THE MEN'S EATING AND LIVING (MEAL) STUDY: A RANDOMIZED TRIAL OF DIET TO ALTER DISEASE PROGRESSION IN PROSTATE CANCER PATIENTS ON ACTIVE SURVEILLANCE

Description of the Protocol

This study 70807 opened on 01 December 2010. Since study activation, there have been 9 updates to the original protocol.

Amendments and Statistical Impact

The primary endpoint in this prevention trial is time to clinical progression; each subject will be followed for two years.

The Alliance Data and Safety Monitoring Board (DSMB) have been reviewing interim analyses of the time to progression endpoint every 6 months. Following the first interim analysis (May 2014), which was conducted when 80 subjects experienced clinical progression or completed their 2 years of follow-up, it was noted that the clinical progression rate in these active surveillance patients was higher than expected. At the time of that first interim analysis, the protocol specified that serum prostate-specific antigen (PSA) doubling time was to be calculated using the 3 most recent PSA values at each 3-month assessment, starting with study month 6. This method was found to be overly sensitive to local fluctuations of PSA values during a short period of time. Consequently, the study was updated (**Amendment 9**) to calculate PSA doubling time as in the Toronto cohort [1], i.e. using all available PSA values starting from the study month at which the subject has 3 PSA values.

At the time of study activation, 12 and 24 study month prostate biopsies were required. Based on **Amendment 6**, the prostate biopsy at 12 months was no longer required. Because pathology information (e.g. Gleason scores) obtained from the prostate biopsies are used, in part, for determining clinical progression, this study update is noted here.

No other study updates impacted the statistical considerations of the protocol.

Description of the Clinical Trial

This is a randomized, phase III clinical trial designed to determine if a telephone-based dietary intervention compared with no intervention will decrease clinical progression in active surveillance (AS) prostate cancer patients. Patients must have a biopsy-proven adenocarcinoma of the prostate diagnosed within 24 months prior to pre-registration; the biopsy showing diagnosis of prostate cancer is used for the purposes of determining eligibility. Furthermore, men ≤ 70 years old and men > 70 years old must have a biopsy Gleason score ≤ 6 and ≤ (**3 + 4**) = **7**, respectively, and a baseline serum PSA < 10 ng/mL to be eligible for the study. A total of approximately 464 patients will be randomized to this study.

Eligible patients will be randomized with equal probability to receive the dietary intervention (experimental arm) or dietary information (control arm); randomization will be stratified on the basis of age (≤ 70; > 70 years), race (African American; Other), and baseline prostate biopsy (0-12; 13-24 months prior to registration).

After randomization, all patients will participate in a 5-10 minute telephone orientation conducted by study staff from the Moores UCSD Cancer Center explaining the randomization results and the next study-related events. The orientation call for the experimental arm participants will briefly explain the counseling program, the dietary targets, and the scientific rationale supporting these targets. The UCSD study staff will mail all participants (experimental and control arms) a copy of the Prostate Cancer Foundation Booklet entitled "Nutrition, Exercise and Prostate Cancer." Additionally, the experimental arm participants will also be mailed a copy of the study-specific "Lifestyle Intervention Manual"; the manual outlines the dietary targets, offers supporting information on strategies to achieve these targets, supplies reference tools to help participants accurately estimate servings of target foods, and offers recipes and articles about diet and prostate cancer. Furthermore, each participant on the experimental arm will be assigned a personal counselor.

Each patient will be followed for 24 months, and serum PSA will be evaluated every 3 months starting from baseline (i.e. prior to randomization). Prostate biopsies are taken at baseline, 12, and 24 months.

Power and Sample Size

Clinical criteria for progression are any of the following occurring:

a) PSA doubling time < 3 years

b) PSA ≥ 10 ng/mL at any time

c) ⩾ 25% of biopsy tissue cores positive for cancer or > 50% of any one biopsy tissue core positive for cancer

d) For men < 70 years old, Gleason sum on repeat biopsy ⩾ 7

   For men ≥ 70 years old, Gleason sum on repeat biopsy ≥ (**4 + 3**) = **7**

PSA doubling time (in years) will be calculated as the natural logarithm of 2 divided by the estimated slope obtained from fitting a linear regression of the natural logarithm of PSA on time (in years) [3]; the PSA doubling time will be calculated at the Alliance Statistics and Data Center. The first 3 PSA measurements will be used at the 6 month assessment (i.e. baseline, and at months 3 and 6); from the 9 month assessment onwards, all available PSA measurements will be used to calculate PSA doubling time, as long as the participant has at least 3 measurements.

Centralized pathology review will be conducted on tissue specimens collected at diagnosis and at study months 12 and 24 to determine changes in tumor volume and Gleason scores; additionally, centralized pathology review will be conducted on tissue specimens based on

repeat biopsies (i.e. additional biopsies) that occurred at the discretion of the treating physician during the 24-month study period. The centrally reviewed changes in tumor volume and Gleason scores will be used in lieu of the local changes in tumor volume and Gleason scores, which are captured on the case report forms, in the definition of clinical progression. Stated differently, the centrally reviewed pathology will be incorporated in all relevant analyses described in the statistical analysis plan; however, in the event that centrally reviewed pathology is unavailable for a particular patient (and time point), the corresponding local pathology information captured on the case report form (if available) will be used. Dr. Donna Hansel, Pathology Chair, will review the prostate biopsy data, prepare the salient results in an Excel spreadsheet, and the results will be sent as an external data transfer courtesy of Linda McCart at the Alliance biorepository at Ohio State. **Post-note:** Because of the challenges with resolving the myriad follow-up data discrepancies in the central pathology spreadsheet, the study team decided not to incorporate these data in the primary analysis (see the memo-to-file in Appendix A); rather the statistical report generated from the analysis plan detailed in this document used the follow-up pathology data collected on the follow-up forms using the case report forms; these data had been monitored, cleaned, and validated.

This study was designed to demonstrate superiority of the experimental arm compared with the control arm in time to progression (TTP). TTP is defined as the length of time from the date of random assignment to progression, as defined above; participants who die from any cause without experiencing disease progression will be censored at the time of death. Additionally, participants who elect to pursue treatment during the study despite not meeting the criteria for progression will be censored at the time of withdrawal for treatment.

Data from the Toronto cohort [1] indicated that approximately 80% of AS patients will not experience clinical progression at 24 months. We hypothesize that 90% of AS patients in the experimental arm of this study will not experience clinical progression at 24 months. Using a two-sided 0.05-level log-rank test, a sample size of 418 participants (209 per arm) would provide at least 80% power to reject the null hypothesis of no treatment difference in TTP at 24 months. Fifty-seven events are the total number events that must be observed in the two arms combined to achieve the specified power for the test comparing survival in the two arms. Freedman's formula for the required number of events was used [2]. It is recognized that some participants will elect to pursue treatment during the study despite not meeting the criteria for progression. These participants will be censored at the time of withdrawal for treatment. Assuming a 10% dropout rate, including those patients who elect to receive treatment before progression, the targeted enrollment is 464 patients.

Objectives

The primary objective of this study is to determine if a telephone-based dietary intervention (experimental arm) compared to no intervention (control arm) will decrease clinical progression in AS patients.

There are 3 secondary objectives:

1. To compare the incidence of active treatment (surgery, irradiation, local ablation, or androgen deprivation) between the two arms.
2. To compare prostate cancer-related anxiety between the two arms based on the Memorial Anxiety Scale for Prostate Cancer (MAX-PC).
3. To compare health-related quality of life between the two arms based on the Expanded Prostate Cancer Composite Index 26 (EPIC-26).

The correlative science objectives pertaining to plasma carotenoid levels will be addressed in a separate statistical analysis plan and, furthermore, selected statistical analyses will be performed at the University of California San Diego (UCSD). Specifically, the group at Moores UCSD Cancer Center will analyze the carotenoid concentrations to assess if diet intervention changes were achieved within the study and whether any changes differed between the two study arms; these analyses support protocol correlative study objective 10.1.2.1. The statistical analyses to support the remaining correlative study objectives will be performed at the Alliance Statistical and Data Center.

Lastly, UCSD will also analyze the information obtained from the dietary recall assessments collected interactively via telephone interview at baseline and at 12 and 24 months post baseline and will not be reflected in this statistical analysis plan.

Analysis Populations

The following analysis populations will be used:

1. Intent-to-Treat

   The intent-to-treat population will include all randomized subjects.

2. Modified Intent-to-Treat

   The modified intent-to-treat population will include all randomized subjects, however, subjects who later become ineligible by centralized pathology review of their baseline tissue specimens will be excluded; we anticipate that no more than 10% of subjects will become ineligible for the study following central pathology review [4].

Demographic and Baseline Characteristics

The summaries of demographic and baseline characteristics will be tabulated for the intent-to-treat and the modified intent-to-treat populations by study arm and overall. For categorical data, frequencies and percentages will be provided and, for continuous data, descriptive statistics, including sample size (n), mean, median, standard deviation, and range of values (i.e. minimum and maximum values) will be provided. No inferential statistics will be presented.

The following demographics and baseline characteristics will be summarized in these two presentations: **age** (years), **age group** (≤ 70; > 70), **race/ethnicity** (Non-Hispanic White; Black or African-American; Hispanic or Latino; Asian; Native Hawaiian or Pacific Islander; American-Indian or Alaska Native; Not Reported; Unknown), **race** (African American; Other), **region**

(Midwest; North East; South; West), **body mass index**, **time since diagnostic prostate biopsy** (0-12 months; 13-24 months prior to registration), **baseline clinical stage** (T1c; T2a), **baseline PSA ng/mL categories** (0 - 2.5; > 2.5 - 5; >5 but less than 10) and **baseline Gleason sum** (6 and 7).

Also, a separate tabular presentation of the number and percentage of subjects within each baseline clinical stage by age (≤ 70; > 70) and study arm will be generated, as well as a tabular presentation of the baseline PSA categories by age (≤ 70; > 70) and study arm. Additionally, a cross tabular presentation of baseline Gleason score with baseline PSA categories by study arm and age (≤ 70; > 70) will be generated. No inferential statistics will be presented.

In a separate tabular presentation, patient responses to the 8 questions on the personal habits questionnaire administered at baseline only will be summarized by study arm. The questionnaire addresses a number of generic health behavior questions (e.g. cigarette smoking, physical activity). No inferential statistics will be presented.

Reasons for Discontinuation

The reasons for discontinuing the study protocol will be summarized in a table overall and by study arm.

Primary Efficacy Analysis

The primary efficacy analysis will be based on the modified intent-to-treat analysis population.

TTP percentages, standard errors, and intervention effect will be obtained from the Kaplan-Meier method, Greenwood's formula, and log-rank test, respectively.

Supportive Analysis

1. Based on the modified intent-to-treat analysis population, a Cox proportional hazards regression model will be used to estimate relative risk and 95% confidence interval for the intervention comparison, adjusting for the following covariates: age group (≤ 70; > 70), race (African American; Other), and time since diagnostic prostate biopsy (0-12 months; 13-24 months prior to registration).

2. We will repeat the univariate analysis based on the Kaplan-Meier method using the intent-to-treat analysis population.

Additional Analysis

1. We will repeat the primary efficacy analysis based on the Kaplan-Meier method where clinical progression only considers Gleason score in its definition (i.e. for men < 70 years old, Gleason sum on repeat biopsy ≥ 7; for men ≥ 70 years old, Gleason sum on repeat biopsy ≥ (**4 + 3**) = **7**); in other words, the definition of clinical progression in this analysis will ignore serial PSA values and changes in tumor volume.

2. Similarly, we will repeat the supportive Cox proportional hazards regression analysis above where clinical progression only considers Gleason score in its definition.

3. Based on the modified intent-to-treat analysis population, we will analyze progression-free survival, defined as time to clinical progression or death, whichever occurs first, based on the Kaplan-Meier method where clinical progression is defined as in the primary efficacy analysis.

Secondary Efficacy Analysis

The secondary analysis will be based on the modified intent-to-treat analysis population.

It is recognized that some participants will elect to pursue treatment during the study despite not meeting the criteria for progression. Time to treatment percentages, standard errors, and intervention effect will be obtained from the Kaplan-Meier method, Greenwood's formula, and log-rank test, respectively. For this analysis, subjects who do not withdrawal from the study to pursue treatment will be censored at the time of clinical progression, death, or their last follow-up visit, whichever occurs first. Additionally, we will report the number and percentage of subjects who withdrew from the study to pursue treatment within the two study arms; a Fisher's exact test of independence will be conducted to assess whether the proportions are different between the study arms.

Exploratory Analysis

Weight (kg) and height (cm) are measured at baseline and at study months 6, 12, 18, and 24. Weight and body mass index (weight (in kilograms) divided by height (in meters) squared) will be summarized longitudinally for each study arm within the modified intent-to-treat population; descriptive summary statistics will include sample size (n), mean, median, standard deviation, and range of values (i.e. minimum and maximum values). Additionally, weight and body mass index measured serially will be plotted. No inferential statistics will be presented.

Although data of whether or not MRI-guided prostate biopsy was performed was not collected as part of the protocol, these data will be obtained externally. Because the MRI-guidance technology was not widely available when the study was activated in 2010, we expect that very few biopsies would have been performed in this manner. Interest is in the frequency of use of MRI-guided prostate biopsies in the two arms and overall. Therefore, the number and proportion of MRI-guided prostate biopsies performed at baseline and at study months 12 and 24 will be presented in a table within each study arm and overall.

Within the intervention arm, patients were administered the Nutrition Self-Efficacy Scale, which assesses the degree to which individuals are confident that they can control their nutrition. Each of the 5 items rated on a 5-point Likert scale ranging from "very confident" to "not confident at all" will be summarized within the intervention arm at baseline and at study months 6, 12, 18, and 24; descriptive summary statistics will include sample size (n), mean, median, standard deviation, and range of values (i.e. minimum and maximum values). No inferential statistics will be presented.

Because quality of life or anxiety has not been formally evaluated in an AS population or in the setting of a randomized clinical trial among AS patients, a battery of scores will be assessed longitudinally in the current study. These analyses are largely descriptive and exploratory in nature; any inferential statistics calculated should be interpreted with care. Tabulated descriptive statistics will include n, mean (standard deviation), and median [min, max]. The median of the total summary scores obtained from each quality of life / anxiety instruments will be plotted serially over time (i.e. baseline and at study months 6, 12, 18, and 24). All available longitudinal data will be summarized within the modified intent-to-treat population; no data imputation will be performed in these exploratory analyses.

The 4 quality of life / anxiety instruments that will be summarized are the Memorial Anxiety Scale for Prostate Cancer (MAX-PC), the Expanded Prostate Cancer Index Composite 26 (EPIC-26), the Functional Assessment of Cancer Therapy Scale – Prostate (FACT-P), and the International Prostate Symptom Score (IPSS). Appendix B provides the scoring algorithms for the battery of instruments used in the study.

MAX-PC

The 3 subscale summary scores and the total summary score on the MAX-PC will be summarized at each protocol defined time point (i.e. at baseline and at study months 6, 12, 18, and 24); additionally, the corresponding within-subject change from baseline will be summarized at post-baseline study months. To assess evidence against the null hypothesis that the median scores are the same in the two study arms, two-sided p-values obtained from the Wilcoxon rank-sum test will be calculated and provided at each study time point.

EPIC-26

The summary scores obtained from the 5 health-related quality of life domains and the total summary score on the EPIC-26 will be summarized at each protocol defined time point (i.e. at baseline and at study months 6, 12, 18, and 24); additionally, the corresponding within-subject change from baseline will be summarized. To assess evidence against the null hypothesis that the median scores are the same in the two study arms, two-sided p-values obtained from the Wilcoxon rank-sum test will be calculated and provided at each study time point.

FACT-P

All subscale summary scores (for the core FACT-G and for the FACT-P) and the total summary score on the FACT-P quality of life questionnaire will be summarized at each protocol defined time point (i.e. at baseline and at study months 6, 12, 18, and 24); additionally, the corresponding within-subject change from baseline will be summarized. To assess evidence against the null hypothesis that the median scores are the same in the two study arms, two-sided p-values obtained from the Wilcoxon rank-sum test will be calculated and provided at each study time point.

IPSS

The IPSS total summary score and the quality-of-life-due-to-urinary-symptoms score will be summarized at each protocol defined time point (i.e. at baseline and at study months 6, 12, 18, and 24); additionally, the corresponding within-subject change from baseline will be summarized. To assess evidence against the null hypothesis that the median scores are the same in the two study arms, two-sided p-values obtained from the Wilcoxon rank-sum test will be calculated and provided at each study time point.

General Analysis Issues

Significance Level

The primary efficacy hypothesis test will be performed using a 5% overall significance level. For the secondary efficacy analyses, as well as any supportive or additional analyses, hypothesis tests will be performed individually at the 5% significance level and there will be neither adjustment for multiple tests nor adjustment for multiplicity of endpoints. All hypothesis tests will be performed with two-sided alternative hypotheses. Any two-way interaction effects assessed, e.g. in the multivariable Cox proportional hazards regression models, will be tested at a 15% significance level.

Missing Data

In the event that we do not have a subject's centrally reviewed Gleason score at a protocol defined time point, the corresponding local Gleason score recorded on the study case report form will be used.

For the exploratory analyses, no missing data will be imputed. All subjects will have the time to event endpoint for the primary and secondary efficacy analyses. In the event that a subject is missing a baseline covariate included in the multivariable Cox proportional hazards regression models, the missing-indicator method will be used to preserve the full analysis population used (e.g. modified intent-to treat analysis population) [5].

APPENDIX A

Memo-to-File

Protocol Number: 70807

Protocol Title: The Men's Eating and Living (MEAL) Study: A Randomized Trial of Diet to Alter Disease Progression in Prostate Cancer Patients on Active Surveillance

Study Chair: J. Kellogg Parsons, M.D., MHS

Study Statisticians: David Zahrieh and Heshan Liu

Statistical Programmer Analyst: Libby Storrick

Purpose: To document the study team's decision **not** to resolve the remaining data discrepancies in the central pathology external database; the external database is an Excel spreadsheet and has been saved in the project workspace. Further, the purpose is to document that the ineligible cases we identified immediately post-database freeze (04APR2018) - based on review of the central pathology spreadsheet - remained ineligible and that the handful of potential ineligible cases that the Alliance Statistics and Data Center (SDC) had identified while performing further quality control on the external spreadsheet were indeed still eligible. In other words, no changes to the primary analysis population described in the statistical summary (*CALGB 70807 Statistical Sumary20180504.docx*) were needed and thus no updates to the final analyses on the primary endpoint were needed.

Background

Centralized pathology review was conducted on tissue specimens collected at diagnosis and at study months 12 and 24 to determine changes in tumor volume and Gleason scores; additionally, centralized pathology review was conducted on tissue specimens based on repeat biopsies (i.e. additional biopsies) that occurred at the discretion of the treating physician during the 24-month study period. The centrally reviewed changes in tumor volume and Gleason scores was to be used in lieu of the local changes in tumor volume and Gleason scores, which were captured on the case report forms, in the definition of clinical progression. Stated differently, the centrally reviewed pathology was to be incorporated in all relevant analyses described in the statistical analysis plan; however, in the event that centrally reviewed pathology was unavailable for a particular patient (and time point), the corresponding local pathology information captured on the case report form (if available) was to be used. Dr. Donna Hansel, Pathology Chair, reviewed the prostate biopsy data, prepared the salient results in an Excel spreadsheet, and the results were sent as an external data transfer courtesy of Linda McCart at the Alliance biorepository at Ohio State.

Summary

Resolving the myriad data discrepancies in the central pathology spreadsheet, which were identified by the members of the Alliance SDC, has been a struggle since the final data freeze (04APR2018) and was further confounded by the departure of the Pathology Chair, Dr. Donna Hansel. Therefore, it was jointly decided that no further resolution of the discrepancies would be pursued. However, the study team did focus their efforts on carefully reviewing the baseline central pathology results contained in the external spreadsheet in order to facilitate the determination of ineligible cases; per the protocol, patients who later become ineligible by centralized pathology review were to be excluded from the primary analysis population. The ineligible cases we identified immediately post-database freeze and re-reviewed/discussed on 21MAR2019 remained and the handful of potential ineligible cases that the SDC identified were indeed still eligible. In other words, no changes to the primary analysis population described in the statistical summary (and saved out to our project workspace) were needed and thus no updates to the final analyses on the primary endpoint were needed.

APPENDIX B

MAX-PC, a prostate cancer-specific measure to assess patient anxiety due to prostate cancer, comprises 3 subscales that measure general prostate cancer anxiety (11 items), anxiety related to PSA levels (3 items), and fear of recurrence (4 items). The sum of the scores on all 3 domains constitutes the total summary score of MAX-PC (18 items in total). Subscale summary scores will be calculated as the average value of each subscale domain, and the total summary score will be the average value across all 18 items.

FACT-P, a prostate cancer specific quality of life questionnaire, is a 39-item questionnaire consisting of 5 domains: physical well-being (7 items), social/family well-being (7 items), emotional well-being (6 items), functional well-being (7 items) and additional concerns (12 items). Scores can range between 0 and 156. A subscale summary score can be generated for each domain. The sum of the scores on the first 4 domains constitutes the FACT-G (27 core quality of life measures / items). The sum of the scores across all 5 domains constitutes the FACT-P. The SAS program for scoring the FACT-P can be found in the QOL forms bank.

EPIC-26 contains 5 multi-item health-related quality of life domains relevant to prostate cancer: urinary incontinence (4 items), urinary irritation/obstruction (4 items), bowel (6 items), sexual (6 items) and vitality/hormonal function (5 items); in addition, the EPIC-26 retains the single item measure of overall urinary bother. Refer to the scoring instructions for the EPIC-26 saved in the Alliance team directory; the SAS program for scoring the EPIC-26 can be found in the QOL forms bank.

IPSS, which measures lower urinary tract symptoms, is an 8 question (7 symptom questions + 1 quality of life question). The 7 symptom questions include feeling of incomplete bladder emptying, frequency, intermittency, urgency, weak stream, straining, and nocturia over the last month; each question is assigned a score from 1 to 5 for a total of maximum 35 points. The 8th question of quality of life is assigned a score of 1 to 6.

Note: All subscale summary scores and total summary scores will be transformed into 0 to 100 scales, with higher scores representing favorable health-related quality of life.

REFERENCES

1. Klotz L, Zhang L, Adam L, Nam R, Mamedov A, Loblaw A. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. Journal of Clinical Oncology, 2010; 28(1):126-131.

2. Freedman, L.S. Tables of the number of patients required in clinical trials using the log-rank test. Statistics in Medicine, 2009; 1, 121-129.

3. Memorial Sloan-Kettering Cancer Center cancer information: prostate nomogram. [Accessed November 2017]. Available at: http://www.mskcc.org/mskcc/html/10088.cfm.

4. Fine SW, Epstein JI. A contemporary study correlating prostate needle biopsy and radical prostatectomy Gleason score. Journal of Urology, 2008; 179(4):1335-1339.

5. Groenwold RH, White IR, Donders AR, Carpenter JR, Altman DG, Moons KG. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. CMAJ, 2012; 184(11):1265-9.