

**Official Study Title: Video Images About Decisions for Ethical Outcomes in
the Emergency Department (VIDEO-ED)**

NCT Number: NCT04931797

Statistical Analysis Plan

Document Date: 06/25/2025

Statistical analysis plan

Preliminary analyses will compare characteristics of study participants and patients who were eligible but declined consent in order to understand the representativeness of the study sample. Baseline variables (sociodemographic characteristics and care preferences) will then be tabulated by study arm. The primary outcome, ACP documentation within 3 months of the index ED visit, is ascertained using NLP of the EHR with human review to confirm presence of documentation of ACP; we anticipate ascertainment of this outcome for all participants. Secondary outcomes will be based on patient interviews and EHR, including NLP processed information on ACP documentation within 6 months of the index ED visit. For these outcomes, we will calculate the loss to follow-up rate for each arm and assess for non-response bias by comparing baseline characteristics of patients who did and did not complete follow-up.

Aim 1: *For patients receiving care in the ED, to evaluate the effect of an intervention combining a certified video decision aid, a research clinician conversation, and communication with the PCP and admitting physician.*

The primary outcome will be compared for all patients who undergo randomization using an intention-to-treat analysis regardless of whether patients receive the intended intervention. In a cross-sectional analysis, we will compare the primary outcome (proportion of patients with ACP documentation within 3 months of the ED visit) between intervention and control patients using a chi-square test. The percentage of patients reaching the primary outcome will be compared between intervention and control patients with adjustment for age, gender, and race using a Mantel-Haenszel chi-squared test at the 5% significance level. We will also estimate an adjusted absolute difference with a 95% confidence interval. We will use a Mantel-Haenszel chi-squared test to test for heterogeneity of treatment effect between the three sites. In addition, we will use logistic regression analysis to increase the precision of the estimate of treatment effect by including known predictors of outcomes (sociodemographic characteristics and ED assessment measures) in the models.

This outcome will be assessed at 3 months (primary outcome) and again at 6 months (secondary outcome). Secondary logistic regression analysis will estimate adjusted odds ratios of treatment, sociodemographic characteristics and ED assessment measures including baseline ACP preferences with the primary outcome.

Aim 2: *To compare secondary outcomes assessed at 3 and 6 months: patient-reported ACP conversations, concordance between stated and documented preferences, and hospital utilization.*

We will conduct repeated measures analyses to include data on outcomes collected from both follow-up time points (3 and 6 months) to estimate the treatment effect over time using the Generalized Estimating Equations (GEE) approach.

1. The percentage of patients reporting a conversation with a) a trusted loved one and b) a primary provider within 3 months (and within 6 months in separate analyses) will be compared between treatments using Mantel-Haenszel chi-squared tests at the 5% level. Comparisons will be adjusted for the same covariates as in Aim 1 and covariate-adjusted marginal estimates of absolute differences in proportions between treatment arms will be produced with 95% confidence intervals using nonparametric covariance analysis methods as in Aim 1. Logistic regression will estimate conditional odds ratios for association of variables and reported conversations.

2. Among patients with documented preferences for medical care in the EHR, concordance between the patients' documented preferences and stated preference during phone interviews will be examined at 3 and

6 months. Using responses to questions regarding the 2 specific items of care (CPR and ventilatory support), we will calculate a percentage agreement by comparing stated and documented preferences. Average concordance among patients in each arm will then be compared using a Mantel-Haenszel chi-squared test at the 5% significance level. For the 3 general categories of preferred intensity of medical care (life-prolonging care, limited medical care, and comfort care), data available in the EHR for each patient will be used to determine agreement with the patient's stated preference. These comparisons will also be adjusted for covariates listed for Aim 1.

3. Hospital utilization (i.e., the number of hospital days) will be assessed at 3 and 6 months separately using marginalized zero-inflated negative binomial (MZINB) regression with the logarithm of follow-up days as an offset to compare incidence density rates of hospitalization between treatment arms, adjusting for age, gender and race in addition to sociodemographic characteristics and ED assessment measures that will be selected into the model using stepwise selection with significance defined at the 0.05 level. Unlike standard zero-inflated count regression models that model the mean of a latent class of "at-risk" individuals, the marginalized counterparts model the overall mean (as in Poisson and negative binomial regression) in addition to the probability of being an "excess zero." Sensitivity analyses will evaluate the goodness-of-fit of the MZINB models by also fitting marginalized zero-inflated generalized Poisson regression for comparability based on likelihood statistics.

Aim 3: *To explore the potential for differential intervention effects across important subgroups.*

We will assess heterogeneity of intervention effects on the primary outcome across racial/ethnic subgroups (White non-Hispanic, White Hispanic, Black and other). We will estimate the intervention effect for each subgroup and test the intervention-by-subgroup interaction in logistic regression models.

To assess for differential intervention effects on the primary outcome across racial/ethnic subgroups, we will use logistic regression models that contain terms for the intervention, the subgroup, and the intervention-by-subgroup interaction, controlling for age and gender. If the interaction is significant at the 5% level, we will estimate subgroup-specific treatment effects. Blacks and Hispanic Whites will each separately be compared to non-Hispanic Whites as the reference group. This approach reflects our understanding of the demographics of the study population (i.e., few Hispanic Blacks) and the interest in examining Blacks, Hispanic Whites, and non-Hispanic Whites as three distinct groups rather than, for example, comparing Blacks to Whites with both Hispanic Whites and non-Hispanics Whites grouped together.

Secondary Analyses: Exploratory outcomes will be compared by study arm using logistic regression for dichotomous variables and multivariable linear regression for continuous variables. Differential effects of the intervention will be explored by gender and for patients enrolled vs. not enrolled in a managed care plan using similar interaction tests as described for Aim 3.

Sex as a Biological Variable: We will examine, and report data disaggregated by gender and look for differential effects. We will focus on gender rather than sex, because available evidence suggests ACP differences are primarily influenced by social and cultural factors rather than biological factors.

Sample Size Calculation: There will be no loss to follow-up for our primary outcome because of our use of Natural Language Processing (NLP) to abstract the primary outcome. We estimated 10% of control patients will achieve ACP (the primary outcome) at 3 months from our pilot study. Assuming that the ED-based intervention will increase this rate to 18.2% (8.2% absolute increase), randomizing 300 patients to each group will provide 80% power to detect the 8.2% difference using a two-sided 5% significance level (Table). Table below shows other combinations of rates from control and intervention with 80% power based on total N=600. As shown in the Table, we will have 80% power for a 7% difference with the sample size if the rate in the control group is as low as 6%.

Table. Possible combinations of rates with 80% power based on a total sample size of 600.

Control	Intervention	Difference
10%	18.2%	8.2%
8%	15.6%	7.6%
6%	13.0%	7.0%