

A PHASE III STUDY OF THE IMPACT OF A PHYSICAL ACTIVITY PROGRAM ON
DISEASE-FREE SURVIVAL IN PATIENTS WITH HIGH RISK STAGE II OR STAGE III
COLON CANCER: A RANDOMIZED CONTROLLED TRIAL (CHALLENGE)

CCTG Protocol Number: CO.21

NCT number: NCT00819208

Document date: January 17, 2025

STATISTICAL ANALYSIS PLAN for

Final Analysis of

**A PHASE III STUDY OF THE IMPACT OF A PHYSICAL ACTIVITY PROGRAM
ON DISEASE-FREE SURVIVAL IN PARTICIPANTS WITH HIGH RISK STAGE II OR
STAGE III COLON CANCER: A RANDOMIZED CONTROLLED TRIAL
(CHALLENGE)**

NCIC CTG Protocol Number: CO.21

Dongsheng Tu, PhD
Senior Biostatistician
CCTG

____ January 17, 2025 _____

Date

Reviewed by:

Chris O'Callaghan, DVM MSc PhD
Senior Investigator
CCTG

Date

ABBREVIATIONS

AE	Adverse Event
ALT	Alanine Aminotransferase
BMI	Body Mass Index
CCTG	Canadian Cancer Trials Group
CEA	Carcinoembryonic antigen
CI	Confidence Interval
cm	Centimeter
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DFS	Disease Free Survival
DSMC	Data and Safety Monitoring Committee
ECOG	Eastern Cooperative Cancer Group
EWB	Emotional Well-being
FACIT-F	Functional Assessment of Chronic Illness Therapy – Fatigue
FREQ	Frequency
FWB	Functional Well-being
GHE	General Health Education Materials Only
HADS	Hospital Anxiety and Depression Scale
HADS-A	Hospital Anxiety and Depression Scale Anxiety Subscale
HADS-D	Hospital Anxiety and Depression Scale Depression Subscale
HR	Hazard Ratio
LKA	Last day the participant is Known Alive
LTEQ	Leisure time exercise questionnaire
MPV	Major Protocol Violation
MET	Metabolic Equivalent
METHW	Metabolic Equivalent hours/week
NA	Not Applicable
NC	Not Computed
OS	Overall Survival
PSQI	Pittsburgh Sleep Quality Index
PA	Physical Activity
PAC	Physical Activity Consultant
PA+GHE	Physical Activity Program Plus General Health Education Materials
PWB	Physical Well-being
QOL	Quality of Life
rParQ	Medical Screening Questionnaire
SAS	Statistical Analysis System
SF-36	36-Item Short Form Health Survey
STD	Standard Deviation
SWB	Social and Family Well-being
TMETHW	Total Metabolic Equivalent hours/week
TOI	Trial Outcome Index
TPAQ	Total Physical Activity Questionnaire Recreational Activity Module
VO2 Max	Maximal Oxygen Consumption
WBC	White Blood Cell Count
WHO	World Health Organization

Table of Contents

1.	<i>Introduction</i>	5
2.	<i>Study Description</i>	5
3.	<i>Objectives</i>	5
3.1	Primary	5
3.2	Secondary	6
4.	<i>Endpoints</i>	6
4.1	Primary Efficacy.....	6
4.2	Secondary Efficacy	6
4.3	Safety	7
5.	<i>Sample Size and Power</i>	7
6.	<i>Data Set Descriptions</i>	7
7.	<i>Statistical Analysis</i>	7
7.1	General Methods	7
7.2	Study Conduct	8
7.2.1	Participant Disposition and Follow-up	8
7.2.2	Accrual Patterns	8
7.2.3	Eligibility Violations	8
7.3	Study Population	8
7.3.1	Pretreatment Characteristics of Participants.....	8
7.3.2	Disease and Treatment Characteristics at Baseline.....	9
7.3.3	Baseline Assessments.....	9
7.3.4	Baseline Symptoms/Adverse Events, Lab Tests, and Other Major Medical Problems	10
7.3.5	Concomitant Medications	10
7.3.6	Tobacco Smoking History at Baseline	10
7.4	Intervention.....	10
7.5	Efficacy	11
7.5.1	Disease-Free Survival (DFS)	11
7.5.2	Overall Survival (OS).....	12
7.5.3	Physical Activity Behavior (using the Total Physical Activity Questionnaire Recreational Activity Module (TPAQ))	12
7.5.4	Fitness Testing (Objective Markers of Physical Fitness)	13
7.6	Safety	13
7.6.1	Adverse Events and Serious Adverse Events.....	13
7.6.2	Laboratory Evaluations	14
7.6.3	Other Safety	14
7.7	Other Anti-Cancer Treatments, Concomitant medications, and Tobacco Smoking History During Follow-up	14
7.8	Participant Reported Outcomes.....	15
7.8.1	Scoring Algorithms.....	15
7.8.2	Data Sets.....	19
7.8.3	Compliance	19
7.8.4	Analyses of PRO	19
8.	<i>Appendices</i>	21

Appendix 1: Tables and Figures	21
---------------------------------------	-----------

Tables

Table 1: Participant Disposition	21
Table 2: Follow-up of Participants	21
Table 3: Accrual by Country	22
Table 4: Accrual by Stratification Factor at Randomization	23
Table 5: Eligibility and Reasons for Ineligibility	23
Table 6: Pretreatment Characteristics at Baseline	24
Table 7: Disease and Treatment Characteristics at Baseline	25
Table 8: Baseline Assessments	26
Table 9: Baseline Symptoms/adverse events	28
Table 10: Baseline Hematology	29
Table 11: Baseline Biochemistry	30
Table 12: Major Medical Problems at Baseline	31
Table 13: Concomitant Medications at Baseline	31
Table 14: Tobacco Smoking History at Baseline	32
Table 15: Compliance with Study Intervention	33
Table 16: Summary of DFS Events	34
Table 17: Log rank and Cox Regression Model for Disease Free Survival	35
Table 18: Analysis for Disease Free Survival (DFS)	36
Table 19: Summary of Deaths	37
Table 20: Log rank and Cox Regression Model for Overall Survival	38
Table 21: Analysis for Overall Survival (OS)	39
Table 22: Completion Rate of TPAQ Assessment by Treatment Arm	40
Table 23: Total MET Hours/week (TMETHW) at Baseline and After Randomization	40
Table 24: Fitness Tests: Submaximal Exercise Test	40
Table 25: Fitness Tests: Anthropometrics	41
Table 26: Fitness Tests: Seniors' Fitness Test	42
Table 27: Acute Adverse Event	43
Table 28: Severe Acute Adverse Events	43
Table 29: Acute Adverse Event Related to Physical Activity Program	44
Table 30: Acute Adverse Event Related to Fitness Testing	44
Table 31: Late Adverse Event	45
Table 32: Severe Late Adverse Events	45
Table 33: Late Adverse Event Related to Physical Activity Program	46
Table 34: Late Adverse Event Related to Fitness Testing	46
Table 35: Biochemistry: Worst Grade per Subject over Study	47
Table 36: Other Radiology Investigations	48
Table 37: Hospitalization	48
Table 38: Non-Protocol Anti-Cancer Treatment	49
Table 39: Concomitant Medication	49
Table 40: Tobacco Smoking	50
Table 41: Compliance Rate of PRO Assessments by Treatment Arms	51
Table 42: Summary of Baseline PRO Scale Scores	52
Table 43: Summary of Change Scores from Baseline for PRO Scale/Domain/Item at Each Time of Assessment	53
Table 44: Summary of Analysis from Linear Mixed Models for Change Scores	54

Figures

Figure 1: Accrual by Calendar Time	22
Figure 2: Kaplan-Meier Curves for DFS	34
Figure 3: Kaplan-Meier Curves for OS	37

1. Introduction

This analysis plan is to describe the final analysis performed by the Canadian Cancer Trials Group (CCTG) for the CO.21 trial. It will be used for the writing of the CCTG study report.

2. Study Description

CCTG CO.21 is a randomized clinical trial to determine if participation in a 36-month physical activity (PA) program designed to increase recreational PA by at least 10 MET (metabolic equivalent) hours/week after adjuvant therapy for high-risk stage II or stage III colon cancer improves disease free survival (DFS). The primary objective of this trial is to compare the DFS between participants randomized to receive a 36-month PA program plus general health education materials (PA+GHE) or general health education materials only (GHE). Participants were stratified by centre, ECOG performance status (0 vs 1), disease stage (II vs. III), and BMI (≤ 27.5 vs. > 27.5) and then randomized by using the dynamic minimization method to receive PA+GHE or GHE in a 1:1 ratio. This study planned to accrue 962 participants to observe a total of 380 DFS events.

This trial was activated on December 3, 2008. Four interim analyses were planned before the final analysis of DFS would be performed: one for feasibility, one for futility and the last two assessing efficacy. First analysis on the feasibility of accrual was assessed and reported to the CCTG Data and Safety Monitoring Committee (DSMC) at its Fall 2011 meeting, while the futility analysis assessing the feasibility for PA behavior change was performed on all participants randomized before March 1, 2014 and reported to DSMC at its Spring 2015 meeting. DSMC recommended continuation of the trial after reviewing the data from both of these analyses. Two efficacy interim analyses were scheduled when respectively around one third and two thirds of the required number of events for final analysis of DFS (i.e. 125 and 250 events) were observed. The first efficacy analysis was performed on November 7, 2019 on a database locked on November 1, 2019 which included all the data observed on or before October 30, 2019, the date of clinical cutoff, and a total of 132 DFS events. After reviewing the report of this first interim analysis, DSMC recommended the trial should continue because the prespecified criteria for stopping were not met.

Because of slow accrual and lower than expected event rate, in its fall meeting of 2023, DSMC approved a request from the trial committee to set an accrual deadline of December 31, 2023, and a target for declaring a clinical cut-off date for final statistical analysis at the end of 2024 conditional on observation of 200 confirmed DFS events. Registration of participants stopped on December 21, 2023 and the final participant was randomized on January 24, 2024 with a total of 889 participants randomized. 231 DFS events were observed on August 28, 2024. It was decided that the final analysis will be performed after the data observed from all participants before August 29, 2024, the date of clinical cut off, are reviewed and cleaned. This analysis plan describes the analyses performed by the Canadian Cancer Trials Group (CCTG) for this final analysis.

3. Objectives

3.1 Primary

To compare the disease-free survival (DFS) of participants randomized to two treatments.

3.2 Secondary

3.2.1 To compare between two treatment groups in

- Overall survival (OS);
- Participant Reported Outcomes including QOL using the SF-36, FACIT-F, PSQI and HADS;
- Objective markers of physical fitness using BMI, hip and waist circumference, submaximal exercise testing, and the Seniors' Fitness Test;
- Physical activity behavior using the Total Physical Activity Questionnaire Recreational Activity Module (TPAQ);
- Safety profile as assessed by Common Toxicity Criteria for Adverse Events (CTCAE);
- Serum levels of insulin, IGF-1, IGF-2, IGFBP3;
- Cytokine levels of IL-1 β , IL-6, IL-2, IL-4, IL-8, IL-10, IL-12, TNF- α , IFN- γ , and GM-CSF and C-reactive protein;

3.2.2 To perform economic evaluations including cost-effective and cost-utility analyses;

3.2.3 To identify predictors of physical activity adherence using the Social-Cognitive Determinants of Exercise Measure;

3.2.4 To evaluate in all randomized participants the potential prognostic associations of the following variables:

- Serum levels of insulin, IGF-1, IGF-2, IGFBP3 and blood glucose, and cytokine levels of IL-1 β , IL-6, IL-2, IL-4, IL-8, IL-10, IL-12, TNF- α , IFN- γ , and GM-CSF and C-reactive protein with DFS, OS, level of physical activity and level of fatigue
- Age, gender, country, incremental increase in physical activity and change in cardiovascular fitness with DFS, OS, level of fatigue and QOL.

3.2.5 To establish a comprehensive specimen bank linked to a clinical database for the further study of molecular markers of colon cancer.

Note: The analysis plan to address correlative and economic analyses will be developed at a later date.

4. Endpoints

4.1 Primary Efficacy

Disease-free survival (DFS) is the primary efficacy endpoint for this analysis.

4.2 Secondary Efficacy

The secondary efficacy endpoints for this analysis include:

- Overall survival (OS)
- Objective markers of physical fitness (using BMI, hip and waist circumference, submaximal exercise testing, and the Seniors' Fitness Test)
- Physical activity behaviour (using the Total Physical Activity Questionnaire Recreational

Activity Module (TPAQ))

- Participant reported outcomes (using SF-36, FACIT-F, PSQI and HADS)

4.3 Safety

The safety endpoints are serious and non-serious adverse events, during the interventions and follow-up.

5. Sample Size and Power

The trial was powered to detect a hazard ratio (HR) of 0.75 for DFS between participants randomized respectively to two treatment arms (intention-to-treat population), which corresponds to an increase in 3 years DFS from 75% on GHE to 80.6% on PA+GHE. To detect such a hazard ratio with a power of 80% and a two-tailed 0.05, we need to observe 380 events before the final analysis. With at least 200 events observed in the final analysis, the study will have 80% power to detect a hazard ratio of 0.67 (corresponding to an increase in 3 year DFS from 75% on GHE to 82.5% on PA+GHE) at two-sided 0.05 level.

6. Data Set Descriptions

All Randomized Participants:

All participants who have been randomized in the study with the treatment arm as randomized.

All Treated Participants:

All participants who have received at least one session of intervention (receipt of general health education material and/or with provision of instruction about PA) with the treatment arm as actually received.

7. Statistical Analysis

7.1 General Methods

Continuous and ordinal categorical variables are summarized using the mean with standard deviation or median with minimum and maximum values and, when appropriate, compared using the Wilcoxon test. Discrete variables are summarized with the number of proportion of participants falling into each category and compared using Fisher's exact test when appropriate. Time to event variables are summarized using Kaplan-Meier plots. Primary estimates of the treatment differences are obtained with the hazard ratios and confidence intervals from stratified Cox regression models using treatment arm as the single factor. All confidence intervals are computed based on normal approximations except those for rates, which will be computed based on the exact method.

Percentages given in the summary tables will be rounded and may therefore not always add up to exactly 100%. Listings, tabulations, and statistical analyses will be carried out using the SAS (Statistical Analysis System, SAS Institute, North Carolina, USA) software.

Unless otherwise specified, date of randomization and stratification factors will be taken from the Centralized Randomization File.

Baseline evaluations will be those collected on CRF Eligibility Worksheet and Baseline Report and closest to, but no later than, the first day of study treatment for treated participants and

closest to, but no later than, the date of randomization, for participants who were randomized but who never received treatment.

Laboratory results, adverse events, and other symptoms are coded and graded using the CTCAE when available.

When converting a number of days to other units, the following conversion factors will be used:

1 year = 365.25 days

1 month = 30.4375 days

When either day or month of a date is missing, the missing day and/or month will be imputed by the midpoint within the smallest known interval. For example, if the day of the month is missing for any date used in a calculation, the 15th of the month will be used to replace the missing day. If the month and day of the year are missing for any date used in a calculation, the first of July of the year will be used to replace the missing date.

7.2 Study Conduct

All randomized participants are included in the analyses of study conduct. Information will be tabulated by randomized treatment (unless otherwise indicated) and pooled treatments.

7.2.1 Participant Disposition and Follow-up

- Number of participants randomized, had intervention (on intervention, off intervention, reason off intervention) (**Table 1**)
- Number of alive participants (**Table 2**)
- Median (estimated by Kaplan-Meier method) and range (minimum and maximum) (**Table 2**) of the follow-up time (months) defined as time from the day of randomization (as recorded in centralized randomization file) to the last day the participant is known alive (LKA) as defined in 7.5.1 or censored at the time of death and calculated as

$$[(\text{date of death or LKA} - \text{date of randomization}) + 1]/30.4375.$$

7.2.2 Accrual Patterns

- Number of participants randomized by participating country (**Table 3**)
- Number of participants by stratification factor (except centre) at randomization (**Table 4**)
- Accrual of participants by calendar time pooled across two treatment arms (**Figure 1**)

7.2.3 Eligibility Violations

Eligibility violations of inclusion or exclusion criteria are centrally reviewed by CCTG.

- Number of participants eligible, not eligible (**Table 5**)
- Reasons for ineligibility (**Table 5**)

7.3 Study Population

All randomized participants are included in the study population analyses. Information will be tabulated by randomized treatment (unless otherwise indicated) and pooled treatments.

7.3.1 Pretreatment Characteristics of Participants

- Gender: female, male (**Table 6**)

- Age (years): median, minimum, maximum values; <65, ≥65 (**Table 6**)
- BMI: median, minimum, maximum values; ≤27.5, >27.5 (**Table 6**)
- ECOG Performance Status: 0, 1 (**Table 6**)

7.3.2 Disease and Treatment Characteristics at Baseline

- Time from first diagnosis of histopathologically confirmed adenocarcinoma of the colon to randomization (years): median, minimum, maximum values (**Table 7**)
- Disease stage: High risk stage II, stage III (**Table 7**)
- Clinical T-stage: Tx, T1, T2, T3, T4 (**Table 7**)
- Histological grade (WHO): I, II, III, IV (**Table 7**)
- Histology result of lymph nodes: positive, negative (**Table 7**)
- Number of positive lymph nodes: median, minimum, maximum values; 0, 1+ (**Table 7**)
- Tumour deposits: present, absent (**Table 7**)
- Time from complete resection of adenocarcinoma of the colon to randomization (years): median, minimum, maximum values (**Table 7**)
- Type of chemotherapy: Fluorouracil-based, oxaliplatin, etc. (**Table 7**)
- Time from last dose of chemotherapy to randomization (months): median, min-max (**Table 7**)

7.3.3 Baseline Assessments

- Answers to Medical Screening Questionnaire (rParQ): All “no”, at least one “yes” and investigators confirms participant is suitable for participation on study (**Table 8**)
- Answers to Leisure time exercise questionnaire (LTEQ) at pre-registration (**Table 8**)
 - Number of minutes per week of light/mild exercise: median, min-max (**Table 8**)
 - Number of minutes per week of moderate exercise: median, min-max (**Table 8**)
 - Number of minutes per week of vigorous/strenuous exercise: median, min-max (**Table 8**)
 - Total number of minutes per week of moderate to vigorous/strenuous exercise: median, min-max (**Table 8**)
- Answers to Leisure time exercise questionnaire (LTEQ) at pre-diagnosis (**Table 8**)
 - Number of minutes per week of light/mild exercise: median, min-max (**Table 8**)
 - Number of minutes per week of moderate exercise: median, min-max (**Table 8**)
 - Number of minutes per week of vigorous/strenuous exercise: median, min-max (**Table 8**)
 - Total number of minutes per week of moderate to vigorous/strenuous exercise: median, min-max (**Table 8**)
- Results of submaximal exercise test at the 2nd last stage completed (**Table 8**)
 - Speed: median, min-max (**Table 8**)
 - % grade: median, min-max (**Table 8**)
 - Heart rate: median, min-max (**Table 8**)
- Results of submaximal exercise test at the last stage completed (**Table 8**)
 - Speed: median, min-max (**Table 8**)
 - % grade: median, min-max (**Table 8**)
 - Heart rate: median, min-max (**Table 8**)
- VO2 max: median, min-max (**Table 8**)
- Anthropometrics
 - Hip circumference (cm): median, min-max (**Table 8**)

- Waist circumference (cm): median, min-max (**Table 8**)
- Results of Senior's fitness test at baseline (**Table 8**)
 - Number of repetitions of Sit to Stand: median, min-max (**Table 8**)
 - Number of repetitions of Arm Curls: median, min-max (**Table 8**)
 - 6-minute walk total (meters): median, min-max (**Table 8**)
 - Chair sit and reach result (cm): median, min-max (**Table 8**)
 - Back Scratch result (cm): median, min-max (**Table 8**)
 - 8-foot Up and Go result (seconds): median, min-max (**Table 8**)

7.3.4 Baseline Symptoms/Adverse Events, Lab Tests, and Other Major Medical Problems

- Baseline symptom/adverse event (**Table 9**)
- Baseline hematology: hemoglobin, WBC, absolute granulocytes, platelets (**Table 10**)
- Blood transfusion or an erythropoietic stimulating agent within six weeks prior to date of hemoglobin: yes, no (**Table 10**)
- Baseline serum chemistry: ALT, alkaline phosphatase, total bilirubin, serum Creatinine, CEA, fasting glucose (**Table 11**)
- Number of participants with major medical problems ongoing at baseline (**Table 12**)
- Type of major medical problem: diabetes, hypertension, etc. (**Table 12**)

7.3.5 Concomitant Medications

- Number of participants with concomitant medication within 14 days prior to the date of randomization (**Table 13**)
- Type of concomitant medication: ASA, NSAIDS, statin, etc. (**Table 13**)

7.3.6 Tobacco Smoking History at Baseline

- Current or past smoker: yes, no, unknown (**Table 14**)
- Currently smoking: yes, no, unknown (**Table 14**)
- Current average number of cigarettes smoked per day: median, min-max (**Table 14**)
- Time from quitting smoking to randomization: median, min-max (**Table 14**)
- Age at the time of quitting smoking: median, min-max (**Table 14**)
- Smoking history: pipe or cigar smoker only, 100 or fewer cigarettes during lifetime, greater than 100 cigarettes during lifetime, unknown (**Table 14**)
- Time from beginning of smoking cigarettes to randomization: median, min-max (**Table 14**)
- Total number of years smoked cigarettes: median, min-max (**Table 14**)
- Average number of cigarettes smoked per day: median, min-max (**Table 14**)
- Pack years: median, min-max (**Table 14**)

7.4 Intervention

Participants in both arms of this study will receive general health materials and those randomized to PA+GHE will receive a physical activity program with 3 phases: an intensive intervention for 6 months in phase 1, reduced intervention for months 6-12 months in phase 2, and a minimal intervention for months 12-36 months in phase 3.

- Received general health education materials in phase 1: yes, no (**Table 15**)
- Reason the general health education materials were not provided in phase 1: participant

refused, etc. (**Table 15**)

For participants on PA+GHE only:

- Number of mandatory behavioral support sessions attended in phases 1-3: median, min-max values (**Table 15**)
- Number of mandatory supervised physical activity sessions attended in phase 1: median, min-max values (**Table 15**)
- Number of recommended supervised physical activity sessions attended in phases 1-3: median, min-max values (**Table 15**)
- Reason for non-attendance of mandatory sessions: recurrent disease, etc. (**Table 15**)
- Number of additional sessions attended: median, min-max values (**Table 15**)
- Reason for attendance of additional sessions: participant request, PAC request (**Table 15**)
- Received ‘tool box’ resource: yes, no (**Table 15**)
- Type of tool box resources received: exercise videos, etc. (**Table 15**)

7.5 Efficacy

7.5.1 Disease-Free Survival (DFS)

Disease-free survival, the primary endpoint of this study, will be calculated for all participants from the day of randomization until the earliest time of the following dates: (1) earliest of suspected and definite dates of local recurrence on CRF Recurrence/New Primary malignancy Report when both dates are non-missing or definite date of local recurrence when the suspected date of local recurrence is missing but definite date of local recurrence is not missing; (2) earliest of suspected and definite dates of distant recurrence on CRF Recurrence/New Primary malignancy Report when both dates are non-missing or definite date of distant recurrence when the suspected date of distant recurrence is missing but definite date of distant recurrence is not missing; (3) earliest of suspected and definite dates of new primary malignancy on CRF Recurrence/New Primary malignancy Report when both dates are non-missing or definite date of new primary malignancy when the suspected date of new primary malignancy is missing but definite date of new primary malignancy is not missing; (4) date of the suspected local recurrence if a patient had a definite date of distant recurrence but only a suspected date of local recurrence; (5) date of death on CRF Death Report. If there is no definite date of local recurrence, distant recurrence, or new primary malignancy on CRF Recurrence/New Primary malignancy Report or date of death on CRF Death Report, DFS will be censored on the last day the participant is known alive (LKA), defined as the latest of randomization date, form dates of CRF Intervention Report (Form 3), Follow-up Year 4 and 5 Report (Forms 5A, 5B, 5C, 5D), and Short (Form 5S) and Minimum (Form 5M) Follow-up, and dates of lost to follow-up and consent withdraw. Reports unless a participant is ineligible with a reason “malignancy at baseline” where the randomization date will be used to censor DFS.

All participants randomized to the study will be included in the analysis in the arm they are randomized. A Kaplan-Meier curve for disease free survival in each treatment arm will be displayed (Figure 2). A frequency table will be provided describing DFS events (with definite date of local or distant recurrence, new primary malignancy, or death) observed. A patient with a definite date of distant recurrence but only a suspected date of local recurrence will be considered having both local and distant recurrences (**Table 16**). The difference between the

two treatment arms will be tested using the log-rank test stratified by ECOG performance status of 0 or 1, disease stage (II vs. III) and BMI (< 27.5 vs. > 27.5) at the time of randomization. In addition, as an exploratory analysis, a stratified Cox regression model adjusting for some other prognostic factors will be applied to verify the impact of the prognostic factors on the treatment effect. The stratified Cox regression model will include the following prespecified variables: treatment (PA+GHE vs. GHE), age (<60 years vs. ≥60 years), sex (female vs. male), country (Canada vs. others), chemotherapy regimen (FOLFOX vs. others), pre-diagnosis total minutes of moderate to vigorous/strenuous exercise (<150 vs. ≥150), baseline total MET Hours/week (<10 vs. ≥10), baseline VO2 max (<30 vs. ≥30), baseline 6-minute walk total (<500 meters vs. ≥500 meters) (**Table 17**). Analyses will also be performed for the subgroups defined by the above variables, except the treatment, plus three stratification factors at baseline (**Table 18**).

7.5.2 Overall Survival (OS)

Overall survival is calculated as the time from randomization to the time of death from any cause. Participants who are alive at the time of the final analysis or who have become lost to follow-up will be censored at their last day the participant is known alive (LKA) as defined above. All analyses for DFS will also be performed for OS, using similar methodology (**Figure 3, Table 19-Table 21**).

7.5.3 Physical Activity Behavior (using the Total Physical Activity Questionnaire Recreational Activity Module (TPAQ))

The Total Physical Activity Questionnaire (TPAQ) is one of the questionnaires that describe the physical activity of a participant in the past month (i.e. four weeks). It is completed by participants at baseline, every 6 months during 36 months of protocol treatment and every 12 months after protocol treatment. It is not required after recurrence/new malignancy or 5 years post randomization. TPAQ consists of three components: (1) Employment & Volunteer activities; (2) Household & Do-it-yourself activities and (3) Recreation & Leisure activities.

For each recreation or leisure activity listed, its METHW (MET hours/week) is calculated as

$$(\text{MET} \times \text{HOURS} \times \text{FREQ})$$

if frequency (FREQ) of the activity (in days) was recorded as week (i.e., period=W) or

$$(\text{MET} \times \text{HOURS} \times \text{FREQ}) / 4.33$$

if frequency (FREQ) of the activity (in days) was recorded as month (i.e., period=M), where HOURS is the total hours per day for the activity and MET is the specific MET value for the activity. The total METHW (TMETHW) for a participant will be the sum of METHWs for all activities of this participant. If a participant didn't have any leisure activity filled out, the total MET will be considered as 0. The following information will be summarized:

- Number of participants with non-missing TMETHW at baseline and each of post-baseline assessment and reason for missing TMETHW (**Table 22**)
- Mean and standard deviation of the TMETHW at baseline and each of post-baseline assessments by treatment arm with p-values from Wilcoxon tests for the comparisons between two intervention groups (**Table 23**)

7.5.4 Fitness Testing (Objective Markers of Physical Fitness)

Objective markers of physical fitness, which include BMI, submaximal exercise test, measurements of hip and waist circumferences, and senior's fitness test, are used to assess the fitness of participants in this study. The improvement from baseline to each of assessment times in these markers will be compared between two treatment arms by a Wilcoxon test.

7.5.4.1 Submaximal Exercise Test

The heart rates between 2nd last stage and last stage tests will be calculate as VO2 Max. For each post-baseline timepoint of assessment, participants with either a negative VO2 Max result at baseline or at this timepoint and also participants that have treatment type = beta blocker in the concomitant medications table of either their baseline or intervention report at this timepoint will be excluded from the analysis.

- Mean and standard deviation of the changes in VO2 Max from baseline to 6, 12, 24, and 36 months post randomization by treatment arm with p-values (**Table 24**)

7.5.4.2 Anthropometrics

- Mean and standard deviation of the changes in the BMI and hip and waist circumferences from baseline to 6, 12, 24, and 36 months by treatment arm with p-values (**Table 25**)

7.5.4.3 Seniors Fitness test

- Mean and standard deviation of the changes in each result of seniors' fitness test from baseline to 6, 12, 24, and 36 months by treatment arm with p-values (**Table 26**)

7.6 Safety

The safety analyses will be based on the treated population defined in Section 6. Adverse events are graded and categorized using the CTCAE.

7.6.1 Adverse Events and Serious Adverse Events

Adverse Events (AEs) sections of CRF Intervention Report (Form 3) and CRF Follow Up Year 4 and 5 Report (Form 5A) record adverse events observed during respectively intervention and follow-up between 4 and 5 years post randomization, both prior to recurrence/new malignancy. AEs reported on CRF Intervention Report are defined as acute AEs and those reported on CRF Follow Up Year 4 and 5 Report as late AEs. Severe adverse events are those events reported with a CTCAE Grade of 3 or higher. Physical activity program or fitness testing related adverse events are those events with a relation to protocol therapy of 3=possible, 4=probable or 5=definite.

The following variables are summarized:

- Acute adverse events: worst CTCAE grade (**Table 27**)
- Severe acute adverse events: worst CTCAE grade (**Table 28**)
- Acute adverse events related to physical activity program: worst CTCAE grade (**Table 29**)
- Acute adverse events related to fitness testing: worst CTCAE grade (**Table 30**)
- Late adverse events: worst CTCAE grade (**Table 31**)
- Severe late adverse events: worst CTCAE grade (**Table 32**)
- Late adverse events related to physical activity program: worst CTCAE grade (**Table 33**)

- Late adverse events related to fitness testing: worst CTCAE grade (**Table 34**)

7.6.2 Laboratory Evaluations

CEA was assessed every 6 months during intervention and every 12 months at 4 and 5 years post randomization and prior to recurrence/new malignancy, while fasting glucose were assessed every 12 months only during intervention.

- CEA: Normal (all lower than or equal to upper lower limit), high (at least one higher than the upper lower limit, not reported (no any post-baseline values) (**Table 35**)
- Fasting glucose: Normal (all lower than or equal to upper lower limit), high (at least one higher than the upper lower limit, not reported (no any post-baseline values) (**Table 35**)

7.6.3 Other Safety

- Any radiologic investigations during intervention (reported on CRF Intervention Report): yes, no (**Table 36**)
- Results of radiologic investigations during intervention: all not malignant, at least one malignant (**Table 36**)
- Any radiologic investigations during 4-5 year follow-up (reported on CRF Follow Up Year 4 and 5 Report): yes, no (**Table 36**)
- Results of radiologic investigations during 4-5 year follow-up: all normal, at least one abnormal but all non-malignant, all abnormal and malignant (**Table 36**)
- Any hospitalization during intervention (reported on CRF Intervention Report): yes, no (**Table 37**)
- Any hospitalization during 4-5 year follow-up but before recurrence/new primary malignancy (reported on CRF Follow Up Year 4 and 5 Report): yes, no (**Table 37**)
- Any hospitalization after recurrence or new primary malignancy but within 5 years from randomization (reported on CRF Short Follow-up Report): yes, no (**Table 37**)

7.7 Other Anti-Cancer Treatments, Concomitant medications, and Tobacco Smoking History During Follow-up

The information on the non-protocol anti-cancer treatment was collected after recurrence or new primary malignancy, on concomitant medications during intervention and 4-5 year follow-up but before recurrence or new primary malignancy, and smoking history before and after recurrence or new primary malignancy prior to 5 years from randomization.

- Any non-protocol anti-cancer treatment after recurrence or new primary malignancy but within 5 years from randomization (reported on Short Follow Up Report): yes, no (**Table 38**)
- Type and reason for non-protocol anti-cancer treatment after recurrence or new primary malignancy but within 5 years from randomization (reported on Short Follow Up Report) (**Table 38**)
- Any Concomitant medication during intervention: yes, no (**Table 39**)
- Type of concomitant medication during intervention (**Table 39**)
- Any Concomitant medication during 4-5 year follow-up but before recurrence/new primary malignancy: yes, no (**Table 39**)
- Type of concomitant medication during 4-5 year follow-up but before recurrence/new primary malignancy (**Table 39**)
- Ever smoked during intervention (reported on CRF Intervention Report): yes, no (**Table 40**)

- Average number of cigarettes smoked per day during intervention: median, min-max (Table 40)
- Ever quit smoking during intervention: yes, no (Table 40)
- Ever smoked during 4-5 year follow-up but before recurrence/new primary malignancy (reported on CRF Follow Up Year 4 and 5 Report): yes, no (Table 40)
- Average number of cigarettes smoked per day during 4-5 year follow-up but before recurrence/new primary malignancy: median, min-max values (Table 40)
- Ever quit smoking during 4-5 year follow-up but before recurrence/new primary malignancy: yes, no (Table 40)
- Ever smoked after recurrence or new primary malignancy but within 5 years from randomization (reported on CRF Short Follow-up Report): yes, no (Table 40)
- Average number of cigarettes smoked per day after recurrence or new primary malignancy but within 5 years from randomization (reported on CRF Short Follow-up Report): median, min-max values (Table 40)
- Ever quit smoking after recurrence or new primary malignancy but within 5 years from randomization (reported on CRF Short Follow-up Report): yes, no (Table 40)

7.8 Participant Reported Outcomes

The participant report outcomes (PROs) of participants in this study include SF-36, FACIT-F, PSQI, and HADS and are assessed before randomization (baseline), every 6 months during intervention and every 12 months 4 and 5 years post randomization and prior to recurrence/new malignancy. SF 36 is also assessed after recurrence/new primary malignancy q6 monthly from randomization to 36 months and then annually until 5 years. The following are the scoring algorithms for SF-36, FACIT-F, PSQI, and HADS with Q_i denoting the i -th question in the PRO questionnaire.

7.8.1 Scoring Algorithms

7.8.1.1 SF36

There are eight health dimensions defined below that can be derived from SF-36: physical health, role function-physical, bodily pain, general health, vitality, social function, role functioning-emotional, mental health.

The MOS SF-36 scoring manual is used to calculate the scale scores for these dimensions. The raw scores for the following questions will be first recoded based on the following scheme, so the high score for each dimension will represent better quality of life:

	Raw score	Recoded score
Questions 1 & 6	1	5.0
	2	4.4
	3	3.4
	4	2.0
	5	1.0
Question 7	1	6.0
	2	5.4
	3	4.2
	4	3.1

	5	2.2
	6	1.0
Question 8	1	6.0 if question 7 is answered with raw score=1 or question 7 is not answered; 5.0 if question 7 is answered with raw score higher than 1;
	2	4.0 if question 7 is answered; 4.75 if question 7 is not answered;
	3	3.0 if question 7 is answered; 3.5 if question 7 is not answered;
	4	2.0 if question 7 is answered; 2.25 if question 7 is not answered;
	5	1.0
Question 9a, 9d, 9e & 9h	1	6.0
	2	5.0
	3	4.0
	4	3.0
	5	2.0
	6	1.0
Question 11b and d	1	5.0
	2	4.0
	3	3.0
	4	2.0
	5	1.0

The following are the questions which are used to define each of the dimensions:

Dimension	Defining questions	Lowest possible raw score	Possible raw score range
Physical health	3a-3j	10	20
Role function-physical	4a-4d	4	4
Bodily pain	7, 8	2	10
General health	1, 11a-11d	5	20
Vitality	9a, 9e, 9g, 9i	4	20
Social function	6, 10	2	8
Role functioning-emotional	5a-5c	3	3
Mental health	9b-9d, 9f, 9h	5	25

The raw scores or recoded raw scores as required of all questions defining a dimension will then be added together to calculate the raw dimensional score if at least 50% of these questions were answered. Otherwise, the raw dimensional score for this dimension will be set as missing. The raw dimensional score will then be transformed to a 0 to 100 scale using the following formula

and the lowest possible raw dimensional score and possible raw dimensional score range as specified in the table below:

Dimensional scale=[(raw dimensional score-lowest possible raw dimensional score)/possible raw dimensional score range] x 100.

These dimensional scales are used to calculate physical component summary and mental component summary by the following algorithm:

1. Calculate standardized dimensional scale is calculated for each dimension:

Standardized scale for physical health (PF_Z)=(dimensional scale-84.52404)/22.89490;

Standardized scale for role function-physical (RP_Z)=(dimensional scale-81.19907)/33.79729;

Standardized scale for bodily pain (BP_Z)=(dimensional scale-75.49196)/23.55879;

Standardized scale for general health (GH_Z)=(dimensional scale-72.21316)/20.16964;

Standardized scale for vitality (VT_Z)=(dimensional scale-61.05453)/20.86942;

Standardized scale for social function (SF_Z)=(dimensional scale-83.59753)/22.37642;

Standardized scale for role functioning-emotional (RE_Z)=(dimensional scale-81.29467)/33.02717;

Standardized scale for mental health (MH_Z)=(dimensional scale-74.84212)/18.01189.

2. Calculate raw scores of physical component summary (RPCS) and mental component summary (RMCS):

RPCS=(PF_Z * .42402)+(RP_Z * .35119)+(BP_Z * .31754)+(SF_Z * (-.00753))+
(MH_Z * (-.22069))+(RE_Z * (-.19206))+(VT_Z * .02877)+(GH_Z * .24954);

RMCS=(PF_Z * (-.22999))+(RP_Z * (-.12329))+(BP_Z * (-.09731))+(SF_Z * .26876)+
(MH_Z * .48581)+(RE_Z * .43407)+(VT_Z * .23534)+(GH_Z * (-.01571)).

3. The final scores of physical component summary (PCS) and mental component summary (MCS) are then calculated respectively:

PCS = (RPCS*10) + 50;

MCS = (RMCS*10) + 50.

7.8.1.2 FACIT-F

The FACIT-F questionnaire is a 40-item instrument consisting of the 27-item cancer-specific QoL instrument FACT-G and 13 items that assess fatigue and its impact (the FACIT-Fatigue). The FACT-G consists of four subscales: (1) physical well-being (PWB); (2) social and family well-being (SWB); (3) emotional well-being (EWB); and functional well-being (FWB). There is only one subscale, fatigue (F) subscale, in FACIT-Fatigue.

The FACIT-F Trial Outcome Index (TOI) is the sum of the PWB, FWB, and Fatigue subscales. Individual scores for each subscale, FACT-G score, the trial outcome index (TOI), and the total FACIT-F score will be scored according to FACIT-F Scoring Guidelines as below with a

subscale in which less than half of the items are completed treated as missing. The higher the score, the better the QoL.

- PWB subscale score = $(28 - GP1 - GP2 - CP3 - GP4 - GP5 - GP6 - GP7) * 7 / (\text{number of items answered})$
- SWB subscale score = $(GS1 + GS2 + CS3 + GS4 + GS5 + GS6 + GS7) * 7 / (\text{number of items answered})$
- EWB subscale score = $(20 - GE1 + GE2 - GE3 - GE4 - GE5 - GE6) * 6 / (\text{number of items answered})$
- FWB subscale score = $(GF1 + GF2 + CF3 + GF4 + GF5 + GF6 + GF7) * 7 / (\text{number of items answered})$
- Fatigue (F) subscale score = $(40 - HI7 - HI12 - An1 - An2 - Ab3 - An4 + An5 + An7 - An8 - An12 - An14 - An15 - An16) * 13 / (\text{number of items answered})$
- FACT-G score = PWB score + SWB score + EWB score + FWB score
- FACIT-F TOI = PWB score + FWB score + F score
- FACIT-F total score = PWB score + SWB score + EWB score + FWB score + F score

7.8.1.3 PSQI

The PSQI is a 19-item self-report measure that provides scores in 7 subcategories (subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction) as well as a total score with higher scores indicating worse sleep quality. The raw scores for the following questions will be first recoded based on the following scheme:

	Raw score	Recoded score
Question 2	< 15 minutes	0
	16-30 minutes	1
	31-60 minutes	2
	> 60 minutes	3
Question 4	> 7 hours	0
	6-7 hours	1
	5-6 hours	2
	< 5 hours	3

The following algorithm is then used to calculate the scores for each subcategories and the total score:

- Subjective sleep quality Q9-1
- Sleep latency =0 if sum of (Q5a-1) and recoded Q2 is 0; =1 if sum is 1 or 2; =2 if sum is 3 or 4; =3 if sum is 5 or 6
- Sleep duration Recoded score of Q4
- Sleep efficiency =0 if Habitual sleep efficiency defined as $((Q4 \text{ raw score}) / (Q3 - Q1 \text{ in hours})) > 0.85$; =1 if 0.75-0.84; =2

- Sleep disturbance if 0.65-0.74; =3 if <0.65
=0 if sum of (Q5b-1) to (Q5j-1) is 0; =1 if sum is 1-9; =2 if sum is 10-18; =3 if sum is 19-27
- Use of sleep medication Q6-1
- Daytime dysfunction =0 if sum of (Q7-1) and (Q8-1) is 0; =1 if sum is 1 or 2; =2 if sum is 3 or 4; =3 if sum is 5 or 6
- Total Score Sum of the scores of above 7 subcategories

7.8.1.4 HADS

The HADS is a 14-item measure with two subscales, the HADS-A (Anxiety subscale) and the HADS-D (Depression subscale), and a total score. For each participant and at each timepoint, the score of HADS-A and HADS-B were calculated and included with the answers to each question. The total score is calculated as:

- HADS Total score Sum of HADS-A score and HADS-D score

For all the above PRO questionnaires, missing items in a multi-item scale will be handled by the following methods: Values will be imputed for missing items by “assuming that the missing items have values equal to the average of those items which are present” for the total score in which at least half the items are completed. The total score will be treated as missing if less than half of the items are completed.

7.8.2 Data Sets

The analyses of PRO data will be restricted to treated participants who have completed at least one PRO questionnaire.

The form dates in the SAS databases for each questionnaire will be matched with completion date of each questionnaire in respective, CRF Baseline report (Form F1), CRF INTERVENTION REPORT (Form 3), and CRF FOLLOW UP YEAR 4 AND 5 REPORT (Form 5A), and CRF Short Follow-up form (Form 5S) (for SF-36 only) to determine the timepoints of evaluations for baseline and change score analysis. Specifically, the evaluations with form dates matched with the completion dates in CRF Baseline report will be considered as baseline evaluations, the time points for those with form dates matched with the completion dates of other forms will be defined by the *Time Elapsed Since Randomization* in these forms.

7.8.3 Compliance

Compliance with each PRO questionnaire will be described for each treatment arm, at each time of evaluation, by the number and percentage of expected participants who filled out a questionnaire (per subject, at least one question answered) in time of evaluation. The expected number used in calculating the percentage will be all alive participants who are required to complete the assessment in each treatment arm at each assessment time point (Table 41).

7.8.4 Analyses of PRO

7.8.4.1 Baseline and Change Score Analysis

Descriptive statistics for SF-36, FACIT-F, PSQI, and HADS scores (mean, standard deviation)

will be presented for each scale and summary scale at baseline for each treatment group (**Table 42**). The change scores from baseline will be generated at each time of post-baseline evaluation and compared between two arms by Wilcoxon test (**Table 43**).

7.8.4.2 Profile Analysis

The profile of change scores over time between two treatment arms will be compared using linear mixed models with the treatment effect, time of assessment from randomization (in months), and their interaction as fixed effects and the intercept and time as two random effects. The estimates for the terms of fixed effect with standard errors and p-values will be presented (**Table 44**).

8. Appendices

Appendix 1: Tables and Figures

Table 1: Participant Disposition

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE	GHE	Total
Randomized	N=***	N=***	N=***
Had intervention	*** (**)	*** (**)	*** (**)
On intervention	*** (**)	*** (**)	*** (**)
Off intervention	*** (**)	*** (**)	*** (**)
Reasons			
Intervention program completed as per protocol	*** (**)	*** (**)	*** (**)
Recurrent Disease	*** (**)	*** (**)	*** (**)
New primary malignancy	*** (**)	*** (**)	*** (**)
Participant choice	*** (**)	*** (**)	*** (**)
Physician discretion	*** (**)	*** (**)	*** (**)
Death	*** (**)	*** (**)	*** (**)
Adverse event (s) related to physical activity program or fitness testing	*** (**)	*** (**)	*** (**)

Table 2: Follow-up of Participants

Dataset: All Intention-To-Treat Participants			
	Number of participants (%)		
	PA+GHE	GHE	Total
Number of participants alive	*** (*)	*** (*)	*** (*)
Follow-up (months)			
Median	**	**	**
Minimum-maximum	**_**	**_**	**_**

Table 3: Accrual by Country

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N=***	GHE N=***	Total N=***
Canada	*** (**)	*** (**)	*** (**)
Australia	*** (**)	*** (**)	*** (**)
USA	*** (**)	*** (**)	*** (**)
...	*** (**)	*** (**)	*** (**)

Figure 1: Accrual by Calendar Time

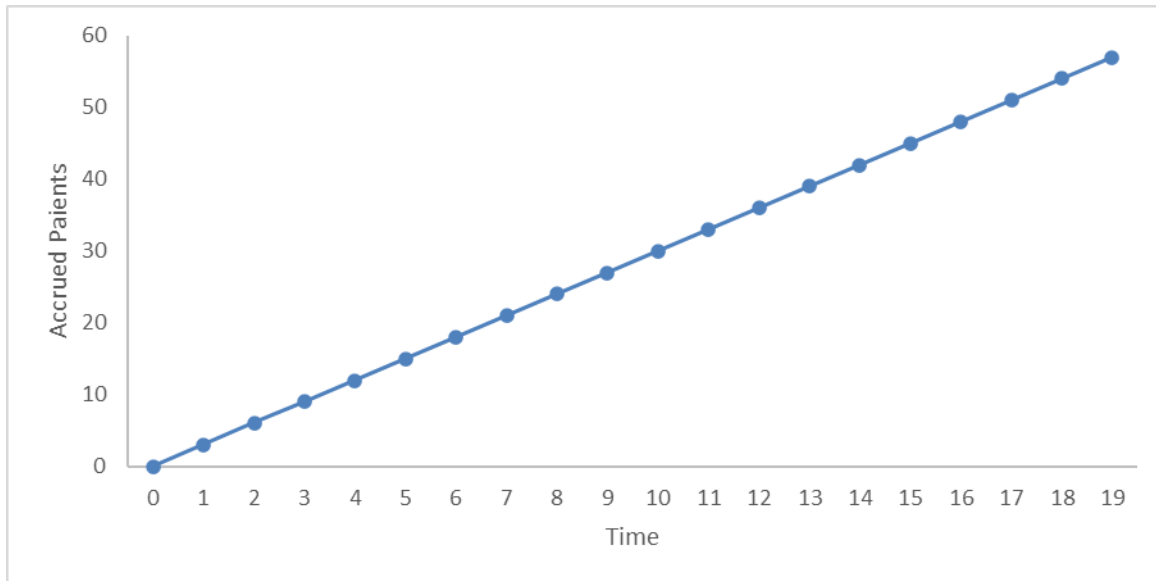


Table 4: Accrual by Stratification Factor at Randomization

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N=***	GHE N=***	Total N=***
Disease stage			
High risk stage II	** (**)	** (**)	** (**)
Stage II	** (**)	** (**)	** (**)
ECOG Performance Status			
0	** (**)	** (**)	** (**)
1	** (**)	** (**)	** (**)
BMI			
≤ 27.5	** (**)	** (**)	** (**)
>27.5	** (**)	** (**)	** (**)

Source: Centralized Randomization File

Table 5: Eligibility and Reasons for Ineligibility

Data set: All Randomized Participants			
	Number of Participants (%)		
	PA+GHE N=***	GHE N=***	Total N=***
Eligible	*** (**)	*** (**)	*** (**)
Not Eligible	*** (**)	*** (**)	*** (**)
Reason for ineligibility			
<Reason 1>	**	**	**
<Reason 2>	**	**	**
...	**	**	**

Table 6: Pretreatment Characteristics at Baseline

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N=***	GHE N=***	Total
Gender			
Female	** (**)	** (**)	** (**)
Male	** (**)	** (**)	** (**)
Age (years)			
N	**	**	**
Median	**	**	**
Min - Max	** - **	** - **	** - **
< 65	** (**)	** (**)	** (**)
≥ 65	** (**)	** (**)	** (**)
ECOG Performance Status			
0	** (**)	** (**)	** (**)
1	** (**)	** (**)	** (**)
BMI			
≤27.5	**(**)	** (**)	** (**)
> 27.5	** (**)	** (**)	** (**)

Table 7: Disease and Treatment Characteristics at Baseline

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N=***	GHE N=***	Total
Years from First Diagnosis to Randomization			
N	**	**	**
Median	**	**	**
Min - Max	** _ **	** _ **	** _ **
Disease stage			
High risk stage II	**(**)	**(**)	** (**)
Stage III	** (**)	** (**)	** (**)
Clinical T-stage			
TX	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)
Histological grade (WHO)			
I	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)
Histology result of lymph nodes			
Positive	** (**)	** (**)	** (**)
Negative	** (**)	** (**)	** (**)
Number of positive lymph nodes			
0	** (**)	** (**)	** (**)
1+	** (**)	** (**)	** (**)
Tumour deposits			
Present	** (**)	** (**)	** (**)
Absent	** (**)	** (**)	** (**)
Years from complete resection of adenocarcinoma of the colon to Randomization			
N	**	**	**
Median	**	**	**
Min - Max	** _ **	** _ **	** _ **
Type of chemotherapy			
fluoropyrimidine-based	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)
Months from last dose of chemotherapy to randomization			
N	**	**	**
Median	**	**	**
Min - Max	** _ **	** _ **	** _ **

Table 8: Baseline Assessments

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N=***	GHE N=***	Total
Answers to Medical Screening Questionnaire (rParQ)			
All “no”	**(**)	**(**)	** (**)
At least one “yes”	** (**)	** (**)	** (**)
Answers to Leisure time exercise questionnaire (LTEQ) at pre-registration			
Number of minutes per week of light/mild exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of moderate exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of vigorous/strenuous exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of moderate to vigorous/strenuous exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Answers to Leisure time exercise questionnaire (LTEQ) at pre-diagnosis			
Number of minutes per week of light/mild exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of moderate exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of vigorous/strenuous exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of minutes per week of moderate to vigorous/strenuous exercise			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Results of submaximal exercise test at the 2nd last stage completed			

Speed (m/min)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
% grade (%)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Heart rate (bpm)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Results of submaximal exercise test at the last stage completed			
Speed (m/min)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
% grade (%)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Heart rate (bpm)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
VO2 max			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Anthropometrics			
Hip circumference			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Waist circumference			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Results of Senior's fitness test			
Number of repetitions of Sit to Stand			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Number of repetitions of Arm Curls			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
6-minute walk total (meters)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
Chair sit and reach result (cm)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)

Back Scratch result (cm)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)
8-foot Up and Go result (seconds)			
N	***	***	***
Mean (STD)	** (**)	** (**)	** (**)

Table 9: Baseline Symptoms/adverse events

Dataset: All Intention-To-Treat Participants on PA+GHE						
	Number of participants (%)					
	N=***					
	Grade					Any grade
	1	2	3	4	5	
Participants with any symptom/ adverse event at baseline	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with particular symptom/adverse event, within body system:						
Body System 1 ⁽¹⁾	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 2	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 3	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Body System 2 ⁽¹⁾						
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
	()	**(**)	**(**)	**(**)	**(**)	**(**)

(1) Participants may have more than one event within a body system

Note: Same table to be made for GHE.

Table 10: Baseline Hematology

Data set: All Randomized Participants			
	Number of Participants (%)		
	PA+GHE N = **	GHE N = **	Total N=**
Hemoglobin			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
WBC			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
Absolute granulocytes			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
Platelet			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
Blood transfusion or an erythropoietic stimulating agent within six weeks prior to date of hemoglobin			
Yes	** (**)	** (**)	** (**)
No	** (**)	** (**)	** (**)

⁽¹⁾ Not done or outside the 14-day window prior to randomization

Table 11: Baseline Biochemistry

Data set: All Randomized Participants			
	Number of Participants (%)		
	PA+GHE N = **	GHE N = **	Total N=**
Total bilirubin			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
Alkaline phosphatase			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
ALT			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
Serum Creatinine			
Grade 0	** (**)	** (**)	** (**)
Grade 1	** (**)	** (**)	** (**)
Grade 2	** (**)	** (**)	** (**)
Grade 3	** (**)	** (**)	** (**)
Grade 4	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)
CEA			
N	**	**	**
Mean (SD)	** (**)	** (**)	** (**)
Fasting glucose			
Normal	** (**)	** (**)	** (**)
High ⁽²⁾	** (**)	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)	** (**)

(1) Not done or outside the 14-day window prior to start of randomization

(2) High than upper lower limit

Table 12: Major Medical Problems at Baseline

Dataset: All Intention-To-Treat Participants			
	Number of participants (%)		
	PA+GHE N = ***	GHE N = ***	Total N=***
Participants with at least one past or current major medical problem	** (**)	** (**)	** (**)
Medical Problem ⁽¹⁾			
Problem 1	** (**)	** (**)	** (**)
...			

(1) participants may report more than one medical problem.

Table 13: Concomitant Medications at Baseline

Data set: All Randomized Participants			
	Number of participants (%)		
	PA+GHE N = **	GHE N = **	Total N=**
Any concomitant medication ⁽¹⁾			
No	** (**)	** (**)	** (**)
Yes	** (**)	** (**)	** (**)
Concomitant medication ⁽²⁾			
ASA	** (**)	** (**)	** (**)
...			

(1) Any medication taken within 14 days prior to randomization.

(2) Participants may report more than one concomitant medication.

Table 14: Tobacco Smoking History at Baseline

Dataset: All Intention-To-Treat Participants			
	Number of participants (%)		
	PA+GHE N = ***	GHE N = ***	Total N=***
Ever smoked any tobacco product			
No	** (**)	** (**)	** (**)
Yes	** (**)	** (**)	** (**)
Current smoker			
No	** (**)	** (**)	** (**)
Yes	** (**)	** (**)	** (**)
Current average cigarettes per day			
N	**	**	**
Median (Min-max)	** (** - **)	** (** - **)	** (** - **)
Years from quitting smoking to randomization			
N	**	**	**
Median (Min-max)	** (** - **)	** (** - **)	** (** - **)
Smoking history			
Pipe or cigar smoker only	** (**)	** (**)	** (**)
100 or fewer cigarettes during lifetime	** (**)	** (**)	** (**)
Greater than 100 cigarettes during lifetime	** (**)	** (**)	** (**)
Number of years smoked cigarettes			
N	**	**	**
Median (Min-max)	** (** - **)	** (** - **)	** (** - **)
Average number of cigarettes per day			
N	**	**	**
Median (Min-max)	** (** - **)	** (** - **)	** (** - **)
Pack Years			
N	**	**	**
Median (Min-max)	** (** - **)	** (** - **)	** (** - **)

Table 15: Compliance with Study Intervention

Data Set: All Treated Participants			
	Number of participants (%)		
	PA+GHE (N=***)	GHE (N=***)	
Received general health education materials			
Yes	** (**)	** (**)	
No	** (**)	** (**)	
Reason general health education materials were not provided			
Participant refused	** (**)	** (**)	
...			
Treated Participants on PA+GHE only (N=***)			
	Phase 1	Phase 2	Phase 3
Number of mandatory behavioral support sessions attended			
N	***	***	***
Median	**	**	**
Min-Max	** ** —	** ** —	** ** —
Number of mandatory supervised physical activity sessions attended			
N	***	NA	NA
Median	**	NA	NA
Min-Max	** ** —	NA	NA
Number of recommended supervised physical activity sessions attended			
N	***	***	***
Median	**	**	**
Min-Max	** ** —	** ** —	** ** —
Reason for non-attendance of mandatory or recommended sessions			
Recurrent disease	** (**)	** (**)	** (**)
...			
Number of additional behavioral support sessions attended			
N	***	***	***
Median	**	**	**
Min-Max	** ** —	** ** —	** ** —
Number of additional supervised physical activity sessions attended			
N	***	NA	NA
Median	**	NA	NA
Min-Max	** ** —	NA	NA
Reason for attendance of additional sessions			
Participant request	** (**)	** (**)	** (**)
...			
Received any ‘tool box’ resources		** (**)	
Tool box resource		** (**)	
Exercise videos		** (**)	
...		** (**)	

NA: Not Applicable

Table 16: Summary of DFS Events

Data set: All Randomized Participants		
	Number of Participants (%)	
	PA+GHE N=***	GHE N=***
Participants who had an event	*** (**)	*** (**)
Recurrence*	**	**
Local colon recurrence	**	**
Distant recurrence	**	**
Liver	**	**
Lung	**	**
Omental	**	**
Other	**	**
New primary malignancy*	**	**
Colon	**	**
Rectum	**	**
Other	**	**
Death without recurrence or new primary malignancy	**	**
Participants who were censored	*** (**)	*** (**)
Malignancy at baseline	**	**
Consent withdrawal	**	**
Lost to follow-up	**	**
Alive without recurrence or new primary malignancy	**	**

* Participants may have more than one site of recurrence and new primary malignancy.

Figure 2: Kaplan-Meier Curves for DFS

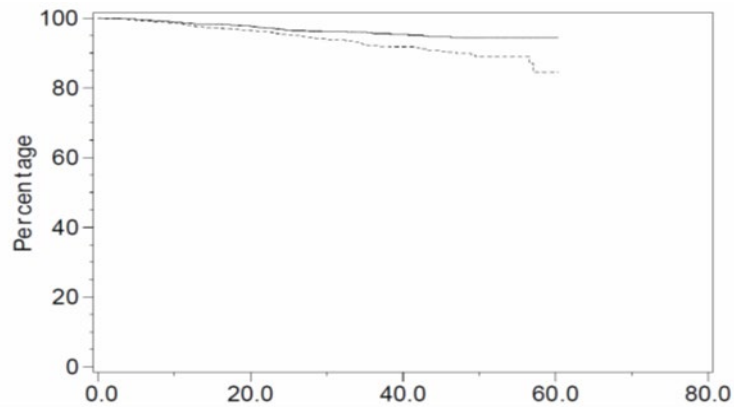


Table 17: Log rank and Cox Regression Model for Disease Free Survival

Data set: All Randomized Participants					
Treatment Arm/ Prognostic Factors at Baseline	Univariate Analysis ⁽¹⁾		Multivariate Analysis ⁽²⁾		
	3 year DFS (95% CI) (%)	Hazard Ratio ⁽⁴⁾ (95% CI)	Log-rank p-value ⁽¹⁾	Hazard Ratio ⁽⁴⁾ (95% C.I.)	P-vlaue from Cox regression
Treatment arm			0.***		0.***
<i>PA+GHE</i>	*** (***-***)	***		***	
<i>GHE</i>	*** (***-***)	(***,***)		(***,***)	
Age			0.***		0.***
< 60	*** (***-***)	NC ⁽³⁾		***	
≥ 60	*** (***-***)			(***,***)	
Sex			0.***		0.***
<i>Female</i>	*** (***-***)	NC		***	
<i>Male</i>	*** (***-***)			(***,***)	
Country			0.***		0.***
<i>Canada</i>	*** (***-***)	NC		***	
<i>Others</i>	*** (***-***)			(***,***)	
Chemotherapy			0.***		0.***
<i>FOLFOX</i>	*** (***-***)	NC		***	
<i>Others</i>	*** (***-***)			(***,***)	
Pre-diagnosis total minutes of moderate to rigorous/strenuous exercise			0.***		0.***
<150	*** (***-***)	NC		***	
≥150	*** (***-***)			(***,***)	
Baseline total MET hours/week			0.***		0.***
<10	*** (***-***)	NC		***	
≥10	*** (***-***)			(***,***)	
Baseline VO2 max			0.***		0.***
<30	*** (***-***)	NC		***	
≥30	*** (***-***)			(***,***)	
Baseline 6 minute walk total (meters)			0.***		0.***
<30	*** (***-***)	NC		***	
≥30	*** (***-***)			(***,***)	

(1) Stratified

(2) Stratified Cox regression with all factors included

(3) NC = not computed

(4) Hazard ratio of first category over second category

Table 18: Analysis for Disease Free Survival (DFS)

Factors	Value	PA+GHE		GHE		Hazard Ratio ⁽¹⁾ (95% C.I.)	P-value
		N	3-year DFS	N	3-year DFS		
Intention-to-treat participants	All	**	***	**	***	*** (*-**-*) ⁽²⁾	0.** ⁽²⁾
Age	<60	**	***	**	***	*** (*-**-*)	0.**
	≥60	**	***	**	***	*** (*-**-*)	0.**
Sex	Female	**	***	**	***	*** (*-**-*)	0.**
	Male	**	***	**	***	*** (*-**-*)	0.**
Country	Canada	**	***	**	***	*** (*-**-*)	0.**
	Australia	**	***	**	***	*** (*-**-*)	0.**
	Others	**	***	**	***	*** (*-**-*)	0.**
Disease stage	High risk stage II	**	***	**	***	*** (*-**-*)	0.**
	Stage III	**	***	**	***	*** (*-**-*)	0.**
ECOG Performance Status	0	**	***	**	***	*** (*-**-*)	0.**
	1	**	***	**	***	*** (*-**-*)	0.**
Chemotherapy Regimen	FOLFOX	**	***	**	***	*** (*-**-*)	0.**
	Others	**	***	**	***	*** (*-**-*)	0.**
BMI	<30	**	***	**	***	*** (*-**-*)	0.**
	≥30	**	***	**	***	*** (*-**-*)	0.**
Pre-diagnosis total minutes of moderate to vigorous/strenuous exercise	<150	**	***	**	***	*** (*-**-*)	0.**
	≥150	**	***	**	***	*** (*-**-*)	0.**
Baseline total MET Hours/week	<10	**	***	**	***	*** (*-**-*)	0.**
	≥10	**	***	**	***	*** (*-**-*)	0.**
Baseline VO2 max	<30	**	***	**	***	*** (*-**-*)	0.**
	≥30	**	***	**	***	*** (*-**-*)	0.**
Baseline 6-minute walk total (meters)	<500	**	***	**	***	*** (*-**-*)	0.**
	≥500	**	***	**	***	*** (*-**-*)	0.**

(1) Between PA+GHE and GHE;

(2) From two-side log-rank test stratified by stratification factors except correlative group at randomization.

Table 19: Summary of Deaths

Data set: All Randomized Participants		
	Number of Participants (%)	
	PA+GHE N=***	GHE N=***
Participants who died	*** (**)	*** (**)
Cause of death		
Colon cancer	**	**
Complication from protocol treatment (i.e. physical activity program or fitness testing)	**	**
Complication from another (i.e. not part of this protocol) treatment	**	**
Other primary malignancy	**	**
Other condition or circumstance	**	**
Participants who were censored	*** (**)	*** (**)
Consent withdraw	**	**
Lost to follow-up	**	**
Still Alive	**	**

Figure 3: Kaplan-Meier Curves for OS

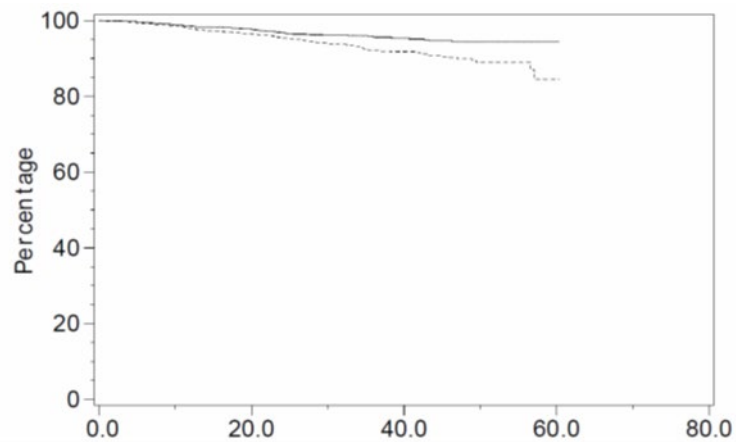


Table 20: Log rank and Cox Regression Model for Overall Survival

Data set: All Randomized Participants					
Treatment Arm/ Prognostic Factors at Baseline	Univariate Analysis ⁽¹⁾		Multivariate Analysis ⁽²⁾		
	3 year OS (95% CI) (%)	Hazard Ratio ⁽⁴⁾ (95% CI)	Log-rank p-value ⁽¹⁾	Hazard Ratio ⁽⁴⁾ (95% C.I.)	P-value from Cox regression
Treatment arm			0.***		0.***
<i>PA+GHE</i>	*** (***-***)	***		***	
<i>GHE</i>	*** (***-***)	(**.*,***.**)		(**.*,***.**)	
Age			0.***		0.***
< 60	*** (***-***)	NC ⁽³⁾		***	
≥ 60	*** (***-***)			(**.*,***.**)	
Sex			0.***		0.***
<i>Female</i>	*** (***-***)	NC		***	
<i>Male</i>	*** (***-***)			(**.*,***.**)	
Country			0.***		0.***
<i>Canada</i>	*** (***-***)	NC		***	
<i>Others</i>	*** (***-***)			(**.*,***.**)	
Chemotherapy			0.***		0.***
<i>FOLFOX</i>	*** (***-***)	NC		***	
<i>Others</i>	*** (***-***)			(**.*,***.**)	
Pre-diagnosis total minutes of moderate to rigorous/strenuous exercise			0.***		0.***
<150	*** (***-***)	NC		***	
≥150	*** (***-***)			(**.*,***.**)	
Baseline total MET hours/week			0.***		0.***
<10	*** (***-***)	NC		***	
≥10	*** (***-***)			(**.*,***.**)	
Baseline VO2 max			0.***		0.***
<30	*** (***-***)	NC		***	
≥30	*** (***-***)			(**.*,***.**)	
Baseline 6 minute walk total (meters)			0.***		0.***
<30	*** (***-***)	NC		***	
≥30	*** (***-***)			(**.*,***.**)	

(1) Stratified

(2) Stratified Cox regression with all factors included

(3) NC = not computed

(4) Hazard ratio of first category over second category

Table 21: Analysis for Overall Survival (OS)

Factors	Value	PA+GHE		GHE		Hazard Ratio(1) (95% C.I.)	P-value
		N	3-year OS	N	3-year OS		
Intention-to-treat participants	All	**	***	**	***	*** (*-**-*) ⁽²⁾	0.** ⁽²⁾
Age	<60	**	***	**	***	*** (*-**-*)	0.**
	≥60	**	***	**	***	*** (*-**-*)	0.**
Sex	Female	**	***	**	***	*** (*-**-*)	0.**
	Male	**	***	**	***	*** (*-**-*)	0.**
Country	Canada	**	***	**	***	*** (*-**-*)	0.**
	Australia	**	***	**	***	*** (*-**-*)	0.**
	Others	**	***	**	***	*** (*-**-*)	0.**
Disease stage	High risk stage II	**	***	**	***	*** (*-**-*)	0.**
	Stage III	**	***	**	***	*** (*-**-*)	0.**
ECOG Performance Status	0	**	***	**	***	*** (*-**-*)	0.**
	1	**	***	**	***	*** (*-**-*)	0.**
Chemotherapy Regimen	FOLFOX	**	***	**	***	*** (*-**-*)	0.**
	Others	**	***	**	***	*** (*-**-*)	0.**
BMI	<30	**	***	**	***	*** (*-**-*)	0.**
	≥30	**	***	**	***	*** (*-**-*)	0.**
Pre-diagnosis total minutes of moderate to vigorous/strenuous exercise	<150	**	***	**	***	*** (*-**-*)	0.**
	≥150	**	***	**	***	*** (*-**-*)	0.**
Baseline total MET Hours/week	<10	**	***	**	***	*** (*-**-*)	0.**
	≥10	**	***	**	***	*** (*-**-*)	0.**
Baseline VO2 max	<30	**	***	**	***	*** (*-**-*)	0.**
	≥30	**	***	**	***	*** (*-**-*)	0.**
Baseline 6-minute walk total (meters)	<500	**	***	**	***	*** (*-**-*)	0.**
	≥500	**	***	**	***	*** (*-**-*)	0.**

(1) Between PA+GHE and GHE;

(2) From two-side log-rank test stratified by stratification factors except correlative group at randomization

Table 22: Completion Rate of TPAQ Assessment by Treatment Arm

	PA+GHE N = ***	GHE N = ***
	N (%)	N (%)
6 month TPAQ assessment		
Yes	** (**)	** (**)
No	** (**)	** (**)
Reason		
Progressed before 6 months	** (**)	** (**)
...		
...		

Table 23: Total MET Hours/week (TMETHW) at Baseline and After Randomization

	PA+GHE	GHE	P-value*
Total MET Hours/week (TMETHW)			
Baseline			
N	***	***	
Mean	***	***	
STD	***	***	
6 months after randomization			
N	***	***	
Mean	***	***	
STD	***	***	
...			

Table 24: Fitness Tests: Submaximal Exercise Test

	PA+GHE	GHE	P Value**
VO2 Max			
Change from baseline to 6 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
...			

** Wilcoxon rank sum test

Table 25: Fitness Tests: Anthropometrics

	PA+GHE	GHE	P Value**
BMI			
Change from baseline to 6 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
...			
Hip circumference			
Change from baseline to 6 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
..			
Waist circumference			
Change from baseline to 6 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
..			

** Wilcoxon rank sum test

Table 26: Fitness Tests: Seniors' Fitness Test

	PA+GHE	GHE	P Value**
Number of repetitions of sit to stand			
Change from baseline to 6 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
...			
Number of repetitions of arm curls			
Change from baseline to 12 months after randomization			.**
N	***	***	
Mean	***	***	
STD	***	***	
...			
...			

** Wilcoxon rank sum test

Table 27: Acute Adverse Event

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 28: Severe Acute Adverse Events

Data set: All Treated Participants on PA+GHE				
	Number of participants (%) N=***			
	Worst grade			Any grade 3 or higher AE
	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)
Participants with AE within category				
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 29: Acute Adverse Event Related to Physical Activity Program

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 2	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 3	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Category 2 ⁽¹⁾	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)

(1) Participants may have more than one event within a category.

Table 30: Acute Adverse Event Related to Fitness Testing

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 2	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 3	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Category 2 ⁽¹⁾	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
Event 1	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)
...	**(**)	**(**)	**(**)	**(**)	**(**)	**(**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 31: Late Adverse Event

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 32: Severe Late Adverse Events

Data set: All Treated Participants on PA+GHE				
	Number of participants (%) N=***			
	Worst grade			Any grade 3 or higher AE
	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)
Participants with AE within category				
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 33: Late Adverse Event Related to Physical Activity Program

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE

Table 34: Late Adverse Event Related to Fitness Testing

Data set: All Treated Participants on PA+GHE						
	Number of participants (%) N=***					
	Worst grade					Any grade
	1	2	3	4	5	
Participants with any AE	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Participants with AE within category						
Category 1 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 2	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 3	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Category 2 ⁽¹⁾	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
Event 1	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)
...	** (**)	** (**)	** (**)	** (**)	** (**)	** (**)

(1) Participants may have more than one event within a category.

Note: Same table will be made for participants on GHE.

Table 35: Biochemistry: Worst Grade per Subject over Study

Data set: All Treated Participants		
	Number of participants (%)	
	PA+GHE N=***	GHE N=***
CEA		
Normal ⁽¹⁾	** (**)	** (**)
High ⁽²⁾	** (**)	** (**)
Not reported ⁽³⁾	** (**)	** (**)
Fasting glucose		
Normal ⁽¹⁾	** (**)	** (**)
High ⁽²⁾	** (**)	** (**)
Not reported ⁽¹⁾	** (**)	** (**)

(1) Lower than or equal to upper lower limit for all post-baseline assessments;

(2) High than upper lower limit at least one post-baseline assessment;

(3) No any value reported after baseline.

Table 36: Other Radiology Investigations

Data set: All Treated Participants		
	Number of participants (%)	
	PA+GHE N = ***	GHE N = ***
Any radiology investigation during intervention		
No	** (**)	** (**)
Yes	** (**)	** (**)
At least one malignant	** (**)	** (**)
All not malignant	** (**)	** (**)
Any radiology investigation 4 and 5 years from randomization but before the recurrence/new primary		
No	** (**)	** (**)
Yes	** (**)	** (**)
At least one abnormal and clinically significant	** (**)	** (**)
At least one abnormal but none clinically significant	** (**)	** (**)
All normal	** (**)	** (**)

Table 37: Hospitalization

Data set: All Treated Participants		
	Number of participants (%)	
	PA+GHE N = ***	GHE N = ***
Any hospitalization during intervention		
No	** (**)	** (**)
Yes	** (**)	** (**)
Any hospitalization during 4-5 year follow-up but before recurrence/new primary malignancy		
No	** (**)	** (**)
Yes	** (**)	** (**)
Any hospitalization after recurrence or new primary malignancy but within 5 years from randomization		
No	** (**)	** (**)
Yes	** (**)	** (**)

Table 38: Non-Protocol Anti-Cancer Treatment

Data set: All Treated participants		
	Number of participants (%)	
	PA+GHE N = ***	GHE N = ***
Any non-protocol anti-cancer treatment after recurrence or new primary malignancy but within 5 years from randomization	*** (**)	*** (**)
Treatment type		
Chemotherapy ⁽¹⁾	*** (**)	*** (**)
Treatment 1 ...	*** (**)	*** (**)
Radiotherapy ⁽¹⁾	*** (**)	*** (**)
...	*** (**)	*** (**)

(1) Participants could have more than one type of systemic and radiation treatment.

Table 39: Concomitant Medication

Data set: All Treated Participants		
	Number of participants (%)	
	PA+GHE N = ***	GHE N = ***
Any Concomitant medication during intervention		
No	** (**)	** (**)
Yes	** (**)	** (**)
Type of concomitant medication		
Medication 1	** (**)	** (**)
...		
Any Concomitant medication during 4-5 year follow-up but before recurrence/new primary malignancy		
No	** (**)	** (**)
Yes	** (**)	** (**)
Type of concomitant medication		
Medication 1	** (**)	** (**)
...		

Table 40: Tobacco Smoking

Data set: All Treated Participants		
	Number of participants (%)	
	PA+GHE N = ***	GHE N = ***
Ever smoked during intervention		
No	** (**)	** (**)
Yes	** (**)	** (**)
Average Number of cigarettes smoked per day during intervention		
N	***	***
Median	**	**
Min-Max	** **	** **
Ever quit smoking during intervention		
No	** (**)	** (**)
Yes	** (**)	** (**)
Ever smoked during 4-5 year follow-up but before recurrence/new primary malignancy		
No	** (**)	** (**)
Yes	** (**)	** (**)
Average Number of cigarettes smoked per day during 4-5 year follow-up but before recurrence/new primary malignancy		
N	***	***
Median	**	**
Min-Max	** **	** **
Ever quit smoking during 4-5 year follow-up but before recurrence/new primary malignancy		
No	** (**)	** (**)
Yes	** (**)	** (**)
Ever smoked after recurrence or new primary malignancy but within 5 years from randomization		
No	** (**)	** (**)
Yes	** (**)	** (**)
Average Number of cigarettes smoked per day after recurrence or new primary malignancy but within 5 years from randomization		
N	***	***
Median	**	**
Min-Max	** **	** **
Ever quit smoking after recurrence or new primary malignancy but within 5 years from randomization		
No	** (**)	** (**)
Yes	** (**)	** (**)

Table 41: Compliance Rate of PRO Assessments by Treatment Arms

	PA+GHE		GHE	
	Expected	Received (%)	Expected	Received (%)
SF-36				
Prior to randomization	***	** (**)	***	** (**)
6 months from randomization	***	** (**)	***	** (**)
12 months from randomization	***	** (**)	***	** (**)
18 months from randomization	***	** (**)	***	** (**)
24 months from randomization	***	** (**)	***	** (**)
36 months from randomization	***	** (**)	***	** (**)
4 years from randomization	***	** (**)	***	** (**)
5 years from randomization	***	** (**)	***	** (**)
FACIT-F				
Prior to randomization	***	** (**)	***	** (**)
6 months from randomization	***	** (**)	***	** (**)
12 months from randomization	***	** (**)	***	** (**)
18 months from randomization	***	** (**)	***	** (**)
24 months from randomization	***	** (**)	***	** (**)
36 months from randomization	***	** (**)	***	** (**)
4 years from randomization	***	** (**)	***	** (**)
5 years from randomization	***	** (**)	***	** (**)
PSQI				
Prior to randomization	***	** (**)	***	** (**)
6 months from randomization	***	** (**)	***	** (**)
12 months from randomization	***	** (**)	***	** (**)
18 months from randomization	***	** (**)	***	** (**)
24 months from randomization	***	** (**)	***	** (**)
36 months from randomization	***	** (**)	***	** (**)
4 years from randomization	***	** (**)	***	** (**)
5 years from randomization	***	** (**)	***	** (**)
HADS				
Prior to randomization	***	** (**)	***	** (**)
6 months from randomization	***	** (**)	***	** (**)
12 months from randomization	***	** (**)	***	** (**)
18 months from randomization	***	** (**)	***	** (**)
24 months from randomization	***	** (**)	***	** (**)
36 months from randomization	***	** (**)	***	** (**)
4 years from randomization	***	** (**)	***	** (**)
5 years from randomization	***	** (**)	***	** (**)

Table 42: Summary of Baseline PRO Scale Scores

	PA+GHE	GHE	P value*
SF-36			
Physical health			
N	***	***	0.***
Mean (STD)	*** (*.*)	*** (*.*)	
...	
FACIT-F			
PWB			
N	***	***	0.***
Mean (STD)	*** (*.*)	*** (*.*)	
...	
PSQI			
Subjective sleep quality			
N	**	**	0.***
Mean (STD)	*** (*.*)	*** (*.*)	
....	
HADS			
HADS-A			
N	**	**	0.***
Mean (STD)	*** (*.*)	*** (*.*)	
...	

* From Wilcoxon test.

Table 43: Summary of Change Scores from Baseline for PRO Scale/Domain/Item at Each Time of Assessment

Scale/Domain/Item*	PA+GHE	GHE	P Value**
6 Months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
12 Months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
18 months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
24 months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
30 months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
36 months from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
4 years from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	
5 years from randomization			.**
N	***	***	
Mean (STD)	** (**)	** (**)	

* Table will be provided for each scale/domain/item. ** From Wilcoxon rank sum test

Table 44: Summary of Analysis from Linear Mixed Models for Change Scores

	Estimate	Standard Error	P value
SF-36			
Physical health			
Treatment	***	***	0.**
Month	***	***	0.**
Treatment*Month	***	***	0.**
...			
FACIT-F			
PWB			
Treatment	***	***	0.**
Month	***	***	0.**
Treatment*Month	***	***	0.**
...			
PSQI			
Subjective sleep quality			
Treatment	***	***	0.**
Month	***	***	0.**
Treatment*Month	***	***	0.**
...			
HADS			
HADS-A			
Treatment	***	***	0.**
Week	***	***	0.**
Treatment*Month	***	***	0.**
...			