Supportive Care in Cancer
https://doi.org/10.1007/s00520-020-05473-2

ORIGINAL ARTICLE

A randomized, double-blind, placebo-controlled trial of testosterone for treatment of postmenopausal women with aromatase inhibitor-induced arthralgias: Alliance study A221102

Elizabeth Cathcart-Rake[1] · Paul Novotny[2] · Roberto Leon-Ferre[1] · Jennifer Le-Rademacher[2] · Elizabeth M. Storrick[2] · Araba A. Adjei[1] · Shelby Terstriep[3] · Rebecca Glaser[4] · Armando Giuliano[5] · William R. Mitchell[6] · Seth Page[7] · Colleen Austin[8] · Richard L. Deming[9] · Margaret A. Ferreira[8] · Jacqueline M. Lafky[1] · Stephen N. Birrell[10] · Charles L. Loprinzi[1,11]

Received: 19 November 2019 / Accepted: 14 April 2020

Abstract

Purpose To evaluate the efficacy of testosterone supplementation for improving aromatase inhibitor musculoskeletal symptoms (AIMSS).

Methods Postmenopausal women experiencing moderate-to-severe arthralgias while taking adjuvant aromatase inhibitors for breast cancer were enrolled in this trial. Initially, patients were randomly allocated to receive either a subcutaneous testosterone pellet versus a placebo pellet. Due to slow accrual, the protocol was modified such that additional participants were randomized to receive either a topical testosterone gel or a placebo gel. Changes in patient-reported joint pain were compared between patients receiving testosterone and those receiving placebo using a two-sample *t* test. Changes in hot flashes and other vasomotor symptoms were also analyzed. Further analyses were conducted to evaluate whether 27 single nucleotide polymorphisms (SNPs) in 14 genes previously associated with AIMSS were associated with testosterone supplementation benefit.

Results While 64% of patients reported an improvement in joint pain at 3 months, there were no significant differences in average pain or joint stiffness at 3 or 6 months between testosterone and placebo arms. Patients receiving testosterone did report improvements in strength, lack of energy, urinary frequency, and stress incontinence ($p < 0.05$). The subset of patients receiving subcutaneous testosterone also experienced improvements in hot flashes and mood swings. An inherited variant (rs7984870 CC genotype) in *TNFSF11* was more likely to be associated with improvements in hot flashes in patients receiving testosterone.

Conclusion The doses of testosterone supplementation used in this study did not significantly improve AIMSS.

Trial registration ClinicalTrials.gov Identifier: NCT01573442

Clinical relevance This study evaluates whether testosterone therapy improves the widespread clinical problem of aromatase inhibitor musculoskeletal symptoms.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00520-020-05473-2) contains supplementary material, which is available to authorized users.

✉ Charles L. Loprinzi
cloprinzi@mayo.edu

[1] Mayo Clinic, Rochester, MN, USA

[2] Alliance Statistics and Data Center, Mayo Clinic, Rochester, MN, USA

[3] Sanford NCORP of the North Central Plains, Sioux Falls, SD, USA

[4] Wright State University Boonshoft School of Medicine, Dayton, OH, USA

[5] Cedars-Sinai Medical Center, Los Angeles, CA, USA

[6] Southeast Clinical Oncology Research Consortium NCORP, Winston-Salem, NC, USA

[7] Wichita NCI Community Oncology Research Program, Wichita, KS, USA

[8] Northside Hospital, Atlanta, GA, USA

[9] Mercy Medical Center, Des Moines, IA, USA

[10] The Breast and Endocrine Centre, Toorak Gardens, South Australia, Australia

[11] Department of Medical Oncology, Mayo Clinic, 200 First St SW, Rochester, MN, USA

Published online: 06 May 2020

Keywords Aromatase inhibitor musculoskeletal symptoms · Testosterone · Toxicity of endocrine therapy · Hot flashes

Introduction

Aromatase inhibitors, effective agents for treating breast cancer [1, 2], can cause musculoskeletal symptoms (AIMSS) in up to 50% and vasomotor symptoms, including hot flashes and night sweats, in 20–60% of women [2–4] taking them, limiting treatment adherence, and affecting life quality [5, 6]. There is a critical need to find new strategies to mitigate these side effects.

Patients with AIMSS typically experience symmetrical joint pains, often affecting hands, wrists, and knees [5]. AIMSS generally start within 1–2 months after aromatase inhibitor initiation and peak at approximately 6 months [6].

Genetic differences may be associated with the likelihood of acquiring AIMSS and vasomotor symptoms. Initial genetic studies on patients with AIMSS reported that single nucleotide polymorphisms (SNPs) involved in drug metabolism and transport, including *CYP19A1* (aromatase gene), *ABCB1* (ATP-binding cassette sub-family B member 1-transporter gene), and *ESR1* (estrogen receptor 1 gene), were associated with AIMSS [7–9]. More recently, a genome-wide association study (GWAS) identified four linked SNPs in the *TCL1A* (T Cell Leukemia/Lymphoma 1A) gene to be significantly associated with AIMSS [10]. SNPs in genes involved in bone regeneration and remodeling, namely *TNFRSF11B* (tumor necrosis factor [TNF] receptor superfamily member 11B, *OPG*) and *TNFSF11* (TNF superfamily member 11, *RANKL*) have also been studied in association with AIMSS [11]. SNPs in hormone metabolizing genes, such as CYP1A1 and CYP19A1, have previously been associated with the likelihood of acquiring vasomotor symptoms [12–14], and a recent GWAS has identified SNPs in *TACR3* (tachykinin receptor gene) to be significantly associated with vasomotor symptoms [12, 15].

Multicenter, randomized controlled trials have reported that antidepressants, such as venlafaxine, lessen vasomotor symptoms in breast cancer survivors [16]. However, at the time that the current trial protocol was written, there were no definitive randomized trial data reporting effective treatments for alleviating AIMSS. Estrogen replacement therapy has been shown to improve arthralgias and joint health in postmenopausal women [17], likely because hormonal changes associated with menopause contribute to the development of arthralgias [18]. Estrogen is thought to have pain-modulating effects through opioid pain fibers, and is important in maintaining a healthy synovium, which expresses estrogen receptors [19, 20]. Joint cartilage turnover and subsequent damage may accelerate in the absence of estrogen [21, 22]. Moreover, the balance between androgens and estrogen, mediated by the aromatase enzyme, appears to be pivotal in maintaining joint health [23]. Androgens, especially testosterone and dihydrotestosterone, appear to be important in countering the pro-inflammatory cytokines leading to joint pain and damage [23–26]. One cohort study reported that postmenopausal women with higher dehydroepiandrosterone-sulfate (DHEAS) concentrations, in particular, experienced fewer AIMSS [27]. As such, it has been hypothesized that testosterone supplementation may relieve AIMSS.

Testosterone supplementation for the treatment of AIMSS was suggested to be effective in a small, double-blind, phase II clinical trial [28], wherein ninety women on adjuvant anastrozole for breast cancer were randomized to one of three arms: placebo, 40 mg of oral testosterone undecanoate (TU), or 80 mg of TU. The higher dose of TU was associated with a significant improvement in pain scores without significant side effects. Pain reduction at 3 months was observed in all three arms of the study: 35% with placebo, 43% with testosterone 40 mg ($p = 0.06$), and 70% with testosterone 80 mg ($p = 0.04$). Testosterone levels stabilized within a physiologic range at 3 months. Importantly, participants receiving testosterone supplementation did not experience an elevation in estradiol levels. A subsequent study evaluated symptoms of hormonal depletion in women treated with combination subcutaneous pellets of testosterone 120 mg with anastrozole 8 mg. The combination implant provides continuous release of bioavailable testosterone as well as simultaneous release of low-dose anastrozole, in an effort to prevent local aromatization of the testosterone. Women reported improved symptoms including arthralgias and hot flashes; testosterone concentrations were within a therapeutic range, and estradiol concentrations were not significantly increased [28, 29]. The results from these trials led to the conduct of the present trial, A221102, by the Alliance for Clinical Trials in Oncology (Alliance), to determine whether testosterone supplementation reduced AIMSS. Patient reports of vasomotor symptoms were also evaluated, as there were rumors, in the last century, that androgens decreased hot flashes in women.

Methods

Eligibility criteria

Postmenopausal women with estrogen and progesterone receptor-positive (> 26% or Allred score ≥ 5, for both) primary breast cancers experiencing moderate-to-severe arthralgias (rated ≥ 5 in a 10-point scale, with higher scores reflecting greater pain) attributed to anastrozole or letrozole were trial eligible. Patients were required to have been receiving anastrozole or letrozole for ≥ 21 days prior to registration

and to have plans to continue it throughout the duration of the study. Additional eligibility criteria included body mass index (BMI) between 18 and 35 kg/m^2, Eastern Cooperative Oncology Group (ECOG) performance status 0–2, and adequate laboratory parameters including hemoglobin, white blood cells, platelets, creatinine, and AST.

Patients were not allowed on trial if they had residual or recurrent cancer, glucose intolerance, coronary artery disease, or venous thromboembolism; were receiving any estrogen therapy, cyclosporine, anticoagulants, insulin, vitamin D doses > 4000 IU/day, prolonged (> 2 weeks) systemic corticosteroid treatment, concurrent chemotherapy, or radiation therapy; or were receiving any other investigational agent.

Study design and oversight

The protocol was approved per US federal guidelines and patients provided IRB-approved informed written consent. Patients were randomly assigned at a 1:1 ratio to receive testosterone or placebo using the Pocock and Simon dynamic randomization procedure. Stratification factors included baseline pain score (5–6 versus 7–10) and age (< 50 years versus 50–60 years versus > 60 years). The trial was monitored by the Alliance Data and Safety Monitoring Committee.

Interventions

At study entry, history, physical examination, and laboratory tests were obtained. Study participants were asked to complete a Hot Flash Diary [30] daily for 7 days prior to study treatment and then daily for 2 months. Additional baseline and 6 follow-up monthly questionnaires were administered to study participants to assess several quality of life measures, including the following: (1) AI-induced arthralgia and associated joint symptoms [modified Brief Pain Inventory for Aromatase Inhibitor Arthralgia (BPI-AIA) [31], (2) mood [Profile of mood states (POMS)] [32], (3) libido [Menopause Specific Quality of Life Questionnaire (MENQOL)] [33], and (4) hot flashes [Hot Flash Diary and Hot Flash Related Daily Interference Scale (HFRDIS)] [30]. Potential testosterone-associated toxicities were evaluated using a symptom experience questionnaire. Adverse events were assessed at baseline and once monthly until the end of the study, 6 months after enrollment.

After baseline questionnaire completion, the initial participants were randomly allocated to receive two surgically implanted pellets containing either a combination of testosterone 120 mg and anastrozole 8 mg or a placebo. Treatment assignments were blinded to both the patient and medical professional. As such, after patient randomization, registration personnel assigned the patient a pellet from the Alliance Research Base Pharmacy, which was marked with a de-identified kit number and the label of "testosterone OR placebo." Pellets were to be implanted at two time points: at the end of the first week on-study (following completion of the hot flash baseline week ascertainment) and 3 months later.

With this design, study accrual was slow, which was attributed to the need for a minor surgical procedure with the subcutaneous pellet preparation. In response, the protocol was amended on January 15, 2016, to change the route of delivery from subcutaneous pellet implants to a topical application of a gel containing either testosterone 10.4 mg (without anastrozole) or placebo. The gel was applied to the skin once daily for 6 months utilizing an AccuPen Dispensing Device for accuracy of dosing. After completion of the 6-month active trial period, patients could choose to continue being followed for an additional 6-month observation period.

Pharmacogenetic studies

All patients on study had DNA samples genotyped for 27 SNPs in 14 genes which have been associated with aromatase inhibitors and testosterone biotransformation including *ABCB1*, *CYP11A1*, *CYP17A1*, *CYP19A1*, *CYP27B1*, *ESR1*, *ESR2*, *MAP4K4*, *TCL1A*, *TNFRSF11B (OPG)*, *TNFSF11 (RANKL)*, *TRAM2-ASI*, *TUBB1*, and *VDR*. Replicates of the samples ($n = 10$) and 3 hapmap CEU samples in triplicates ($n = 9$) served as controls for genotyping. The genotyping was performed on the sequenom platform in the Genotyping Core of the Mayo Clinic Medical Genome Facility (MGF).

Analysis

Patient baseline characteristics were summarized by mean (standard deviation) or median (range) for continuous variables and frequency (percentage) for categorical variables.

Primary endpoint

The primary endpoint of the study was the intra-patient change in joint pain at 3 months from baseline, as measured by the participant's average pain on a scale from 0 to 10 (with 10 reflecting the worst amount of pain) on the modified BPI-AIA (item no. 3). If a patient was missing their baseline BPI-AIA pain score, their on-study pain score was used, instead. If the patient was missing their month 3 pain score, that pain score was imputed using their last reported value. The two-sample, two-sided t test with unequal variances was applied for comparison of the changes of joint pain at 3 months between testosterone and placebo arms. Based on results from the ART 2 trial, the standard deviations for change from baseline were estimated as 2.9 for the placebo arm and 1.9 for the testosterone arm. An absolute difference of joint pain between placebo and testosterone arms of 1.0 point was considered to be a clinically meaningful effect size. Based on the primary analysis, using a two-sample t test, it was estimated that a total

sample size of 194 patients (97 per arm) would provide 80% power to detect the effect size at the 5% significance level. This sample size was inflated to a total of 224 patients (112 patients per arm) to account for 15% non-evaluable patients due to ineligibility, cancel, or major violations.

In addition, item no. 3 of the BPI-AIA was also utilized to evaluate three additional points. (1) The proportion of women with improvement in joint pain, measured by a reduction in pain of at least 10% at 3 months compared with baseline, measured by the Fisher's exact test; (2) The intra-patient change in joint pain at 6 months from baseline (measured by the two-sample *t* test, as above); and (3) the intra-patient changes in joint pain at each month from baseline, measured by the repeated measures analysis of variance (RM-ANOVA) model.

Secondary endpoints

The secondary endpoints were pre-specified and were evaluated by additional items on the BPI-AIA, including the monthly change from baseline of worst pain, least pain, current pain, stiffness, and interference in activities. All of these analyses were completed using the RM-ANOVA model. Furthermore, the changes in hot flashes from baseline to the first 2 months were evaluated by the Hot Flash Diary. The area-under-the-curve of hot flash score and frequency were compared by two-sample *t* tests between testosterone and placebo arms. The monthly change of libido and menopause-specific quality of life were measured by the MENQOL and POMS monthly. The RM-ANOVA model was used to compare monthly changes from baseline for these measures.

The safety and tolerability of testosterone were measured using the CTCAE 4.0 criteria, in addition to questionnaires inquiring about self-reports of alopecia, acne, and hirsutism. Descriptive statistics and statistical plots, including frequency and histogram, were used to summarize safety and tolerability data.

Genotype and allele frequencies were calculated, and then SNP frequencies were compared with that of MA.27 control samples, a prior study which included women on aromatase inhibitors, some of whom did and some of whom did not experience any AIMSS [10]. Two-sample, two-sided *t* tests were used for this comparison. Fisher's exact tests were used to assess the associations between genetic markers and having any improvement on the testosterone arm. Adjustments for multiple testing were done using the method of Benjamini-Hochberg [34].

Data collection and statistical analyses were conducted by the Alliance Statistics and Data Center. All analyses were based on the study database frozen on 1/23/2018. Data quality was ensured by review of data by the Alliance Statistics and Data Center and by the study chairperson, following Alliance policies.

Results

Baseline characteristics

This study accrued 227 patients between September 10, 2013, and November 29, 2017 (55 patients prior to the January 15, 2016, amendment; 172 thereafter) from 71 sites. Nineteen patients canceled (Fig. 1). Baseline patient characteristics (Table 1) were relatively well balanced between arms.

Baseline symptomatology and quality of life measures at study entry were balanced between the two arms, with the exception of patients assigned to placebo being more significantly bothered by breast tenderness, dissatisfaction with their personal life, feeling depressed or anxious, and by changes in skin appearance, texture, or tone (Supplementary Table S2). The majority of participants (79%) had been on an AI for 6 or more months at study entry, with no substantial differences in AI duration between the study arms.

Patients remained on study an average of 23.2 weeks (range 2.0–45.9 weeks; includes additional observation period) with no differences between the testosterone and placebo arms (mean 24.0 and 22.4 weeks, respectively, $p = 0.19$).

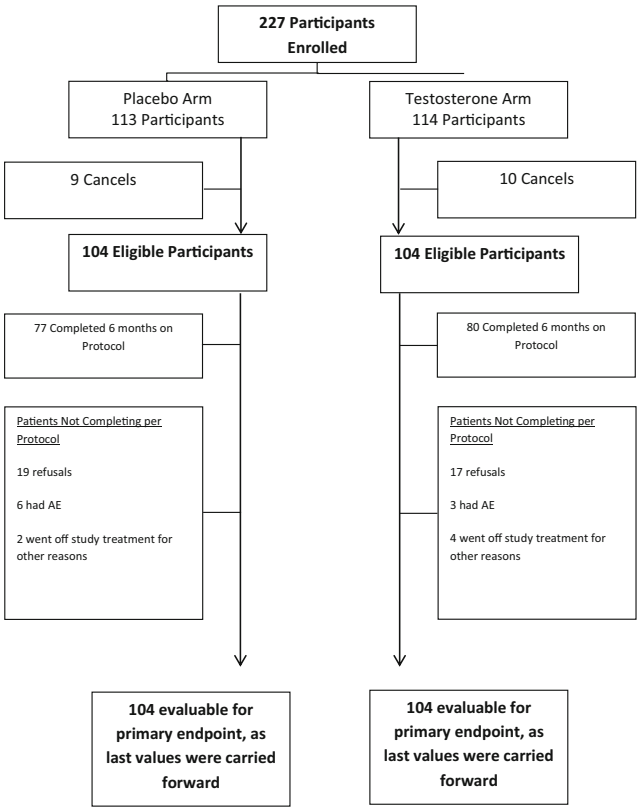

Fig. 1 CONSORT diagram

Table 1 Baseline patient characteristics

	Placebo ($N = 104$)	Testosterone ($N = 104$)
Age, years [mean (SD)]	60.1 (9.7)	59.9 (8.7)
Race		
White	94 (90%)	100 (96%)
Black or African American	4 (4%)	2 (2%)
Asian	3 (3%)	2 (2%)
Not reported or unknown	3 (3%)	0 (0%)
Ethnicity		
Hispanic	7 (7%)	9 (9%)
Non-Hispanic	96 (92%)	93 (89%)
Not reported or unknown	1 (1%)	2 (1%)
AI duration		
< 6 months	27 (26%)	14 (14%)
6–12 months	24 (23%)	26 (25%)
> 12 months	52 (50%)	63 (61%)
Missing	1	1
Baseline weight [mean (SD)]		
Weight in kg	71.3 (12.2)	75.5 (12.8)
BMI[1]	26.8 (4.2)	28.2 (4.2)
ECOG PS[2]		
0	81 (78%)	81 (78%)
1	23 (22%)	20 (19%)
2	0 (0%)	3 (3%)
Baseline average pain score [mean (SD)]	5.5 (1.8)	5.4 (1.7)

[1] Body mass index

[2] Eastern Cooperative Oncology Group Performance Status

Average pain over time

Per the BPI-AIA questionnaire, there were no significant differences between the two arms in the average joint pain at 3 months compared with baseline (Fig. 2. Pain scores were reduced by a mean (95% confidence interval (CI)) of 2.0 (1.5, 2.5) points on testosterone compared with a mean reduction of 1.9 (1.3, 2.4) points on placebo ($p = 0.50$). Sixty-four percent of patients reported an improvement of at least one point in BPI-AIA average pain at 3 months, but this was not significantly different between arms. There was also no difference between the study arms in BPI average pain from baseline to month 6: an average (±SD) pain decrease of -1.9 ± 2.2 was reported in the testosterone arm compared -2.2 ± 2.7 in the placebo arm ($p = 0.67$). Patients on testosterone had slightly more pain reduction at 1 month, but there were no significant differences at any other time.

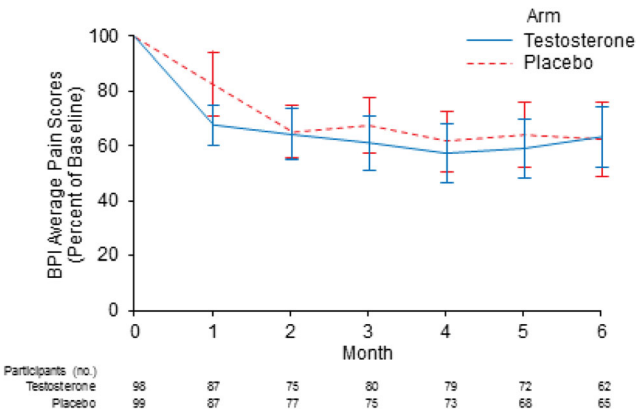

Fig. 2 Mean percent change from baseline in BPI average pain scores over time

Using repeated measures models, the BPI-AIA average pain score was significantly lower each month, from baseline ($p < 0.01$), independent of treatment arm, age, race, or AI duration. There were no significant differences in the changes from baseline in the scores of any other BPI-AIA questions.

Hot flashes and quality of life

Hot flash scores were not significantly different between arms (Fig. 3a), nor were there significant differences between hot flash frequencies. Participants who received testosterone reported more favorable relations with others compared with placebo with a mean (CI) change from baseline to month 3

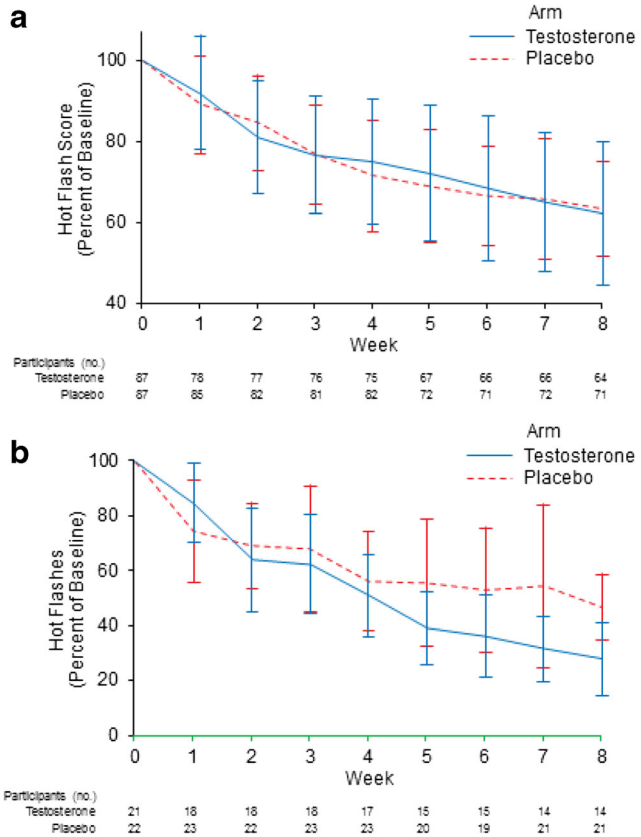

Fig. 3 Mean percent of baseline hot flash score for the entire treatment group (**a**) and the subset of patients who received the subcutaneous testosterone preparation (**b**)

of − 1.3 (− 1.8, − 0.8) for patients on testosterone compared with − 0.3 (− 0.9, 0.3) for patients on placebo, $p = 0.05$. Otherwise, there were no significant differences in the mean HFRDIS score at 3 months, or in any of its items evaluating the impact of hot flashes in quality of life.

Regarding MENQOL data at 3 months (Supplemental Table S3), testosterone patients had significantly more improvement in "Decrease in Physical Strength" and "Lack of Energy." There were no significant differences on any of the items evaluated in the POMS instrument.

Testosterone toxicity evaluation

There were no significant differences between treatment arms in any of the following symptoms at 3 or 6 months, on the Symptom Experience Diary: stomach pain or cramps, nausea, diarrhea, dizziness, decrease in appetite, abnormal sweating, trouble sleeping, mood swings, trouble concentrating, hot flashes, deepening of voice, unwanted weight gain, acne, hand or feet swelling, or undesirable hair growth. Patients on testosterone did report less fatigue at 6 months, with a mean (±SD) fatigue score of 5.4 ± 2.6 in the testosterone arm, as compared with a mean fatigue score of 6.1 ± 2.7 with placebo (lower scores representing less fatigue; $p = 0.04$).

Genetic polymorphisms, distribution of variants, and genetic associations

A total of 191 out of 207 patients on study had DNA samples genotyped. The 191 DNA samples genotyped comprised 178 Caucasians, 6 African-Americans, 4 Asians, and 3 with unknown race. The distribution of genotypes, allele frequencies, and frequency comparisons for the Caucasian patients on the study is shown in Supplementary Table S1-A. Except for 5 SNPs, all other SNPs genotyped were in Hardy Weinberg Equilibrium (HWE) (Table S1-A).

When SNP frequencies in this study were compared with that for Caucasian control samples from the MA.27 AIMSS study, 4 SNPs (namely rs11632698 and rs900798 (*CYP11A1*), rs2073618 (*TNFRSF11B*), and rs7984870 (*TNFSF11*)) showed significant association with AIMSS (unadjusted $p < 0.05$; Supplementary Table S1-B). The two *CYP11A1* SNPs did not show any association after adjusting for multiple comparisons. The GG genotypes of *TNFRSF11B* rs2073618 and *TNFSF11* rs7984870 remained significantly associated with AIMSS after adjusting for multiple comparisons ($p < 0.01$ for both). The allele frequency for *TCL1A* rs11849538 SNP was increased compared with the MA.27 control but its association with AIMSS was marginal (unadjusted p value for GG genotype = 0.08, Table S1-B).

In testosterone-treated patients on study (Table 2), the *TNFSF11* rs7984870 CC genotype was associated with greater reduction in hot flash frequency than the CG+GG genotype ($p = 0.04$).

Subset analysis of patients who received subcutaneous treatment

A subset analysis was performed, evaluating the 55 patients enrolled on the study prior to the January 15, 2016, amendment, who were randomized to receive surgically implanted pellets containing either testosterone/anastrozole or placebo. Of these, 4 patients canceled (2 on testosterone and 2 on placebo), leaving 51 patients who started subcutaneous treatment (25 on testosterone and 26 on placebo).

The median age of these patients was 59 ± 7.1 years. Baseline patient characteristics were well balanced between the two arms, with the exception that patients assigned to subcutaneous placebo reported more backache ($p = 0.05$).

While patients on subcutaneous testosterone/anastrozole had more reduction in the BPI-AIA average pain scores during the first month ($p = 0.04$), there were no significant differences in the subsequent months. Patients on subcutaneous testosterone had significantly more reduction in the percent of baseline hot flash frequency and score after 8 weeks (Fig. 3b). Patients on subcutaneous testosterone/anastrozole also reported significantly less nausea ($p = 0.02$), fatigue (p = 0.04), mood swings ($p = 0.03$), hand/feet swelling ($p = 0.01$),

Table 2 *TNFSF11* (RANKL) rs7984870 SNP: association with improvement in study endpoints for patients on testosterone

Event	Genotype (*N*)	*N* events (% between allele groups)	Odds ratio (95% CI)	Fisher's exact *p* value
Any hot flash frequency reduction from baseline to week 8	CC (14)	4 (9.1)	3.53 (0.88–14.12)	0.08
	GC+GG (78)	40 (90.9)	Reference[1]	
Any hot flash frequency reduction from baseline to week 8	CC (16)	4 (8.7)	4.32 (1.12–16.70)	0.04[3]
	GC+GG (80)	42 (91.3)	Reference[2]	
Any hot flash score reduction from baseline to week 8	CC (16)	5 (10.2)	3.52 (0.94–13.22)	0.07
	GC+GG (80)	44 (89.8)	Reference[2]	
Any joint pain reduction from baseline to month 3	CC (16)	7 (12.3)	2.86 (0.83–9.81)	0.10
	GC+GG (80)	50 (87.7)	Reference[2]	

A total of 104 patients were treated with testosterone and 8 of them did not have any genetic information

TNFSF11 (RANKL): osteoprotegerin ligand and a member of the tumor necrosis factor (TNF) superfamily 11

[1] Caucasian samples alone

[2] All patients on testosterone

[3] Association reaching statistical significance, $p < 0.05$

stress urinary incontinence ($p = 0.04$), and changes in appearance, texture, or tone of their skin ($p = 0.01$), than did patients on subcutaneous placebo (Table 3).

Discussion

In contrast to the prior double-blind, phase II clinical trial, which suggested testosterone undecanoate may improve AIMSS, in the current study, testosterone supplementation did not improve the average pain or joint stiffness, compared with placebo [28]. The discrepancy between the prior phase II and the current trial may be partly explained by the differences in the testosterone preparations between the two studies. The current trial did not evaluate systemic testosterone or DHEAS concentrations of participants (for instance, in patients who received the subcutaneous versus topical testosterone preparations, or versus the testosterone administered in the previous phase II trial), which may explain some of the differences observed. The topical preparation may have led to lower systemic androgen exposure and decreased antiarthralgia efficacy. The potential need for higher systemic doses is supported by the fact that the 80-mg daily dose in the prior phase II study appeared to be superior to the placebo, but the 40-mg daily dose only achieved borderline statistical significance. Also, the improvement in the mean BPI pain score at month one in the subcutaneous implant group may be explained by the higher levels of testosterone released earlier in the implant cycle.

Data from the current study, which demonstrated that protocol patients experienced a consistent improvement in pain every month independently of treatment assignment, confirms

Table 3 Endpoint Comparisons for subcutaneous placebo vs subcutaneous testosterone (means and 95% confidence intervals)

Endpoint	Placebo ($N = 26$)	Testosterone ($N = 25$)	*p* value
Change in BPI-AIA average pain score from baseline to month 1	− 0.8 (− 1.6, − 0.1)	− 2.0 (− 2.8, − 1.2)	0.04
Percent of baseline hot flash frequency at week 8	59.9% (44%, 76%)	35.8% (20%, 52%)	0.03
Percent of baseline hot flash scores at week 8	48.1% (35%, 61%)	28.1% (15%, 41%)	0.03
Maximum SED nausea at month 3	3.0 (1.8, 4.1)	1.1 (0.5, 1.7)	0.02
Maximum SED fatigue at month 3	7.0 (5.9, 8.0)	5.3 (4.2, 6.5)	0.04
Maximum SED mood swings at month 3	5.4 (4.2, 6.7)	3.5 (2.3, 4.7)	0.03
Maximum SED hand or feet swelling at month 3	4.2 (2.9, 5.5)	1.7 (0.8, 2.6)	0.01
MENQOL incidence of stress urinary incontinence at month 3	42% (22%, 63%)	10% (1%, 32%)	0.04
MENQOL incidence of changes in appearance, texture, or tone of skin at month 3	46% (26%, 67%)	10% (1%, 29%)	0.01

BPI-AIA, modified Brief Pain Inventory for Aromatase Inhibitor Arthralgia; *SED*, symptom experience diary; *MENQOL*, Menopause Specific Quality of Life Questionnaire

a phenomenon observed in recently reported studies, in which patients in both placebo and intervention arms reported AIMSS improvement over time [35–37].

Interestingly, the subset analysis of patients receiving subcutaneous testosterone/anastrozole in the current trial reported improvements in hot flashes, as well as several other menopausal symptoms including fatigue, mood swings, urinary incontinence, and skin appearance, tone, and texture. These results suggest that subcutaneous testosterone may have a potential benefit in the relief of hormone deficiency symptoms. Testosterone's therapeutic effect is dose dependent [38, 39], and an observational study using higher doses of subcutaneous testosterone (169 mg ± 32 mg) in breast cancer survivors reported significant improvements in psychological, somatic, and urogenital symptoms [40].

The improvements in untoward urogenital symptoms (polyuria, stress incontinence, and vaginal dryness) in the patients who received testosterone are consistent with data from a small randomized trial of vaginal testosterone to treat vaginal dryness [41], and data from studies on vaginal DHEA, a testosterone precursor, which was also shown to help vaginal symptoms [42, 43].

The previously described association between *TCL1A* rs11849538 SNP and AIMSS [10] was not replicated in the present study; however, a marginal trend for an association with AIMSS was observed in Caucasian patients, which disappeared after adjustment for multiple comparisons. Our observation that the GG genotypes for both *TNFSF11* rs7984870 and *TNFRSF11B* rs2073618 are significantly associated with increase in AIMSS in our study of Caucasian patients conflicts with an Asian cohort study [11] that reported that the GG genotypes of rs7984870 and rs2073618 are protective for AIMSS, thus making the CC genotype the at-risk genotype for AIMSS in the Asian study. In another study with a mostly Caucasian AI-treated patient cohort [44], a nominal protective association was observed for the *TNFSF11* rs7984870 GG genotype with AIMSS but no association was seen for the *TNFRSF11B* rs2073618 SNP. However, our observation for *TNFRSF11B* rs2073618 GG genotype is in agreement with a European study [45] where carriers of the CG and GG genotypes of *TNFRSF11B* rs2073618 exhibited increased risk of AIMSS. Finally, after testosterone supplementation for AIMSS patients, those carrying the T*NFSF11* rs7984870 CC genotype showed improvement in hot flashes reduction. These results warrant further studies with a larger set of patients to validate the subset of patients who might benefit from testosterone supplementation for hot flashes.

Since the time that the current trial was developed, other therapies have shown promise in alleviating the prominent clinical problem of AIMSS. Exercise was shown to decrease AIMSS pain scores when compared with usual care in a randomized controlled trial [46]. Zoledronic acid appeared to improve AIMSS in a nonrandomized pilot study, but no subsequent trial has been done to confirm this finding [47]. A third trial, evaluating omega-3 fatty acids versus placebo, did not find benefit in all protocol patients but suggested benefit over placebo in the subset of women with a body mass index over 30 [36, 48]. Even more encouragingly, the favorable results of pilot studies evaluating acupuncture and duloxetine led to definitive placebo-controlled clinical trials, which have established these two strategies as efficacious treatments for AIMSS [37, 49].

Funding information The research reported in this publication was financially supported by the National Cancer Institute of the National Institutes of Health under the Award Number UG1CA189823 (Alliance for Clinical Trials in Oncology NCORP Grant), UG1CA232760, U24CA196171, UG1CA189858, UG1CA189997, U10CA180820, and UG1CA189861 (ECOG-ACRIN), U10CA180868 (NRG), and U24CA196171, as well as the Alliance biorepository resource.

Compliance with ethical standards

Conflict of interest The US NCI provided funding for this trial.

The authors have no financial relationship with any private company regarding this research with the exception that Dr. Birrell reports personal fees from Havah Therapeutics, that is developing androgen-based therapies for women, and has a patent AU2005905768A0; additionally, Dr. Glaser has a patent (US 10,071,104 B2) related to this topic.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Baum M, Budzar AU, Cuzick J, Forbes J, Houghton JH, Klijn JG, Sahmoud T, ATAC Trialists' Group (2002) Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. Lancet 359(9324):2131–2139
2. Howell A, Cuzick J, Baum M, Buzdar A, Dowsett M, Forbes JF, Hoctin-Boes G, Houghton J, Locker GY, Tobias JS, ATAC Trialists' Group (2005) Results of the ATAC (Arimidex, tamoxifen, alone or in combination) trial after completion of 5 years' adjuvant treatment for breast cancer. Lancet 365(9453):60–62
3. Harris PF, Remington PL, Trentham-Dietz A, Allen CI, Newcomb PA (2002) Prevalence and treatment of menopausal symptoms among breast cancer survivors. J Pain Symptom Manag 23(6):501–509
4. Kligman L, Younus J (2010) Management of hot flashes in women with breast cancer. Curr Oncol 17(1):81–86
5. Crew KD, Greenlee H, Capodice J, Raptis G, Brafman L, Fuentes D, Sierra A, Hershman DL (2007) Prevalence of joint symptoms in postmenopausal women taking aromatase inhibitors for early-stage breast cancer. J Clin Oncol 25(25):3877–3883
6. Henry NL, Giles JT, Ang D, Mohan M, Dadabhoy D, Robarge J, Hayden J, Lemler S, Shahverdi K, Powers P, Li L, Flockhart D, Stearns V, Hayes DF, Storniolo AM, Clauw DJ (2008) Prospective characterization of musculoskeletal symptoms in early stage breast cancer patients treated with aromatase inhibitors. Breast Cancer Res Treat 111(2):365–372

7. Henry NL, Skaar TC, Dantzer J, Li L, Kidwell K, Gersch C, Nguyen AT, Rae JM, Desta Z, Oesterreich S, Philips S, Carpenter JS, Storniolo AM, Stearns V, Hayes DF, Flockhart DA (2013) Genetic associations with toxicity-related discontinuation of aromatase inhibitor therapy for breast cancer. Breast Cancer Res Treat 138(3):807–816
8. Gervasini G, Jara C, Olier C, Romero N, Martínez R, Carrillo JA (2017) Polymorphisms in ABCB1 and CYP19A1 genes affect anastrozole plasma concentrations and clinical outcomes in postmenopausal breast cancer patients. Br J Clin Pharmacol 83(3):562–571
9. Leyland-Jones B et al (2015) ESR1 and ESR2 polymorphisms in the BIG 1-98 trial comparing adjuvant letrozole versus tamoxifen or their sequence for early breast cancer. Breast Cancer Res Treat 154(3):543–555
10. Ingle JN, Schaid DJ, Goss PE, Liu M, Mushiroda T, Chapman JAW, Kubo M, Jenkins GD, Batzler A, Shepherd L, Pater J, Wang L, Ellis MJ, Stearns V, Rohrer DC, Goetz MP, Pritchard KI, Flockhart DA, Nakamura Y, Weinshilboum RM (2010) Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. J Clin Oncol 28(31):4674–4682
11. Wang J, Lu K, Song Y, Zhao S, Ma W, Xuan Q, Tang D, Zhao H, Liu L, Zhang Q (2015) RANKL and OPG polymorphisms are associated with aromatase inhibitor-related musculoskeletal adverse events in Chinese Han breast cancer patients. PLoS One 10(7): e0133964
12. Crandall CJ, Crawford SL, Gold EB (2006) Vasomotor symptom prevalence is associated with polymorphisms in sex steroid-metabolizing enzymes and receptors. Am J Med 119(9 Suppl 1): S52–S60
13. Fontein DYB, Houtsma D, Nortier JW (2014) Germline variants in the CYP19A1 gene are related to specific adverse events in aromatase inhibitor users: a substudy of Dutch patients in the TEAM trial. Breast Cancer Res Treat 144(3):599–606
14. Johansson H, Gray KP, Pagani O (2016) Impact of CYP19A1 and ESR1 variants on early-onset side effects during combined endocrine therapy in the TEXT trial. Breast Cancer Res 18(1):110
15. Crandall C et al (2017) Association of genetic variation in the tachykinin receptor 3 locus with hot flashes and night sweats in the women's health initiative study. Menopause 24:252–261
16. Bordeleau L, Pritchard KI, Loprinzi CL, Ennis M, Jugovic O, Warr D, Haq R, Goodwin PJ (2010) Multicenter, randomized, cross-over clinical trial of venlafaxine versus gabapentin for the management of hot flashes in breast cancer survivors. J Clin Oncol 28(35):5147–5152
17. Cirillo DJ, Wallace RB, Wu LL, Yood RA (2006) Effect of hormone therapy on risk of hip and knee joint replacement in the Women's Health Initiative. Arthritis Rheum 54(10):3194–3204
18. Neugarten BL, Kraines RJ (1965) "Menopausal symptoms" in women of various ages. Psychosom Med 27(3):266–273
19. Dietrich W, Haitel A, Holzer G, Huber JC, Kolbus A, Tschugguel W (2006) Estrogen receptor-β is the predominant estrogen receptor subtype in normal human synovia. J Soc Gynecol Investig 13(7): 512–517
20. Dawson-Basoa M, Gintzler AR (1997) Involvement of spinal cord delta opiate receptors in the antinociception of gestation and its hormonal simulation. Brain Res 757(1):37–42
21. Sniekers Y et al (2008) Animal models for osteoarthritis: the effect of ovariectomy and estrogen treatment–a systematic approach. Osteoarthr Cartil 16(5):533–541
22. Oestergaard S, Sondergaard BC, Hoegh-Andersen P, Henriksen K, Qvist P, Christiansen C, Tankó LB, Karsdal MA (2006) Effects of ovariectomy and estrogen therapy on type II collagen degradation and structural integrity of articular cartilage in rats: implications of the time of initiation. Arthritis Rheum 54(8):2441–2451
23. Schmidt M et al (2006) Inflammation and sex hormone metabolism. Ann N Y Acad Sci 1069:236–246
24. Islander U, Jochems C, Lagerquist MK, Forsblad-d'Elia H, Carlsten H (2011) Estrogens in rheumatoid arthritis; the immune system and bone. Mol Cell Endocrinol 335(1):14–29
25. Cutolo M, Villaggio B, Seriolo B, Montagna P, Capellino S, Straub RH, Sulli A (2004) Synovial fluid estrogens in rheumatoid arthritis. Autoimmun Rev 3(3):193–198
26. Cutolo M et al (2006) Anti-TNF and sex hormones. Ann N Y Acad Sci 1069(1):391–400
27. Gallicchio L, MacDonald R, Wood B, Rushovich E, Helzlsouer KJ (2011) Androgens and musculoskeletal symptoms among breast cancer patients on aromatase inhibitor therapy. Breast Cancer Res Treat 130(2):569–577
28. Birrell S, Tilley W (2009) Testosterone undecanoate treatment reduces joint morbidities induced by anastrozole tin postmenopausal women with breast cancer: results of a double-blind, randomized phase II trial. AACR
29. Glaser RL Subcutaneous testosterone-anastrozole implant therapy in breast cancer survivors. ASCO Breast Cancer Conference September 2010: p Abstract # 221
30. Carpenter JS (2001) The hot flash related daily interference scale: a tool for assessing the impact of hot flashes on quality of life following breast cancer. J Pain Symptom Manag 22(6):979–989
31. Bauml J, Chen L, Chen J, Boyer J, Kalos M, Li SQ, DeMichele A, Mao JJ (2015) Arthralgia among women taking aromatase inhibitors: is there a shared inflammatory mechanism with co-morbid fatigue and insomnia? Breast Cancer Res 17(1):89
32. Biehl B, Landauer A (1975) Das Profile of Mood States (POMS). Mannheim (Unveröffentlichtes Manuskript)
33. Hilditch JR, Lewis J, Peter A, van Maris B, Ross A, Franssen E, Guyatt GH, Norton PG, Dunn E (1996) A menopause-specific quality of life questionnaire: development and psychometric properties. Maturitas 24(6):161–175
34. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. J R Stat Soc B 57:289–300
35. Henry NL, Unger JM, Schott AF, Fehrenbacher L, Flynn PJ, Prow DM, Sharer CW, Burton GV, Kuzma CS, Moseley A, Lew DL, Fisch MJ, Moinpour CM, Hershman DL, Wade JL III (2018) Randomized, multicenter, placebo-controlled clinical trial of duloxetine versus placebo for aromatase inhibitor-associated arthralgias in early-stage breast cancer: SWOG S1202. J Clin Oncol 36(4):326–332
36. Hershman DL, Unger JM, Crew KD, Awad D, Dakhil SR, Gralow J, Greenlee H, Lew DL, Minasian LM, Till C, Wade JL III, Meyskens FL, Moinpour CM (2015) Randomized multicenter placebo-controlled trial of omega-3 fatty acids for the control of aromatase inhibitor-induced musculoskeletal pain: SWOG S0927. J Clin Oncol 33(17):1910–1917
37. Hershman DL et al Randomized blinded sham- and waitlist-controlled trial of acupuncture for joint symptoms related to aromatase inhibitors in women with early stage breast cancer (SWOG 1200). Oral Presentation at: San Antonio Breast Cancer Symposium December 2017
38. Huang G, Basaria S, Travison TG, Ho MH, Davda M, Mazer NA, Miciek R, Knapp PE, Zhang A, Collins L, Ursino M, Appleman E, Dzekov C, Stroh H, Ouellette M, Rundell T, Baby M, Bhatia NN, Khorram O, Friedman T, Storer TW, Bhasin S (2014) Testosterone dose-response relationships in hysterectomized women with or without oophorectomy: effects on sexual function, body composition, muscle performance and physical function in a randomized trial. Menopause 21(6):612–623
39. Glaser R, York AE, Dimitrakakis C (2011) Beneficial effects of testosterone therapy in women measured by the validated Menopause Rating Scale (MRS). Maturitas 68(4):355–361

40. Glaser RL, York AE, Dimitrakakis C (2014) Efficacy of subcutaneous testosterone on menopausal symptoms in breast cancer survivors. J Clin Oncol 32(Suppl 2):109
41. Melisko ME, Goldman ME, Hwang J, de Luca A, Fang S, Esserman LJ, Chien AJ, Park JW, Rugo HS (2017) Vaginal testosterone cream vs estradiol vaginal ring for vaginal dryness or decreased libido in women receiving aromatase inhibitors for early-stage breast cancer: a randomized clinical trial. JAMA Oncol 3(3): 313–319
42. Labrie F, Archer DF, Koltun W, Vachon A, Young D, Frenette L, Portman D, Montesino M, Côté I, Parent J, Lavoie L, BSc AB, Martel C, Vaillancourt M, Balser J, Moyneur É, members of the VVA Prasterone Research Group (2018) Efficacy of intravaginal dehydroepiandrosterone (DHEA) on moderate to severe dyspareunia and vaginal dryness, symptoms of vulvovaginal atrophy, and of the genitourinary syndrome of menopause. Menopause 25(11):1339–1353
43. Barton DL, Shuster LT, Dockter T, Atherton PJ, Thielen J, Birrell SN, Sood R, Griffin P, Terstriep SA, Mattar B, Lafky JM, Loprinzi CL (2018) Systemic and local effects of vaginal dehydroepiandrosterone (DHEA): NCCTG N10C1 (Alliance). Support Care Cancer 26(4):1335–1343
44. Dempsey JM, Xi J, Henry NL, Rae JM, Hertz DL (2018) Attempted replication of SNPs in RANKL and OPG with musculoskeletal adverse events during aromatase inhibitor treatment for breast cancer. Physiol Genomics 50(2):98–99
45. Lintermans A, Neven P (2015) Safety of aromatase inhibitor therapy in breast cancer. Expert Opin Drug Saf 14(8):1201–1211
46. Irwin ML, Cartmel B, Gross CP, Ercolano E, Li F, Yao X, Fiellin M, Capozza S, Rothbard M, Zhou Y, Harrigan M, Sanft T, Schmitz K, Neogi T, Hershman D, Ligibel J (2015) Randomized exercise trial of aromatase inhibitor-induced arthralgia in breast cancer survivors. J Clin Oncol 33(10):1104–1111
47. Santa-Maria CA, Bardia A, Blackford AL, Snyder C, Connolly RM, Fetting JH, Hayes DF, Jeter SC, Miller RS, Nguyen A, Quinlan K, Rosner GL, Slater S, Storniolo AM, Wolff AC, Zorzi J, Henry NL, Stearns V (2018) A phase II study evaluating the efficacy of zoledronic acid in prevention of aromatase inhibitor-associated musculoskeletal symptoms: the ZAP trial. Breast Cancer Res Treat 171(1):121–129
48. Shen S, Unger JM, Crew KD, Till C, Greenlee H, Gralow J, Dakhil SR, Minasian LM, Wade JL III, Fisch MJ, Henry NL, Hershman DL (2018) Omega-3 fatty acid use for obese breast cancer patients with aromatase inhibitor-related arthralgia (SWOG S0927). Breast Cancer Res Treat 172(3):603–610
49. Smith EM et al (2013) Effect of duloxetine on pain, function, and quality of life among patients with chemotherapy-induced painful peripheral neuropathy: a randomized clinical trial. JAMA 309(13): 1359–1367