SHAPE Test for Preoperative Risk Stratification

NCT05743673

Date of Documentation (IRB modification approval): 7/25/2025

Revised Statistical Analysis Plan

Statistical Analysis Plan

Below is a **protocol-based statistical analysis plan** drafted from the modified SHAPE-HF study document and updated to reflect the current expanded sample size of **N=371**.

1. Overview

This is a **prospective, single-site, non-randomized observational study** evaluating the SHAPE-HF submaximal cardiopulmonary exercise testing system in adults older than 60 years undergoing preoperative evaluation for **moderate- to high-risk surgery**.

The study was originally developed as a **technical proof-of-concept study** with **42 participants**, then modified to a **feasibility study** with **101 participants**, and has now been expanded to **371 participants** to better evaluate the association between preoperative submaximal cardiopulmonary exercise testing results and **early and major perioperative complications**.

The aim of the study is to determine whether preoperative rapid submaximal cardiopulmonary exercise testing can detect early postoperative morbidity, major postoperative complications and further characterize novel variables derived from cardiopulmonary exercise testing that may aid in enhancing risk detection or resource allocation.

The statistical objectives are to:

- describe recruitment and feasibility
- compare SHAPE-HF-derived measures with conventional functional capacity measures
- evaluate associations between SHAPE-HF performance and short-term postoperative morbidity
- evaluate associations between SHAPE-HF performance and 30-day postoperative early and major complications, including cardiovascular events, length of stay, and hospital readmission after surgery.

2. Analysis Populations

2.1 Enrolled Population

All consented participants who are enrolled in the study.

2.2 Tested Population

All enrolled participants who undergo the SHAPE-HF testing session.

2.3 Evaluable Population

All participants with analyzable SHAPE-HF data and available postoperative follow-up data sufficient for endpoint assessment.

2.4 Safety Population

All participants who begin the SHAPE-HF testing procedure will be included in safety summaries for study test-related adverse events.

3. Endpoints

3.1 Feasibility Endpoint (Original Protocol Primary Endpoint)

The original protocol-defined primary endpoint is study feasibility, operationalized as enrollment rate among eligible candidates. In the feasibility phase, success was defined as enrollment of at least 25% of eligible candidates.

For the expanded study, feasibility remains an important operational endpoint and will be summarized descriptively, including number screened, number eligible, number approached, number consented, proportion enrolled among eligible participants, and proportion completing SHAPE-HF testing.

3.2 Primary Clinical Outcome Endpoint for the Expanded Predictive Analysis

The primary clinical outcome endpoint for the expanded outcome-oriented analysis will be the Comprehensive Complication Index (CCI) within 30 days after surgery.

CCI will be calculated from all eligible postoperative complications occurring within the 30-day postoperative window using the Clavien-Dindo severity framework and the standard CCI scoring approach. CCI ranges from 0 to 100, where 0 indicates no postoperative complication and higher values indicate greater cumulative severity-weighted complication burden.

The primary estimand is the association between preoperative SHAPE-HF-derived submaximal cardiopulmonary exercise variables and 30-day cumulative postoperative morbidity burden as measured by CCI.

3.3 Key Secondary Clinical Outcome Endpoints

Key secondary clinical endpoints will include major postoperative complication defined as Clavien-Dindo grade III or higher within 30 days after surgery, any postoperative complication within 30 days, POMS-positive morbidity at prespecified postoperative time points, and persistent POMS morbidity through postoperative day 7 or discharge when available.

Clavien-Dindo grade III or higher will be used as the major morbidity anchor because it identifies complications requiring surgical, endoscopic, radiologic, intensive-care, or life-threatening-level intervention. If grade V mortality occurs, it will be included within the grade III or higher endpoint and summarized separately.

3.4 POMS and Early Postoperative Morbidity Endpoints

Early postoperative morbidity will be assessed using the Postoperative Morbidity Survey (POMS) at protocol-defined time points, currently postoperative days 1, 3, and 5, with postoperative day 7 and discharge included if captured in the final dataset.

POMS will not be treated as a single unidimensional severity scale. Instead, it will be reported as a structured profile of organ-system morbidity. Analyses will include any POMS-positive morbidity at each time point, POMS-positive morbidity by domain at each time point, number of positive POMS domains as a descriptive burden measure, and persistence or resolution of POMS morbidity over time.

3.5 Domain-Level and Mechanistic Outcome Endpoints

Domain-level outcomes will be reported to describe the clinical pattern of postoperative morbidity associated with SHAPE-HF performance. Domains will include pulmonary, infectious, cardiovascular, renal, gastrointestinal, wound, pain, hematologic, neurologic or delirium if available, thromboembolic, readmission, ICU escalation, and other clinically prespecified domains available in the case report forms.

Domain-level analyses are intended to be supportive and mechanistic unless a domain is explicitly prespecified as a key secondary endpoint before database lock. The primary inferential conclusion will be based on the 30-day CCI endpoint, with Clavien-Dindo grade III or higher serving as the principal major-complication endpoint.

3.6 Health Care Utilization and Recovery Endpoints

Health care utilization and recovery endpoints will include hospital length of stay, unplanned ICU admission or escalation of care, hospital readmission within 30 days, days alive and out of hospital at 30 days if derivable, and discharge disposition if available.

3.7 Endpoint Hierarchy

The endpoint hierarchy for the expanded analysis is as follows: primary clinical endpoint, 30-day CCI; key secondary endpoint, Clavien-Dindo grade III or higher within 30 days; supportive secondary endpoints, any complication, POMS-positive morbidity at prespecified days, and persistent POMS morbidity; exploratory endpoints, individual POMS domains, complication categories, length of stay, readmission, ICU escalation, and other recovery measures.

4. Variables

4.1 Baseline Variables

Baseline characteristics will be summarized, including:

- age
- sex
- height
- weight
- body mass index
- planned surgical procedure
- subjective METs
- Duke Activity Status Index
- Revised Cardiac Risk Index
- other protocol-recorded preoperative clinical characteristics

4.2 SHAPE-HF Variables

SHAPE-HF variables will include available device-derived cardiopulmonary performance parameters recorded during the submaximal exercise protocol, including objective measures of aerobic fitness such as estimated VO_2-related parameters, measures of gas-exchange derived pulmonary capacitance, heart rate response, recovery variables, and other recorded test outputs.

4.3 Postoperative Variables

Postoperative variables will include clinically meaningful postoperative morbidity, complications, and recovery measures:

30-day Comprehensive Complication Index score

individual postoperative complications mapped to Clavien-Dindo grade

major postoperative complication defined as Clavien-Dindo grade III or higher within 30 days

POMS-defined early postoperative morbidity by time point and by domain

persistent POMS morbidity through postoperative day 7 or discharge, if available

major adverse cardiovascular events, including myocardial infarction, stroke, congestive heart failure, arrhythmia or atrial fibrillation, angina, and other prespecified cardiovascular events

pulmonary, infectious, renal, gastrointestinal, wound, pain, hematologic, neurologic or delirium, and thromboembolic morbidity domains when available

hospital readmission within 30 days

length of stay

unplanned ICU admission or escalation of care

days alive and out of hospital at 30 days, if derivable

5. General Statistical Principles

All statistical tests will be two-sided with a significance level of 0.05, unless otherwise specified. Continuous variables will generally be summarized using mean and standard deviation for approximately normally distributed variables and median and interquartile range for skewed variables. Categorical variables will be summarized using counts and percentages.

The primary clinical analysis will focus on the association between SHAPE-HF-derived variables and 30-day CCI. Secondary and exploratory analyses will be interpreted in the context of the prespecified endpoint hierarchy.

Because several secondary and domain-level endpoints will be evaluated, results for POMS domains and individual complication categories will be interpreted as supportive or exploratory unless explicitly prespecified as key secondary endpoints. Multiplicity will be addressed by prespecification of the endpoint hierarchy, cautious interpretation of p-values, and, where appropriate for domain-level families of tests, false-discovery-rate-adjusted p-values or unadjusted estimates presented with confidence intervals but without formal claims of confirmatory significance.

Normality and model assumptions will be assessed graphically and, where helpful, with standard diagnostic procedures. For skewed or zero-inflated outcomes such as CCI or length of stay, robust, transformed, rank-based, quantile, bootstrap, or two-part modeling approaches may be used as sensitivity analyses.

6. Descriptive Analyses

Participant flow will be summarized in a CONSORT-style screening and enrollment table, even though the study is not randomized. Baseline characteristics will be presented overall and, where useful, stratified by:

- occurrence of postoperative morbidity
- occurrence of major adverse cardiovascular events
- SHAPE-HF performance categories

Feasibility metrics will be summarized with proportions and 95% confidence intervals.

7. Primary Feasibility Analysis

The primary feasibility analysis will estimate the proportion of eligible participants who enroll in the study.

This will be presented as:

- enrollment proportion
- exact or Wald 95% confidence interval
- comparison of observed enrollment rate against the protocol benchmark of **25%**

If a formal hypothesis test is desired, a one-sample proportion test may be used to evaluate whether the observed enrollment rate differs from the pre-specified benchmark. The main emphasis, however, should remain on estimation rather than hypothesis testing.

8. Comparison of SHAPE-HF With Conventional Functional Measures

The relationship between SHAPE-HF-derived objective measures and conventional preoperative functional assessment measures, including subjective METs and Duke Activity Status Index, will be evaluated using:

- Pearson correlation for approximately normally distributed continuous variables
- Spearman rank correlation when non-normality or outliers are present

Agreement and discrimination may also be explored using:

- scatterplots with fitted lines
- subgroup summaries across categories of subjective functional capacity
- linear regression models with SHAPE-HF variables as dependent or independent variables, depending on the analytic question

Where clinically meaningful cut points are defined for SHAPE-HF variables, group comparisons may be performed using:

- **t-tests** or **Wilcoxon rank-sum tests** for continuous outcomes
- **chi-square** or **Fisher exact tests** for categorical outcomes

9. Analysis of Postoperative Morbidity and POMS Endpoints

Early postoperative morbidity measured by POMS will be analyzed as a structured morbidity profile rather than as a single severity scale.

For each POMS time point, the number and proportion of participants with any POMS-positive morbidity will be summarized with 95% confidence intervals. POMS-positive morbidity by domain will also be summarized at each available time point.

The primary POMS-based supportive endpoint will be any POMS-positive morbidity on postoperative day 5, unless final data availability supports a more clinically appropriate time point such as postoperative day 7 or discharge. Persistent POMS morbidity will be defined before database lock, preferably as any POMS-positive morbidity present at postoperative day 5 or later, postoperative day 7, or discharge, depending on the available assessment schedule.

Domain-level POMS results will be displayed as absolute risks by time point. Where comparisons or associations are modeled, adjusted risk differences, adjusted risk ratios, or odds ratios with 95% confidence intervals will be reported. Modified Poisson regression with robust standard errors may be used to estimate adjusted risk ratios for common binary domain outcomes; logistic regression may be used when event rates or model stability favor odds-ratio reporting.

Longitudinal POMS analyses may use generalized estimating equations or mixed-effects logistic regression to account for repeated measures within participants. These models may include time point, SHAPE-HF predictor, and a SHAPE-HF-by-time interaction when clinically meaningful.

The number of positive POMS domains at each time point may be summarized descriptively as a morbidity burden measure. If modeled, count models, ordinal models, or nonparametric approaches may be considered, recognizing that POMS domains are not interchangeable measures of one latent severity construct.

POMS domain-level analyses will be interpreted as explanatory and mechanistic unless prespecified as key secondary outcomes. Particular attention will be given to pulmonary, cardiovascular, infectious, and renal domains because these are biologically plausible manifestations of reduced physiologic reserve measured by submaximal cardiopulmonary exercise testing.

Analyses will distinguish expected routine postoperative care from morbidity where possible. Sensitivity analyses may exclude items that reflect planned postoperative management rather than unexpected morbidity, such as expected oxygen therapy, nasogastric tubes, urinary catheters, or protocol-driven antibiotics when these are routine for a given procedure.

10. Analysis of 30-Day Comprehensive Complication Index and Major Postoperative Complications

The primary clinical endpoint for the expanded predictive analysis will be 30-day CCI. CCI will be summarized using mean, standard deviation, median, interquartile range, minimum, maximum, and the proportion with CCI greater than 0.

Because CCI may be right-skewed and may include a substantial number of zero values, the primary model will estimate the association between prespecified SHAPE-HF predictors and CCI using regression methods appropriate to the observed distribution. Candidate approaches include linear regression with robust or bootstrap confidence intervals, generalized linear models, quantile regression, rank-based methods, or two-part models separating any complication from positive CCI burden.

The preferred primary reporting scale will include an adjusted mean difference or other prespecified effect estimate in CCI units with 95% confidence intervals. If the final distribution makes mean-based modeling inappropriate, median difference, quantile regression estimates, or bootstrap estimates will be reported as the primary interpretable effect.

Major postoperative complications will be analyzed using Clavien-Dindo grade III or higher within 30 days after surgery. This endpoint will be summarized as counts and percentages and modeled using logistic regression or modified Poisson regression with robust standard errors. Results will be reported as adjusted odds ratios or adjusted risk ratios with 95% confidence intervals.

Any postoperative complication within 30 days will be summarized as a supportive binary endpoint. Individual complications and complication categories will be summarized descriptively and modeled only when event counts are sufficient.

Complication categories may include cardiovascular, pulmonary, infectious, renal, gastrointestinal, wound, neurologic or delirium, thromboembolic, bleeding or transfusion-related, reoperation, readmission, and ICU escalation endpoints.

If event timing is sufficiently available and accurate, time-to-event methods such as Kaplan-Meier estimation and Cox proportional hazards regression may be used in supplementary analyses. However, because the protocol describes 30-day chart-extracted outcomes rather than a dedicated time-to-event framework, binary 30-day and CCI burden analyses will be considered primary.

11. Major Adverse Postoperative Events and Predictive Model Assessment

Major adverse postoperative events will be evaluated as clinically meaningful secondary and exploratory endpoints rather than as the sole primary clinical endpoint. The major clinical outcomes analysis will therefore focus on three complementary layers: cumulative morbidity burden using CCI, major morbidity using Clavien-Dindo grade III or higher, and early postoperative organ-system morbidity using POMS.

The major adverse cardiovascular event composite will include myocardial infarction, stroke, atrial fibrillation or clinically significant arrhythmia, congestive heart failure, angina, and other prespecified cardiovascular complications captured in the case report forms. The exact component list and coding rules should be finalized before database lock.

Multivariable regression will assess the association between preoperative SHAPE-HF variables and major adverse postoperative events. Models will adjust for clinically relevant baseline covariates expected to include age, sex, body mass index, RCRI, subjective METs, DASI, surgical risk category or procedure category, and other prespecified perioperative factors as supported by event counts.

Model performance will be summarized using measures appropriate to the outcome. For binary outcomes, this may include odds ratios or risk ratios with 95% confidence intervals, c-statistic or area under the ROC curve, calibration

assessment, and decision-curve or net-benefit summaries if justified. For CCI, performance may include explained variation, calibration plots for predicted morbidity burden, mean absolute error, or bootstrap optimism-corrected performance estimates where feasible.

Nested model comparisons may be used to evaluate whether SHAPE-HF adds predictive value beyond conventional assessment alone. For example, a base model may include age, RCRI, subjective METs, and DASI, with an expanded model adding prespecified SHAPE-HF variables.

Incremental value may be summarized using likelihood ratio tests, change in c-statistic for binary major morbidity endpoints, integrated discrimination improvement or net reclassification improvement where appropriate, and calibration or prediction error comparisons. These analyses will be interpreted cautiously given the observational design and sample-size constraints.

- sex
- body mass index
- RCRI
- subjective METs
- Duke Activity Status Index
- procedure category
- surgical risk category

If sample size permits, exploratory subgroup analyses may be performed by:

- age strata
- sex
- surgical category
- higher versus lower RCRI category within the eligible range
- higher versus lower SHAPE-HF performance category

Subgroup analyses will be considered exploratory and interpreted cautiously.

13. Missing Data

The protocol includes a section for handling missing data, but the operational approach should be stated explicitly in the SAP.

The extent and pattern of missing data will be summarized for all major variables and endpoints.

The primary approach will be:

- use available-case analysis for descriptive summaries
- use complete-case analysis for primary regression models if missingness is minimal

If missing covariate data are more substantial and plausibly missing at random, multiple imputation may be used in sensitivity analyses, especially for multivariable outcome models.

Participants without adequate postoperative follow-up to determine the endpoint will be excluded from endpoint-specific regression analyses but will remain in enrollment and testing summaries.

14. Outliers and Data Quality

Continuous SHAPE-HF and clinical variables will be reviewed for implausible values, outliers, and data entry errors prior to database lock.

Outliers that appear physiologically plausible will generally be retained. Clearly erroneous values may be corrected using source documentation when available or otherwise set to missing with documentation.

Derived variables and endpoint coding rules should be finalized before database lock.

15. Safety Analysis

Adverse events related to the SHAPE-HF testing session will be summarized in the safety population.

Safety summaries will include:

- number and proportion of participants with any test-related adverse event
- type of adverse event
- severity, where available
- whether testing was stopped early due to predefined stopping criteria

Given the noninvasive and abbreviated exercise nature of the intervention, safety analyses are expected to be descriptive.

16. Interim Analysis

No formal interim efficacy analysis is described in the protocol text reviewed. Unless added in a later amendment, no interim hypothesis-testing analysis is planned. Operational or recruitment monitoring may occur during study conduct, but such monitoring will not involve formal alpha-spending adjustments unless explicitly introduced later.

17. Statistical Software

Analyses will be performed using validated statistical software (**R, SPSS)**

The final study report will specify the exact software version used.

18. Tables, Listings, and Figures

Planned outputs should include participant flow summary, baseline characteristics table, SHAPE-HF and conventional functional measure summary table, postoperative endpoint summary table, and adverse events table.

The postoperative endpoint summary table should include 30-day CCI, Clavien-Dindo maximum grade, Clavien-Dindo grade III or higher, any complication, individual major complications, POMS positivity by time point, POMS domains by time point, length of stay, readmission, ICU escalation, and days alive and out of hospital if derivable.

Planned figures should include correlation plots for SHAPE-HF versus DASI and subjective METs, distribution plots of CCI overall and by SHAPE-HF risk category, boxplots or violin plots of CCI by SHAPE-HF categories, ROC curves for binary major morbidity models if applicable, calibration plots if feasible, and forest plots of adjusted associations for key secondary endpoints.

A POMS domain-by-time heatmap should be generated to display absolute risks or adjusted effect estimates across postoperative days and domains. This figure will support interpretation of whether SHAPE-HF variables are associated with broad persistent morbidity or with specific organ-system patterns.

Domain-level findings will be presented with emphasis on effect sizes and confidence intervals rather than p-values alone.

19. Protocol-Aligned Statistical Summary

In protocol terms, the statistical framework remains anchored in three linked questions.

First, can eligible older adults be recruited and tested successfully in the preoperative clinic.

Second, how do SHAPE-HF objective measures compare with conventional functional capacity assessment.

Third, are SHAPE-HF-derived preoperative performance measures associated with postoperative morbidity burden, major 30-day complications, and early postoperative organ-system morbidity.

With the expansion to N=371, the study moves beyond proof-of-concept and feasibility alone into an outcome-oriented observational analysis. The clinical endpoint framework will therefore use 30-day CCI as the primary cumulative morbidity burden endpoint, Clavien-Dindo grade III or higher as the major morbidity endpoint, POMS as the early recovery and physiologic morbidity profile, and domain-level reporting to explain the clinical pattern of postoperative morbidity associated with submaximal cardiopulmonary exercise performance.

This structure is intended to avoid overreliance on a single binary complication endpoint while preserving a clear inferential hierarchy and a clinically interpretable story of risk prediction.