Project SUPPORT
Statistical Analysis Plan

Clinicaltrials.gov ref#: NCT02232074
Principal Investigator: Tracy A. Battaglia, MD MPH

Patient-Centered Outcomes

Revised 02/07/18

Aim 1-To measure the impact of patient navigation enhanced by MLP intervention on patient-reported outcomes: distress, needs, and satisfaction

Patient Reported Surveys are:
- Distress Thermometer (DT)
- Communication and Attitudinal Self Efficacy-Case Cancer (Case Cancer)
- Cancer Needs Distress Inventory (CaNDI)
- Patient Satisfaction with Interpersonal Relationship with Navigator (PSN-I)

Analysis
The proposed analysis for patient centered outcomes will be used to compare the intervention and control group with regards to changes in patient-reported distress and psychosocial needs at follow up. And also examine the intervention effects on questionnaire total scores.

Study population
The primary analysis will include only those study participants who have legal issues at baseline screening (defined by responding positively to one or more questions on the baseline I-HELP general legal screening questionnaire, which will take place within one month of new diagnosis). Patients who have withdrawn consent are excluded from the analysis.

1.1 DISTRESS THERMOMETER (DT)

Dependent Variable	Distress Thermometer
Independent Variable	intervention and control
Hypothesis/Outcome	Navigation that aims to reduce socio legal barriers to care and facilitate timely care may mitigate distress.

Description of the Measure
- Collected at Baseline, 3, 6 & 12 months
- A one-item measure of global distress; rate distress levels during past week - this one item instrument utilizes a figure and asks individuals to rate distress levels during the past week.
- Main measure: Could be treated as either a continuous or categorical measure
- Measured on a 10-point scale ranging from 0 (none) to 10 (extreme distress).
- It can be categorized as an ordinal categorical variable as: 0-4: no or low levels of distress, 5-6: moderate levels of distress, 7-10: distressed
- It can be dichotomized as 0-4 as not distressed and 5-10 as distressed (will need to justify choice of dichotomization before viewing the data
- We will stratify all of the proposed analysis below by cancer site (lung/breast).

Analysis of Distress Thermometer

- The analysis will be used to compare intervention and control groups on changes in distress at follow up at 3, 6 and 12 months.
- Summary statistics: (the following points could be separated into a separate section, since it is the first part of the analysis). Using the 10 point scale for the DT: Univariate summary statistics: mean and standard deviation, median and inter-quartile range (IQR) overall and for each of the intervention and control groups.
- Examine histograms of the distress thermometer values at each time point for each study group for normality. If non-normal (skewed and/or with outliers), it is better to report median and inter-quartile range (IQR) than the mean and standard deviation, or both.
- If DT scores are dichotomized or categorized, we can calculate the number and percent in each category. We can also compare between the intervention and control groups using a chi-squared test.

Analysis (a)	Analysis (b)	Analysis (c)
Simple linear regression to examine association of intervention with the change in Distress from baseline to primary endpoint at 6 months	A multivariate linear regression	Longitudinal modeling of Distress Thermometer (DT outcome measured at 0, 3, 6 and 12 months).

Models and covariates of interest to be use as a reference for all PCO

Model 1:
A simple linear regression model to test whether there is an association between the intervention group (covariate - 0=control, 1= intervention) with the distress thermometer (outcome) at the primary endpoint (six months)
We adjust for baseline DT (at Time 0) Repeat using DT at three and 12 months as a secondary outcomes

Model 2:
A multivariate regression model to test whether there is an association between the intervention group (covariate - 0=control, 1= intervention) with the distress thermometer (outcome) at the primary endpoint (six months)
In this analysis we will adjust for the effects of the same set of covariates as for the clinical outcomes. See Table ** below

Model 3:
Linear mixed model with each patient having a random effect value. All data is included at all-time points, even those patients missing some data at one or more time points. Intervention group is the key covariate, baseline Distress Thermometer as a covariate. Time points included as separate fixed covariates.

Model 4:
As for model (3) and also including the same covariates as for the clinical outcome. A plot showing the mean Distress Thermometer at each time point for each intervention group.

Covariates	Categories	Reference Group
Intervention group	Intervention/Control	Control
Age-group	Continuous	NONE
Insurance Status:	Public/Private	Private
Race	White/Hispanic/Black/Other (where Other combines Asian and Other).	White
Cancer stage	0,I, II, III, IV.	Stage 0
Health literacy	Adequate/Not adequate (which combines inadequate and marginal literacy)	Adequate

Note: language is a collinear variable with race and will not be used as a covariate. Can be included in a stratified analysis to see if there is a difference since language can affect literacy.

1.2 CASE Cancer Communication and Attitudinal Self Efficacy

Dependent Variable	CASE-Cancer questionnaire score
Independent Variable	Treatment (intervention and control) group
Hypothesis/Outcome	Navigational intervention that targets socio-legal needs will reduce perceived needs overtime.

The Communication and Attitudinal Self-Efficacy (CASECANCER) instrument measured each participant's perception of his or her ability to fulfill three treatment related roles: understanding and participating in care, maintaining a positive attitude, and seeking and obtaining information.

Description of the Measure

- Measure of patient's confidence in understanding and participating, maintaining a positive attitude and seeking and obtaining information surrounding their cancer care
- Collected at Baseline, 6 & 12 months
- 12 items (each measured on a scale 1 to 4). For each item a score of 1 means a low score while a score of 4 is a high score
- A total score of 48 is the maximum score representing best positive attitude and strong self-efficacy, while the minimum score is 12 representing poor self-efficacy
- 3 Sub domains (For each subdomain: minimum score = 4, maximum score = 16):
  - Domain A: understanding and participating in care (includes items 1-4)
  - Domain B: maintaining a positive attitude (includes items 5-8)
  - Domain C: seeking and obtaining information (includes items 9-12)
- Add responses to create score for each subscale; higher score=higher self-efficacy
- We will conduct analyses for the total score and for each of the three subdomains.

The CASE-cancer scale

Understand and participate in care
1 I know that I will be able to deal with any unexpected health problems.
2 I am confident in my ability to understand cancer materials.
3 I am confident in my ability to understand my doctor's instructions.
4 It is easy for me to actively participate in decisions about my treatment.

Maintain a positive attitude
5 I won't let cancer get me down.
6 It is easy for me to keep a positive attitude.
7 It is easy for me to maintain a sense of humor.
8 I am confident that I can control my negative feelings about cancer.

Seek and obtain information
9 If I don't understand something, it is easy for me to ask for help.
10 It is easy for me to ask nurses questions.
11 It is easy for me to ask my doctor questions.
12 It is easy for me to get information about cancer.

See Table One above for Regression models and plot

Analysis (a)	**Analysis (b)**	**Analysis (c)**
Simple linear regression to examine association of intervention with the change in CASE-cancer Total score to primary endpoint at 6 months	A multivariate linear regression	Longitudinal modeling of CASE-cancer Total score outcome measured at baseline, 6 and 12 months.

1.3 CANDI Cancer Needs Distress Inventory- CaNDi

Dependent Variable	Cancer Needs Distress Inventory
Independent Variable	Treatment (intervention and control) group
Hypothesis/Outcome	Navigational intervention that targets socio-legal needs will reduce distress

Description of the Measure

- Collected at Baseline, 6 & 12 months
- CaNDI is a self-reported cancer-needs assessment instrument
- Validated 39 items survey: One question (#34) removed, thus a total of 38 questions will be included.
- Each item measured on a scale from 1 (not a problem) to 5 (very severe problem)
- Items are summed to create a total distress score - with a minimum score of 38 and a maximum score of 190
- Patients rate the extent of their concern in the past two weeks
- Two sub-scales:
  - Anxiety (questions 30,31 with a minimum score of 2 and maximum score of 10)
  - Depression sub-scales (questions 29, 32, 33, 36 with a minimum score of 4 and maximum score of 20)
- Questions range across the 2 domains (healthcare, and practical).

  HEALTHCARE= item # 5, 11, 13, 15
  PRACTICAL= item # 1, 2, 3, 4, 9, 14

- Item scores were summed to create an average total distress score. Items with responses of 'Prefer not to answer' and 'Do not know' were coded as missing.
- Total scores were calculated as a mean of the sum of completed items. Subscales for depression (four items) and anxiety (two items) were computed in the same manner
- Higher score = Higher distress

See Table One above for Regression models and plot

Analysis (a)	Analysis (b)	Analysis (c)
Simple linear regression to examine association of intervention with the change in CANDI Total score to primary endpoint at 6 months	A multivariate linear regression	Longitudinal modeling of CANDI Total score outcome measured at 0, 6 and 12 months).

1.4 Patient Satisfaction with Interpersonal Relationship with Navigator (PSN-I)

Dependent Variable	Patient Satisfaction with Interpersonal Relation with Navigator score at 6 and 12 months
Independent Variable	Treatment (intervention and control) group
Hypothesis/Outcome	Navigational intervention that targets socio-legal needs will reduce perceived needs overtime.

The Patient Satisfaction Survey, PSN-1

A validated tool that indicates the extent to which patients healthcare experience match their expectations which has been linked to health status, quality of life and adherence to recommended treatment.

Description of the Measure

- 9-item questions
- Response is summed to create total score (min score =12, max score=45)
- Higher score = higher satisfaction

Analysis (a)	Analysis (b)
A simple linear regression model	A multivariate linear regression

Clinical Outcomes

Revised 01/28/18

Aim 2: To measure the impact of patient navigation enhanced by MLP intervention on clinically relevant outcomes: initiation of timely and quality cancer care

1.1 MAIN CLINICAL OUTCOME

Receipt of timely treatment within 90 days of date of diagnosis, (Dichotomous yes=1/no=0)

Hypothesis
Compared to study participants receiving standard navigation, study participants receiving enhanced navigation will be more likely to initiate their cancer treatment in a timely manner.

Study population
The primary analysis will include only those study participants who have legal issues at baseline screening (defined by responding positively to one or more questions on the baseline I-HELP general legal screening questionnaire, which will take place within one month of new diagnosis). Patients who have withdrawn consent are excluded from the analysis

Primary analysis
Assess the impact of the navigation intervention on the receipt of their primary treatment. The clinically relevant outcome will be based on the metric of timely initiation of the primary treatment, a dichotomous outcome (1=yes,0=no). The clinical outcome will be derived from the participant's medical record.

Model 1.1a
A univariate logistic regression to examine the association between intervention and control groups (enhanced legal patient navigation versus standard patient navigation). The binary outcome measured on each patient is whether treatment was initiated within 90 days after patient's date of tissue diagnosis (1=yes,0=no).
Model 1.1b
1. We will conduct a multivariate Cox regression model, which follows a similar approach to Model 1.1a, where we adjust for clinically relevant covariates. (see below)

Table 1.1.1 LIST OF COVARIATES FOR THE MULTIVARIATE MODEL WITH REFERENCE GROUPS

Covariates	Categories	Reference Group
Intervention group	Intervention/Control	Control
Age-group	Continuous - 10 Year increments	NONE
Insurance Status:	Public*/Private *Including Public Medicare, Medicaid and other.	Private
Race	White/Hispanic/Black/Other (where Other combines Asian and Other).	White
Cancer stage	0,I, II, III, IV & missing stage	Stage 0
Health literacy	Adequate/Not adequate (which combines inadequate and marginal literacy)	Adequate

***Note**: Will not include gender, marital status or primary language.*

Cancer Site: this analysis will be done on breast. But we will conduct an exploratory analysis on the lung data.

1.2 SECONDARY CLINICAL OUTCOME

Time to initiation of first treatment (Continuous, number of days). We will conduct a survival analysis to assess the impact of the navigation intervention on the time to first treatment.

Hypothesis
Compared to study participants receiving standard navigation, study participants receiving enhanced navigation will be more likely to initiate their cancer treatment in a timely manner.

Study population
As defined above.

Primary Analysis
We consider the start date to be the date of tissue diagnosis (clinical diagnosis). The date that a patient experiences their first treatment is their "event" date (event = 1). This can take any value from 0 days to 365 days (study period). Once a patient experiences their first treatment we no longer need further data on them for this particular analysis.

Model 1.2a
1. A univariate Cox regression model, which uses the time from start date until the time of event or censoring as the outcome for each patient. 2. The only covariate of interest will be intervention group (control or intervention). 3. We will get a hazards ratio to describe the increased (or decreased) "hazard" of experiencing timely treatment in the next short time interval for the intervention group (compared to the control or reference group). In our setting, a hazard ratio greater than 1 is a good result. 4. We will also draw Kaplan-Meier curves and conduct a log-rank test (with p-value) to show and compare the differences between the control and intervention groups in their times to first treatment over the course of the study. These approaches take into consideration the censoring that some patients will experience (see Table 1.1 below for details)

Model 1.2b
2. We will conduct a multivariate Cox regression model, which follows a similar approach to Model 1, where we adjust for clinically relevant covariates. (see below)

Table 1.2.1 LIST OF COVARIATES FOR THE MULTIVARIATE MODEL WITH REFERENCE GROUPS

Covariates	Categories	Reference Group
Intervention group	Intervention/Control	Control
Age-group	Continuous - 10 Year increments	NONE
Insurance Status:	Public*/Private *Including Public Medicare, Medicaid and other.	Private
Race	White/Hispanic/Black/Other (where Other combines Asian and Other).	White
Cancer stage	0,I, II, III, IV & missing stage	Stage 0
Health literacy	Adequate/Not adequate (which combines inadequate and marginal literacy)	Adequate

***Note**: Will not include gender, marital status or primary language*
*We will conduct an exploratory analysis on the lung data.

Missing Data and Survival Analysis Guidelines

Missing Data

The following guidelines are for the clinically relevant binary outcome for the timely initiation of the primary treatment, within 90 days from the date of tissue diagnosis. Ensuring the correct code is used in all the likely scenarios listed below. The clinical binary outcome will be derived from the participant's medical record.

For primary binary outcome:
Timely=Yes (1)
Not timely = No (0)
Unable to verify treatment= Missing (m)
Table 1.0

Scenarios of clinically verified treatment If a Person:	**Timely Outcome Code**
Died	
Received treatment within 90 days	1
No timely treatment within 90 days	0
during the first 90 days on study and receive no treatment	m
Transferred Care	
If it is clinically verified from either BMC or at their transfer institution that they received treatment within 90 days	1
If it is clinically verified that they did not initiate care within 90 days	0
If unable to confirm clinical verification of treatment receipt within 90 days	m
Withdrew (Among those who withdrew from the EMR review portion of the study)	
Anyone who does not give EMR review consent	m
Lost to Follow-up (LTFU)	
If it is clinically verified that they received treatment within 90 days	1
If it is clinically verified that they did not initiate care within 90 days	0
If unable to clinically verify receipt of treatment within 90 days	m

* Note: Participants have the option to withdraw from 3 different parts of the project. (Survey completion, patient navigation services and access to medical records.) If they withdraw from everything except medical records review the aforementioned would apply.

^ This could happen if patient transferred and we are unable to obtain patients permission or medical records from another institution.

Missing Data and Survival Analysis Guidelines

Survival Analysis

The following guidelines are for the clinically relevant continuous outcome for the timely initiation of the primary treatment from 0 days to 365 days (study period).

The following scenarios are applicable to participants for whom to verification of treatment initiation was not possible or "event date", requiring a censor date for the analysis.

Table 1.1

If a person	Censor Date*
Died	
Without initiating treatment (clinically verified)	The date of death or 365 days, whichever is earlier
Transferred Care	
If unable to clinically verify receipt of treatment at BMC or outside institution	The date of transfer
If clinically verified NO treatment was received at BMC or outside institution	365 days, whichever is earlier
Withdrew*	
Data dropped from analysis	No censor date
Lost to Follow-up (LTFU)	
If unable to clinically verify receipt of treatment prior to LTFU	Last known clinical contact (any documented clinical personnel visit or call) or 365 days, whichever is earlier

*The censor date will be adjusted to 183 days for PCORI analysis and 365 for ACS.

*Anyone who withdraws from the study and emr abstraction their clinical data will not be included in the analysis.

Note:

All study participants who we are unable to clinically verify initiation of treatment by the end of a) 183 or b) 365 days, they will have a censored date of 1 year. Otherwise (if less than 1 year), their censoring date is the date of last known clinical contact

The date that a patient experiences their first treatment is their "event" date (event = 1). This can take any value from 0 days to 365 days (study period).