Statistical Analysis Plan

Text-based intervention to minimize the time burden of routine cancer care
NCT05134636
Last Updated 3/1/2022

1 Statistical Plan

1.1 Sample Size and Power Determination

AIM 1 Pilot study: This protocol describes a cross-sectional study designed to investigate the operating characteristics of a text-based e-triage to capture patients experiencing immune-related adverse events against the gold standard of documented toxicity assessments in the EMR. Due to the granular nature of our e-triage questionnaire we believe a sensitivity of 95-100% is very reasonable to anticipate.

Results from the interim analysis including 48 patients are shown in Table 1. Median age was 68.5 (IQR 58-72), 58% were female, 88% white, and most patients either had thoracic (25%), genitourinary (27%), head/neck (15%), or skin (15%) cancers. 53.2% of patients had at least 1 toxicity related adverse event documented in the EMR. Sensitivity to identify at least one toxicity related adverse event was 100% (95% CI 0.863-1.00) aligning with our primary endpoint. Despite being at our target sensitivity we identified two underperforming questions (pain and rash) and therefore optimized the instrument creating a version 2 (see appendix). Assuming the sensitivity and prevalence are same as those in the preliminary data, a sample of 47 patients will produce adequate precision for the estimate for sensitivity, with the half width of the 95% confidence interval no larger than 7%.

AIM 2 Randomized control trial: We plan to enroll 176 patients and expect to have 160 patients left assuming a 10% drop out rate (e.g., due to disease progression and need to change therapy). Based on the preliminary data recently published by this group measuring total care time of patients treated at the Penn Abramson Cancer Center[8], we expect the total care time in the control group to have a mean of 130 hours with a SD of 118. A sample size of 80 per group will provide 79% power to detect a 40% reduction in the total care time (130 vs. 78) assuming a type I error rate of 0.05.

1.2 Statistical Methods

Aim 1 Pilot study: Sensitivity and specificity will be calculated along with exact 95% confidence intervals (CIs). Sensitivity will be calculated as the proportion of patients with toxicity captured by the text-based e-triage among patients with toxicity as documented by the provider in the EMR. Specificity will be calculated as the proportion of negative e-triage screens among patients without symptoms as documented by the provider in the EMR. We will additionally be performing a series of psychometric analyses including response distribution to assure responses are not skewed, item-scale correlation, internal consistency, floor and ceiling effects and item reduction as well as response rates.

Table 2. Power to detect reduction in total care time (N=160)

Reduction in care time	Power
35%	68%
40%	79%
45%	88%
50%	93%

AIM 2 Randomized control study: Standard descriptive statistics will be calculated to compare patient characteristics for the intervention arm versus the control arm, such as medians and interquartile ranges for continuous variables and frequencies and proportions for categorical variables. These characteristics will be compared between the two groups using the Wilcoxon rank sum test for continuous variables and the chi squared or Fisher's exact test for categorical variables. We expect these characteristics will be similarly distributed between the intervention groups due to randomization, but in the case of imbalance, they may be adjusted in the multivariable model of the outcome.

The primary analyses will be Intent-to-Treat, with the primary outcome (total care time) compared using a two-sample t test, or Wilcoxon test if normal distribution assumption is violated. A multivariate linear regression will be used to adjust potential confounders including demographics (age, gender, race, SES), comorbidities (Elixhauser comorbidity index), treatment intent (curative versus palliative), and performance status. The final model will be established using stepwise forward regression based on 10% change in effect estimate for intervention. As a secondary analysis, "as treated" analyses will be performed to account for permissive non-adherence to the triage assignment. Standard descriptive statistics will be used to describe patient factors associated with more frequent triage non-adherence.

Secondary endpoints include: total wait time measured, total number of hospitalization/emergency department encounters, patient satisfaction [The Patient Satisfaction Questionnaire Short Form (PSQ-18)][1], provider satisfaction (Net-Promoter Score)[2,3], and health related quality of life [Functional Assessment of Cancer Therapy-General (FACT-G)[4]]. Total wait time and number of hospitalizations will be compared between groups using Wilcoxon rank sum test. The rest of the secondary endpoints will be measured at multiple time points and longitudinal linear mixed-effects model will be used to compare the group difference over time. Additional secondary analyses exploring changes in visit time across the duration of follow-up will also be performed.

1.3 Control of Bias and Confounding

Patients will be randomized 1:1 to control or intervention arm of the study. Despite the benefits of randomization to minimize the risks of confounding we will additionally perform a multivariate linear regression to adjust potential confounders including demographics as describe in section 6.2.

1.3.1 Baseline Data

Standard descriptive statistics will be calculated to compare patient characteristics for the intervention arm versus the control arm, such as medians and interquartile ranges for continuous variables and frequencies and proportions for categorical variables. These characteristics will be compared between the two groups using the Wilcoxon rank sum test for continuous variables and the chi squared or Fisher's exact test for categorical variables. We expect these characteristics will be similarly distributed between the intervention groups due to randomization, but in the case of imbalance, they may be adjusted in the multivariable model of the outcome.

1.3.2 Analysis of Primary Outcome of Interest

The primary analyses will be Intent-to-Treat, with the primary outcome (total care time measured over three months) compared using a two-sample t test, or Wilcoxon test if normal distribution assumption is violated (see section 6.2).