ORIGINAL ARTICLE

Genome-wide association study of aromatase inhibitor discontinuation due to musculoskeletal symptoms

Daniel L. Hertz[1] · Julie A. Douglas[2,3] · Robert M. Miller[2] · Kelley M. Kidwell[4] · Christina L. Gersch[5] · Zeruesenay Desta[6] · Anna Maria Storniolo[6] · Vered Stearns[7] · Todd C. Skaar[6] · Daniel F. Hayes[5] · N. Lynn Henry[5] · James M. Rae[5]

Abstract

Objective Aromatase inhibitors (AIs) are commonly used to treat hormone receptor positive (HR +) breast cancer. AI-induced musculoskeletal syndrome (AIMSS) is a common toxicity that causes AI treatment discontinuation. The objective of this genome-wide association study (GWAS) was to identify genetic variants associated with discontinuation of AI therapy due to AIMSS and attempt to replicate previously reported associations.

Methods In the Exemestane and Letrozole Pharmacogenetics (ELPh) study, postmenopausal patients with HR + non-metastatic breast cancer were randomized to letrozole or exemestane. Genome-wide genotyping of germline DNA was conducted followed by imputation. Each imputed variant was tested for association with time-to-treatment discontinuation due to AIMSS using a Cox proportional hazards model assuming additive genetic effects and adjusting for age, baseline pain score, prior taxane treatment, and AI arm. Secondary analyses were conducted within each AI arm and analyses of candidate variants previously reported to be associated with AIMSS risk.

Results Four hundred ELPh participants were included in the combined analysis. Two variants surpassed the genome-wide significance level in the primary analysis (p value $< 5 \times 10^{-8}$), an intronic variant (rs79048288) within *CCDC148* (HR = 4.42, 95% CI: 2.67–7.33) and an intergenic variant (rs912571) upstream of *PPP1R14C* (HR = 0.30, 95% CI: 0.20–0.47). In the secondary analysis, rs74418677, which is known to be associated with expression of *SUPT20H*, was significantly associated with discontinuation of letrozole therapy due to AIMSS (HR = 5.91, 95% CI: 3.16–11.06). We were able to replicate associations for candidate variants previously reported to be associated with AIMSS in this cohort, but were not able to replicate associations for any other variants previously reported in other patient cohorts.

Conclusions Our GWAS findings identify several candidate variants that may be associated with AIMSS risk from AI generally or letrozole specifically. Validation of these associations in independent cohorts is needed before translating these findings into clinical practice to improve treatment outcomes in patients with HR + breast cancer.

Keywords Pharmacogenetic · Genome-wide association study (GWAS) · Aromatase inhibitor · Musculoskeletal adverse effects · Treatment discontinuation · Adherence · Breast cancer

Introduction

Depleting systemic estrogen concentrations by inhibiting the CYP19A1 (aromatase)-mediated conversion of androgens to estrogens is an effective treatment strategy for hormone receptor positive (HR +) breast cancer. Third-generation aromatase inhibitors (AIs) (anastrozole, letrozole,

or exemestane) are first-line treatment in postmenopausal patients with HR + breast cancer [1, 2]. However, AI treatment is associated with characteristic toxicities that cause treatment non-persistence, which increases risk of breast cancer recurrence and death [3]. One of the toxicities that most often causes treatment discontinuation is AI-induced musculoskeletal syndrome (AIMSS)[4–7]. AIMSS is characterized by joint pain and stiffness, myalgias, carpal tunnel syndrome, tenosynovitis, and/or reduced grip strength and is experienced by approximately half of AI-treated patients [8–10].

✉ Daniel L. Hertz
DLHertz@med.umich.edu

Extended author information available on the last page of the article

AIMSS risk may be affected by clinical factors such as body mass index (BMI), prior chemotherapy (especially taxanes), and prior tamoxifen. However, these clinical factors alone do not accurately predict which patients will necessitate treatment discontinuation due to AIMSS [10–13]. Inherited germline variants in candidate genes including *CYP19A1* [14–16] and *ESR1* [17, 18] have been reported to affect AIMSS risk [19]. Additionally, a genome-wide association study (GWAS) identified variants in *TCL1A* that may predict AIMSS; however, this has not been successfully validated in other cohorts [20, 21]. Further pharmacogenetic discovery and replication are necessary to identify variants that may predict which patients are at increased risk of AIMSS and AI treatment discontinuation.

A substudy of the Exemestane and Letrozole Pharmacogenetics (ELPh) study was conducted to document the clinical course of AIMSS and identify clinical and genetic risk factors, primarily using candidate gene approaches [22, 23]. We recently performed genome-wide genotyping within ELPh to validate that *CYP2A6* genetics was a major determinant of letrozole pharmacokinetics [24]. The objective of the current study was to identify new genetic variants associated with discontinuation of AI therapy due to AIMSS in ELPh participants using a genome-wide approach and, secondarily, replicate previously reported pharmacogenetic associations.

Materials and methods

ELPh patients and treatment

ELPh was a prospective, open-label study that enrolled postmenopausal women with stage 0–III HR + breast cancer considering AI therapy. AI could be given as front-line treatment or following completion of local (i.e., surgery and/ or radiation) and systemic (tamoxifen and/or chemotherapy) treatment [25]. Patients were randomized 1:1 to receive oral exemestane (25 mg/day) or letrozole (2.5 mg/day) for 2 years, and the arms were stratified by prior treatment with bisphosphonate, tamoxifen, and chemotherapy. Patients were enrolled from August 2005 to July 2009 from Indiana University Melvin and Bren Simon Comprehensive Cancer Center, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, and the University of Michigan Rogel Cancer Center. Due to the small number of non-white patients in the ELPh study, all analyses and results were restricted to self-reported white patients. All patients provided written informed consent prior to enrollment and the study was approved by the Institutional Review Boards of each site.

Collection of treatment discontinuation due to AIMSS

Reasons for discontinuation of AI treatment, including the specific side effect, were prospectively recorded by study coordinators at each site. The primary end point for this pharmacogenetic analysis was the time to discontinuation of the initial AI medication due to musculoskeletal toxicity (i.e., AIMSS), defined as arthralgias, myalgias, joint pain or stiffness, tendonitis, numbness or tingling, and/or carpal tunnel syndrome. We have previously reported analyses of clinical and candidate genetic predictors of this endpoint [4, 23, 26].

Genome-wide genotyping and imputation

Methods for germline DNA isolation, genome-wide genotyping, imputation, and quality control have been previously reported [27]. Briefly, pre-treatment germline DNA [28] underwent genome-wide genotyping on the Infinium Global Screening Array V2.0 (> 650,000 variants) at the University of Michigan Advanced Genomics Biomedical Research Core. All sample call rates were > 98%. Variants with call rates (< 95%) or large departure from Hardy Weinberg Equilibrium ($p < 10^{-6}$) were removed. Imputation was performed to generate allelic doses for > 16 million variants using the Michigan Imputation Server [29], SHAPEIT, and Eagle (v2.4), with removal of variants with imputation $r^2 < 0.20$.

Statistical methods

Cox proportional hazards models were used to test the association between each genotyped or imputed variant and time to discontinuation of AI therapy due to AIMSS. Patients who discontinued for reasons other than AIMSS were censored at the time of discontinuation. Associations were analyzed under an additive genetic model using the "gwasurvivr" package in R [Rizvi et al. 2019] and adjusted for clinical variables that were previously reported to be associated with discontinuation due to AIMSS, including age (≤ 55 years), prior taxane chemotherapy (Yes vs. No), pre-treatment pain score on the visual analog scale (continuous), and treatment arm (exemestane vs. letrozole) [4]. As a secondary analysis, we repeated all association analyses by AI treatment arm (exemestane versus letrozole) to identify AI-specific associations. For variants that were significantly associated in one AI treatment group but not the other, we also tested the significance of an interaction term in the combined cohort.

Variants with minor allele frequencies (MAF) less than 0.025 in the combined cohort or 0.05 in the treatment-specific arms were excluded. Unless noted otherwise, results are reported for imputed allelic dosages. *P* values less than

5×10^{-8} were considered genome-wide significant. For the twenty-five variants in 11 genes that were previously reported to be associated with AIMSS risk [16, 18–21, 30–33], p values less than 0.05 were considered statistically significantly. Where noted below, associations for a subset of these variants were previously reported in the ELPh cohort [22, 23].

Similar to our previous GWAS, dbSNP and LDlink [34] were used to annotate variants of interest and assess patterns of linkage disequilibrium. Genetic variants that were significantly associated with discontinuation of treatment were also functionally interrogated using publicly available databases, including Genotype-Tissue Expression (GTEx) [35], RegulomeDB [36, 37], and the mQTL database [38]. Unless specified otherwise, all analyses were carried out using a combination of R and Python programs.

Results

Of the 503 patients enrolled on the ELPh study, 400 self-reported white patients who initiated AI treatment and had genome-wide genotype data were included in the analysis
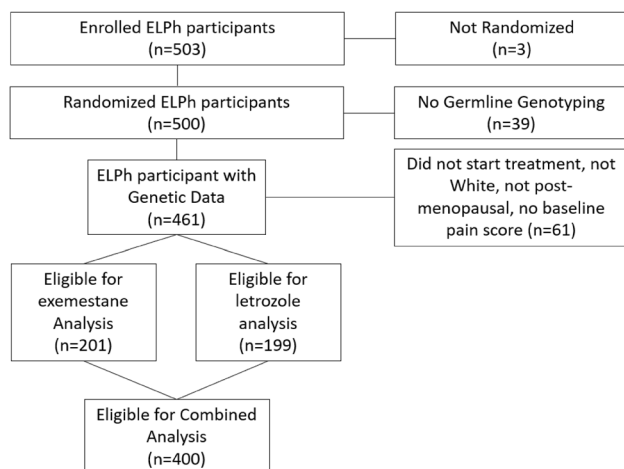

Fig. 1 CONSORT diagram of patient flow from ELPh Trial to this AIMSS GWAS

(Fig. 1). Demographic and clinical data for these patients are presented in Table 1 [28]. One hundred patients discontinued their treatment due to AIMSS, with a median time of 6.3 months (interquartile range of 8.3 months). Clinical variables previously reported to be associated with AIMSS risk in ELPh [4] had consistent findings in this analysis (i.e., increased AIMSS risk with younger age, prior taxane chemotherapy, higher baseline pain score, and exemestane treatment) (data not shown).

In the primary analysis of the combined cohort of 400 AI-treated patients, two imputed variants (rs79048288 and rs912571) were associated with AIMSS-related treatment discontinuation risk after adjustment for clinical covariates (Table 2, Fig. 2). Patients carrying the "T" allele at rs79048288 had increased risk of AIMSS (HR = 4.42, 95% CI: 2.67–7.33, $p = 7.69 \times 10^{-9}$, Supplementary Fig. 1). Rs79048288 is an intronic variant within *CCDC148*, the gene encoding the coiled-coil domain containing 148 protein. According to GTEx or RegulomeDB, this variant is not known to affect expression of *CCDC148* or any other genes or protein. The RegulomeDB score for this variant is 0.35 (scale 0.0–1.0), with higher scores indicating increased likelihood of being a regulatory variant. There was no interaction between rs79048288 and treatment arm ($p = 0.30$), indicating that the association between rs79048288 and AIMSS risk is not AI-specific.

Carriers of the "G" allele at rs912571 were protected from AIMSS in the combined AI therapy GWAS (HR = 0.30, 95% CI: 0.20–0.47, $p = 4.74 \times 10^{-8}$, Supplementary Fig. 2). Rs912571 is in an intergenic region upstream of *PPP1R14C*, the gene encoding protein phosphatase 1 regulatory inhibitor subunit 14C. The RegulomeDB score for rs912571 is 0.43 (0.0–1.0). According to GTEx, rs912571 is also an expression quantitative trait locus (eQTL) for *PPP1R14C* in sun exposed skin, with higher expression for the variant G allele ($p = 4.8 \times 10^{-5}$). Again, there was no interaction between rs912571 and the AI treatment arm ($p = 0.87$).

Secondary GWAS analyses were conducted within each drug separately. In the letrozole analysis, two variants in (rs1324052 and rs74418677) were associated with increased

Table 1 Characteristics of breast cancer patients by AI treatment

	Letrozole ($n = 199$)	Exemestane ($n = 201$)
Age at enrollment (years)	60 (13.5)	58 (11)
Body mass index (kg/m^2)	28.9 (8.9)	28.8 (7.0)
Prior taxane chemotherapy treatment	68 (34%)	66 (33%)
Pre-treatment pain score on visual analog scale	2.0 (3.6)	2.3 (3.8)
Discontinuation of AI due to AIMSS	44 (22%)	56 (28%)
Discontinuation of AI due to other reasons	22 (11%)	35 (17%)
Time to discontinuation of AI due to AIMSS (months)	9.0 (8.3)	5.9 (6.4)

AI aromatase inhibitor; *AIMSS* AI-induced musculoskeletal symptoms

Data are median (interquartile range) or number (percentage)

Table 2 Variants associated with time to discontinuation of AI treatment due to AI-induced musculoskeletal symptoms

Treatment group	Variant	Chromosome	Position[a]	Nearest gene	Alleles[b]	EAF[c]	r^2	HR (95% CI)[d]	*P* value
Combined	rs79048288	2	159,271,033	CCDC148	C > T	0.026	0.95	4.42 (2.67–7.33)	7.69×10^{-9}
Combined	rs912571	6	150,440,290	Intergenic	C > G	0.93	0.92	0.30 (0.20–0.47)	4.74×10^{-8}
Letrozole	rs1324052	13	37,841,344	Intergenic	G > A	0.091	0.91	5.91 (3.16–11.06)	2.80×10^{-8}
Letrozole	rs74418677	13	37,846,201	Intergenic	G > C	0.091	0.91	5.91 (3.16–11.06)	2.8×10^{-8}

AI aromatase inhibitor; *EAF* effect allele frequency; r^2 imputation r-squared; *HR* hazard ratio; *CI* confidence interval.

[a]Position based on genome build 37.

[b]Effect allele is second allele.

[c]EAF in treatment group.

[d]Hazard ratio based on Cox proportional hazards model assuming additive genetic effects and adjusted for age (under 55 years), baseline pain score on visual analog scale, prior taxane chemotherapy treatment, and (for combined treatment group) drug (exemestane).
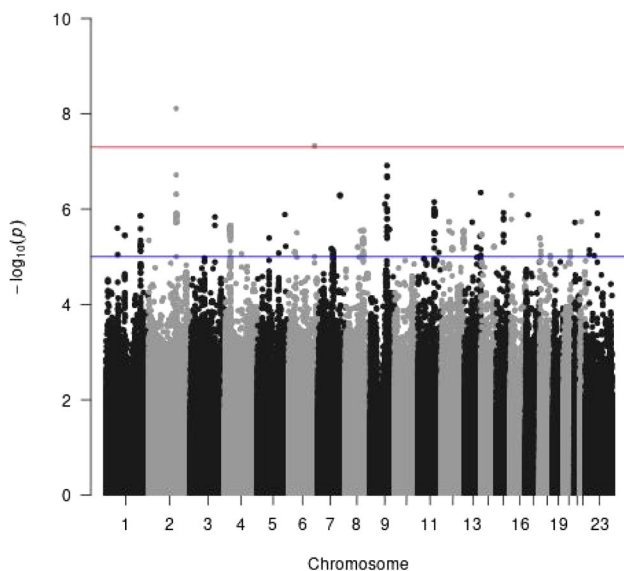

Fig. 2 Associations between all imputed variants and discontinuation of AI therapy due to AIMSS after adjustment for age (under 55 years), baseline pain score on visual analog scale, prior taxane chemotherapy treatment, and drug (exemestane). Two variants (rs79048288 and rs912571) were significantly associated with AIMSS risk in the combined cohort (*p* value < 5×10^{-8}, indicated by red horizontal line)

AIMSS risk (5.91, 95% CI: 3.16–11.06, $p = 2.80 \times 10^{-8}$, Table 2, Supplementary Figs. 3 and 4). These variants, which are in perfect linkage disequilibrium ($R^2 = 1$), are in an intergenic region downstream of a long noncoding RNA known as LOC105377814. Each variant is associated with expression of *SUPT20H* in testis tissue in GTEx, with lower gene expression in carriers of the risk variant (i.e., rs74418677 C allele) (both $p = 2.3 \times 10^{-5}$), possibly via a cis-acting methylation effect on the nearby methylation site cg19272349 (mQTL in adults β = −0.504, $p = 1.99 \times 10^{-9}$). Based on RegulomeDB, there is stronger evidence that rs74418677 is a regulatory variant than rs1324052 (score of 0.59 versus 0.35). No variants were significantly associated with AIMSS risk in the exemestane analysis (Supplementary Fig. 3). A list of all imputed variants with suggestive associations ($p < 1 \times 10^{-5}$) in the combined, letrozole, or exemestane analyses can be found in Supplementary Table 1. Six candidate variants, rs912571, rs9322336, and rs2347868 in *ESR1*, rs7984870 in *RANKL*, and rs2369049 and rs11849538 in *TCL1A*, were significantly associated with AIMSS risk in either the combined cohort or the exemestane-treated group (*p* value < 0.05) (Table 3 and Supplementary Table 2). However, these associations have been previously reported in similar analyses of the ELPh cohort [22, 23]. None of the variants that were previously reported to be associated with AIMSS that had not been previously tested in ELPh was associated with AIMSS in this analysis (Supplementary Table 2).

Discussion

Musculoskeletal symptoms are a common AI treatment-emergent toxicity that often leads to treatment discontinuation [4–7]. Prior studies, primarily using candidate gene approaches, have reported several genetic variants that affect AIMSS risk [14–19]. Additionally, a genome-wide association study (GWAS) identified variants in *TCL1A* that may predict AIMSS [20]; however, this has not been successfully validated in other cohorts [20, 21]. Our primary GWAS identified two new variants (rs79048288 and rs912571) that were associated with AIMSS risk in the ELPh cohort.

In the primary GWAS analysis, the T allele of the intronic rs79048288 variant within *CCDC148* was associated with a 4.4-fold higher AIMSS risk and the G allele of the intergenic rs912571 upstream of *PPP1R14C* was associated with a 3.3-fold lower AIMSS risk. In silico analyses did not reveal any obvious functional impact of the intronic variant within *CCDC148* (rs79048288), and little is known about the physiological function of this protein,

Table 3 Candidate variants associated with time to discontinuation of AI treatment due to AI-induced musculoskeletal symptoms

Treatment group	Variant	Chromosome	Position[a]	Nearest gene	Alleles[b]	EAF[c]	r^2	HR (95% CI)[d]	P value
Combined	rs912571	6	150,440,290	*ESR1*	C > G	0.765	0.98	0.61 (0.44–0.83)	0.002
Combined	rs9322336	6	152,200,430	*ESR1*	C > T	0.774	1.00	0.64 (0.46–0.88)	0.007
Exemestane	rs9322336	6	152,200,430	*ESR1*	C > T	0.769	1.00	0.61 (0.39–0.99)	0.033
Exemestane	rs2347868	6	152,251,568	*ESR1*	T > C	0.775	0.98	0.58 (0.38–0.88)	0.011
Combined	rs7984879	13	43,146,482	*RANKL*	G > C	0.447	1.00	1.42 (1.06–1.89)	0.018
Exemestane	rs7984879	13	43,146,482	*RANKL*	G > C	0.460	1.00	1.61 (1.07–2.42)	0.022
Exemestane	rs2369049	14	96,171,851	*TCL1A*	A > G	0.157	1.00	0.53 (0.29–1.00)	<0.05
Exemestane	rs11849538	14	96,175,978	*TCL1A*	C > G	0.145	1.00	0.50 (0.25–0.98)	0.042

AI aromatase inhibitor; *EAF* effect allele frequency; r^2 imputation r-squared; *HR* hazard ratio; *CI* confidence interval.

[a]Position based on genome build 37.

[b]Effect allele is second allele.

[c]EAF in treatment group.

[d]Hazard ratio based on Cox proportional hazards model assuming additive genetic effects and adjusted for age (under 55 years), baseline pain score on visual analog scale, prior taxane chemotherapy treatment, and (for combined treatment group) drug (exemestane).

except that dysregulated expression has been reported in several cancer types [39, 40]. In silico analyses revealed that the G allele of rs912571 was associated with higher expression of *PPP1R14C* in sun-exposed skin, indicating that this variant could be functionally consequential. This gene encodes a signal-transducing protein phosphatase, also referred to as KEPI, that is an inhibitor of myosin phosphatase and regulates smooth muscle contraction, providing further suggestive evidence that this variant could be functionally associated with AIMSS [41–43].

Two additional variants of interest, rs1324052 and rs74418677, were found to be associated with AIMSS in the letrozole-only analysis. In silico analyses suggest that rs74418677 is a regulatory variant that affects methylation of cg19272349 and expression of *SUPT20H*, also referred to as P38IP, a protein that is known to be involved in cell cycle regulation and cellular autophagy [44, 45]. Intriguingly, a nonsense variant (p.Lys25X) in this gene was identified as the likely causal variant in hereditary rheumatoid arthritis [46]. We speculate that patients with subclinical arthritis-like conditions are at increased risk of clinically overt musculoskeletal pain when administered letrozole, similar to the identification of hereditary neuropathy genes as predictors of taxane-induced neuropathy [47]. However, we are not aware of any prior studies that have investigated or reported an association for these or other polymorphisms in these genes (i.e., *CCDC148*, *PPP1R14C*, or *SUPT20H*) with AIMSS or any other AI treatment outcome.

Our attempted replication of variants previously reported to be associated with AIMSS found only two significant associations in our combined cohort, both of which have been previously reported in ELPh. The association between rs9322336 in *ESR1* and AIMSS risk in exemestane-treated patients was previously reported by our group in 2013 [23].

Interestingly, another group recently reported that this variant was also associated with lower AIMSS risk in an independent cohort of 196 patients treated with letrozole or anastrozole [18]. While these two studies provide consistent and suggestive evidence of association, the association between rs9322336 and AIMSS risk awaits validation in additional studies. In fact, such a validation was recently attempted in the racially diverse ECOG E1Z11 cohort of anastrozole-treated patients using rs2347868, which is modestly correlated with rs9322336 (linkage disequilibrium R^2 of 0.31), and no association was detected for this variant or any of the other nine variants tested [48]. The other association for *RANKL* (rs7984870) was previously reported by our group in 2019 [22] and was itself an attempt to replicate a previously reported association by Wang et al. [49]. This variant was not included in the E1Z11 analysis and was not successfully replicated in our recent analysis of an independent cohort of 143 AI-treated women [50]. Taken together, there is weak evidence that any of these candidate variants in *CYP17A1* [14], *CYP19A1* [14–16], *ESR1* [17, 18], or *TCL1A* [20, 21] are associated with AIMSS risk.

A genetic biomarker of AIMSS could be useful in clinical decision-making. Patients carrying a variant that increases risk of AIMSS for all AI may be candidates for enhanced toxicity monitoring [51] or evidence-based interventions such as exercise, yoga, duloxetine, and acupuncture [52–54]. The identification of a genetic variant that increases risk for only one or two of the AI's would be even more clinically useful. These patients could be switched within the AI class, since all third-generation AI's are similarly effective [55] and switching within the drug class can improve treatment tolerability and persistence [4]. This study indicates that none of the previously reported genetic biomarkers is sufficiently robust

for clinical use. The variants identified in our GWAS, particularly the rs74418677 variant that may increase risk of letrozole-induced musculoskeletal toxicity through expression of *SUPT20H*, should be prioritized for future validation studies, e.g., within the E1Z11 cohort [48]. Convincing evidence of clinical validity is warranted prior to further investigation of the causal mechanism for these candidate variants and is necessary before these variants can be used for clinical decision-making.

This study had several strengths, including the use of hypothesis-agnostic genome-wide association approach in a large prospectively accrued cohort of patients with a well-documented clinical outcome. However, there were also some limitations that should be considered. Though the ELPh cohort is similar in size to a prior AIMSS GWAS [20], pharmacogenetic sample sizes like these are still orders of magnitude smaller than those used in disease genetics GWAS [56–58], which limits power to detect associations with smaller effect sizes or for uncommon variants (i.e., MAF < 0.025), or pathways that are enriched in the detected associations [59]. Finally, we were not able to attempt validation of an AIMSS polygenic risk score reported by another group due to a lack of information in their publication with which to recapitulate their 70-variant signature [60]. While we were not in a position to functionally characterize these variants in preclinical models, whether rs74418677 plays a role in regulating SUPT20H expression should investigated.

In conclusion, we identified several new variants that were associated with AIMSS risk in our cohort of AI-treated patients, including rs912571 (*PPP1R14C*) and rs74418677 (*SUPT20H*), that should be prioritized for attempted replication in independent cohorts of AI-treated patients. Successful validation of these associations is necessary prior to prospective studies that use genetic biomarkers to inform clinical decision making to reduce AIMSS and enhance AI treatment persistence to improve clinical outcomes in patients with HR + breast cancer.

Author contribution DLH was involved in designing this analysis and writing the manuscript. JAD and RMM conducted the statistical analysis. KMK cleaned the data for analysis and helped design the analysis. CLG helped generate the genetic data. ZD and TS helped design the initial trial and generate data. AMS, VS, DFH, and NLH designed the initial clinical trial and enrolled participants. JMR helped design this analysis and oversaw the genetic data generation and analysis.

Data availability Data and material are available upon reasonable request to the corresponding author.

Code availability Code is available upon reasonable request to the corresponding author.

Declarations

Ethics approval This was a retrospective secondary analysis of a previously reported prospective clinical study. The prospective study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of each of the three participating Universities.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish The authors affirm that human research participants provided informed consent to participate. No individual data was included in this manuscript.

Conflict of interest This work was supported in part by Pfizer and Novartis Pharma AG. Dr. Stearns has received research funding from Abbvie, Biocept, Celgene, Merck, Novartis, Medimmune, Pfizer, and Puma Biotechnology. Dr. Stearns is on an advisory board for Novartis (10/25/2021, is a member of the Data Safety Monitoring Board for Immunomedics, Inc, and Chair of the Data Safety Monitoring Board for AstraZeneca, and has received non-financial support from Foundation Medicine Study Assays. Dr. Henry has received research funding to conduct pharmaceutical sponsored clinical trials from Abbvie, Innocrin Pharmaceuticals, Pfizer, and Blue Note Therapeutics. Dr. Hayes reports research funding from Merrimack Pharmaceuticals, Eli Lilly, Menarini Silicon Biosystems, Puma Biotechnology, Pfizer, and Astra Zeneca in the last 24 months. He also reports consulting fees from Cepheid, Freenome, Artiman Ventures, Agendia, Lexent Bio, Epic Sciences, and Salutogenic Innovations. He is the named investigator of a patent held by the University of Michigan which is licensed to Menarini Silicon Biosystems, from whom he receives annual royalties. He holds stock options in Oncimmune LLC and InBiomotion.

References

1. Goss PE, Ingle JN, Martino S et al (2003) A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med 349:1793–1802

2. Early Breast Cancer Trialists' Collaborative G, Dowsett M, Forbes JF et al (2015) Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet (London, England) 386:1341–1352

3. Hershman DL, Shao T, Kushi LH et al (2011) Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. Breast Cancer Res Treat 126:529–537

4. Henry NL, Azzouz F, Desta Z et al (2012) Predictors of aromatase inhibitor discontinuation as a result of treatment-emergent symptoms in early-stage breast cancer. J Clin Oncol : Off J Am Soc Clin Oncol 30:936–942

5. Henry NL, Speth K, Lintermans A et al (2017) Associations between patient and anthropometric characteristics and aromatase inhibitor discontinuation. Clin Breast Cancer 17:350-355.e354. https://doi.org/10.1016/j.clbc.2017.1003.1002

6. Moscetti L, Agnese Fabbri M, Sperduti I et al (2015) Adjuvant aromatase inhibitor therapy in early breast cancer: what factors lead patients to discontinue treatment? Tumori 101(469–473):4. https://doi.org/10.5301/tj.5000376

7. Chim K, Xie SX, Stricker CT et al (2013) Joint pain severity predicts premature discontinuation of aromatase inhibitors in breast cancer survivors. BMC Cancer 13:401. https://doi.org/10.1186/1471-2407-1113-1401

8. Felson DT, Cummings SR (2005) Aromatase inhibitors and the syndrome of arthralgias with estrogen deprivation. Arthritis Rheum 52:2594–2598

9. Hadji P, Jackisch C, Bolten W et al (2014) COMPliance and arthralgia in clinical therapy: the COMPACT trial, assessing the incidence of arthralgia, and compliance within the first year of adjuvant anastrozole therapy. Ann Oncol 25(372–377):3. https://doi.org/10.1093/annonc/mdt1513

10. Mao JJ, Stricker C, Bruner D et al (2009) Patterns and risk factors associated with aromatase inhibitor-related arthralgia among breast cancer survivors. Cancer 115:3631–3639

11. Sestak I, Cuzick J, Sapunar F et al (2008) Risk factors for joint symptoms in patients enrolled in the ATAC trial: a retrospective, exploratory analysis. Lancet Oncol 9(866–872):8. https://doi.org/10.1016/S1470-2045(1008)70182-70187

12. Crew KD, Greenlee H, Capodice J et al (2007) Prevalence of joint symptoms in postmenopausal women taking aromatase inhibitors for early-stage breast cancer. J Clin Oncol 25(3877–3883):38. https://doi.org/10.1200/JCO.2007.3810.7573

13. Beckwée D, Leysen L, Meuwis K, Adriaenssens N (2017) Prevalence of aromatase inhibitor-induced arthralgia in breast cancer: a systematic review and meta-analysis. Support Care Cancer 25(1673–1686):16. https://doi.org/10.1007/s00520-00017-03613-z

14. Garcia-Giralt N, Rodriguez-Sanz M, Prieto-Alhambra D et al (2013) Genetic determinants of aromatase inhibitor-related arthralgia: the B-ABLE cohort study. Breast Cancer Res Treat 140:385–395

15. Fontein DB, Houtsma D, Nortier JW et al (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 144:599–606

16. Leyland-Jones B, Gray KP, Abramovitz M et al (2015) CYP19A1 polymorphisms and clinical outcomes in postmenopausal women with hormone receptor-positive breast cancer in the BIG 1–98 trial. Breast Cancer Res Treat 151:373–384

17. Wang J, Lu K, Song Y et al (2013) Indications of clinical and genetic predictors for aromatase inhibitors related musculoskeletal adverse events in Chinese Han women with breast cancer. PLoS ONE 8:e68798

18. Borrie AE, Rose FA, Choi YH et al (2020) Genetic and clinical predictors of arthralgia during letrozole or anastrozole therapy in breast cancer patients. Breast Cancer Res Treat 6:020–05777

19. Hertz DL, Henry NL, Rae JM (2017) Germline genetic predictors of aromatase inhibitor concentrations, estrogen suppression and drug efficacy and toxicity in breast cancer patients. Pharmacogenomics 18:481–499

20. Ingle JN, Schaid DJ, Goss PE et al (2010) Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. J Clin Oncol : Off J Am Soc Clin Oncol 28:4674–4682

21. Umamaheswaran G, Kadambari D, Muthuvel SK et al (2021) Polymorphisms of T- cell leukemia 1A gene loci are not related to the development of adjuvant letrozole-induced adverse events in breast cancer. PLoS One 16:e0247989. https://doi.org/10.1371/journal.pone.0247989 (**eCollection 0242021**)

22. 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 50:98–99

23. Henry NL, Skaar TC, Dantzer J et al (2013) Genetic associations with toxicity-related discontinuation of aromatase inhibitor therapy for breast cancer. Breast Cancer Res Treat 138:807–816

24. Hertz DL, Douglas JA, Kidwell KM et al (2021) Genome-wide association study of letrozole plasma concentrations identifies non-exonic variants that may affect CYP2A6 metabolic activity. Pharmacogenet Genomics 31:116–123. https://doi.org/10.1097/FPC.0000000000000429

25. Henry NL, Jacobson JA, Banerjee M et al (2010) A prospective study of aromatase inhibitor-associated musculoskeletal symptoms and abnormalities on serial high-resolution wrist ultrasonography. Cancer 116:4360–4367

26. Kadakia KC, Snyder CF, Kidwell KM et al (2016) Patient-Reported Outcomes and Early Discontinuation in Aromatase Inhibitor-Treated Postmenopausal Women With Early Stage Breast Cancer. Oncologist 21:539–546

27. Hertz DL, Douglas JA, Kidwell KM, et al. Genome-wide association study of letrozole plasma concentrations identifies non-exonic variants that may affect CYP2A6 metabolic activity. Pharmacogenetics and Genomics. 9000;Publish Ahead of Print.

28. Desta Z, Kreutz Y, Nguyen AT et al (2011) Plasma letrozole concentrations in postmenopausal women with breast cancer are associated with CYP2A6 genetic variants, body mass index, and age. Clin Pharmacol Ther 90:693–700

29. Das S, Forer L, Schönherr S et al (2016) Next-generation genotype imputation service and methods. Nat Genet 48(1284–1287):12. https://doi.org/10.1038/ng.3656

30. 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(562–571):5. https://doi.org/10.1111/bcp.13130

31. Niravath P, Chen B, Chapman JW et al (2018) Vitamin D Levels, Vitamin D Receptor Polymorphisms, and Inflammatory Cytokines in Aromatase Inhibitor-Induced Arthralgias: An Analysis of CCTG MA.27. Clin Breast Cancer 18:78–87. https://doi.org/10.1016/j.clbc.2017.1010.1009

32. Romero SAD, Su HI, Satagopan J et al (2020) Clinical and genetic risk factors for aromatase inhibitor-associated arthralgia in breast cancer survivors. Breast 49:48–54. https://doi.org/10.1016/j.breast.2019.1010.1008

33. Umamaheswaran G, Kadambari D, Muthuvel SK et al (2020) Association of CYP19A1 gene variations with adjuvant letrozole-induced adverse events in South Indian postmenopausal breast cancer cohort expressing hormone-receptor positivity. Breast Cancer Res Treat 182(147–158):1. https://doi.org/10.1007/s10549-10020-05656-10549

34. Machiela MJ, Chanock SJ (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking

correlated alleles of possible functional variants. Bioinformatics 31(3555–3557):35. https://doi.org/10.1093/bioinformatics/btv3402

35. (2013) The Genotype-Tissue Expression (GTEx) project. Nat Genet 45:580–585. https://doi.org/10.1038/ng.2653.

36. Boyle AP, Hong EL, Hariharan M et al (2012) Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22(1790–1797):17. https://doi.org/10.1101/gr.137323.137112

37. Dong S, Boyle AP (2019) Predicting functional variants in enhancer and promoter elements using RegulomeDB. Hum Mutat 40(1292–1298):12. https://doi.org/10.1002/humu.23791

38. Gaunt TR, Shihab HA, Hemani G et al (2016) Systematic identification of genetic influences on methylation across the human life course. Genome Biol 17:61. https://doi.org/10.1186/s13059-13016-10926-z

39. Dufresne J, Bowden P, Thavarajah T et al (2018) The plasma peptides of ovarian cancer. Clin Proteomics 15:41. https://doi.org/10.1186/s12014-12018-19215-z

40. Thomas J, Leufflen L, Chesnais V et al (2020) Identification of Specific Tumor Markers in Vulvar Carcinoma Through Extensive Human Papillomavirus DNA Characterization Using Next Generation Sequencing Method. J Low Genit Tract Dis 24:53–60. https://doi.org/10.1097/LGT.0000000000000498

41. Liu QR, Zhang PW, Zhen Q, Walther D, Wang XB, Uhl GR (2002) KEPI, a PKC-dependent protein phosphatase inhibitor regulated by morphine. J Biol Chem 277:13312–13320

42. Drgonova J, Zimonjic DB, Hall FS, Uhl GR (2010) Effect of KEPI (Ppp1r14c) deletion on morphine analgesia and tolerance in mice of different genetic backgrounds: when a knockout is near a relevant quantitative trait locus. Neuroscience 165(882–895):8. https://doi.org/10.1016/j.neuroscience.2009.1010.1007

43. Lang I, Virk G, Zheng DC et al (2020) The Evolution of Duplicated Genes of the Cpi-17/Phi-1 (ppp1r14) Family of Protein Phosphatase 1 Inhibitors in Teleosts. Int J Mol Sci 21(5709):57. https://doi.org/10.3390/ijms21165709

44. Webber JL, Tooze SA (2010) Coordinated regulation of autophagy by p38alpha MAPK through mAtg9 and p38IP. EMBO J 29:27–40. https://doi.org/10.1038/emboj.2009.1321

45. Liu X, Xiao W, Wang XD, Li YF, Han J, Li Y (2013) The p38-interacting protein (p38IP) regulates G2/M progression by promoting α-tubulin acetylation via inhibiting ubiquitination-induced degradation of the acetyltransferase GCN5. J Biol Chem. 288:36648–36661. https://doi.org/10.31074/jbc.M36113.486910

46. Veyssiere M, Perea J, Michou L et al (2019) A novel nonsense variant in SUPT20H gene associated with Rheumatoid Arthritis identified by Whole Exome Sequencing of multiplex families. PLoS One 14:e0213387. https://doi.org/10.1371/journal.pone.0213387 (eCollection 0212019)

47. Chua KC, Kroetz DL (2017) Genetic advances uncover mechanisms of chemotherapy-induced peripheral neuropathy. Clin Pharmacol Ther 101:450–452

48. Stearns V, Jegede O, Chang VT-S et al (2021) Prospective validation of genetic predictors of aromatase inhibitor-associated musculoskeletal symptoms (AIMSS) in a racially diverse cohort: Results from ECOG-ACRIN E1Z11. J Clin Oncol 39:12003–12003

49. Wang J, Lu K, Song Y et al (2015) RANKL and OPG Polymorphisms Are Associated with Aromatase Inhibitor-Related Musculoskeletal Adverse Events in Chinese Han Breast Cancer Patients. PLoS ONE 10:e0133964

50. Hertz DL, Smith KL, Zong Y et al (2021) Further Evidence That OPG rs2073618 Is Associated With Increased Risk of Musculoskeletal Symptoms in Patients Receiving Aromatase Inhibitors for Early Breast Cancer. Front Genet 12:662734. https://doi.org/10.3389/fgene.2021.662734

51. Basch E, Deal AM, Dueck AC et al (2017) Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 318:197–198

52. Gupta A, Henry NL, Loprinzi CL (2020) Management of Aromatase Inhibitor-Induced Musculoskeletal Symptoms. JCO Oncol Pract 16(733–739):7. https://doi.org/10.1200/OP.1220.00113

53. Henry NL, Unger JM, Schott AF et al (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 36:326–332

54. Santa-Maria CA, Bardia A, Blackford AL et al (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(121–129):1. https://doi.org/10.1007/s10549-10018-14811-10541

55. Burstein HJ, Prestrud AA, Seidenfeld J et al (2010) American Society of Clinical Oncology clinical practice guideline: update on adjuvant endocrine therapy for women with hormone receptor-positive breast cancer. J Clin Oncol : Off J Am Soc Clin Oncol 28:3784–3796

56. Siddiq A, Couch FJ, Chen GK et al (2012) A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. Hum Mol Genet 21:5373–5384

57. Ghoussaini M, French JD, Michailidou K et al (2016) Evidence that the 5p12 Variant rs10941679 Confers Susceptibility to Estrogen-Receptor-Positive Breast Cancer through FGF10 and MRPS30 Regulation. Am J Hum Genet 99:903–911

58. Michailidou K, Lindstrom S, Dennis J et al (2017) Association analysis identifies 65 new breast cancer risk loci. Nature 551:92–94

59. de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 11:e1004219

60. Reinbolt RE, Sonis S, Timmers CD et al (2018) Genomic risk prediction of aromatase inhibitor-related arthralgia in patients with breast cancer using a novel machine-learning algorithm. Cancer Med 7(240–253):2. https://doi.org/10.1002/cam1004.1256

Authors and Affiliations

Daniel L. Hertz[1] · Julie A. Douglas[2,3] · Robert M. Miller[2] · Kelley M. Kidwell[4] · Christina L. Gersch[5] · Zeruesenay Desta[6] · Anna Maria Storniolo[6] · Vered Stearns[7] · Todd C. Skaar[6] · Daniel F. Hayes[5] · N. Lynn Henry[5] · James M. Rae[5]

Julie A. Douglas
jdougla3@skidmore.edu

Robert M. Miller
rmiller2@skidmore.edu

Kelley M. Kidwell
kidwell@umich.edu

Christina L. Gersch
clgersch@med.umich.edu

Zeruesenay Desta
zdesta@iu.edu

Anna Maria Storniolo
astornio@iu.edu

Vered Stearns
vstearn1@jhmi.edu

Todd C. Skaar
tskaar@iu.edu

Daniel F. Hayes
hayesdf@med.umich.edu

N. Lynn Henry
norahh@med.umich.edu

James M. Rae
jimmyrae@med.umich.edu

[1] Department of Clinical Pharmacy, University of Michigan College of Pharmacy, 428 Church St., Room 3054, Ann Arbor, MI 48109-1065, USA

[2] Department of Mathematics and Statistics, Skidmore College, Saratoga Springs, NY 12866, USA

[3] Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI, USA

[4] Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI, USA

[5] Department of Internal Medicine, Division of Hematology/Oncology, University of Michigan Medical School, Ann Arbor, MI, USA

[6] Indiana University School of Medicine, Indianapolis, IN, USA

[7] Johns Hopkins School of Medicine, Baltimore, MD, USA