STATISTICAL ANALYSIS PLAN

Metabolic Reprogramming, Immune Modulation, and Palliative Outcomes Associated With ET-SPACE Whole-Body Infrared Hyperthermia in Patients With Advanced Solid Tumors: An Ambispective Observational Cohort Study

Standalone analysis plan for the HEAIS-009 cohort

Protocol Identifier: HEAIS-009

ClinicalTrials.gov Identifier: NCT Pending assignment

Sponsor: Zhejiang Hospital

Study Site: Sandun Campus, Zhejiang Hospital, Hangzhou, Zhejiang, China

Document Version: 1.0

Document Date: August 25, 2026

Public registry document - no participant-identifying information

Document Control

Official title	Metabolic Reprogramming, Immune Modulation, and Palliative Outcomes Associated With ET-SPACE Whole-Body Infrared Hyperthermia in Patients With Advanced Solid Tumors: An Ambispective Observational Cohort Study
Protocol identifier	HEAIS-009
ClinicalTrials.gov identifier	NCT Pending assignment
SAP version	1.0
SAP date	August 25, 2026
Related protocol	HEAIS-009 Study Protocol, Version 1.0, August 25, 2026
Status	Initial registry-ready SAP; to be finalized before database lock and before outcome analyses

Purpose and Principles

This Statistical Analysis Plan specifies the confirmatory and supportive analyses for HEAIS-009 before database lock. It separates patient-centered, biomarker, safety, tumor, and survival analyses; defines the analytic populations and derived variables; and prespecifies methods for repeated measurements, confounding, missing data, multiplicity, and sensitivity analyses.

- The study is observational and uncontrolled. Estimates will be described as changes and associations, not causal treatment effects.
- The participant is the primary unit of analysis; the WBH session is an additional unit for feasibility and safety analyses.
- The primary estimand is the mean within-participant change in EORTC QLQ-C30 global health status from baseline to Week 4.
- All analysis decisions made after database lock will be identified as post hoc.
- Effect estimates, confidence intervals, denominators, missingness, and model diagnostics will be emphasized over isolated P values.

Study Design Summary

HEAIS-009 is a single-center ambispective cohort of consecutive adults with advanced solid tumors who initiated at least one ET-SPACE WBH session from January 1, 2022 through December 31, 2028. The index date is the first eligible session. Retrospective and prospective participants share the same core eligibility and outcome definitions, but availability of PROMs and specialized biomarkers may differ.

Primary time contrast	Baseline within 7 days before index WBH versus Week 4 plus or minus 7 days
Acute biomarker window	24 to 72 hours after index WBH
Sustained outcome window	Week 12 plus or minus 14 days
Survival horizon	Up to 12 months after index WBH
Primary PROM target	At least 85 participants with evaluable baseline and Week 4 EORTC QLQ-C30 data

Analysis Populations

Population	Definition
Source cohort	All records identified from ET-SPACE treatment logs or prospective screening, before final eligibility adjudication.
Full analysis cohort	All eligible participants who initiated at least one ET-SPACE WBH session. Used for cohort description and all outcomes supported by available data.
Primary PROM cohort	Eligible participants with a valid baseline EORTC QLQ-C30 and at least one valid post-baseline assessment in the Week 4 window. A mixed-model analysis may additionally use valid Week 12 observations.
Laboratory-evaluable cohort	Participants with an eligible baseline value and at least one eligible post-baseline value for the specific analyte. Defined separately for each analyte.
Imaging-evaluable cohort	Participants with measurable or evaluable disease, an adequate baseline assessment, and at least one post-baseline tumor assessment, or documented clinical progression or death before assessment.
Safety cohort	All eligible participants whose first WBH session was initiated, including sessions that were interrupted or terminated early.
Survival cohort	All eligible participants with a known index date and follow-up information sufficient to determine survival status or censoring.
Session-level cohort	All initiated eligible WBH sessions linked to an eligible participant. Used for target temperature, interruption, and acute safety outcomes.

General Data Conventions

- Calendar day is the basic time unit. Day 0 is the index WBH session.
- Baseline is the value closest to and before the index session within 7 days. If more than one value is equally close, the value obtained latest in clock time will be used.
- The acute window is 24 to 72 hours after the end of the index session. The value closest to 48 hours will be selected for single-time-point summaries.
- The Week 4 window is Day 21 through Day 35. The value closest to Day 28 will be selected when a single value is required.

- The Week 12 window is Day 70 through Day 98. The value closest to Day 84 will be selected when a single value is required.
- Measurements obtained after documented progression, acute infection, major surgery, or hospitalization remain in the main real-world analysis but will be flagged for sensitivity analyses.
- For values below or above an assay limit, the original qualifier will be retained. Substitution rules, if needed, will be prespecified by analyte before analysis.
- Implausible values will be queried against source data and not automatically deleted. Corrected, confirmed extreme, and unresolved values will be documented.

Derived Variables

Variable	Derivation
BMI	Weight in kg divided by height in meters squared.
Percentage weight change	100 times post-baseline weight minus baseline weight, divided by baseline weight.
PNI	Serum albumin in g/L plus 5 times absolute lymphocyte count in 10^9/L.
SII	Platelet count times neutrophil count divided by lymphocyte count, using counts from the same specimen.
NLR	Absolute neutrophil count divided by absolute lymphocyte count.
PLR	Platelet count divided by absolute lymphocyte count.
LMR	Absolute lymphocyte count divided by absolute monocyte count.
OMEDD	Total daily opioid dose converted to oral morphine equivalent mg/day using the protocol-approved conversion table.
EORTC QLQ-C30	Raw items scored and linearly transformed to 0-100 according to the official manual. Higher global or functioning scores are better; higher symptom scores are worse.
ESAS-r total	Sum of nine symptom ratings, each 0-10, yielding 0-90; higher is worse.
ECOG change	Improved if decreased by at least 1 grade; stable if unchanged; deteriorated if increased by at least 1 grade.
ORR	Complete response or partial response as best overall response by RECIST 1.1, divided by the prespecified imaging-evaluable denominator.
DCR	Complete response, partial response, or stable disease as best overall response by RECIST 1.1, divided by the prespecified imaging-evaluable denominator.
PFS	Time from index date to first documented progression or death from any cause, whichever occurs first.
OS	Time from index date to death from any cause.
Thermal target achievement	Maximum observed core temperature at or above the participant-specific target documented before treatment. An additional standardized threshold analysis will use 38.5 degrees C.
Session completion	Completed planned maintenance and cooling phases without unplanned early termination; protocol deviations and interruption reasons will be reported.

Primary Estimand

Population	Primary PROM cohort
Variable	EORTC QLQ-C30 global health status score, 0-100
Treatment or exposure	Initiation of ET-SPACE WBH during routine care
Summary measure	Mean within-participant change from baseline to Week 4
Intercurrent events	Concomitant anticancer treatment, hospitalization, infection, progression, and incomplete WBH will not automatically exclude observations; each will be flagged and addressed in supportive analyses
Interpretation	Observed post-exposure change in the cohort, not a causal effect attributable solely to WBH

Primary Analysis

The primary analysis will use a linear mixed-effects model with EORTC QLQ-C30 global health status as the dependent variable, categorical time as a fixed effect, and a participant-specific random intercept. Baseline and Week 4 are required for the primary contrast; valid Week 12 values may be retained in the model to improve estimation of the covariance structure but do not alter the prespecified Week 4 contrast.

The estimand will be the model-based adjusted mean difference between Week 4 and baseline with a two-sided 95 percent confidence interval and P value. Restricted maximum likelihood will be used. An unstructured covariance matrix will be attempted; if convergence fails, compound symmetry will be used and the change documented.

The primary model will include time and cohort component, retrospective or prospective. A supportive adjusted model will add prespecified baseline prognostic factors subject to data support: age, sex, broad tumor group, locally advanced versus metastatic disease, baseline ECOG performance status, baseline PNI, comorbidity burden, concurrent treatment class, corticosteroid exposure, nutritional intervention, and calendar period. Continuous covariates will be retained continuously and modeled with restricted cubic splines when nonlinearity is clinically plausible and data are sufficient.

Primary Sensitivity Analyses

- Paired-sample analysis of participants with both baseline and Week 4 scores, using a paired t test or Wilcoxon signed-rank test according to distribution.
- Prospective-component-only analysis.

- Analysis excluding assessments after acute infection, major surgery, unplanned hospitalization, or documented progression before Week 4.
- Multiple-imputation analysis under MAR and delta-adjusted pattern-mixture sensitivity analysis under MNAR.
- Robust regression or bootstrap confidence interval if residual assumptions remain materially violated after transformation or model refinement.

Secondary and Exploratory Analyses

Outcome class	Primary summary	Main method	Time
Week 12 QLQ-C30	Mean change	Linear mixed-effects model	Baseline, Week 4, Week 12
Pain NRS	Mean change	Linear mixed model; ordinal sensitivity	Baseline to Week 4
ESAS-r total	Mean change	Linear mixed model or robust model	Baseline to Week 4
ECOG status	Improved/stable/deteriorated	Ordinal mixed model or multinomial model	Baseline to Week 4
PNI	Mean change	Linear mixed model	Baseline to Week 4
SII, NLR, PLR, CRP	Geometric mean change or ratio	Log-linear mixed model	Baseline, 24-72 h, Week 4
CD4/CD8 ratio, NK cells	Mean or geometric mean change	Mixed model by analyte	Baseline, 24-72 h, Week 4
PG-SGA category	Category transition	Descriptive and ordinal model	Baseline to Week 12
Weight and BMI	Mean and percentage change	Linear mixed model	Baseline to Week 12
OMEDD	Median and mean change	Two-part or rank-based method if zero inflated	Baseline to Week 4
WBH-related AE	Participant and session counts	Exact confidence intervals; mixed logistic model exploratory	Through 30 days after last WBH
Thermal target	Proportion achieved	Logistic regression; session clustering	Each session
ORR and DCR	Proportion	Wilson 95 percent confidence interval	Through Week 12 assessment
PFS and OS	Median and landmark rates	Kaplan-Meier; Cox exploratory	Up to 12 months
Metabolic biomarkers	Change by analyte	Separate mixed model for each analyte	Baseline, 24-72 h, Week 4

Descriptive Statistics

- Continuous variables: number observed, number missing, mean, standard deviation, median, interquartile range, minimum, and maximum.
- Categorical variables: count and percentage using a stated denominator.
- Time-to-event variables: number of events, total person-time, Kaplan-Meier median, and prespecified landmark estimates with 95 percent confidence intervals.

- Baseline tables will be presented overall and by cohort component. Standardized differences, not significance tests, will describe imbalance across clinically relevant groups.
- Participant flow will show records identified, exclusions, index treatments, available assessments, and outcome-specific analysis denominators.

Safety Analyses

Adverse events will be tabulated by CTCAE term, maximum grade, seriousness, attribution, and action taken. Participant-level incidence will count each participant once per event type at the maximum grade. Session-level incidence will use the number of initiated WBH sessions as the denominator. Recurrent events may be modeled using generalized estimating equations or mixed-effects Poisson or negative-binomial models, with participant clustering.

Treatment feasibility will include target temperature achievement, session completion, interruption, maximum temperature, time to target, maintenance duration, and total duration. Predictors of target achievement and intolerance will be explored using multivariable mixed logistic regression, with a random participant intercept when repeated sessions are analyzed.

Tumor and Survival Analyses

ORR and DCR will be reported with exact denominators and Wilson confidence intervals. Participants without an adequate post-baseline assessment will not be silently classified as nonresponders; a separate conservative sensitivity analysis will classify death or clinical progression before assessment as nonresponse and all other missing assessments as non-evaluable.

PFS and OS will be summarized from the index date. Participants without an event will be censored at the last adequate disease assessment for PFS and last known alive date for OS. Cox proportional-hazards models will report hazard ratios only as exploratory associations. Proportional hazards will be assessed using Schoenfeld residuals and time interactions when needed.

Analyses relating number of WBH sessions to PFS or OS will not classify participants by future total session count at baseline. Cumulative session count will be modeled as a time-dependent covariate, or a prespecified Week 12 landmark analysis will include only participants alive and event free at the landmark.

Confounding Strategy

Confounder selection will be based on clinical knowledge and a prespecified causal diagram rather than univariable P-value screening. The following domains will be considered:

- Demographics: age and sex.

- Cancer prognosis: broad tumor group, recurrent or metastatic status, metastatic sites, measurable disease, line of therapy, and baseline tumor burden proxy.
- Clinical status: ECOG performance status, BMI, comorbidity burden, baseline laboratory values, recent infection, hospitalization, and corticosteroid use.
- Treatment: systemic therapy class and timing, radiotherapy, surgery or local procedure, analgesic therapy, nutritional support, and palliative intervention.
- Study process: retrospective or prospective component, calendar year, and availability of baseline assessment.
- WBH delivery: achieved temperature, maintenance duration, interruption, and cumulative sessions, using time-aware definitions.

The number of parameters will be constrained by effective sample size. If the full model is unstable, covariates will be reduced by prespecified domain priority, not data-driven stepwise selection. Penalized regression may be used for exploratory prediction but will not replace the prespecified primary model.

Missing Data

1. Report missingness for every primary and secondary outcome at each time point and for every candidate covariate.
2. Compare baseline characteristics of participants with and without the Week 4 primary outcome using standardized differences.
3. Use mixed-effects models under MAR for repeated continuous outcomes, retaining all valid observations from participants in the relevant analysis cohort.
4. For covariate missingness and supportive analyses, use multiple imputation by chained equations when the proportion and pattern justify imputation. The imputation model will include outcomes, time points, exposure measures, cohort component, predictors of missingness, and auxiliary variables.
5. Generate at least 20 imputed datasets and increase the number when the fraction of missing information is high. Combine estimates using Rubin rules.
6. Do not impute participant identity, eligibility, index date, death, or definitively known progression events.
7. Perform complete-case and prospective-only analyses and, for the primary outcome, a delta-adjusted pattern-mixture sensitivity analysis representing worse unobserved Week 4 scores among participants lost to follow-up.

Multiplicity

The primary outcome has a two-sided alpha of 0.05. There is one confirmatory primary contrast. Secondary outcomes are supportive and will be interpreted using effect sizes and 95 percent confidence intervals. Within each biomarker family, immune-inflammatory and metabolism-related, the Benjamini-Hochberg procedure will control the false discovery rate at 0.05. Exploratory and subgroup analyses will not be described as confirmatory regardless of nominal P values.

Subgroup and Effect-Modification Analyses

- Retrospective versus prospective component.
- Broad tumor group, restricted to groups with adequate sample size; otherwise descriptive only.
- Locally advanced versus metastatic disease.
- Baseline ECOG performance status 0-1 versus 2 or higher.
- Baseline nutritional risk according to PG-SGA or prespecified PNI category.
- Concurrent treatment class: immunotherapy, chemotherapy, targeted therapy, endocrine therapy, radiotherapy, supportive care only, or combinations.
- Achieved versus not achieved standardized core temperature threshold of 38.5 degrees C.
- One versus multiple WBH sessions, assessed only with time-dependent or landmark methods.

Subgroup effects will be evaluated by interaction terms rather than separate within-group significance tests. Estimates will be shown with confidence intervals, and sparse strata may be combined or suppressed to protect privacy and avoid unstable inference.

Model Diagnostics

- Inspect distributions, residuals, influence, leverage, and variance structure.
- Assess linearity of continuous predictors and use restricted cubic splines when supported.
- Assess multicollinearity using correlation matrices and variance inflation factors.
- Use robust standard errors when residual heteroscedasticity persists.
- Evaluate mixed-model convergence and covariance assumptions; document any simplified structure.
- For Cox models, evaluate proportional hazards using Schoenfeld residuals and time interactions.
- For logistic and survival models, report calibration and discrimination only when the analysis is explicitly predictive and adequately powered.

Sensitivity Analyses

1. Prospective participants only.
2. Retrospective participants only for outcomes with comparable measurement.
3. Participants with complete paired primary-outcome data.
4. Exclude participants with acute infection, major surgery, or unplanned hospitalization between baseline and Week 4.
5. Exclude outcomes measured after documented progression before the assessment window.
6. Restrict to participants whose concomitant systemic therapy remained unchanged from 14 days before index through Week 4.
7. Alternative outcome windows: Week 4 defined as Day 18-38 and Week 12 as Day 63-105.
8. Multiple imputation under MAR and delta-adjusted MNAR analyses.
9. Robust or nonparametric methods for highly skewed outcomes.
10. Time-dependent and landmark analyses for cumulative session exposure.

Interim Analyses and Data Review

No formal interim efficacy analysis or early-stopping rule is planned. Periodic descriptive reviews may assess recruitment, completeness, range violations, and safety signals. These reviews will not estimate confirmatory treatment effects or change the primary outcome without a documented protocol and SAP amendment.

Statistical Software and Reproducibility

- Analyses will be conducted using validated versions of R, version 4.4 or later, and, if required, SAS version 9.4 or later.
- A frozen data dictionary, analysis dataset specification, and version-controlled code repository will be retained.
- Random seeds will be set for imputation, bootstrap, and resampling procedures.
- Tables and figures will be generated directly from validated analysis datasets whenever feasible.
- All output will identify the analysis population, denominator, time window, and data-cut date.

Planned Tables and Figures

- Figure 1: participant flow diagram.
- Table 1: baseline characteristics overall and by cohort component.
- Table 2: WBH session characteristics and feasibility.
- Table 3: primary and patient-centered outcomes.
- Table 4: nutritional, immune-inflammatory, and metabolic biomarkers.
- Table 5: adverse events by term, grade, and attribution.
- Table 6: tumor response and survival outcomes.
- Figure 2: EORTC QLQ-C30 trajectories with confidence intervals.
- Figure 3: pain, ESAS-r, ECOG, and PNI changes.
- Figure 4: standardized biomarker changes by domain.
- Figure 5: Kaplan-Meier curves for PFS and OS.
- Supplement: missingness maps, model diagnostics, subgroup interactions, and sensitivity analyses.

Database Lock and SAP Deviations

Before database lock, the analysis team will finalize eligibility, outcome windows, derivation code, censoring dates, duplicate resolution, and the list of protocol deviations. The SAP will be signed off institutionally according to local procedures. Deviations from this plan will be tabulated with date, rationale, timing relative to database lock, and effect on interpretation.

References

1. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30. J Natl Cancer Inst. 1993;85(5):365-376.
2. Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual. 3rd ed. EORTC; 2001.
3. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol. 1998;16(1):139-144.
4. Eisenhauer EA, Therasse P, Bogaerts J, et al. Revised RECIST guideline, version 1.1. Eur J Cancer. 2009;45(2):228-247.
5. White IR, Royston P, Wood AM. Multiple imputation using chained equations. Stat Med. 2011;30(4):377-399.
6. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data. BMJ. 2009;338:b2393.
7. Benjamini Y, Hochberg Y. Controlling the false discovery rate. J R Stat Soc Series B. 1995;57(1):289-300.
8. von Elm E, Altman DG, Egger M, et al. The STROBE statement. Lancet. 2007;370(9596):1453-1457.
9. Benchimol EI, Smeeth L, Guttmann A, et al. The RECORD statement. PLoS Med. 2015;12(10):e1001885.

Appendix: Outcome-Specific Reporting Checklist

- State the exact analysis population and denominator.
- State the baseline and post-baseline time window.
- Report the number observed and number missing at every time point.
- Report the effect estimate, unit, 95 percent confidence interval, and P value when applicable.
- Identify whether the estimate is unadjusted, adjusted, imputed, or sensitivity analysis.
- For scales, state the range and direction of better or worse scores.
- For biomarkers, state assay units and any transformation.
- For tumor outcomes, state RECIST evaluability and reasons for missing assessment.
- For survival, state events, censoring rule, data-cut date, and follow-up distribution.
- For safety, report both participant-level and session-level denominators.