Introducing Palliative Care (PC) within the Treatment of End Stage Liver Disease (ESLD): A Cluster Randomized Controlled Trial

Statistical Analysis Plan (SAP)

Sponsor: Thomas Jefferson University

Funding: PCORI (Patient Centered Outcomes Research Institute)

Authors:
Yang Yue, PhD
Jeffrey D. Leimberger, PhD
Andrzej S. Kosinski, PhD

Version 3.0

Date: April 3, 2025

Table of Revisions

Section	Date	Revision Summary
2.3	10/292021	Sample size re-computation (with approval from PCORI)
2.2	4/12/2024	Added non-inferiority analysis (with approval from PCORI)

TABLE OF CONTENTS

ABBREVIATIONS

AUDIT	Alcohol Use Disorders Identification Test
BMI	Body Mass Index
CI	Confidence Interval
CTP	Child-Turcotte Pugh Score
DCRI	Duke Clinical Research Institute
DSMB	Data and Safety Monitoring board
DT	Distress Thermometer
EC	Executive Committee
ECOG	Eastern Cooperative Oncology Group
ESAS	Edmonton Symptom Assessment Scale
ESLD	End Stage Liver Disease
FACT-Hep	Functional Assessment of Cancer Therapy – Hepatobiliary
FAMCARE-P	Family Satisfaction with Care
FCS	Fully Conditional Specification
ICC	Intraclass Correlation Coefficient
ITT	Intention to Treat
MAR	Missing at Random
MCID	Minimal Clinically Important Difference
MELD	Model for End Stage Liver Disease
MI	Multiple Imputations
mITT	Modified Intention to Treat
PCORI	Patient Centered Outcomes Research Institute
RCT	Randomized Controlled Trial
PC	Palliative Care
PHQ-9	Personal Health Questionnaire
PROMIS-29	Patient Reported Outcomes Measurement Information System
SAP	Statistical Analysis Plan

SD	Standard Deviation
SOP	Standard Operations Procedures
QOL	Quality of Life
VA	Veterans Affairs
ZBI-12	Zarit Burden Interview

1 PURPOSE

This statistical analysis plan (SAP) briefly overviews the study design and objectives, outlines the types of analyses and data presentations that address study objectives. The SAP describes, in detail, the statistical methods for efficacy analyses specified in the study protocol (Protocol titled "Title: Introducing Palliative Care (PC) within the Treatment of End Stage Liver Disease (ESLD): A Cluster Randomized Controlled Trial", Version 4.0). Duke Clinical Research Institute (DCRI) will conduct all statistical analyses described in this SAP.

2 STUDY OVERVIEW

This section summarizes the study objectives and design as background for the statistical methods presented in the SAP. For definitive details of the study objectives and design, consult the study protocol. Each participating patient/caregiver dyad (caregiver if available) will participate for 3 months of study intervention, and 9 months of follow up for a total of 12-months of study participation. Survival data will be collected for 12 months of participation. All patient-caregiver dyads will receive the intervention at the initial visit, and at 1, 2, and 3 months from the initial visit. Initial/ first visit will occur within 6 weeks after informed consent. All study visits can occur in person, or remotely (by phone or telehealth with the institutional secure platform). Ascertainment of baseline assessments will occur within 1 week before the scheduled in person initial visit. 1, 2 and 3 month visits can occur within a window period of 1 week (before/after) of their respective times from the initial visit. All caregiver visits can occur in person or by phone, as preferred by the caregivers. The initial caregiver visit can occur on the same day or within 1 week of patient's initial visit. Study follow up data collection will be conducted at 6, 9 and 12

months, by phone, paper or online, within 1 week (before/after) of the due date. Originally, we planned to enroll 1260 patients, but the trial sample size was modified to enrollment of 936 patients at 19 participating sites.

2.1 Study Design

This is a two-armed cluster randomized controlled trial (RCT). Randomization occurred at the level of clinical centers, and was stratified by VA (Veterans Affairs) vs non-VA. Each arm had initially 7 clinical centers.

Clinical centers are randomized to one of two models of palliative care (PC) delivery: Model 1- Consultative PC (i.e., PC offered by a board-certified or board-eligible PC provider) or Model 2- Trained hepatologist led PC (i.e., PC offered by a trained hepatologist). An additional 5 centers were randomized (resulting in one in Model 1 and four in Model 2) and the study utilized 19 centers. Standardized protocols (including visit agenda) were be followed at each of the clinical centers to maintain intervention fidelity across sites in both models. For additional details, please refer to the study protocol.

2.2 Study Objectives

The primary objective of the PAL-LIVER study is to assess the comparative effectiveness of two PC delivery models for patients with ESLD in improving the disease specific quality of life from baseline to 3 months as assessed using FACT-Hep (Functional Assessment of Cancer Therapy- Hepatobiliary). The primary superiority hypothesis is that compared to consultative PC, ESLD patients receiving trained hepatologist led PC will have higher QOL (Quality of Life) change scores at 3 months after baseline. The superiority hypothesis will be assessed first and only if such superiority is not demonstrated (i.e., when superiority p-value > 0.05) then a non-inferiority (NI) analysis of the trained hepatologist led PC (Model 2) will be performed. The NI margin is pre-specified as DELTA = 4. If the lower limit of the two-sided 95% Confidence Interval (CI) of the estimated treatment effect (Model 2 minus Model 1) is above negative DELTA, then the non-inferiority of Model 2 will be demonstrated.

The secondary objectives are:

1) To compare the effects of the two above-mentioned models of PC delivery (as a change from baseline to 3 months) on:

- Overall quality of Life, assessed using PROMIS-29 (Patient Reported Outcomes Measurement Information System);
- Patient's symptom burden, assessed using modified ESAS (Edmonton Symptom Assessment Scale);
- Patient's depression severity, as assessed by PHQ-9 (Personal Health Questionnaire);
- Patient distress, as assessed by Distress Thermometer (DT) ;
- Patient satisfaction with care, assessed using FAMCARE-P13 (Family Satisfaction with Care);
- Caregiver burden, assessed by Zarit Burden Interview (ZBI-12);
- Caregiver quality of life, assessed using PROMIS-29;
- Goal concordant care measures assessed using advance directive and advance care planning documentation, a survey questionnaire and qualitative interviews with patients and caregivers.

2) To compare the effects of the two above models of PC delivery on:

- Healthcare utilization from the time of initial visit (30 and 90-day scheduled and unscheduled hospital admission rate, unscheduled office visits within 1 year, ER visits within 1 year).

3) Conduct semi structured interviews to explore and compare patient, caregiver, and provider experiences with the two models.

- Qualitative interviews of patients' and caregivers' experiences with receiving PC as a part of routine care provided by hepatologists versus palliative care providers;
- Qualitative interviews of palliative care and hepatology providers' experiences and confidence with providing PC to ESLD population.

4) The exploratory objective is to compare the effects of the two comparative models on survival over 1 year. We will use the date of informed consent as the first day in the trial.

2.3 Sample Size Determination and Statistical Power

2.3.1. Sample Size Determination

Patient-reported quality of life, as measured by FACT-Hep total score, is the primary outcome measure in this study. Published estimates show the average performance on this measure for typical patients with ESLD is 143.0 points (SD 20.6) and a difference of 9 points on this measure reflects a minimal clinically meaningful difference (MCID) in QOL [1,2]. However, the power estimates below are based on SD=19.1 for change in FACT-Hep total score from baseline to 3 months observed in the study at the time of sample size re-computation (**n=936**). Patient level data analysis will be considered with adjustment for clustering, i.e., adjustment for correlation of patients' outcomes within a cluster (site). This correlation is quantified by the intraclass correlation coefficient (ICC). At the time of sample size re-computation ICC=0.09 was noted in the study.

2.3.2. Statistical Power for Superiority Hypothesis

Power for superiority hypothesis is **83.2%** and it is based on the final enrollment in the sites and estimated (as of September 25, 2024) counts of patients evaluable for change in FACT-Hep total score from baseline to 3 months.

Power was derived via simulations utilizing multivariate normal distribution with the earlier mentioned SD=19.1 and ICC= 0.09. Each simulation consists of generating, for each of the 19 sites, the evaluable patient count of realizations from this distribution. Since only difference in means between the two arms matters, for simplicity, the consultative PC sites were considered to have zero mean and the Hepatologist led PC sites to have mean equal to MCID=9. For each simulation a mixed model $\text{change}_{ij} = \alpha + \beta trt_{ij} + s_j + e_{ij}$ was fit to the simulated data (with trt=1 for Hepatologist led PC-Model 2 and trt=0 for Consultative PC-Model 1) where response change_{ij} is change in FACT-Hep total score from baseline to 3 months for the i-th patient in the j-th site (j=1,2,…,K=19). Random effects $s_j \sim N(0, \sigma_s^2)$ inducing correlation within sites are assumed to be independent across the sites and also independent from independent random

errors $e_{ij} \sim N(0, \sigma_e^2)$. Proportion (out of 100,000 simulations) of treatment variable p-values less than 0.05 provided estimate of the power for superiority.

2.3.3. Statistical Power for Non-inferiority Hypothesis

Note that "non-inferiority" means that the true mean FACT-Hep total score change (CHANGE_H) for the Hepatologist led PC is allowed to be less than the true mean FACT-Hep total score change (CHANGE_C) for the Consultative PC, but by not more than the so-called NI margin (DELTA). Namely, we hope to reject null hypothesis H0: CHANGE_H ≤ CHANGE_C – DELTA in favor of alternative hypothesis H1: CHANGE_H > CHANGE_C – DELTA (or equivalently H1: β = CHANGE_H - CHANGE_C > – DELTA). This is accomplished by first computing a lower two-sided 95% confidence limit for β. Subsequently, if this limit is above negative DELTA, then non-inferiority (with NI margin DELTA) of the Hepatologist led PC as compared to the Consultative PC is declared. The NI margin is pre-specified at DELTA = 4.

Power for non-inferiority (NI) of Model 2 (Hepatologist led PC) is **79.2%** and was obtained via simulations as outlined above although for the NI power computations we assumed that Hepatologist led PC performs better than the Consultative PC only by half of MCID=9 (i.e., 4.5) specified in superiority hypothesis. For each simulation, the lower two-sided 95% CI limit (L) for β was computed. Proportion (out of 100,000 simulations) with L > -DELTA provided power for non-inferiority of Model 2 (Hepatologist led PC).

2.4 Analysis Populations

There are two analysis populations defined for this protocol:

1. ITT (Intention to Treat) population including all randomized patients.
2. mITT (modified Intention to Treat) population is as ITT population with the exception that patients who undergo liver transplantation or hospice transfer prior to the 3 month end point are excluded. The reason for consideration of the mITT population is that the

primary quality of life FACT-Hep total score outcome is measured at 3 months and liver transplant is a curative treatment and patients in hospice undergo a much more intensive palliative care. The mITT population will be utilized in the primary analysis.

2.5 Study Endpoints

2.5.1 Primary Efficacy Endpoint

Change in quality of life from baseline to 3 months, as assessed by FACT-Hep total score (3 month score minus baseline score) in the mITT population. Total score is calculated from the FACT-Hep eCRF values according to the FACIT scoring guidelines version 4 (www.facit.org). A detailed model with adjustment for the baseline FACT-Hep total score and Veterans Affairs status will be discussed later in the document (section 5.1).

3.5.2 Secondary Efficacy Endpoints

The endpoints inset below will be assessed as a change from baseline to 3 months (3-month score minus baseline score) in the mITT population:

- Overall quality of life (PROMIS-29, separate T scores)
- Symptom burden (modified ESAS, point score for each item, potentially total score for ESAS-r (9 components) and total score for the 13 components from this study)
- Depression severity (PHQ-9, total score)
- Distress (Distress Thermometer, total score)
- Satisfaction with care (FAMCARE-P, total score)
- Caregiver burden (ZBI-12, total score)
- Caregiver quality of life (PROMIS-29, separate T scores)
- Goal concordant care (Patient and Caregiver Questionnaire, total score)

Healthcare utilization in the mITT population:

- Proportion of patients with at least one scheduled or unscheduled hospital admissions within 30 days from the initial visit (based on data available from 1 month visit).
- Proportion of patients with at least one scheduled or unscheduled hospital admissions within 90 days from the initial visit (based on data available from 3 months visit).
- Rate of ER visits within 1 year from the initial visit.
- Rate of unscheduled office visits within 1 year from the initial visit.

3.5.3 Exploratory Endpoint

Survival over 1 year in the ITT population. We will use the date of informed consent as the first day in the trial. Or those alive, follow-up time will be censored at the time of loss to follow-up or at the time of completion of the study.

3 PATIENT DISPOSITION

Patient disposition will be summarized for the ITT population with counts and percentages for all patients and will include the following:

- Patient who completed study (enrollment to 1 year follow-up).
- Evaluable patients (patients for whom FACT-Hep baseline to 3 months change score can be calculated).
- Patients who underwent liver transplantation or hospice transfer prior to the 3 month end point assessment.
- Patients who discontinued the study and reasons for study discontinuation.

4 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

4.1 Demographics

Patient and caregiver demographics will be summarized separately for both ITT and mITT populations. Demographics will be summarized as mean (SD), median (25, 75 percentiles), min

and max for continuous and counts (percentages) for categorical variables. Demographics will include age, sex, race, ethnicity, education level, employment status, marital status, insurance status, smoking status, height, weight and BMI (Body Mass Index)

4.2 General Medical History

Medical history will be summarized as counts (percentages) for the patients and caregivers in both ITT and mITT populations. Medical history will be presented by the diseases pre-specified in the eCRF.

4.3 Current Liver Conditions

Current liver conditions pre-specified on the appropriate eCRF will be summarized as counts (percentages) for the patients in both ITT and mITT populations.

5 EFFICACY ANALYSES

5.1 Primary Efficacy Analyses

The primary analysis will be performed on the mITT population. To account for the cluster-based randomization we will utilize linear mixed model with the treatment arm indicator (binary variable) as a fixed effect and center (cluster) as the random effect. We will also adjust for the baseline FACT-Hep total score and the binary randomization strata variable (VA vs. non-VA). Namely we will consider the model $Y_{ij} = \beta_0 + \beta_1 \text{Treatment} + \beta_2 \text{Baseline FACTHep} + \beta_3 \text{Stratum} + a_j + e_{ij}$ where response Y_{ij} is change in FACT-Hep total score from baseline to 3 months for the i-th patient in the j-th center (j=1,2,…,19). The random effect of center j is $s_j \sim N(0, \sigma_s^2)$ and is assumed to be independent from random errors $e_{ij} \sim N(0, \sigma_e^2)$. Binary variable "Treatment" is coded as 1 for the hepatologist led PC and 0 for the Consultative PC. Binary variable "Stratum" is coded 1 for VA and 0 for non-VA.

The superiority null hypothesis is H_0: $\beta_1=0$ and alternative hypothesis H_1: $\beta_1 \neq 0$. We will reject the null hypothesis if two-sided P-value is < 0.05 and conclude that treatment has significant effect on change in FACT-Hep total score.

Only if superiority is not demonstrated (i.e., when superiority p-value ≥ 0.05) then the non-inferiority (NI) analysis will be performed. Non-inferiority of Model 2 (Hepatologist led PC) will be demonstrated if the lower two-sided 95% CI limit (L) for β will be greater than negative pre-specified DELTA margin, i.e., when L > - 4.0.

In addition to the statistical hypothesis testing, 95% confidence intervals descriptively summarizing the difference in outcome between the two treatment arms, as well as outcome in each randomized arm will be computed.

A complete case analysis will be performed if the primary analysis has ≤ 10% of patients with missing outcome information. If there are more than 10% of patients with missing outcome information, we will utilize multiple imputations (MI) (see section 12.4 for more details). Variables used in multiple imputations will include FACT-Hep total score values measured at baseline, 1 month, 2 months, and 3 months as well as treatment assignment, stratum variable (VA vs. Non-VA), and baseline demographics (including age, gender, education level, employment status, marital status, and smoking status). Change in FACT-Hep total score will be computed with baseline and 3 month FACT-Hep total score values resulting from MI.

5.2 Primary Efficacy Sensitivity Analyses

Missingness of the primary outcome may also occur due to mortality prior to the 3 month follow-up. It is not likely mortality prior to the 3 month follow-up will bias treatment effect estimation because of randomization and we do not expect different survival due to treatment by that time. However, we will perform two sensitivity analyses. The first sensitivity analysis will be performed only if there are more than 10% of missing records and will utilize the mITT population but additionally exclude patients who died prior to the 3 month follow-up (MI will be utilized). The second sensitivity analysis will utilize the whole mITT population but we will set change in FACT-Hep total score (baseline to 3 months) for those who died prior to 3 months to the average of the 5% largest reductions in FACT-Hep total score among those alive by 3 months. MI will be utilized if there are more than 10% of missing records.

5.3 Primary Efficacy Supplementary Analysis

A supplementary analysis will involve multivariable analysis with the mixed linear regression model as in the primary analysis model but with inclusion of additional baseline covariates as fixed effects: age, sex, race, underlying diagnosis, comorbid conditions, any non-liver malignancy, and potentially other baseline scores. Multiple imputations will be used if needed.

5.4 Secondary Efficacy Analyses

The secondary endpoints will be presented as descriptive statistics by visit.

Where appropriate (i.e., change from baseline can be calculated), models similar to the primary endpoint superiority analysis may be utilized for secondary endpoint analyses.

The healthcare utilization endpoints will be analyzed as follows. For scheduled and unscheduled hospital admissions within 30 and 90 days of enrollment, proportions will be analyzed using random effects logistic regression. Counts of hospital admissions, ER visits, and unscheduled office visits over 1 year will be analyzed using Poisson regression with robust standard errors and non-identity dispersion via generalized estimating equations.

5.5 Exploratory Analyses

Survival will be summarized with Kaplan-Meier curves. Cox regression will be used for multivariable models, and robust standard errors will be used to account for within center clustering. ITT population will be used for the survival analyses.

Longitudinal analysis of FACT-Hep total score at 1, 2, 3, 6, 9 and 12 months will be performed utilizing the ITT population. Measurements of Fact-Hep will be set to missing after patients receive liver transplant or are transferred to hospice. A mixed model will be considered with center as a random effect and unstructured covariance structure (alternatively autoregressive covariance structure in case of a convergence difficulty) for repeated measurements of FACT-Hep over time within a patient. As in the primary analysis, the model will have treatment effect and adjustment for baseline FACT-Hep total score and VA status.

Additional covariates with be considered for multivariable modelling. MI will be utilized if more than 10% of missing records are present.

6 SUB-GROUP ANALYSES

For subgroup analyses, the following groups will be evaluated: transplant eligible vs. ineligible, presence of HCC vs. no HCC, and disease severity as assessed by MELD score <20 vs. ≥20. Using multivariable linear mixed models, a potentially differential impact of the intervention will be assessed across complementary subgroups by testing for interaction of the treatment variable with a subgrouping variable by adding terms for the subgroup and its interaction with treatment to the model from the primary analysis. In addition to the test of interaction, we will present estimated treatment effects and the corresponding 95%CI for subgroup treatment effects from this model. These descriptive summaries will be interpreted in conjunction with the formal interaction tests. The main subgroup analyses will be conducted for the primary endpoint.

7 LABORATORY MEASURES

MELD (Model for End Stage Liver Disease) labs will be summarized as mean (SD), median (25, 75 percentiles), min and max by visit for the baseline visit, and 3, 6, 9, 12 month visits.

8 MEDICATIONS RELATED ANALYSIS

8.1 Concomitant Medications

Specific medication types prescribed as a result of the intervention through 3 months will be summarized as counts (percentages) for the patients in the ITT and mITT population. Medication types include medications for pain (opioids, OTC analgesics, herbal supplements), depression/anxiety, anti-psychotic, insomnia medications, itching medications, diuretics, beta-blockers, and Rifaximin.

9 ADDITIONAL NON-ENDPOINT ASSESSEMENTS

Three additional assessments will be summarized. ECOG (Eastern Cooperative Oncology Group) will be summarized as counts (percentages) by visit for the initial, 1, 2, 3-month visits. The individual components that make up the CTP (Child-Turcotte Pugh Score) and the CTP class will be summarized as counts (percentages) and the CTP score will be summarized as mean (SD), median (25, 75 percentiles), min and max for the baseline, 3, 6, 9,12 month visits. The individual components of the AUDIT (Alcohol Use Disorders Identification Test) will be summarized as counts (percentages) and the AUDIT partial and total scores will be summarized as mean (SD), median (25, 75 percentiles), min and max for the baseline and 3 month visit.

10 DATA SOURCES

SAS datasets created from the PAL-LIVER eCRF (Electronic Case Report Form) contains data entered by participating sites, as well as data entered by patients and caregivers.

11 TIMING OF ANALYSES

11.1 DSMB Analyses

DSMB (Data and Safety Monitoring board) will have planned reviews of data approximately every 6 months starting after enrollment through the end of trial follow-up. Unscheduled data review meetings may be called by the DSMB or held at the request of the EC (Executive Committee).

11.2 Final Analyses

The final analyses will be carried out after the last enrolled patient has completed the 6th month follow-up assessment, the database has been cleaned and database lock has occurred. A preliminary analysis of the primary endpoint will take place after the month 3 visit data has been entered and substantially cleaned. Data will be exported from the clinical database and archived for the month 3 analyses.

12 STATISTICAL CONSIDERATIONS

12.1 General Analysis Conventions/Rules

Continuous variables will be summarized as number of observations, mean, standard deviation (SD), minimum, 25th percentile, median, 75th percentile, and maximum. Categorical variables will be summarized as frequency counts and percentages. Statistical comparisons will be performed using two-sided significance tests.

12.2 Statistical Software

The majority of the statistical analyses will be performed using SAS®, version 9.4 or higher (SAS Institute Inc., Cary, NC) and R statistical software. Additional statistical software will be utilized as needed.

12.3 Verification of Programming Codes

All tables, listings, and graphs will be verified and reviewed before being considered final. The verification process will ensure that the numbers are produced by a statistically valid method and that the execution of the computations is correct. Suitably qualified personnel who have not been previously involved in the production of the original programming code will perform the verification procedures. Methods of verification will include independent programming of all analysis datasets as specified in the DCRI Statistical Standard Operations Procedures (SOPs). Tables will be reviewed for accuracy, consistency with this analysis plan, consistency within tables, and consistency with corresponding output. Once verification is complete, all documentation of the verification process will be filed in the study statistical documentation repository for PAL-LIVER as required by the DCRI statistical SOPs of the Duke Clinical Research Institute (DCRI).

12.4 Handling of Missing Data

Every effort will be made to minimize the extent of missing values by proactive strategies, inevitably we will need to deal with some missing value problems [3]. To prevent and limit

missing data other than due to patient's death or curative therapy prior to completion of the study we will perform data checks and consistency checks during the study and collaborate with sites to complete missing or incorrect data items. If ≤up to 10% of patients have missing data for a particular analysis, we will conduct complete case analysis utilizing only non-missing data. However, if more missing data (> 10%) is present we will utilize multiple imputations (MI) to mitigate potential bias of the complete case analysis and appropriately reflect the uncertainty due to imputations [4]. The MI method assumes that the data are missing at random (MAR), meaning that the probability of missingness depends on observed data but not on unobserved data. An imputation model will be constructed including all variables that are part of the analysis model and additional auxiliary variables (including age, gender, education level, employment status, marital status, and smoking status) that may be predictive of the missingness. Both continuous and categorical variables will be included. Multiple imputation will be performed using the PROC MI procedure in SAS 9.4 utilizing fully conditional specification (FCS) method for data sets with arbitrary missing patterns [5]. A total of 1000 imputations will be performed. Each of the imputed datasets will be analyzed separately, and the results will be combined using SAS PROC MIANALYZE to obtain valid statistical inferences that reflect the uncertainty due to imputations of missing data and account for the variability both within and between the imputed datasets.

12.5 Addressing not answered survey questions

In some cases, patients fail to answer all of the survey questions, leading to inaccurate calculated raw scores which consider only the answered questions. Some assessment surveys provide mechanisms for addressing partially answered questionnaires. For assessments, which do not provide explicit recommended mechanisms, we will use the following approach to adjust the raw score.

If at least 70% of the questions are answered, we will compute an adjusted score according to the following formula: adjusted score = (raw score) * (number of questions in a survey) / (number of questions answered). This is equivalent to imputing the average value of answered questions for the not answered questions.

If less than 70% of the questions are answered, the survey score will be set to missing.

12.6 Multiple Comparisons

We will not perform multiple comparison adjustment in this study.

13 REFERENCES

1. Steel, J. L., Eton, D. T., Cella, D., Olek, M. C., & Carr, B. I. (2006). Clinically meaningful changes in health-related quality of life in patients diagnosed with hepatobiliary carcinoma. Annals of oncology, 17(2), 304-312.
2. Kavalieratos, D., Corbelli, J., Zhang, D., Dionne-Odom, J. N., Ernecoff, N. C., Hanmer, J., ... & Morton, S. C. (2016). Association Between Palliative Care and Patient and Caregiver Outcomes: A Systematic Review and Meta-analysis. JAMA, 316(20), 2104-2114.
3. Little, R. J. A., D'Agostino, R., Dickersin, K., Emerson, S. S., Farrar, J. T., Frangakis, C., ... & Rotnitzky, A. (2010). The Prevention and Treatment of Missing Data in Clinical Trials. Pannel on Handling Missing Data in Clinical Trials. The National Academies Press.
4. Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons Inc., New York. http://dx.doi.org/10.1002/9780470316696
5. Van Buuren, S. (2007). “Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification.” Statistical Methods in Medical Research 16:219–242.