8.7%) (7.0–10.7)
4–8 hours N=490 (25.1%)	29/30 (96.7%) (82.8–99.9)	146/376 (38.8%) (33.9–44.0)	146/147 (99.3%) (96.3–100)	29/259 (11.2%) (7.6–15.7)
0–8 hours N=1859 (95.5%)	112/114 (98.2%) (93.8–99.8)	632/1745 (36.2%) (34.0–38.5)	632/634 (99.7%) (98.9–100)	112/1225 (9.1%) (7.6–10.9)
8–12 hours N=84 (4.3%)	5/6 (83.3%) (35.9–99.6)	35/78 (44.9%) (33.6–56.6)	35/36 (97.2%) (85.5–99.9	5/48 (10.4%) (3.5–22.7)

b. Other clinical supportive data (when a. is not applicable):

Not applicable

4. Clinical cut-off:

Refer to assay cut-off.

5. Expected values/Reference range:

The expected values from 695 healthy donors ranging in age from 18 to 80 in the U.S. population who do not have acute injury to the head were determined in accordance with the CSLI guideline C28-A3c. The mean (SD) age was 39.4 (15) years. The mean (SD) concentration for GFAP was 21 (37) pg/mL, and the median was 10 pg/mL. The mean (SD) concentration for UCH-L1 was 134 (175) pg/mL, and the median was 80 pg/mL.

There were 95 healthy donors who tested positive for GFAP only and 30 were positive for both GFAP and UCH-L1. The results summarized in the table below show that 82% have a negative Banyan BTI assay result and 18% have a positive Banyan BTI assay result. The medians for both GFAP and UCH-L1 did not differ significantly across gender and race. A slight trend of increasing levels was observed with age, especially over the age of 60 years. It is the responsibility of each laboratory to establish its own reference ranges for the population of patients it serves, as expected values may be affected by different factors including age.

Banyan UCH-L1 Kit result (relative to cutoff)[A]	Banyan GFAP Kit result (relative to cutoff)[B]	Banyan BTI result	All subjects (N=695)
Below	Below	Negative	570 (82.0%)
Above	Above	Positive	30 (4.3%)
Below	Above	Positive	95 (13.7%)
Above	Below	Positive	0 (0.0%)

[A]Above = the UCH-L1 concentration is equal to or above 327 pg/mL; Below means the UCH-L1 concentration is below 327 pg/mL
[B]Above = the GFAP concentration is equal to or above 22 pg/mL; Below means the GFAP concentration is below 22 pg/mL

M. Proposed Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Parts 801 and 809 and the special controls for this device type.

N. Patient Perspectives:

This submission did not include specific information on patient perspectives for this device.

O. Identified Risks to Health and Identified Mitigations:

Identified Risks to Health	Identified Mitigations
Inaccurate test results that provide false positive or false negative results	General controls and special control (1)
Failure to correctly interpret test results can lead to false positive or false negative results	General controls and special control (2)

P. Benefit/Risk Determination:

Summary	
Summary of the Benefit(s)	The device provides a rapid blood test result. Clinicians may use the test in conjunction with clinical examination and traditional imaging. There are no currently FDA-cleared or approved blood tests for the detection of intracranial bleeds.
Summary of the Risk(s)	The risk of the device is a false negative. The clinical trial showed a false negative rate of 2.5%. Individuals with false-positive results would undergo CT imaging of the head and be exposed to some radiation.
Summary of Other Factors	None
Conclusions Do the probable benefits outweigh the probable risks?	The device is intended to be an aid in the evaluation of mild traumatic brain injury. As such, the assay provides clinicians with an additional assessment tool for a heterogeneously presenting condition. The device displays a high sensitivity for detection of intracranial bleeds with little risk of adverse events or subject burden. Therefore, the benefits appear to outweigh the probable risk in light of the special controls established for this device and in combination with general controls.

Q. Conclusion:

The information provided in this *de novo* submission is sufficient to classify this device into class II under regulation 21 CFR 866.5830. FDA believes that the stated special controls, along with the applicable general controls provide reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

Product Code: QAT
Device Type: Brain trauma assessment test
Class: II (special controls)
Regulation: 21 CFR 866.5830

(a) *Identification*. A brain trauma assessment test is a device that consists of reagents used to detect and measure brain injury biomarkers in human specimens. The measurements aid in the evaluation of patients with suspected mild traumatic brain injury in conjunction with other clinical information to assist in determining the need for head imaging per current standard of care.

(b) *Classification*. Class II (special controls). A brain trauma assessment test must comply with the following special controls:

(1) The 21 CFR 809.10(b) compliant labeling must include detailed descriptions of and results from performance testing conducted to evaluate precision, accuracy, linearity, analytical sensitivity, interference, and cross-reactivity. This information must include the following:

(i) Performance testing of device precision must, at minimum, use one unmodified clinical specimen from the intended use population with concentration of the brain injury biomarker(s) near the medical decision point. Contrived specimens that have been generated from pooling of multiple samples or spiking of purified analyte to cover the measuring range may be used, but the contrived samples must be prepared to mimic clinical specimens as closely as possible. This testing must evaluate repeatability and reproducibility using a protocol from an FDA-recognized standard.

(ii) Device performance data must be demonstrated through a clinical study and must include the following:

(a) Data demonstrating clinical validity including the clinical sensitivity and specificity, and positive and negative predictive value of the test in the intended use population of patients with suspected mild traumatic brain injury (i.e., Glasgow Coma Score (GCS) of 13–15), or equivalent standard of care for determination of severity of Traumatic Brain Injury (TBI).

(b) Study must be performed using the operators and in settings that are representative of the types of operators and settings for which the device is intended to be used.

(c) All eligible subjects must meet the well-defined study inclusion and exclusion criteria that define the intended use population. The prevalence of diseased or injured subjects in the study population must reflect the prevalence of the device's intended use population, or alternatively, statistical measures must be used to account for any bias due to enrichment of subpopulations of the intended use population.

(d) All eligible subjects must have undergone a head CT scan or other appropriate clinical diagnostic standard used to determine the presence of an intracranial lesion as part of standard of care and must also be evaluated by

the subject device. All clinical diagnostic standards used in the clinical study must follow standard clinical practice in the U.S.

(e) Relevant demographic variables and baseline characteristics including medical history and neurological history. In addition, head injury characteristics, neurological assessments, and physical evidence of trauma must be provided for each subject. This information includes but is not limited to the following: time since head injury, time from head injury to CT scan, time from head injury to blood draw, GCS score or equivalent, experience of loss of consciousness, presence of confusion, episodes of vomiting, post-traumatic amnesia characteristics, presence of post traumatic seizures, drug or alcohol intoxication, mechanism of injury, acute intracranial lesion type, neurosurgical lesion, and cranial fracture.

(f) Each CT scan or other imaging result must be independently evaluated in a blinded manner by at least two board-certified radiologists to determine whether it is positive or negative as defined by the presence or absence of acute intracranial lesions. This independent review must be conducted without access to test results of the device. Prior to conducting the review, the criteria and procedures to be followed for scoring the images must be established, including the mechanism for determining consensus.

(g) All the clinical samples must be tested with the subject device blinded to the TBI-status and the neurological-lesion-status of the subject.

(h) Details on how missing values in data are handled must be provided.

(i) For banked clinical samples, details on storage conditions and storage period must be provided. In addition, a specimen stability study must be conducted for the duration of storage to demonstrate integrity of archived clinical samples. The samples evaluated in the assay test development must not be used to establish the clinical validity of the assays.

(iii) Performance testing of device analytical specificity must include the most commonly reported concomitant medications present in specimens from the intended use population. Additionally, potential cross-reacting endogenous analytes must be evaluated at the highest concentration reported in specimens from the intended use population.

(iv) Expected/reference values generated by testing a statistically appropriate number of samples from apparently healthy normal individuals.

(2) The 21 CFR 809.10(a) and 809.10(b) compliant labeling must include the following limitations:

(i) A limiting statement that this device is not intended to be used a stand-alone device but as an adjunct to other clinical information to aid in the evaluation of patients who

are being considered for standard of care neuroimaging.

(ii) A limiting statement that reads "A negative result is generally associated with the absence of acute intracranial lesions. An appropriate neuroimaging method is required for diagnosis of acute intracranial lesions."

(iii) As applicable, a limiting statement that reads "This device is for use by laboratory professionals in a clinical laboratory setting."