STATISTICAL ANALYSIS PLAN

Understanding Factors in Decision Making for Children With Medical Complexity

NCT number: NCT05913206

Last reviewed: August 29, 2025

PI: Jody Lin, MD, MS

General Considerations
This randomized factorial experiment examines the effectiveness of uncertainty communication statements and parent narratives for parents of children with medical complexity (CMC). The experiment was nested within a video vignette of a family's visit to a pediatric orthopaedic surgeon regarding potential surgery treatment of their child's neuromuscular scoliosis (NMS). The trial includes 16 possible scenarios based on 4 factors (ambiguity, complexity, normalizing and explanatory language, and narratives) with 2 levels each (present, absent). Participants underwent stratified randomization by language, site, ethnicity, and race.

Study Objectives and Hypotheses
The main objective of the study is to evaluate the effect of various uncertainty communication statements and narratives on decision readiness and additional cognitive, behavioral, and affective outcomes related to decision quality and risk perception.

We hypothesize that participants who receive uncertainty statements and/or narratives will report higher decision readiness without experiencing detrimental effects of increased worry or decreased hope. We hypothesized that these effects would be summative across factors.

Study Design
All participants were recruited from two sites. Eligible parents were at least 18 years old and had a child 0-21 years of age with medical complexity. Randomization was conducted using REDCap, a HIPAA compliant remote data capture system. Participants were not explicitly informed of their group assignment.

The specific trial arms are:

*Control condition***:** Participants in the control condition received a vignette reflective of a typical orthopaedic surgery clinic visit to discuss neuromuscular scoliosis treatment without the addition of any uncertainty communication statements or exposure to parent narratives.

Ambiguity only: This intervention group received a vignette with ambiguity statements embedded in the video.

Complexity only: This intervention group received a vignette with complexity statements embedded in the video.

Normalizing and explanatory language only: This intervention group received a vignette with normalizing and explanatory language about uncertainty embedded in the video.

Narrative only: This intervention group received the control condition vignette followed by a set of parent narratives to read and/or listen to.

Ambiguity and complexity only: This intervention group received a vignette with ambiguity and complexity statements embedded in the video.

Ambiguity and normalizing and explanatory language only: This intervention group received a vignette with ambiguity statements and normalizing and explanatory language about uncertainty embedded in the video.

Ambiguity and narratives only: This intervention group received a vignette with ambiguity statements embedded in the video followed by a set of parent narratives to read and/or listen to.

Ambiguity, normalizing and explanatory language and narratives only: This intervention group received a vignette with ambiguity statements and normalizing and explanatory language about uncertainty embedded in the video followed by a set of parent narratives to read and/or listen to.

Complexity and normalizing and explanatory language only: This intervention group received a vignette with complexity statements and normalizing and explanatory language about uncertainty embedded in the video.

Complexity and narratives only: This intervention group received a vignette with complexity statements embedded in the video followed by a set of parent narratives to read and/or listen to.

Complexity, normalizing and explanatory language, and narratives only: This intervention group received a vignette with complexity statements and normalizing and explanatory language about uncertainty embedded in the video followed by a set of parent narratives to read and/or listen to.

Normalizing and explanatory language and narratives only: This intervention group received a vignette with normalizing and explanatory language about uncertainty embedded in the video followed by a set of parent narratives to read and/or listen to.

Ambiguity, complexity, and narratives only: This intervention group received a vignette with ambiguity and complexity statements embedded in the video followed by a set of parent narratives to read and/or listen to.

Ambiguity, complexity, and normalizing and explanatory language only: This intervention group received a vignette with ambiguity and complexity statements and normalizing and explanatory language about uncertainty embedded in the video.

All: This intervention group received a vignette with ambiguity and complexity statements and normalizing and explanatory language about uncertainty embedded in the video followed by a set of parent narratives to read and/or listen to.

Primary Outcomes

All study measures were categorized into three conceptual domains: affective, cognitive, and behavioral. The primary outcome is decision readiness, measured by the validated PrepDM scale, which is comprised of 10 items each on a 5-point Likert scale from 1=not at all to 5=a great deal. The final score is calculated by summing all responses and transforming to a scale of 0-100 by subtracting 1 from the summed score and multiplying by 25.

The co-secondary outcomes are decisional conflict, intention for NMS treatment (i.e., treatment selected, degree of decision preference), and decision-related knowledge.

Additional exploratory outcomes will also be measured and analyzed (see Table 1).

Table 1. Study Outcomes, Descriptions, and Survey Measure Time Points

Measure	Description	Measure Timepoints Time 1
Primary Outcome		
Preparation for Decision Making[1]	A validated scale which will assess participants' perspectives of the decision aid's usefulness in preparing them to communicate with their clinicians and for Shared Decision Making. These questions are answered on a Likert scale ranging from 1=not at all to 5=a great deal.	X
Secondary Outcomes		
Decision Conflict[2]	Sixteen questions measuring: 1) perceptions of uncertainty in choosing options, 2) feelings of having adequate knowledge and clear values, and 3) effective decision making. All items use a 5-point Likert scale ranging from 0=strongly disagree to 4=strongly agree.	
Intention for treatment decision	Two items: 1) which treatment option is selected and 2) degree of preference for a particular treatment option	
Decision-related knowledge	5 items about decision-related knowledge	
Exploratory Outcomes		
Affective		
Trust[3]	Dugan Physician Trust scale: 5-items on a 5-point Likert scale (1-Strongly Disagree to 5-Strongly Agree). Summary composite outcome variable (range 5-25) with higher scores indicating more trust.	X

Worry	Single unique survey item about the degree of worry that the patient in the vignette will need surgery on a 5-point Likert scale (1- Not at all to 5-Extremely). Based on NCI HINTS.[4]	
Behavioral		
Intention for not delaying treatment	For those participants who selected surgery as a treatment preference, the urgency with which they would like to pursue surgery.	X
Cognitive		
Risk perception	2 unique survey items scored individually on a 5-point Likert scale (1-very low to 5-very high). Based on NCI HINTS.[4]	
Parents' Characteristics and Survey Feedback		
Demographics	Participants indicate their age, gender, race, ethnicity, language preference, relation to patient, rurality, employment status, marital status, education level, and household income. Participants also indicate their child's age, gender, race, and ethnicity.	X
Child medical history	Participants indicate their child's use of subspecialists, technology dependence, neurologic impairment, history of prematurity, history of neuromuscular scoliosis, and history of surgery for neuromuscular scoliosis.	X
Readability assessment of video	3-item Likert-scale assessment of video readability.	X
Intolerance of Uncertainty[5]	Using the IUS-12 validated survey measure to calculate a summary composite variable and associated subdomains of prospective anxiety and inhibitory anxiety. Additional two unique Likert-based items each scored individually, based on NCI HINTS.[4]	X
Health Literacy[6]	Chew Health Literacy Scale: 3-item measure on a Likert scale (0-never to 4-always) with summary composite outcome score. Higher score indicates lower health literacy.	X

Numeracy[7]	Subjective numeracy scale: 8-item measure on a Likert-based scale with a mean composite outcome. Q7 is reverse-coded.	X
Risk taking behavior scale[8]	DOSPERT: 8-item Likert-based scale from 1-very unlikely to 5-very likely. Mean composite outcome.	X
Hope[9]	Herth Hope Index: 12-items on a Likert scale (1-strongly disagree to 4-strongly agree). Summary composite outcome with reverse scaling for items hhi3, hhi6.	X
Preferences and values	6-item ranking question where participates are asked to rank particular values in order of most important to least important	X

Planned Data Analysis as of May 2023

Participants who complete the study will be analyzed in the group to which they were randomized, regardless of fidelity. A two-tailed alpha level of 0.05 will be applied to all statistical tests. For the outcomes, we will use standard psychometric analyses (e.g., Cronbach's alpha) will be used to confirm the internal consistency reliability of self-report measures. For the primary outcome variable PrepDM, we will use a 4-way ANOVA model to describe differences in decision preparedness across each factor (ambiguity, complexity, normalizing and explanatory language, and narratives). We will use post-fit marginal estimation to express effect sizes (adjusted means and adjusted mean differences), which incorporates mean centering to make the main effects interpretable in the presence of factor interactions. Regression models, with main effect and interaction terms, can be fitted that are identical to a 4-way ANOVA model. We will use these to model the secondary outcomes, using binary logistic, ordinal logistic, logistic regression for proportions, and linear regression, depending on the level of measurement of the outcome variable, and add covariates including race and education level to all models. We will select covariates using the 10% rule. If we have sufficient power, we will perform exploratory analyses for populations with and without NMS and by ethnicity and language.

Although multiple outcomes will be considered, we have designated a single primary outcome. We do not plan formal multiple comparison adjustments for secondary and exploratory outcomes. Results for secondary and exploratory outcomes will be interpreted based on the overall pattern of results with the awareness that some nominally significant relationships may be false positive findings in the context of multiple analyses.[10]

We also do not plan formal multiple comparison adjustments for the randomized comparison between intervention groups, as these comparisons address distinct hypotheses and are thus appropriate for evaluation on a comparison wise basis.[10]

Missing data

We will use careful data collection and data management to minimize missing data. If missing data are present in Aim 1 or Aim 3, we will explore patterns of missingness and then impute missing data. We will use the multivariate imputation via chained equations (MICE) approach on both the outcome and predictor variables to create 40 imputed data sets (Graham, et al. 2007)[148], after which coefficients and standard errors are combined using Rubin's rules (Rubin, 1987).[149]

Sample size and power calculation for the primary analysis

In a validation study of PrepDM in orthopedic decision making, the scale had mean±SD of 3.7±1.0 after the decision aid intervention. We will randomize n=91 subjects into each of the 2x2x2x2=16 cells of the 4-way factorial design (each factor at 2 levels), for a total N=728 subjects. This is a sufficient sample size to include 4 main effects, 4 two-way interactions, and 3 3-way interactions, without overfitting the data. This provides 80% power, using a two-sided alpha 0.05 comparison, to detect an adjusted mean difference of 3.7±1.0 vs. 3.4±1.0 (0.3 standardized mean difference) between the two levels of the factor. That is smaller than a 0.5-point difference on the scale of each item and has been achieved in previous studies of decision aids in orthopedics.

References

1. Graham ID, O'Connor AM. Preparation for decision making scale. Ottawa Hospital Research Institute. Updated 2010. Accessed 5/5/20, 2020. http://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_PrepDM.pdf
2. O'Connor AM. Validation of a Decisional Conflict Scale. *Medical Decision Making*. 1995/02/01 1995;15(1):25-30. doi:10.1177/0272989X9501500105
3. Dugan E, Trachtenberg F, Hall MA. Development of abbreviated measures to assess patient trust in a physician, a health insurer, and the medical profession. *BMC health services research*. Oct 3 2005;5:64. doi:10.1186/1472-6963-5-64
4. Institute NC. Health Information National Trends Survey (HINTS). Accessed August 5, 2020, 2020. https://hints.cancer.gov/default.aspx
5. Birrell J, Meares K, Wilkinson A, Freeston M. Toward a definition of intolerance of uncertainty: a review of factor analytical studies of the Intolerance of Uncertainty Scale. *Clin Psychol Rev*. Nov 2011;31(7):1198-208. doi:10.1016/j.cpr.2011.07.009
6. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Family medicine*. Sep 2004;36(8):588-94.
7. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Medical decision making : an international journal of the Society for Medical Decision Making*. Sep-Oct 2007;27(5):672-80. doi:10.1177/0272989x07304449
8. Weber EU, Blais A-R, Betz NE. A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*. 2002;15:263-290. doi:10.1002/bdm.414
9. Herth K. Abbreviated instrument to measure hope: development and psychometric evaluation. *Journal of advanced nursing*. 1992;17(10):1251-1259.
10. Bender R, Lange S. Adjusting for multiple testing--when and how? *J Clin Epidemiol*. Apr 2001;54(4):343-9. doi:10.1016/s0895-4356(00)00314-0