STATISTICAL ANALYSIS PLAN

TESCO Clinical Investigation

TES HT100 Clinical Outcomes – Clinical Validation of the TES HT100 Temnograph
for Brain Injury Screening in Patients with Mild Head Trauma or Suspected Neurological Symptoms

Study Code: CLV-01-TES HT | ClinicalTrials.gov: NCT07490418

Document title	Statistical Analysis Plan (SAP)
SAP version	1.1
SAP date	30 January 2026
Reference protocol	TESCO Protocol v1.2, dated 27 January 2026
Sponsor / Manufacturer	B&B S.r.l., Strada Fienile 1, 80013 Casalnuovo di Napoli (NA), Italy
Investigational device	TES HT100 (Temnography)
Study design	Prospective, multicentre, blinded diagnostic-accuracy clinical investigation
Planned sample size	850 participants: 425 mild head trauma + 425 non-traumatic neurological symptoms
Document status	Standalone SAP for public registry disclosure; extracted from the final protocol and aligned with the methods used for the protocol-defined first interim analysis.

Document status statement

This standalone SAP was prepared after completion of the first interim analysis in order to consolidate, in a single controlled document, the statistical methods already specified in the study protocol and those applied in the first interim evaluation, and to prospectively define the technical implementation of the remaining interim and final analyses. It does not modify the protocol-defined objectives, eligibility criteria, endpoints, reference standard, or overall analytical framework.

Revision History

Version	Date	Description	Author/Owner
1.0	2 Jul 2024	Initial standalone SAP based on Protocol v1.1	B&B S.r.l.
1.1	30 Jan 2026	Review based on Protocol v1.2. (minimal update)	B&B S.r.l.

1. Purpose and Scope

This Statistical Analysis Plan (SAP) describes the statistical principles and procedures for the TESCO clinical investigation (CLV-01-TES HT). It is derived from the final TESCO Protocol v1.1 dated 02 Jul 2024 and updated according to TESCO protocol v1.2 dated 27 Jan 2026 and adds operational statistical detail for protocol-defined interim and final analyses. The SAP is intended to support transparent, reproducible and consistent analysis of diagnostic performance, safety and prespecified secondary analyses across the multicentre study.

The SAP is aligned with the analytical approach used in the first protocol-defined interim analysis, including cross-tabulation of TES HT100 against head CT, Wilson confidence intervals for proportions, log-scale confidence intervals for likelihood ratios, balanced accuracy, Cohen's kappa as a supplementary agreement measure, exploratory subgroup comparisons, and robustness analyses for excluded or non-evaluable cases. The first interim analysis is treated as completed and is not redefined by this document.

2. Study Design and Analysis Framework

TESCO is a prospective, multicentre, blinded diagnostic-accuracy clinical investigation evaluating TES HT100 as a rapid screening and triage-support device in adult Emergency Department patients. The reference standard is the radiologist-interpreted head computed tomography (CT) report. TES operators are blinded to the CT report at the time of TES testing, and radiologists are blinded to the TES result. TES results are made available only after the CT report has been finalized and recorded.

- Cohort A – Mild head trauma: adults referred to the Emergency Department for mild head injury.
- Cohort B – Non-traumatic neurological symptoms: non-critical adults presenting with neurological symptoms such as syncope, spatial disorientation and/or loss of balance, speech disturbance and/or facial hemi-paresis, visual alterations, memory loss and/or confusional state.
- Planned sample size: 425 participants per cohort (850 total), including allowance for approximately 10% dropout/non-evaluable data.
- Primary analysis: pooled multicentre diagnostic performance, with cohort-specific and centre-specific analyses reported as prespecified or exploratory secondary analyses.

3. Objectives and Endpoints

3.1 Primary objectives

- Evaluate the diagnostic accuracy of TES HT100 for identification of intracranial brain lesions or structural brain abnormalities by comparison with head CT.
- Establish the safety of TES HT100 by monitoring the incidence and type of adverse events associated with use of the device.

3.2 Primary diagnostic endpoints

- Sensitivity and specificity of the binary TES HT100 result versus the CT reference standard, each with two-sided 95% confidence intervals.
- Protocol-specified ROC-AUC analysis when a valid continuous TES-derived score is available and suitable for discrimination analysis. The protocol identifies the percentage cranial-bone estimate (Q) as a continuous variable; because Q is also described as a measurement-quality surrogate, any AUC based on Q will be interpreted as supportive and will not supersede the binary TES-versus-CT analysis.
- Device-related adverse-event rate and description of AEs, SAEs, SADEs and device deficiencies.

3.3 Secondary and exploratory endpoints

- Positive predictive value (PPV), negative predictive value (NPV), overall accuracy, balanced accuracy, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and supplementary agreement statistics.
- Diagnostic performance in prespecified subgroups: age, history of prior intracranial injury/brain surgery, neurodegenerative disease, cohort, and study centre.
- Descriptive assessment of the potential clinical-use benefit of TES HT100 in terms of patient routing, operator time, procedures and resource use, where the corresponding data are available.
- Exploratory target-condition analyses separating acute clinically relevant CT abnormalities from chronic/incidental intracranial abnormalities, without changing the protocol-defined primary CT-positive endpoint.

4. Analysis Populations

4.1 Enrolled Set (ES)

All participants considered enrolled according to the protocol after confirmation of eligibility and signed informed consent. The Enrolled Set will be used for participant disposition and protocol-flow reporting.

4.2 Safety Set (SS)

All enrolled participants who undergo any TES HT100 study examination or attempted TES HT100 acquisition. Safety analyses will be performed on the Safety Set according to the treatment/device exposure actually received.

4.3 Paired Diagnostic Accuracy Set (PDAS) – primary diagnostic set

All enrolled participants with an interpretable binary TES HT100 result and an available, classifiable head CT reference result. The primary diagnostic-accuracy analyses at each interim and at final analysis will use the PDAS. Participants with non-evaluable TES examinations or missing/unclassifiable CT will not be imputed into the primary 2×2 table; they will be summarized separately and included in robustness analyses as described in Section 11.

4.4 Per-Protocol Diagnostic Set (PPDS) – supportive

The PDAS excluding participants with major protocol deviations considered capable of materially affecting diagnostic interpretation, such as major eligibility violations, major blinding violations, invalid device acquisition,

or a TES–CT interval outside the protocol-defined 24-hour window when clinically relevant. The PPDS is supportive and does not replace the PDAS.

5. Reference Standard and CT Classification

For the primary diagnostic analysis, head CT is the reference standard. A participant is classified as CT-positive when the CT report identifies at least one protocol-defined intracranial structural abnormality. The protocol-defined CT-positive category includes, among others:

- intraparenchymal, subdural or subarachnoid haemorrhage;
- cerebral contusion with oedema;
- intracranial calcifications;
- cystic or nodular intracranial lesions;
- acute, subacute or previous cerebral ischaemia;
- other acute or previous intracranial structural abnormalities.

Participants without protocol-defined intracranial findings are classified as CT-negative. Extracranial findings alone are not considered CT-positive for the primary analysis. Classification will be based on the finalized radiology report and, where required for exploratory analyses, central clinical/radiological review.

6. General Statistical Principles

- All analyses will be two-sided where inferential testing is performed.
- The nominal significance level for exploratory comparisons will be 0.05. No formal multiplicity adjustment is planned because subgroup and category-specific comparisons are exploratory and no interim decision boundary is based on p-values.
- Diagnostic performance estimates will be accompanied by two-sided 95% confidence intervals.
- Continuous variables will be summarized using n, mean, standard deviation, median, interquartile range, minimum and maximum as appropriate. Categorical variables will be summarized using counts and percentages.
- The primary diagnostic analysis is descriptive estimation of diagnostic performance rather than a sequential hypothesis-testing procedure. Interim analyses are not intended for early stopping for efficacy or futility and therefore no alpha-spending function is applied.
- The exact statistical software, version, packages and analysis scripts used for each data cut will be documented in the corresponding statistical report to ensure reproducibility.

7. Primary Diagnostic Accuracy Analysis

TES HT100 will be cross-tabulated against the CT reference standard in a 2×2 table. The binary TES result will be classified as positive or negative; non-evaluable examinations are excluded from the primary 2×2 table and reported separately.

Table 1. Standard 2×2 diagnostic classification

	CT-positive	**CT-negative**	**Total**
TES-positive	TP	FP	TP+FP
TES-negative	FN	TN	FN+TN
Total	TP+FN	FP+TN	N

7.1 Performance measures and confidence intervals

Sensitivity: TP / (TP + FN)

Specificity: TN / (TN + FP)
PPV: TP / (TP + FP)
NPV: TN / (TN + FN)
Overall accuracy: (TP + TN) / N
Balanced accuracy: (Sensitivity + Specificity) / 2
LR+: Sensitivity / (1 − Specificity)
LR−: (1 − Sensitivity) / Specificity
Youden index (J): Sensitivity + Specificity − 1
Number needed to diagnose (exploratory): 1 / J, when J > 0

Two-sided 95% confidence intervals for sensitivity, specificity, PPV, NPV and accuracy will be calculated using the Wilson score method, consistent with the first interim analysis. Confidence intervals for LR+ and LR− will be derived on the logarithmic scale. Cohen's kappa will be reported as a supplementary measure of agreement between TES HT100 and CT classification.

7.2 Protocol-specified ROC-AUC and supportive modelling

The protocol specifies ROC-AUC analysis and identifies the TES-derived Q variable (percentage cranial-bone estimate) as a continuous quantity. Where Q is available, technically valid and directionally interpretable with respect to CT positivity, a supportive ROC curve and AUC with 95% confidence interval may be calculated. Because the marketed/clinical TES output is binary and Q is also described in the protocol as a surrogate of measurement quality, the binary TES-versus-CT operating point remains the principal diagnostic-performance analysis. Any Q-derived AUC will be clearly labelled supportive.

As an additional supportive analysis, the probability of CT-positive status may be modelled using logistic regression with TES binary outcome as the main predictor and Q as a continuous covariate. In multicentre analyses, study centre may be added as an exploratory covariate where data density is sufficient. Model coefficients, odds ratios, 95% confidence intervals and model-discrimination summaries will be reported as appropriate. Sparse-centre models will not be over-interpreted.

8. Secondary, Subgroup and Exploratory Analyses

8.1 Cohort-specific and centre-specific analyses

Sensitivity, specificity, PPV, NPV, accuracy and likelihood ratios will be reported separately for Cohort A and Cohort B. The protocol-defined primary assessment remains the pooled multicentre analysis. Centre-specific performance will be described when sample size and cell counts are sufficient; small centres or sparse 2×2 tables will be reported descriptively without unstable inferential claims.

8.2 Prespecified subgroups

- Age: <65 years versus ≥65 years. The ≥65 definition is used operationally so that participants aged exactly 65 years are assigned unambiguously, consistent with the first interim reporting.
- Prior intracranial injury and/or prior brain surgery: yes versus no.
- Neurodegenerative disease: yes versus no.
- Clinical cohort: mild head trauma versus non-traumatic neurological symptoms.
- Study centre, as an exploratory stratification factor.

Subgroup diagnostic performance will be estimated with 95% confidence intervals. Two-sided Fisher's exact tests may be used for exploratory comparison of proportions where appropriate, consistent with the first interim analysis. Resulting p-values are exploratory and will not be interpreted as confirmatory evidence of interaction unless a formal interaction model has been prespecified for that analysis.

8.3 Acute versus chronic/incidental CT target-condition analyses

To provide clinical context, CT-positive findings may be classified as (i) acute clinically relevant abnormalities and/or (ii) chronic/incidental intracranial abnormalities. These are target-condition analyses rather than mutually exclusive participant subgroups because acute and chronic findings may coexist in the same CT examination. These analyses are supplementary and do not alter the protocol-defined broad CT-positive primary endpoint.

9. Interim Analysis Plan

9.1 General principles

The protocol specifies interim analyses when each cohort reaches the enrollment thresholds of approximately 150 and 300 participants. Each interim analysis is an estimation exercise intended to assess accumulating diagnostic performance and data quality; it is not a formal group-sequential efficacy or futility analysis and is not used to stop recruitment on the basis of statistical significance.

- A formal data-cut date and data snapshot/freeze will be documented for each interim analysis.
- Participant inclusion in an interim analysis will follow chronological enrollment order and objective evaluability criteria; no participant will be selected or excluded on the basis of TES or CT result.
- The same definitions of CT positivity, TES evaluability and diagnostic metrics will be used at each interim analysis unless a protocol amendment or documented SAP amendment prospectively specifies otherwise.
- Each interim report will include a STARD-style participant flow summary showing screened/evaluated, excluded, enrolled, paired evaluable, non-evaluable and analysed participants by cohort and, where relevant, by centre.

9.2 First interim analysis

The first interim data cut was 21 January 2026. The completed paired diagnostic analysis comprised 301 participants: 151 in Cohort A and the first 150 consecutive eligible and evaluable participants in Cohort B. The statistical methods applied included 2×2 cross-tabulation, sensitivity, specificity, PPV, NPV, accuracy and balanced accuracy with Wilson 95% confidence intervals for proportions; LR+ and LR− with log-scale confidence intervals; Cohen's kappa; exploratory Fisher's exact subgroup comparisons; acute/chronic target-condition analyses; and a conservative robustness/tipping-point analysis for excluded or non-evaluable cases. This SAP preserves that analytical framework for subsequent analyses.

9.3 Second interim analysis – approximately 300 participants per cohort

The second protocol-defined interim analysis will be performed after both cohorts have reached at least 300 enrolled participants and the corresponding data have undergone data cleaning and interim freeze. The primary cohort-balanced interim diagnostic set will comprise the first 300 consecutive eligible and evaluable participants in each cohort, identified strictly by chronological enrollment order. This approach prevents selective inclusion when one cohort or centre temporarily over-enrolls before database freeze.

- Primary second-interim analysis: pooled balanced PDAS of approximately 600 participants (300 per cohort), using the same diagnostic metrics and confidence-interval methods as the first interim.
- Supportive cumulative analysis: all additional eligible/evaluable participants enrolled before the interim data freeze, if any, will be summarized separately rather than selectively discarded; they will remain in the study database and will be included in subsequent cumulative/final analyses.
- Multicentre analysis: overall pooled estimates will be primary. Cohort- and centre-stratified estimates will be reported as secondary/exploratory. Centre may be assessed as a covariate in supportive logistic regression if cell counts are adequate.
- Robustness analysis: non-evaluable or missing paired data will be summarized and evaluated using the sensitivity analyses in Section 11.
- No efficacy/futility stopping boundary will be applied; recruitment will continue toward the protocol-defined final sample unless stopped for safety, operational or regulatory reasons independent of interim significance testing.

9.4 Final analysis

The final analysis will be performed after completion of the planned recruitment target (425 participants per cohort, subject to actual evaluable numbers), completion of data cleaning, resolution of material queries, and final database freeze. The final primary diagnostic analysis will include all eligible participants in the PDAS through the final cut-off, including eligible evaluable participants beyond the nominal target if minor enrollment overshoot occurs. Final reporting will present pooled multicentre performance, cohort-specific estimates, centre-stratified summaries, prespecified subgroup analyses, safety, participant disposition and the protocol-specified supportive ROC/model analyses where appropriate.

10. Sample Size and Precision

The protocol assumes a target diagnostic performance corresponding to an AUC of approximately 0.90 and specifies that approximately 384 subjects are required to achieve a maximum error of about 3 percentage points around a 95% confidence interval for AUC. Allowing for approximately 10% dropout/non-evaluable data, the protocol target is 425 participants per cohort. If a substantial imbalance between CT-positive and CT-negative outcomes is observed, precision will be reassessed and a tighter AUC precision target of approximately 0.02 may be considered as described in the protocol. This SAP does not change the protocol-defined enrollment target.

11. Missing Data, Non-evaluable Tests and Robustness Analyses

11.1 Primary handling

- No statistical imputation will be used for the primary 2×2 diagnostic analysis.
- Participants with an uninterpretable/non-evaluable TES result will be excluded from the primary PDAS and counted separately by reason, including motion-related artifact or technical acquisition failure.
- Participants without an available/classifiable CT reference result will be excluded from the primary PDAS and counted separately by reason.
- Missing covariates will reduce the denominator only for analyses that require that covariate; the primary diagnostic analysis will not be affected if TES and CT are available.

11.2 Robustness / tipping-point analysis

Consistent with the first interim analysis, robustness analyses will evaluate whether reasonable or deliberately conservative assumptions for excluded/non-evaluable cases could materially alter the interpretation of sensitivity, specificity or NPV. Where CT status is known but TES is non-evaluable, a conservative scenario may assign the TES result discordantly with CT. Where CT is missing, extreme-case bounds may be generated by assigning combinations of reference and index results that minimize sensitivity and/or specificity. These analyses are sensitivity analyses only and do not replace the observed-case primary PDAS results. Assumptions and case counts will be explicitly reported.

12. Safety Analysis

Safety will be summarized in the Safety Set. AEs, SAEs, SADEs and device deficiencies will be coded/classified according to Regulation (EU) 2017/745 and ISO 14155:2020 as specified in the protocol. The following will be reported:

- number and percentage of participants with any AE;
- number and percentage with any SAE or SADE;
- number and percentage with an event considered related or possibly related to TES HT100;
- number and type of device deficiencies;
- event listings including seriousness, relationship, action taken and outcome where available.

Safety analyses are descriptive. No formal between-group hypothesis test is planned because all participants receive the same study-device examination as an adjunct to standard care.

13. Data Quality, Data Freeze and Statistical Quality Control

Before each interim and final analysis, the analysis dataset will undergo documented quality-control checks. At minimum, the following will be verified:

- unique study identifier and absence of duplicate records;
- eligibility and cohort assignment;
- informed-consent status;
- TES examination date/time, result and evaluability;
- CT examination date/time, finalized report and CT-positive/negative classification;
- TES–CT interval and major timing deviations;
- blinding deviations, if any;
- centre/site code and enrollment chronology;
- adverse-event and device-deficiency data;
- resolution of material data queries affecting endpoint classification.

A data-cut/freeze memorandum should record the cut-off date, dataset version, number of enrolled participants, number included in each analysis set, unresolved data issues judged non-material, and the statistical software environment. Any post-freeze correction affecting primary results will be documented and the analysis rerun under version control.

14. Planned Statistical Outputs

Each interim and final statistical report should include, as applicable:

- participant-flow diagram by cohort and overall, with centre detail when relevant;
- baseline demographic and clinical characteristics;
- TES/CT evaluability and missing-data summary;
- overall 2×2 confusion matrix;
- overall diagnostic-performance table with 95% CIs;
- cohort-specific performance tables;
- prespecified subgroup analyses;
- centre-stratified descriptive analyses for the multicentre phase;
- acute versus chronic/incidental exploratory target-condition analyses;
- robustness/tipping-point analyses;
- safety summary and listings;
- supportive ROC-AUC/logistic-regression analyses if technically applicable;
- clear statement that interim analyses are descriptive and do not constitute an early-stopping decision.

15. Deviations from the SAP and Amendments

Any deviation from this SAP that materially affects analysis-set definition, endpoint classification, statistical method, or interpretation of the primary diagnostic endpoints will be documented in the corresponding analysis report, including the reason, timing and potential impact. Any planned change made before a future data freeze should be implemented through a dated SAP amendment. Changes made after review of unblinded outcome data will be clearly identified as post hoc or exploratory.

16. References

1. TESCO Study Protocol, CLV-01-TES HT, Version 1.2, dated 27 January 2026.

2. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. BMJ. 2015;351:h5527.

3. Regulation (EU) 2017/745 on medical devices.

4. ISO 14155:2020. Clinical investigation of medical devices for human subjects – Good clinical practice.

Appendix A. Interim Analysis Milestones

Analysis	Trigger	Primary interim set	Status / method	Role in study
Interim 1	≥150 enrolled in each cohort	301 paired evaluable: 151 Cohort A + first 150 consecutive eligible/evaluable Cohort B	Completed; methods described in Section 9.2	Protocol-defined estimation; no early stopping
Interim 2	≥300 enrolled in each cohort	First 300 consecutive eligible/evaluable per cohort; pooled ≈600 primary balanced set	Same methods as Interim 1; multicentre centre-stratified/supportive modelling added	Protocol-defined estimation; no early stopping
Final	Target recruitment complete: 425 per cohort	All eligible/evaluable PDAS through final data cut	Full pooled multicentre + cohort/centre/subgroup/safety/supportive analyses	Final study analysis