Method

Participants

Eligibility criteria.

Inclusion criteria were: (a) aged between 18 and 80 years old; (b) have received a non-metastatic cancer diagnosis for which the adjuvant treatment (except hormone therapy) ended within the past 6 months (given the expected and demonstrated benefits of physical activity during the rehabilitation period; (Spence, Heesch, & Brown, 2009); (c) have insomnia symptoms, as indicated by a score of 8 or greater on the ISI (M. H. Savard et al., 2005); (d) not regularly exercising, i.e., less than 90 minutes of moderate to vigorous intensity aerobic EX or less than 150 minutes of low-intensity EX per week; (e) be able to read and understand French; and (f) have a physician's written permission to engage in an aerobic EX program. Exclusion criteria were: (a) having a sleep disorder other than insomnia (e.g., obstructive sleep apnea); (b) having a severe medical condition that could interfere with exercising (e.g., pulmonary disease, coronary artery disease) or a contraindication to engage in an EX program (e.g., musculoskeletal disease); and (c) having a severe psychiatric disorder (e.g., schizophrenia, bipolar disorder) or severe cognitive impairments (e.g., Alzheimer's disease) as reported by the patient.

Recruitment.

Participants were recruited between June 2012 and August 2014. Forty-one participants, male and female, were recruited at L'Hôtel-Dieu de Québec (CHU de Québec-Université Laval). Participants were mainly approached by a research assistant at the radio-oncology department, who briefly introduced the study and collected contact information of those interested in the project. Other participants were recruited through a letter that was handed by the radio-oncology team to patients who were about to end their radiation therapy. Then, a phone screening was conducted to assess the eligibility criteria and explain the project in detail. The PRIME-MD questionnaire was administered over the phone to assess the presence of a severe psychological disorder (exclusion criterion) (Spitzer et al.,

MCours.com

1994). At this time, participants were also informed that they needed to ask their physicians to deliver a medical clearance in order for them to perform an aerobic EX program of a moderate intensity. Eligible and interested patients then received by mail the written consent form which they were invited to sign, along with the first battery of questionnaires. Figure 1 shows the participants' flowchart and detailed reasons for exclusions. The overall participation rate was 63% (41/65 of eligible patients). The main reason for exclusion was practicing more than 90 minutes of moderate to vigorous intensity aerobic EX per week at baseline (n= 31, representing 25.4% of all exclusions). Six participants (30%) assigned to the EX group and 2 (10%) assigned to the CBT-I group dropped out during the course of the study, for a total dropout rate of 20%. This study was approved by the research ethics committee of CHU de Québec-Université Laval.

Study Design.

This pilot RCT included two experimental conditions (ratio 1:1): self-administered CBT-I (n=21) and home-based exercise (EX) intervention (n=20). Participants were assessed at pre- and post-treatment (about 6 weeks after baseline), as well as at 3 and 6-month follow-ups. A non-inferiority study design was used to investigate the main goal and a superiority design was used for the secondary goals.

Randomization and allocation concealment.

The randomization sequence was prepared by a biostatistician using a random permuted-block (size = 6) procedure with SAS 9.3 (SAS Institute, 2011). Investigators and research assistants were blind to the allocation sequence that was concealed in opaque and sealed envelopes until participants completed all baseline measurements. The graduate student in charge of the the project (JM) or a research assistant opened the envelope in the presence of the participant when his/her eligibility was confirmed (following the physical assessment).

Sample size justification and power analyses.

A priori power analyses were conducted to determine the sample size needed to test the study hypotheses. More specifically, a sensitivity analysis was performed based on the

recommendations of (Hwang & Morikawa, 1999). With a standard 80% power and an alpha level of 5% unilateral (given the unilateral direction of the non-inferiority hypothesis), the analysis showed that a sample of 20 participants per group (total N = 40) would have a sufficient power to test the non-inferiority hypothesis with a clinical margin of 3.5 units on the ISI. For the purpose of the study, a clinical margin of 4 points on the ISI was used, which corresponds to half of the score that is considered to be a clinically significant change (C. M. Morin, Belleville, Belanger, & Ivers, 2011). This appears to be a conservative clinical margin given the general recommendation to use the smallest change value that may be clinically significant (Piaggio, Elbourne, Altman, Pocock, & Evans, 2006). A power (sensitivity) analysis was also performed for superiority hypotheses, for an experimental design of 2 groups x 4 times with a total sample of 40 participants expected. The results showed that a total sample size of 40 would detect a minimal effect size of d = 0.29 for the group X time interaction at a standard power of 80%, which corresponds to a small effect (Cohen, 1988).

Procedure

Pretreatment assessment.

Eligible participants were sent by courier a battery of self-report scales to complete at home, containing a 2-week daily sleep diary and an actigraphic recorder to be worn during 7 consecutive 24-hour periods. After the completion of measures, participants had their physical fitness evaluated by a professional kinesiologist at the University Laval Kinesiology Clinic. None of the patients were excluded at that point. The graduate student in charge of the project (JM) then met the participant at the clinic to proceed with the randomization. When the participant was assigned to CBT-I, the treatment material and relevant explanations were provided and, when the participant was allocated to the home-based EX program, he/she again met the kinesiologist to develop a personalized 6-week exercise program.

Intervention Phase.

The intervention phase lasted 6 weeks. During that phase, all participants completed a daily sleep diary and a daily exercise diary. Moreover, a weekly phone call from a

member of our research team took place to know how participants were doing with the intervention they were assigned to, to answer their questions, if needed, and to enhance treatment adherence (e.g., by reinforcing the importance of pursuing treatment even in the absence of perceived improvements). Participants were also asked to complete the *Treatment Perception Questionnaire* after the first and fifth intervention week.

CBT-I intervention.

The self-administered treatment package is composed of a 60-min video (DVD format) and six booklets. Each week, participants had to watch a video segment (5-20 min each) and read a booklet. CBT-I includes behavioral (i.e., stimulus control therapy, sleep restriction), cognitive (i.e., cognitive restructuring), and educational (i.e., sleep hygiene) strategies. More details about this video-based intervention can be found elsewhere (J. Savard, Villa, Simard, Ivers, & Morin, 2011).

Exercise intervention.

The EX program was based on the recommendations of the American College of Sport Medicine (ACSM) (Schmitz et al., 2010). The general goal was to do 3 to 5 20-30 min sessions per week of aerobic exercise with a gradual increase over time until 150 minutes of EX per week was attained. However, the programs were individualized based on the participant's initial physical condition. For instance, for participants who were fairly active at baseline (i.e., near the maximal cutoff allowed for inclusion in study), the final objective was to increase their practice by 60 minutes by the end of the sixth week. Exercises had to be of at least a moderate intensity, thus corresponding to a perceived exertion rate between 3 and 5 on the modified Borg scale (G. Borg, 1998). Participants were free to choose the type of aerobic EX they wanted in order to maximize their motivation to engage in and adhere to their EX program, including brisk walking, jogging, swimming or a combination of different aerobic exercises. A phone follow-up with the kinesiologist was also conducted around the third intervention week to identify difficulties encountered and to make changes to the EX prescription if necessary.

Post-treatment and follow-up assessments.

At post-treatment, as well as at 3- and 6-month follow-ups, participants completed the same battery of self-report scales at home (including both sleep and exercise diaries for 14 days) and were asked to wear the actigraphic recorder at each time assessment for 7 continuous 24-hour periods. Participants received $20 (Canadian) for each time point completed.

Measures.

Unless otherwise specified, French (Canadian) versions of measures used have been empirically validated or developed by the authors of the original version.

Primary outcome measures (sleep and exercise).

Insomnia Severity Index (ISI) (C. M. Morin, 1993). The ISI includes seven items which evaluate, for the previous two weeks, the perceived severity of difficulties falling asleep, difficulties maintaining sleep and early morning awakenings, as well as the degree of dissatisfaction with current sleep, the degree to which sleep difficulties interfere with daytime functioning, the degree to which the deterioration of functioning related to the sleep problem is noticeable by others, and the level of distress or worry caused by the sleep difficulties (rated on a scale from 0 ["not at all"] to 4 ["very much"]). The ISI was empirically validated among cancer patients and a score of 8 or greater is used to detect clinically significant insomnia (M. H. Savard et al., 2005).

Pittsburgh Sleep Quality Index (PSQI) (Buysse et al., 1989). This questionnaire was developed to assess the subjective sleep quality of the previous month on seven components: sleep latency, sleep duration, daytime dysfunction, sleep disorders, use of sleep medication, habitual sleep efficiency and subjective sleep quality. The questionnaire consists of 19 items using a Likert scale ranging from 0 ("no difficulty") to 3 ("severe difficulties"). The total score ranges from 0 to 21 with a higher score indicating a poorer sleep. A total score > 5 suggests the presence of significant sleep difficulties with a sensitivity of 89.6% and a specificity of 86.5% for distinguishing between good and poor sleepers (kappa = .75, $p < 0.001$). The internal consistency ($\alpha = .83$) and the test-retest reliability (on average 29 days later, $r = .83$) of the scale was supported in the general population. Support for the validity and good psychometric properties in cancer patients is

also available (Beck, Schwartz, Towsley, Dudley, & Barsevick, 2004).

Sleep diary. The following variables were derived from the sleep diary: sleep onset latency (SOL; time from lights out to sleep onset), number of nocturnal awakenings, wake after sleep onset (WASO; time spent awake after initial sleep), total wake time (TWT; sum of all awakenings, from lights out until the last awakening), total sleep time (TST; sum of all sleep periods from initial sleep until last awakening), sleep efficiency (SE; total sleep time divided by total time in bed), and hypnotic usage.

Actigraphy. The Actiwatch-2® (Philips, Respironics, Andover, MA) is a small, waterproof, non-intrusive actigraphy device that is worn on the wrist. Following usual recommendations, patients were instructed to wear the actigraphic recorder on their non-dominant hand for 7 consecutive 24-hr periods at each time assessment (Ancoli-Israel et al., 2003). By calculating orientation and movement, the Actiwatch estimates sleep-wake activity and provides an objective measure of the same sleep parameters as the sleep diary. In the current study, actigraphic data were also used to objectively measure the participants' physical activity level. The validity of actigraphy has been demonstrated, both for the evaluation of sleep in insomnia patients (Sanchez-Ortuno, Edinger, Means, & Almirall, 2010), as well as the level of physical activity (Chen et al., 2003).

Exercise diary. An EX diary was developed specifically for the needs of the current study. Participants of the two groups completed this measure daily during the 6-week intervention phase, as well as for 2 consecutive weeks at post-treatment, 3- and 6-month follow-ups. Specifically, they were asked to document, for each day, the type of exercise they performed, its duration and intensity (perceived effort according to the modified Borg scale). This allowed us to calculate the frequency and the total duration of EX performed per week, and also to assess patients' adherence to the EX program.

Godin Leisure-Time Exercise (GLTEQ) (Godin & Shephard, 1985). The GLTEQ is a 4-item questionnaire that evaluates the habitual physical activity performed, during free time in a typical weekly period, of high, moderate and low intensity, separately. A correlation of r = .35 was obtained between the reported frequency of high intensity exercise on this questionnaire and the VO2 max. The 2-week test-retest reliability of this

instrument is high (r = 0.74 to 0.80). The GLTEQ is commonly used in oncology settings (Amireault, Godin, Lacombe, & Sabiston, 2015)

Secondary outcome measures.

Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983). This is a 14-item questionnaire divided into two sub-scales: depression (HADS-D: 7 items) and anxiety (HADS-A: 7 items) that are rated on a scale from 0 to 3. The HADS contains no somatic items that may be confounded with symptoms of the physical illness (J. Savard, Laberge, Gauthier, Ivers, & Bergeron, 1998).

Fatigue Symptoms Inventory (FSI) (Hann et al., 1998). The FSI is a multidimensional questionnaire of fatigue that was developed and validated in people with cancer. It contains three subscales, for a total of 13 items. The first four questions assess the intensity of fatigue, the following seven evaluate the impact of fatigue on different aspects of quality of life (general activity level, ability to dress and wash oneself, work, concentration, social life, vitality, mood) and the last two items quantify the duration of the fatigue experienced. Each item is rated on a Likert scale of 11 points, ranging from 0 to 10, a higher score indicating worse fatigue or interference, or a longer duration of fatigue. Since this tool was not available in French, a house translation was used. The FSI has good psychometric properties, such as an excellent internal consistency (α = 0.92 to 0.94) and a good convergent validity with the *Profile of Mood States - Fatigue Scale* (POMS-F) (Hann, Denniston, & Baker, 2000).

European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC-QOL) (Aaronson et al., 1993). The French version of this questionnaire was developed by the authors of the original English version. This is a 13-item questionnaire, rated on a scale from 0 to 3. Scores are transformed to give a score ranging from 0 to 100 where a higher score corresponds to a better quality of life. All scales have a good internal consistency ($\alpha \geq .70$) and correlations of .40 or greater were observed between all items and their respective scale (Aaronson et al., 1993).

Other measures included.

Treatment Perception Questionnaire (J. Savard, Villa, et al., 2011). This questionnaire, developed by our research team, assesses participants' satisfaction with the content of the booklets and the video containing CBT-I (13 items), the contact person (8 items), and global satisfaction with the intervention (e.g., interest towards the proposed strategies, global sleep improvement; 7 items). A similar questionnaire was developed for the EX intervention and documented participants' global satisfaction with the program (9 items) and the kinesiologist (7 items). The TPQ was administered at post-treatment only.

Treatment Credibility and Expectancies for Improvement scale. The questionnaire was adapted by our team from the questionnaire developed by Borkovec and Nau (Borkovec & Nau, 1972). It includes 5 questions rated on a Likert scale ranging from 0 ("not at all") to 3 ("very much") and evaluates participants' therapeutic expectancies and the credibility they give to the treatment proposed. Before they were randomized, all patients completed two versions of the questionnaire, one for each treatment. The version corresponding to their group allocation was completed again after the fifth intervention week. One item was added at pre-treatment to ask patients if they had a preference between the two interventions on a scale from -3 ("strong preference for EX"), 0 ("no preference") to 3 ("strong preference for CBT-I").

Demographic and medical data. Demographic data collected on this questionnaire include age, marital status, level of education, socioeconomic status, tobacco, alcohol and caffeine consumption, time since cancer diagnosis, cancer type and treatments received, presence of comorbid psychological or medical conditions, and medication use. The participants' medical records were consulted to corroborate cancer-related data and document the cancer stage.

Statistical analyses.

Data were entered by JM and verified independently by another research assistant. Examination of missing data, outliers and distributions was performed using standard procedures. No missing data was imputed. Analyses were performed using an intent-to-treat approach. All analyses were conducted using the SAS 9.4 software (SAS Institute, Cary, NC, USA) and the alpha level was set at 5%, two-tailed (except for the non-

inferiority analysis, where the alpha level was set at 5% one-tailed). In order to assess changes on study variables within and between conditions, 2 Groups (CBT-I vs. EX) X 4 Times (pre- and post-treatment, 3- and 6-month follow-ups) ANOVA mixed model analyses were completed (SASInstituteInc., 2011). Simple effects were conducted to test temporal changes between time points. Given the exploratory nature of these analyses, no statistical correction for multiple tests was done. Based on the suggestions of (Frigon & Laurencelle, 1993), it was planned to statistically control only for variables showing a moderate association $r \geq .30$ (Cohen, 1988) with at least two of the main dependent variables (ISI, PSQI or SE%). Demographics, cancer characteristics and treatments, medication use, psychological difficulties (past and current) and health-related data were investigated as potential covariates. Since no one was found to meet the above-described criterion, no covariate was included in the analyses. Three categories of EX intensity were calculated using actigraphic data: low intensity (0 to 1534 counts per minute), moderate (1534 to 3959 counts per minute) and vigorous (≥ 3960 counts per minute (Colley & Tremblay, 2011). Finally, effect sizes (Cohen's *d*) for time effects were calculated as the raw difference between assessment points (e.g., pre- vs. post-treatment) divided by the RMSE of the mixed models for all study outcomes.

Results

Participants' Characteristics

All participants were French-Canadian and Caucasian. The participants' demographic and medical characteristics are presented in Table 1. The mean age of the sample was 57.1 years old, and it was mainly composed of women with breast cancer (n = 22; 54%). A majority of participants were married or in a common-law relationship (n = 27; 66%) and had a university degree (n = 24; 59%). All participants received radiation therapy, and a large proportion of them underwent a surgery beforehand (n = 32; 78%). More than a third of the participants also received chemotherapy (n = 15; 37%). Finally, 46% of the sample were hypnotic users at baseline, with an average frequency of use of 3.7 nights per week. No significant between-groups differences were found on any demographic and medical variable at baseline.

Non-inferiority Analyses

The CBT-I group showed a reduction of -4.56 points on the ISI from pre- to post-treatment (from 14.8 to 10.3; *d* = -0.78) as compared to a decrease of -3.91 for the EX group (from 16.0 to 12.1, *d* = -0.67). The between-groups difference on ISI scores change (-0.65) is small but the confidence interval (-4.87 – 3.57) exceeds the established clinical margin of 4 points, thus indicating that the EX intervention was significantly inferior to CBT-I in improving insomnia symptoms as assessed with the ISI.

However, the analyses conducted at follow-up, indicated a non-significant inferiority of EX in producing a reduction of ISI scores. More precisely, the CBT-I group showed a reduction of -2.22 points from pre- to the 3-month follow-up (14.8 to 12.59; *d* = -0.38) as compared to a decrease of -5.58 for the EX group (from 16.0 to 10.42, *d* = -0.96; between-groups difference of 3.36 with a confidence interval of -0.94 – 7.67). At the 6-month follow-up, the CBT-I group showed a reduction of -3.04 points on the ISI relative to pre-treatment (14.8 to 11.77; *d* = -0.52) as compared to a decrease of -6.63 for the EX group (from 16.0 to 9.37, *d* = -1.14; between-groups difference of 3.59 with a confidence interval of -0.24 – 7.42). Since both confidence intervals fall within the clinical margin of 4 points, it is possible to conclude that EX was non-inferior to CBT-I in reducing ISI scores at follow-up.

Treatment effects on sleep parameters at post-treatment and follow-ups (superiority analyses)

Table 2 and Figure 2 show mean scores obtained on subjective sleep parameters at each time point, in both groups. No significant group X time interaction was found on any variable, except on PSQI scores. However, simple effects revealed no significant between-groups difference on this outcome at any time assessment. The interaction was marginally significant on ISI scores (p = .06) but, again, simple effects on between-group differences were all non-significant. Significant main time effects were obtained for all variables (all ps < .05) and simple time effects were significant in both groups on all variables, with the exception of SOL and TIB for EX and EMA and TIB for CBT-I. Pre vs. post-treatment effect sizes of a moderate magnitude were observed (ds > 0.70) for SOL, WASO, TWT and

MCours.com

SE for both interventions. Moreover, pre vs. post-treatment effect sizes obtained on ISI and PSQI scores were superior in CBT-I patients (ISI: d = -0.78 for CBT-I vs. d = -0.67 for EX; PSQI: d = -1.10 for CBT-I vs. d = -0.54 for EX). Finally, the main group effects were not significant for any outcome and the only simple between-groups significant difference was at post-treatment on the total TIB which was shorter in the CBT-I group (p = .03). Interestingly, SOL decreased under the 30-min clinical threshold at post-treatment in both groups, an improvement that was sustained at both follow-ups. In addition, SE increased by more than 8% in both groups at post-treatment, although it was still falling under the 85% threshold. SE continued to improve at follow-up reaching 84% and 85% for CBT-I and EX groups, respectively, at the 6-month evaluation. On the other hand, mean ISI and PSQI scores remained above the clinical cut-off at each time point, in both groups. Regarding the weekly mean days of consumption of prescribed hypnotic and anxiolytic medications during the course of study, no significant group X time interaction was found, nor time or group effects (all ps > .05). However, CBT-I patients more importantly reduced their utilization of these medications from pre- to post-treatment as indicated by the effect size that was superior in CBT-I (d = -0.35 for CBT-I vs. d = 0.03 for EX).

A different pattern of results was obtained on objective sleep variables as assessed with actigraphy (see Table 3). Specifically, none of the group X time interactions, main time effect and main group effect was significant. The only significant time effects from pre- to post-treatment were obtained in the CBT-I group only on EMA (reduction of approximately 5 min; p = .01) and WASO (reduction of approximately 7 min; p < .01). It is noteworthy, that SE was below the clinical threshold of 85% in participants of both groups, at each time point. Effect sizes of pre- vs. post-treatment differences were all of a small magnitude (all ds ≤ 0.40), with the largest effect size obtained on TIB in the CBT-I group.

Treatment effects on secondary outcomes at post-treatment and follow-ups

Table 4 shows mean scores obtained on secondary variables for each group, at each time point. A significant group X time interaction was found on quality of life scores only, and simple effects indicated that the EX group showed a significantly greater quality of life as compared to the CBT-I group at the 6-month follow-up only (p = .04). Also, the pre- vs. post-treatment effect size obtained on quality of life was superior in the EX group (d = 0.79

for EX vs. $d = 0.39$ for CBT-I), as well as post-treatment vs. 6-month follow-up effect sizes ($d = 0.74$ for EX vs. $d = 0.13$ for CBT-I). Significant main time effects were obtained on depressive symptoms ($p < .01$), fatigue ($p < .05$) and quality of life ($p < .01$). Simple time effects revealed a significant improvement of these three variables from pre- to post-treatment, with moderate effect sizes (ds = -0.51 – 0.79). They also indicated a significant reduction of depressive and fatigue symptoms in the CBT-I group, with small to moderate effect sizes (ds = -0.35 to -0.59). The main group effect was not significant for any variable.

Complementary Analyses

Integrity of the interventions.

To assess the integrity of the CBT-I intervention and the possible contamination of the EX group, two behavioral indices were calculated based on sleep diary data: the regularity of the sleep schedule[1] and the number of naps per day. This provided measures of adherence to two stimulus control instructions, that is to keep a sleep/wake schedule as regular as possible and to reduce the number of naps. No group X time interaction was found on any of these indices (see Table 5). Moreover, no significant time effect was found between pre- and post-treatment in CBT-I participants, nor in EX patients, on the regularity of the sleep schedule, while a significant reduction of napping was observed in both intervention groups (significant time effects, ps < .05).

To evaluate the integrity of the EX intervention in increasing physical activity and the possible contamination of the CBT-I group, two indices derived from the EX daily diary were used: 1) the proportion of participants reaching the recommendation of 150 minutes of moderate-vigorous aerobic exercise per week at the end of the program (week 5 or 6); and 2) the proportion of participants who increased their practice by 60 minutes or greater of moderate-vigorous physical activity between intervention week 1 and week 5 or 6. The mean weekly exercise duration attained during the 6-week intervention phase, by each intensity category, is presented in Table 6. Although there was no significant time

[1] The regularity of the sleep schedule was estimated as the standard deviation of bedtime/arising time for the 14 diary days, with the expectation that the variability of the sleep schedule would be reduced at post-treatment in CBT-I, but not in the EX group.

effect and no significant group X time interaction (all *p*s > .05), there was a general increase of self-reported time spent in moderate exercising until week 3 or 4, followed by a decrease in both groups. Chi-square tests were conducted to assess between-groups differences on the attainment of both exercise recommendations at the end of the program. Results indicated that 60.0% of EX participants reached the recommendation of ≥150 minutes of moderate-vigorous aerobic exercise, as compared to 42.9% of CBT-I participants, a difference that was not significant, $\chi^2 = 1.2$, $p = .27$. In addition, 55.0% of EX participants increased their exercise practice by ≥ 60 minutes of moderate-vigorous intensity, as compared to 38.1% of CBT-I participants, again a difference that was not significant, $\chi^2 = 1.18$, $p = .28$. Regarding fitness outcomes, the EX group showed a significant improvement in their cardiorespiratory capacity (estimate VO2peak) at post-treatment, F (1,31) = 4.3, $p = .046$, while no significant time effect was found in CBT-I patients, F (1,31) = 0.02, $p = .89$. However, the group X time interaction was not significant, F (1,31) = 2.5, $p = .12$. More precisely, an increase of 6% on VO2peak was recorded in the EX group at post-treatment (from 26.91 to 28.65) comparatively to a decrease of 0.4% (from 26.21 to 26.10) in CBT-I patients.

Moderating role of treatment preferences at baseline.

At pre-treatment, 60% ($n = 21$) of participants indicated that they had a preference (between "moderate" and "high") for the EX intervention, whereas 6 participants (17%) only had a preference (between "low" and "high") for CBT-I (see Table 7) and 8 participants (23%) had no preference. More patients assigned to the CBT-I group than to the EX condition (12 participants versus 9) initially indicated they would have preferred an EX program. To evaluate the possible contribution (moderator effect) of being assigned to the condition for which the participant expressed having at least a low, moderate or high preference (i.e., being *matched*), a mixed model matched X condition X time ANOVA was performed on the main sleep outcomes (ISI, PSQI and ES) at pre- and post-treatment. Participants with no preference were considered matched by default.

No significant two-way moderator interaction (matched X time for each condition) was found in any group (*p*s > .05; see Table 8 and Figure 3) on the ISI, the PSQI and SE. However, a significant pre vs. post-treatment time effect in the direction of an improvement

was found in the CBT-I *mismatched* subgroup on the three sleep variables while an improvement was found only on PSQI for the CBT-I *matched* subgroup. A different profile of results was observed in EX patients in which no significant pre- vs. post-treatment time effect on any outcome was found in the *mismatched* subgroup. The EX *matched* subgroup showed a significant improvement on ISI scores and SE values.

Given that 51% of the sample reported having a preference for the EX intervention, we further looked at the contribution of being *matched* to the preferred treatment on the attainment of the two EX recommendations described above. No significant moderating role in the attainment of these recommendations was obtained (data not shown; all *p*s > .05). For example, in the CBT-I group, 42.8% (3/7) of *matched* participants reached the recommendation of 150 minutes of moderate-vigorous aerobic exercise per week at the end of the program compared to 50.0% (6/12) of *mismatched* patients (patients who initially preferred EX), $\chi^2 = 0.09$, $p = .76$. Similarly, in the EX group, 61.5% (8/13) of *matched* participants reached this recommendation as compared to 66.7% (2/3) of *mismatched* (patients who initially preferred CBT-I) participants, $\chi^2 = 0.03$, $p = .87$.

Evolution of treatment expectancies for improvement

A significant group X time interaction was found on treatment expectancies scores: $F(1, 31) = 12.25$, $p = .001$, thus indicating that scores obtained on this questionnaire had a different evolution across groups. While CBT-I participants reported an increase in treatment credibility and expectancies scores from baseline (9.9) to the fifth week of intervention (11.3), EX participants obtained decreased scores (from 12.0 to 9.6).

Discussion

The main goal of this non-inferiority RCT, conducted in patients with mixed cancer sites, was to compare the efficacy of a home-based EX intervention to a self-administered CBT-I, considered as the standard, for reducing insomnia severity at post-treatment and at 3- and 6-month follow-ups. Other study goals were to compare, using standard superiority analyses, the effects of these two non-pharmacological home-based interventions on sleep measured subjectively and objectively and other variables that are often associated with insomnia (i.e., anxiety, depression, fatigue and quality of life) at post-treatment and 3- and

6-month follow-ups. Results are mixed and only partially support the initial study hypotheses.

More specifically, results of the non-inferiority analysis showed that EX was significantly inferior to CBT-I in reducing insomnia symptoms at post-treatment as measured with the ISI. Although the between-groups difference in the reduction of mean ISI scores at post-treatment was rather small (- 0.65), the confidence interval exceeded the established clinical margin. This conclusion is consistent with results of the superiority analyses which revealed a time effect of a greater effect size for CBT-I on that variable (pre- vs. post-treatment time effect; d = -0.78 for CBT-I vs. d = -0.67 for EX). However, results of the non-inferiority analysis at follow-up revealed that EX was not significantly inferior to CBT-I in reducing ISI scores. Together, these findings are contrary to our hypotheses which predicted non-inferiority of EX at post-treatment and inferiority at follow-up. They are however consistent with those obtained in a previous non-inferiority trial in cancer in which a mindfulness-based stress reduction (MBSR) intervention was found to be statistically inferior to CBT-I at post-treatment but was non-inferior at a 5-month follow-up (Garland, Carlson, et al., 2014). On the other hand, in another non-inferiority trial conducted in cancer patients, a 12-week Tai Chi Chih (TCC) intervention was found to be non-inferior to CBT-I in leading to an insomnia treatment response at post-treatment (defined by a decrease of $\geq$ 5 points on the PSQI) and this non-inferiority was sustained at 6- and 15-month follow-ups (Irwin et al., 2017).

Results of superiority analyses comparing the two groups on subjective and objective measures of sleep were partly inconsistent with results of the non-inferiority analyses. First, the lack of significant effects on actigraphic data suggested that both interventions had a modest impact on objective sleep of participants. Second, no significant between-groups differences on the improvement of subjective sleep variables was observed either at post-treatment or at follow-up, as reflected by a lack of significant group-by-time interaction on all variables (except for PSQI but no significant simple effects were detected). However, the significant main time effect found on all subjective variables suggests that both interventions significantly improved subjective sleep, as well as the

similar medium to large time effect sizes obtained on many variables in both groups (*d*s > 0.50 for ISI, PSQI, SOL, WASO, TWT and SE).

Nevertheless, scores obtained on self-reported questionnaires (ISI and PSQI) indicated the persistence of clinically significant sleep difficulties in both groups at post-treatment and at follow-up. In fact, remission rates of insomnia at post-treatment, defined by an ISI score < 8, were only 30.7% in the CBT-I group and 35.2% in the EX group, a difference that was not significant. Rates were even weaker when using the PSQI (score ≤ 5), attaining only 9.9% for CBT-I and 11.3% for EX group.

A first possible explanation for these mitigated results is the fact that self-administered (i.e., home-based) forms of both CBT-I and EX were used. For EX, the home-based format is the one that has been used the most frequently in previous studies assessing its effect on sleep of cancer patients (Mercier et al., 2016). Hence, it is difficult to know whether effects would have been greater if a supervised program had been used. Research comparing delivery modes of EX is greatly needed (Buffart et al., 2014). Nonetheless, these results are consistent with the modest effects observed in our meta-analysis (Mercier et al., 2016). It is also possible that an EX intervention requires more time in order to produce a positive effect on insomnia symptoms given the previous literature showing that regular EX is associated with greater sleep improvements (Kredlow et al., 2015). For CBT-I, remission rates found in the current study are weaker than what has previously been reported. Indeed, although the remission rate of insomnia obtained in our previous RCT in patients who received the same self-administered CBT-I intervention (ISI score < 8; 44.3%) was lower than that found in the professionally-administered CBT-I (71.3%), it was higher than the rate obtained in this study (30.7%) (J. Savard et al., 2014). In the same study, an average decrease of 6.2 points on the ISI with an effect size of d = -1.40 was observed (pre- vs. post-treatment), as compared to an average decrease of 4.5 points with an effect size of d = -0.78 in our study. This suggests that CBT-I underperformed in the current study.

Findings of our preference and integrity analyses provided some possible answers as to why CBT-I underperformed in the current study. Participants expressed a greater interest for the EX intervention when entering the study than for CBT-I. This may reflect the fact that the general benefits of EX on health are largely known. In addition, healthcare

providers and society in general also greatly stress the importance to adopt a healthy lifestyle, thus possibly increasing patients' interest in receiving some support in achieving that goal. Moreover, it is easier for patients to imagine what this intervention could represent as compared to a psychotherapy intervention. Our findings indicating no significant between-groups difference on adherence to some CBT-I strategies (regularity of sleep/wake schedule and reduce napping) suggest that, because the motivation to participate in CBT-I was lower, patients did not adhere as well to CBT-I strategies as they usually do in CBT-I trials with no EX condition. This lower adherence may have translated into more limited treatment effects.

The lack of significant between-groups differences on studied outcomes could also be explained by a contamination effect of the EX condition. Indeed, 42.9% of CBT-I participants met at least one of the two exercise recommendations of our EX intervention, as compared to 65.0% in the EX group, a difference that was not significant. In fact, both groups maintained a relatively similar practice of exercise at each time assessment (absence of interaction and group effects) based on GLTEQ scores and actigraphy-based exercise data. This high contamination effect may be due, again, to the fact that most participants (60%) had a favorable bias towards EX at study entry. Despite this high contamination rate, it is interesting to note that the EX group showed a significant improvement of their cardiorespiratory capacity (estimate VO2peak) at post-treatment, while no significant time effect was found in CBT-I patients. Nevertheless, this improvement appears smaller than what was observed in a meta-analysis exploring the effect of supervised EX training on VO2peak in adults with cancer (Jones et al., 2011). More precisely, the authors reported a mean increase of 2.90 on VO2peak as compared to a mean increase of 1.74 in the EX group in our study. However, our intervention was of shorter duration and home-based, which may have contributed to the smaller improvement. Indeed, in a recent RCT comparing supervised interventions of high and low-to-moderate EX intensity on fitness and cancer-related symptoms including sleep, Kampshoff et al. (2015) found a dose-response relationship between EX intensity and VO2peak. A larger improvement of VO2peak was obtained in the high intensity group (mean increase of 2.2 compared to control group), with both interventions being significantly superior than a control condition. However, no significant between-groups difference was observed on a sleep outcome (PSQI scores).

Overall, it is difficult to conclude as to whether improved fitness level mediates the effect of an EX intervention on sleep in the context of cancer. This hypothesis deserves a more rigorous examination in the future (Kredlow et al., 2015; Uchida et al., 2012).

When looking closer at EX profiles of participants during the 6-week intervention (Table 6), it is interesting to note that EX patients reported greater levels of moderate EX at the outset while CBT-I patients reported an increased number of minutes of moderate EX from Week 4. This could be explained by a possible reciprocal relationship between EX and sleep. Bernard et al. recently found, using actigraphy, that a lower WASO, TWT and TST were significantly associated with a greater daily activity count the next day in breast cancer patients (Bernard, Ivers, Savard, & Savard, 2016). Hence, it is possible that CBT-I increased their EX level from Week 4 because of their sleep improvements of Week 1 to 3. This hypothesis, which needs to be confirmed, still stresses the importance of addressing efficaciously sleep difficulties in order to enhance patients' readiness to adhere to an EX program.

Concerning the possible influence of treatment preferences, it was surprising to find that the *mismatched* subgroup of CBT-I participants showed the most consistent improvement across the main sleep outcomes (ISI, PSQI and SE), while EX participants showed more consistent improvements when they were *matched*. This suggests that CBT-I is beneficial even when it is not the preferred option. This is consistent with the finding indicating that CBT-I participants reported significantly increased scores of treatment credibility and expectancies for improvement during the course of the intervention (from baseline to the 5th week) as compared to EX participants who reported decreased treatment expectancies during the same period.

Contrary to what was initially expected, CBT-I participants did not show a better sustainment of treatment gains over time than EX participants. It was initially hypothesized that EX participants would reduce their exercise practice after the intervention phase (given that exercise maintenance is often a challenge, especially in the absence of professional guidance), thus leading to an upsurge in their symptoms. Rather, EX participants reported a non-significant change (no significant time effect) in their exercise practice during the entire study, including follow-up assessments, as assessed with the GLTEQ, as well as with

actigraphy. This hypothesis was also based on the assumption that CBT-I participants would include in their lifestyle the main sleep recommendations proposed during the intervention, thus ensuring a long-term sustainment of treatment gains. However, this hypothesis was not confirmed by the integrity analysis, which showed that CBT-I participants did not adopt a more regular sleep/wake schedule between pre- and post-treatment and both interventions led to a decrease in napping.

It is important to note that more EX participants dropped out of the study (6 vs. 2), 4 of whom (67%) did so during the follow-up periods. Given that patients were not asked to continue to be physically active during the follow-up period, the differential dropout rate across groups is difficult to explain. It is possible that EX patients were more likely to drop out of the study because they got less benefit from the intervention than they desired, which may have biased the follow-up results in favor of EX. Interestingly, results of a meta-analysis including 17 trials (mainly musculoskeletal trials) investigating the impact of preferences on attrition rates and outcomes reported that participants allocated to the less preferred treatment were more likely to complete follow-up assignments relative to indifferent participants (Preference Collaborative Review Group, 2008), which is what we found in CBT-I patients.

With regard to the effects on secondary variables, both treatments led to significant reductions of depression and fatigue from pre- to post-treatment but no significant between-groups differences were found in the improvement of depressive, anxiety and fatigue symptoms. In both groups, small pre- vs. post-treatment effect sizes (d = -0.09 to -0.51) were found for depressive and anxiety, whereas moderate effect sizes ($d \geq 0.50$) were obtained for fatigue symptoms. Interestingly, EX produced a superior effect to CBT-I on global quality of life at the 6-month follow-up. The multiple positive impacts that EX has on various areas of cancer patients' functioning are well recognized (e.g., improved physical functioning and fitness, reduced side effects of cancer treatments, prevention of bone loss and weight gain, decrease of fatigue and depression symptoms), and may translate into improved global quality of life in cancer patients (Courneya & Friedenreich, 2011; Mishra et al., 2012b; Schmitz et al., 2010; Speck et al., 2009).

To our knowledge, this is the first RCT comparing an EX intervention to CBT-I for improving clinical levels of insomnia symptoms in cancer patients. Strengths of this study include methodological aspects such as the randomization, the use of a clinical threshold of insomnia at baseline as an inclusion criterion, the variety of subjective and objective sleep measures administered, and the follow-up assessments to assess the sustainment of treatment gains over time. In addition, as recommended by the ACSM, all participants underwent a personalized assessment of their physical condition before the intervention. On the other hand, the sample was small and comprised a large proportion of women (78%) with a breast cancer diagnosis (54%), thus reducing the statistical power to detect significant differences, a problem in part circumvented by the calculation of effect sizes, and the generalization of the findings. The small sample size of the study may also have influenced the extent of the confidence interval found in non-inferiority analyses due to a greater error risk. The lack of a no-treatment control group is another important limitation that would have made it possible to delineate the effects that are due to non-specific ingredients such as the simple passage of time and regression to the mean.

Conclusion

In conclusion, insomnia is a common problem associated with cancer and its treatment. CBT-I is now considered the gold standard for treating this condition. However, accessibility to this treatment is limited. Results of this pilot RCT suggest that a self-administered CBT-I and a home-based EX program are both efficacious in reducing insomnia symptoms with a slight advantage for CBT-I immediately after the intervention. However, these findings need to be interpreted cautiously given the high risk of contamination between the two groups and the general patients' positive bias towards EX at baseline. In the future, it would be important to conduct a larger scale RCT that would not specify the interventions studied when recruiting patients. An alternative would be to use a preference trial to better distinguish the treatment effects from the preference effects. Overall, this study suggests that CBT-I is still the treatment of choice for insomnia comorbid with cancer, but that an EX program could constitute an acceptable alternative.

MCours.com