SYSTEMATIC REVIEW

Longitudinal Effects of Adjuvant Endocrine Therapy on the Quality of Life of Post-menopausal Women with Non-metastatic ER+ Breast Cancer: A Systematic Review

Hong Xiao¹ · Xinyi Jiang² · Cheng Chen² · Alberto J. Montero³ · Vakaramoko Diaby⁴

Published online: 22 February 2018

© The Author(s) 2018. This article is an open access publication

Abstract

Background Anti-estrogen (ER) endocrine therapy is an effective treatment strategy in reducing breast cancer mortality. This therapy has a better therapeutic index than chemotherapy but can still affect patients' quality of life (QOL) over time.

Objective The objectives of this systematic review were to (1) describe QOL instruments used in ER-positive (ER+) non-metastatic breast cancer trials and (2) document the longitudinal effects of adjuvant endocrine therapy on the QOL of post-menopausal women with ER+ non-metastatic breast cancer.

Methods We searched three electronic bibliographic databases for articles published from inception to October 2017 that described (1) a randomized controlled trial (RCT) of non-metastatic breast cancer containing an adjuvant endocrine regimen in at least one arm; (2) the use of a patient self-report measure assessing general or breast

cancer-specific QOL; and (3) QOL outcomes at multiple time points during follow-up of at least 5 years. All included trials were independently evaluated by two reviewers, and data were extracted using standardized forms.

Results In total, 13 studies met our inclusion criteria and were assessed in this review. The quality of the trials was reasonably good. The top three most commonly used QOL instruments in the trials were the Functional Assessment of Cancer Therapy/Functional Assessment of Chronic Illness Therapy, the Short Form-36 and the Menopause-Specific Quality of Life. Most studies found no differences between tamoxifen and aromatase inhibitor groups in terms of global QOL. QOL data affected treatment regimen recommendations in a few cases. A meta-analysis was not feasible because the RCTs included in our review varied in terms of sample size, comparators, QOL instrument used, and timing of QOL measurement. Additionally, as no search strategy has perfect sensitivity, specificity and accuracy, there is always a chance that potentially relevant articles were missed.

Conclusion This systematic review suggests that the QOL of post-menopausal women is unlikely to be adversely affected by long-term use of adjuvant endocrine therapy. Efforts are needed to improve the quality of QOL reporting in clinical trials.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41669-018-0070-7>) contains supplementary material, which is available to authorized users.

✉ Hong Xiao
hxiao18@ufl.edu

¹ Department of Pharmaceutical Outcomes and Policy, University of Florida College of Pharmacy, HPNP Building Room 3338, 1225 Center Drive, Gainesville, FL 32610, USA

² Department of Pharmaceutical Outcomes and Policy, University of Florida College of Pharmacy, HPNP 2309, 1225 Center Drive, Gainesville, FL 32610, USA

³ Cleveland Clinic Foundation, Taussig Cancer Institute, 9500 Euclid Ave, R35, Cleveland, OH 44195, USA

⁴ Department of Pharmaceutical Outcomes and Policy, University of Florida College of Pharmacy, HPNP 3317, 1225 Center Drive, Gainesville, FL 32610, USA

Key Points for Decision Makers

Thirteen randomized controlled trials (RCTs) assessing the longitudinal effects of adjuvant endocrine therapy on quality of life (QOL) in post-menopausal women with non-metastatic estrogen-positive breast cancer used a variety of QOL instruments; The FACT/FACIT, MENQOL and SF-36 were the most common.

Most studies found no statistically significant differences between tamoxifen and aromatase inhibitor groups in terms of global QOL, although—in a few cases—tamoxifen exhibited a better QOL profile in the early stages of treatment.

The QOL of post-menopausal women is unlikely to be adversely affected by long-term use of adjuvant endocrine therapy.

1 Introduction

Globally, breast cancer is the most common cancer in women and the second leading cause of cancer mortality [1]. There exist three major clinically relevant subtypes: estrogen receptor positive (ER+), human epidermal growth factor receptor (HER)-2 positive and triple negative breast cancer, with ER+ breast cancer being the most prevalent subset, representing more than 75% of all breast cancer [2].

Anti-estrogen endocrine therapy has been purported in several meta-analyses to be effective in reducing breast cancer mortality [3, 4]. Recent estimates suggest that 5 years of adjuvant anti-estrogen therapy reduces cancer mortality by about 33% [4]. There are two major classes of anti-estrogen therapy: (1) selective ER modulators (SERMs), e.g. tamoxifen, and (2) aromatase inhibitors (AIs). SERMs and AIs have major side effects, including vasomotor symptoms, myalgias/arthralgias, sexuality disorder and thrombo-embolic toxicities [5, 6]. Although endocrine therapy has a much more favourable therapeutic index than chemotherapy, the chronicity of endocrine therapy can have a major impact on patients' quality of life (QOL). For this reason, understanding the impact of adjuvant endocrine therapy on patient-reported QOL is clinically relevant and important.

Lemieux et al. [7] published a systematic review in 2011 focusing on QOL measurement. This was an updated review of previous work by Goodwin et al. [8] and included 190 randomized controlled trials (RCTs) conducted

from 2001 to 2009 that measured QOL as an outcome in women of any age with breast cancer receiving any biomedical or non-biomedical intervention. The authors concluded that the reporting of QOL methodology could be improved. In a similar systematic review, Montazeri [9] included 477 papers ranging from descriptive studies to clinical trials that documented health-related QOL (HR-QOL) in patients with breast cancer from 1974 to 2007. The author concluded that a better understanding of HR-QOL in patients with breast cancer could be achieved by more qualitative research focusing on treatment side effects and symptoms and sexual functioning. Although very informative, the evidence synthesized in these studies may be dated and focusses on heterogeneous populations and diverse treatment options.

We performed a systematic review of published prospectively collected QOL data on adjuvant endocrine therapy, either AI or tamoxifen, in post-menopausal women with non-metastatic stage 1–3 ER+ breast cancer, where the therapeutic intent was curative. In the metastatic setting (i.e. stage 4), the therapeutic intent is palliative, i.e. delay of tumour progression, so we limited our literature review to non-metastatic ER+ breast cancer. The objectives of this systematic review were to (1) describe QOL instruments used in ER+ non-metastatic breast cancer trials and (2) document the longitudinal effects (treatment follow-up ≥ 5 years) of adjuvant endocrine therapy on the QOL of post-menopausal women with ER+ non-metastatic breast cancer.

2 Methods

2.1 Literature Search

We performed a systematic review of RCTs reporting the longitudinal effects (≥ 5 years) of adjuvant endocrine therapy with either an AI or tamoxifen in post-menopausal women with non-metastatic stage 1–3 ER+ breast cancer. Relevant studies were identified through an initial search of three electronic data sources from inception until 30 September 2016: Ovid MEDLINE, Cochrane Central Register of Controlled Trials and the US National Library of Medicine's PubMed using controlled vocabulary as well as all applicable search terms. An update search was performed on 30 October 2017. The search was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement criteria [10]. We used the following terms: breast neoplasms, breast cancer, breast carcinoma, anti-estrogen therapy, endocrine therapy, aromatase inhibitor, tamoxifen, letrozole, anastrozole, exemestane, quality of life, value of life, and females. We also used Boolean operators, truncation

Fig. 1 Flow diagram showing the study-selection process. *DCIS* ductal carcinoma in situ, *RCTs* randomized controlled trials, *SDs* standard deviations

and wildcard operators (Appendix 1 in the Electronic Supplementary Material). Additionally, we manually searched the reference lists of included articles and consulted with experts in the breast cancer field to increase the likelihood of identifying further relevant articles.

2.2 Review Selection Process

The bibliographic records obtained via the literature search were exported into Endnote. Figure 1 shows the process by which articles were selected for this systematic review. First, we removed duplicates from our original sample of bibliographic records, then two reviewers (XJ and CC) independently screened the titles and abstracts and excluded those that did not meet the eligibility criteria. XJ and CC then independently assessed the full text of the remaining articles against predetermined inclusion criteria. We tested the inter-rater agreement for the reviewers' assessments of study selection by calculating the Cohen's kappa coefficient of agreement (percentage agreement), κ [11]. Disagreements were resolved via discussion with a third reviewer (VD).

Studies were included only if they described (1) an RCT of non-metastatic breast cancer with at least one arm including an adjuvant endocrine regimen; (2) the use of a patient self-report measure that assessed general or breast

cancer-specific QOL; and (3) QOL outcomes at multiple time points during follow-up of at least 5 years (followed longitudinally). Studies were excluded if they (1) described only a methodological approach; (2) were duplicates using data from the same patients; (3) enrolled patients with stage 0/ductal carcinoma in situ (DCIS) or advanced/metastatic breast cancer.

2.3 Data Extraction

XJ and CC extracted and synthesized (numerically) the data meeting the inclusion criteria using published predefined extraction forms [7]. These forms present criteria that relate to the quality of the RCTs using the Jadad scoring system [12], the summary of the QOL instruments used, and findings from QOL studies. The following criteria were used: trial name and type (superiority, non-inferiority, equivalence), geographical location, treatment arms description, mean age of patient population, QOL endpoint (primary or secondary), QOL instruments, trial and QOL study sample sizes, statistical power for QOL reporting, statistical method for QOL analysis, timing of QOL measurements, QOL study findings, and clinical significance of QOL study findings.

3 Results

3.1 Literature Search

Figure 1 shows the systematic search process. The literature search yielded 1031 bibliographic records from the three databases (871 and 160 bibliographic records obtained, respectively, from the initial and updated database searches). Tables in Appendix 1 show the search strategy and results from Ovid MEDLINE, the Cochrane Central Register of Controlled Trials and PubMed, respectively. We removed 230 duplicates from the 1031 records. Following review of titles and abstracts, 706 studies were rejected as they enrolled patients with stage 0/DCIS or advanced/metastatic breast cancer. Of the remaining records subject to full-text review, 82 were removed for a range of reasons, including that they had a follow-up time of <5 years or were conducted in the metastatic setting and in non-adjuvant therapies. The final set eligible for inclusion included 13 bibliographic records.

Inter-rater agreement for reviewer assessments of study selection were very high: 0.92 for title and abstract screening, indicating almost perfect agreement, and 0.82 for full-text screening, indicating strong agreement [11].

3.2 Quality Assessment of the Randomized Controlled Trials

Table 1 shows the results of the quality assessment of the trials. The Jadad score was calculated for each of the 13 trials based on randomization, blinding and an account of all patients [12]. Seven studies (53.8%) [13–19] scored 3, one (7.7%) [20] scored 4 and five (38.4%) scored 5

Table 2 Summary of quality-of-life instruments used in clinical trials in this systematic review

QOL instrument	N (%) of trials using instrument ^a
Total trials (N)	13
FACT/FACIT	5 (38.4)
SF-36	5 (38.4)
MENQOL	4 (30.7)
CES-D	2 (15.3)
SCL	1 (7.6)
LASA	1 (7.6)
RSCL	1 (7.6)
EORTC/QLQ-C30	1 (7.6)
PACIS	1 (7.6)
VAS	1 (7.6)

CES-D Center for Epidemiological Studies–Depression Scale, *EORTC QLQ* European Organization for the Research and Treatment of Cancer Quality-of-Life Questionnaire, *FACT/FACIT* Functional Assessment of Cancer Therapy/Functional Assessment of Chronic Illness Therapy, *LASA* linear analogue scale assessment, *MENQOL* Menopause-Specific Quality of Life, *PACIS* Personal Adjustment to Chronic Illness Scale, *RSCL* Rotterdam Symptom Checklist, *SCL* Symptom Checklist Depression Scale, *SF-36* Short form-36, *VAS* visual analogue scale

^aSome trials used more than one instrument. Different variants of the same questionnaire (e.g. FACT-G, FACT-B) was counted as one questionnaire

[21–25], indicating that the studies included in this systematic review were of reasonably good overall quality.

3.3 Description of the Quality of Life (QOL) Instruments in the Clinical Trials

Our review included 13 trials [13–25] reporting on different QOL domains in patients with non-metastatic ER+

Table 1 Quality assessment of clinical trials

References	Randomization (max. 2)	Blinding (max. 2)	An account of all pts (max. 1)	Total Jadad scale (max. 5)
Buijs et al. [13]	2	0	1	3
Cella et al. [21]	2	2	1	5
Crivellari et al. [14]	2	0	1	3
Fallowfield et al. [22]	2	2	1	5
Francini et al. [15]	2	0	1	3
Goss et al. [23]	2	2	1	5
Land et al. [16]	2	0	1	3
Mamounas et al. [20]	2	1	1	4
Muss et al. [24]	2	2	1	5
Ohsumi et al. [17]	2	0	1	3
Pagani et al. [18]	2	0	1	3
Takei et al. [19]	2	0	1	3
Whelan et al. [25]	2	2	1	5

max. maximum, pts patients

breast cancer. Table 2 summarizes the QOL instruments used in these trials: Functional Assessment of Cancer Therapy/Functional Assessment of Chronic Illness Therapy (FACT/FACIT), Short Form-36 (SF-36), Menopause-Specific Quality of Life (MENQOL), Center for Epidemiological Studies–Depression Scale (CES-D), Symptom Checklist Depression Scale (SCL), Linear Analogue Self-Assessment (LASA), Rotterdam Symptom Checklist (RSCL), European Organization for the Research and Treatment of Cancer Quality-of-Life Questionnaire (EORTC QLQ), Personal Adjustment to Chronic Illness Scale (PACIS), and the visual analogue scale (VAS). A description of each instrument and the extent to which they were used follows.

The FACT measurement system was developed by the FACIT organization [26] and was intended to score and interpret HR-QOL for people with chronic diseases. It includes FACT-G (27 items) for general QOL, which measures four subdomains (physical well-being, social/family well-being, emotional well-being and functional well-being); FACT-B, which is FACT-G plus the nine-item Breast Cancer Subscale (BCS); and FACT-ES, which is FACT-G plus the 18-item endocrine subscale. The total possible score ranges for FACT-G, FACT-B and FACT-ES are 0–108, 0–36 and 0–72, respectively. Higher scores indicate better QOL. The FACT/FACIT questionnaire, along with its variants, was one of the most commonly used (38.4%) QOL instruments in the selected trials [16, 17, 19, 21, 22].

SF-36 is a 36-item questionnaire that measures overall QOL for the general population [27] across eight physical and emotional domains: physical function (ten items), role-physical (four items), bodily pain (two items), general health (five items), vitality (four items), social functioning (two items), role-emotional (three items) and mental health (five items). Each scale is directly transformed into a scale of 0–100, assuming each question carries an equal weight. Higher scores indicate better QOL. The SF-36 questionnaire was the second of the two most commonly used (38.4%) questionnaires in the trials [13, 16, 23–25].

MENQOL is a tool to assess HR-QOL in the immediate post-menopausal period [28]. It is a 29-item self-administered questionnaire in which each item assesses the impact of one of four domains of menopausal symptoms: vasomotor, psychosocial, physical and sexual. The average for each domain scores between 1 (not at all a problem) and 8 (experiencing symptom in the domain at the highest degree of bother). In this review, MENQOL was used in 30.7% of the trials, making it the second most commonly used QOL instrument [20, 23–25].

CES-D is a 20-item self-administered tool that measures symptoms associated with depression [29]. The total score can range from 0 to 60 and is obtained by summing the

individual score for each item. A responder is said to be at risk for clinical depression if the total score is ≥ 16 . In total, 15.3% of the trials use the CES-D to measure a dimension of QOL [17, 19].

The RSCL is a self-reported questionnaire that measures four main dimensions of QOL in patients with cancer: (1) physical symptom distress, (2) psychological distress, (3) activity level and (4) overall global life quality [30]. The scores for these dimensions range from 23 to 92, 7 to 28, 1 to 8, and 1 to 7 for dimensions 1–4, respectively. High scores correlate with poor QOL, except for activity, for which higher scores correlate with better QOL. The RSCL was used in 7.6% of the trials selected for this systematic review [13].

The EORTC QLQ C30 was developed to assess the QOL of patients with cancer and comprises five functioning scales (i.e., physical functioning, role functioning, cognitive functioning, emotional functioning and social functioning), a global health status/QOL scale, three symptom scales (fatigue, pain and nausea/vomiting), and six single items (dyspnoea, appetite loss, sleep disturbance, constipation, diarrhoea and perceived financial impact). Each scale of the QLQ C30 is converted to a 100-point scale [31]. High scores represent higher response levels. For instance, a high score for a functional scale indicates a high/healthy level of functioning, whereas a high score for a symptom scale corresponds to more frequent or more intense symptoms. The EORTC QLQ C-30 was used in 7.6% of the trials selected in this systematic review [15].

PACIS is a tool used in clinical trials to assess coping in patients with operable breast cancer [32]. The score on this scale ranges from 0 (no effort at all to cope) to 100 (a great deal of effort to cope). Only 7.6% of the selected trials used the PACIS [14].

The VAS is a commonly used 100-millimeter horizontal or vertical scale with verbal anchors at the low and high endpoints. It is used to measure various subjective experiences such as pain on a continuum of values [33]. VAS scores are obtained by measuring the distance between the low endpoint of the line and the point representing the intensity of the responder's experience. A VAS was used in 7.6% of the trials included in this systematic review [13].

LASA is commonly used to measure functioning, overall QOL and other symptoms [34]. It resembles a 100-millimeter thermometer (line) with descriptors at opposite extremes (0 low endpoint or anchor and 100 upper endpoint). For each respondent, a score is obtained by measuring the distance (mm) between the low endpoint and the patient's mark. LASA was used in 7.6% of the trials included in this systematic review [18].

Table 3 Basic information of included studies

Study	Region	Trial name	Arms	Treatment duration (y)	QOL endpoint	QOL instrument ^a	QOL study/trial samples (N/N)	Statistical power for QOL reporting	Statistical method for QOL analysis	Timing of QOL measures
Buijs et al. 2007 [13]	Netherlands	Unknown	High-dose CTX + RTX + TAM vs. conventional CTX + RTX + TAM	5	Secondary	SF-36, VAS, RSCL	885/804	No	Yes	After CTX completion, after RTX completion, every 6 mo thereafter
Cella et al. 2006 [21]	Multinational	ATAC	ANA vs. TAM vs. ANA + TAM	5	Secondary	FACT-B, FACT-ES	1105/9366	Yes	Yes	BL, 3 and 6 mo, every 6 mo thereafter
Crivellari et al. 2000 [14]	Multinational	IBCSG Trial VII	TAM alone vs. TAM + CMF	5	Secondary	PACIS	NR/608	No	No	2 mo, then every 3 mo until 24 mo
Fallowfield et al. 2004 [22]	Multinational	ATAC	ANA vs. TAM vs. ANA + TAM	5	Secondary	FACT-B, FACT-ES	1105/9366	Yes	Yes	BL, 3 and 6 mo, every 6 mo thereafter
Francini et al. 2006 [15]	Unknown	Unknown	2–3 y TAM → EXE vs. 5 y TAM	5	Secondary	EORTC QLQ-C30	NR/60	No	Yes	BL, after 6 and 12 mo of tx
Goss et al. 2016 [23]	Canada, USA	MA-17R	5 y AI + 5 y LET vs. 5 y AI + PL	10	Secondary	SF-36, MENQOL	1630/1918	No	No	BL and 12, 24, 36, 48 and 60 mo
Land et al. 2004 [16]	USA	NSABP B-23	CMF + 5 y TAM vs. CMF + PL vs. AC + TAM vs. AC + PL	5	Secondary	FACT-B, SCL, SF-36	160/2008	Yes	Yes	Beginning of each CTX cycle and at several FU time points
Mamounas et al. 2008 [20]	USA	NSABP B-33	5 y TAM → EXE vs. 5 y TAM → PL	5	Secondary	MENQOL	454/1598	Yes	Yes	BL and through 24 mo of FU
Muss et al. 2008 [24]	Multinational	MA.17	5 y TAM → LET vs. 5 y TAM → PL	5	Secondary	SF-36, MENQOL	3612/5187	No	Yes	BL, 24 mo
Ohsumi et al. 2011 [17]	Japan	N-SAS BC 03	1–4 y TAM → ANA vs. TAM	5	Secondary	FACT-B, FACT-ES, CES-D, FACT-G	694/NR	Yes	Yes	BL, 3 mo, 1 and 2 y
Pagani et al. 2004 [18]	Multinational	IBCSG trials 12-93 and 14-93	5 y TOR vs. 5 y TAM	5	Secondary	LASA	697/1037	No	Yes	BL, mo 3, 6, 9, 12, 18 and 24, yearly thereafter for 6 y
Takei et al. 2012 [19]	Japan	N-SAS BC 04	TAM vs. EXE vs. ANA	5	Secondary	FACT-B, FACT-ES, CES-D, FACT-G	166/NR	Yes	Yes	BL, 3 and 12 mo

Table 3 continued

Study	Region	Trial name	Arms	Treatment duration (y)	QOL endpoint	QOL instrument ^a	QOL study/trial samples (N/N)	Statistical power for QOL reporting	Statistical method for QOL analysis	Timing of QOL measures
Whelan et al. 2005 [25]	Multinational	MA.17	5 y TAM → LET vs. 5 y TAM → PL	5	Secondary	SF-36, MENQOL	3612/5187	No	Yes	Semi-annually during y 1, annually thereafter

AC adriamycin + cytoxan, AI aromatase inhibitors, ANA anastrozole, BL baseline, CMF cyclophosphamide, methotrexate, fluorouracil, CTX chemotherapy, EXE exemestane, FU follow-up, LET letrozole, mo months, NR not reported, PL placebo, QOL quality of life, RTX radiotherapy, TAM tamoxifen, TOR toremifene, tx treatment, y year(s)

^aRefer to Table 2 and the main text for full names of instruments

3.4 Main Findings Regarding QOL in Clinical Trials in Non-metastatic Estrogen Receptor-Positive Breast Cancer

Tables 3 and 4 summarize the characteristics of the included clinical trials and their main findings regarding QOL. All studies involved AI therapy in post-menopausal women with non-metastatic (stage 1–3) breast cancer in at least one treatment arm and were published in or after the year 2000. The RCTs covered a variety of anti-estrogen treatment regimens ranging from monotherapy to combinations of the following treatments: tamoxifen, letrozole, anastrozole and exemestane. Interventions were compared with another anti-estrogen therapy, chemotherapy, radiation therapy, toremifene or placebo. The trials were conducted in North America [16, 20, 23], Europe [13], and Asia [17, 19], as well as in multiple regions [14, 18, 21, 22, 24, 25]. In total, 12 trials followed patients for 5 years, and one study had a follow-up time of 10 years. In all trials, QOL was measured as a secondary endpoint in a sub-population of the trial. The mean age of patients included in the clinical trials ranged from 59 to 65.1 years, and 38.4% of the trials reported power calculations for QOL outcomes. The main findings in terms of QOL outcomes are presented per treatment groups.

3.5 Hormonal Therapy vs. Hormonal Therapy

Eight studies in this systematic review investigated a hormonal therapy compared with another hormonal therapy [15, 17–19, 21–23, 25].

Cella et al. [21] and Fallowfield et al. [22] compared 5 years of anastrozole alone, tamoxifen alone, and anastrozole plus tamoxifen. A slight improvement was seen in the overall QOL of all three groups during the 2-year period. No statistically significant difference was seen across intervention groups across the 5 years. Endocrine scores for both arms decreased initially and recovered partially during the first 2 years after initiation of therapy. However, side effect profiles were reported in different treatment groups. Patients receiving an AI experienced more sexual side effects but fewer cold sweats than those receiving tamoxifen.

Francini et al. [15] investigated tamoxifen compared with AIs and found no statistically significant difference in global health status between groups, measured with the EORTC QLQ-C30.

Ohsumi et al. [17] compared tamoxifen versus AIs and found that total FACT-G scores were better in the tamoxifen group than in the anastrozole group. The scores in the tamoxifen group were stable over time, whereas those in the anastrozole group declined but not significantly. Additionally, FACT-B scores remained stable, without any clinically significant change [17].

Table 4 Main findings regarding quality of life

Study	Mean age (y)	Main QOL findings ^a	Clinical significance of QOL study findings
Buijs et al. 2007 [13]	NA	Shortly after high-dose chemotherapy, HR-QOL was negatively affected compared with conventional-dose chemotherapy. 1 year after random assignment, differences were negligible	Yes
Cella et al. 2006 [21]	63.6	Following an initial worsening of the mean ES score at 3 months in both groups, the ES score stabilized, although it never recovered to baseline levels, even at 5 years	Yes
Crivellari et al. 2000 [14]	NA	All treatment groups showed substantial improvement in QOL scores during adjuvant therapy. Longer initial cytotoxic therapy delayed improvement in QOL scores. Later cytotoxic therapy had transient adverse effects	No
Fallowfield et al. 2004 [22]	63.6	Following an initial worsening of the mean ES score at 3 months in both groups, the ES score stabilized, although it never recovered to baseline levels, even at 5 years	Yes
Francini et al. 2006 [15]	61.5	Compared with baseline, none of the p-values in the functional and symptom scales showed any clear between-group trend at month 6 but, at month 12, there was a trend toward improved physical functioning, global health status and global QOL scores in the exemestane group; the between-group difference was not statistically significant	No
Goss et al. 2016 [23]	65.1	No statistically significant between-group differences were observed in the SF-36 summary scores on any of the four MENQOL symptom subscales	Yes
Land et al. 2004 [16]	50.6% < 49	The effects of tamoxifen arm, surgery, tumour size group, and age were either not statistically significant or were of negligible magnitude	No
Mamounas et al. 2008 [20]	NA	No significant treatment effects in the vasomotor ($p = 0.87$), psychosocial ($p = 0.27$), physical ($p = 0.13$), or sexual ($p = 0.23$) scales	No
Muss et al. 2008 [24]	62.23	There was no difference in QOL at 24 months among letrozole- and placebo-treated patients aged ≥ 70 years	Yes
Ohsumi et al. 2011 [17]	NA	Total scores of FACT-G, FACT-ES, and the FACT-G physical well-being subscale were statistically significantly better in the tamoxifen group than in the anastrozole group ($\alpha = 0.042$, 0.038 , and 0.005 , respectively). Total FACT-G scores declined in the anastrozole group at 1 year and continued to do so until 2 years, whereas the scores of the tamoxifen group were generally stable in all questionnaires	Yes
Pagani et al. 2004 [18]	59	No significant difference between toremifene and tamoxifen emerged on any of the ten items assessed at the different time points, e.g. physical well-being and hot flushes. Physical well-being scores declined initially, possibly due to chemotherapy, but then recovered. LASA scores for hot flushes also declined initially, indicating decreased QOL associated with hot flushes, but began to recover with increasing follow-up after month 12	No
Takei et al. 2012 [19]	63	In the tamoxifen group, FACT-G and BCS scores increased after treatment began. FACT-B scores increased after treatment began and remained significantly higher in the tamoxifen group than in the exemestane or anastrozole groups for 1 year ($p = 0.045$). ES scores were largely unchanged in all three treatment groups, and there was no significant difference between any groups ($p = 0.36$). In all patients assigned to exemestane or tamoxifen, FACT-B scores increased after treatment began and remained significantly higher in the tamoxifen group than in the exemestane group for 1 year ($p = 0.047$)	Yes
Whelan et al. 2005 [25]	62.23	Mean change in QOL scores for minority women who received letrozole demonstrated improved mental health at the 6-month assessment ($p = 0.02$) and less bodily pain at the 12-month assessment ($p = 0.046$)	Yes

ES endocrine subscale, HR-QOL health-related quality of life, NA not applicable, QOL quality of life

^aRefer to Table 2 and the main text for full names of instruments

Whelan et al. [25] compared tamoxifen versus placebo. No statistically significant between-group differences were found in summary scores measured with the

SF-36, although differences between interventions might exist across dimensions measured by the QOL instrument.

Goss et al. [23] studied 5 years of AIs + letrozole compared with 5 years' AI alone using MENQOL. The authors reported no significant treatment effects on domains of MENQOL over time [20, 23, 24].

Pagani et al. [18] investigated toremifene compared with tamoxifen using the LASA. No significant difference between toremifene and tamoxifen was found on any of the ten items assessed. LASA scores for hot flushes declined initially, indicating decreased QOL associated with hot flushes, but began to recover with further follow-up after month 12.

Takei et al. [19] studied tamoxifen compared with two types of AIs. Scores of overall QOL increased after treatment began. The tamoxifen group achieved better QOL than the AI group in the first year. Endocrine-related symptoms remained unchanged in all groups and did not differ between the interventions. In addition, this study used FACT-G scores to measure global QOL among those receiving exemestane versus tamoxifen versus anastrozole and found no significant differences across groups. FACT-G scores increased after treatment began for the tamoxifen group. However, the score change over time for the tamoxifen group was a maximum of four, meaning there were no significant changes for the tamoxifen group over time [19]. The N-SAS BC 04 trial included three arms: two AIs alone versus tamoxifen alone [19]. FACT-B scores increased after treatment began and remained significantly higher in the tamoxifen group than in the AI group for 1 year.

3.6 Hormonal Therapy + Chemotherapy vs. Hormonal Therapy

Crivellari et al. [14] used the PACIS to study QOL with tamoxifen alone compared with tamoxifen + the combination of cyclophosphamide, methotrexate and fluorouracil (CMF). All treatment groups showed substantial improvement in QOL scores during adjuvant therapy. Longer initial cytotoxic therapy delayed improvement in QOL scores.

Mamounas et al. [20] used the MENQOL to study the effects of 5 years' tamoxifen followed by exemestane compared with 5 years' tamoxifen alone. They found no significant treatment effects in the vasomotor ($p = 0.87$), psychosocial ($p = 0.27$), physical ($p = 0.13$) or sexual ($p = 0.23$) scales.

Muss et al. [24] investigated 5 years' tamoxifen followed by exemestane compared with 5 years' letrozole alone using the SF-36 and MENQOL. There was no difference in QOL at 24 months among letrozole- and placebo-treated patients aged ≥ 70 years [24].

3.7 Hormonal Therapy + Chemotherapy vs. Hormonal Therapy + Chemotherapy

Buijs et al. [13] investigated a high-dose chemotherapy regimen with radiotherapy and tamoxifen compared with a conventional chemotherapy regimen with radiotherapy and tamoxifen. The time horizon was 5 years, with QOL measured using the SF-36, VAS and RSCL. The authors found HR-QOL to be negatively affected compared with conventional-dose chemotherapy. Differences were insignificant 1 year post-randomization.

Land et al. [16] conducted a study of CMF + 5 years of tamoxifen versus CMF + placebo versus adriamycin + cyclophosphamide (AC) + tamoxifen compared with AC + placebo. The QOL instruments used were FACT-B, SCL and SF-36. The results of the study showed that the effects of the tamoxifen arm, surgery, tumour size group, and age were either not statistically significant or were of negligible magnitude.

4 Discussion

This study aimed to describe QOL instruments used in ER+ non-metastatic breast cancer trials, documenting the long-term effects of adjuvant endocrine therapy on the QOL of post-menopausal women with ER+ non-metastatic breast cancer. Therefore, a systematic review study design was adopted to retrieve citation data from three major electronic databases from their inception to 2017. A variety of QOL instruments have been used to capture different dimensions of QOL in ER+ non-metastatic breast cancer trials, with FACT/FACIT, MENQOL and SF-36 being the most common. In addition, most studies found no statistically significant differences between tamoxifen and AI groups in terms of global QOL, although—in a few cases—tamoxifen exhibited a better QOL profile in the early stages of treatment.

The results of this systematic review are concordant with those found in previous systematic studies covering similar topics [7, 9], except that our study focused on post-menopausal women with ER+ non-metastatic breast cancer. This systematic review was built on structured search and data extraction strategies that enabled the retrieval of bibliographic records from several electronic databases over an up-to-date and wide-ranging time horizon. Our study is prone to several limitations that warrant highlighting. No search strategy is perfect in terms of sensitivity, specificity and accuracy, meaning there is a chance we missed potentially relevant articles. As acknowledged by Lemieux et al. [7], the ability to clearly establish that QOL results contributed to the “trial authors' decision to recommend the use of” a treatment regimen was not straightforward in all cases. Given the variation among the

RCTs included in our review in terms of sample size, comparators, QOL instrument used, and timing of QOL measurement, a meta-analysis was not feasible.

Despite the above limitations, this review is the first to focus on QOL and its change over time (≥ 5 years) in postmenopausal women with non-metastatic breast cancer who have received adjuvant endocrine therapy with either tamoxifen or AI therapy.

This systematic review suggests that the QOL of postmenopausal women is unlikely to be adversely affected by the long-term use of adjuvant endocrine therapy (AI over tamoxifen). Nonetheless, efforts are needed to harmonize the use of QOL instruments and the reporting of QOL data to enable quantitative assessment of QOL for patients receiving adjuvant endocrine therapy and therefore empower clinicians and patients in their shared decision making.

Author Contributions Hong Xiao, Alberto J. Montero and Vakaramoko Diaby contributed to the conception and design of the study. Xinyi Jiang, Cheng Chen, Vakaramoko Diaby and Hong Xiao contributed to data acquisition, analysis and interpretation. All authors participated in drafting and revising the manuscript and provided final approval of the version to be submitted.

Data Availability Statement All data generated or analysed during this study are included in this published article (and its supplementary information files).

Compliance with Ethical Standards

Funding This study was partially supported by the University of Florida College of Pharmacy PROSPER award.

Conflict of interest Hong Xiao, Xinyi Jiang, Cheng Chen, Alberto J. Montero, and Vakaramoko Diaby have no conflicts of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. American Cancer Society. Breast cancer facts & figures 2015–2016. Atlanta: American Cancer Society, Inc.; 2015. <http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-046381.pdf>.
2. Blows FM, Driver KE, Schmidt MK, Brooks A, Van Leeuwen FE, Wesseling J, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med.* 2010;7(5):e1000279.
3. Group EBCTC. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet.* 2011;378(9793):771–84.
4. Group EBCTC. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet.* 2005;365(9472):1687–717.
5. Winer EP, Hudis C, Burstein HJ, Wolff AC, Pritchard KI, Ingle JN, et al. American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for postmenopausal women with hormone receptor-positive breast cancer: status report 2004. *J Clin Oncol.* 2005;23(3):619–29.
6. Riggs BL, Hartmann LC. Selective estrogen-receptor modulators—mechanisms of action and application to clinical practice. *N Engl J Med.* 2003;348(7):618–29.
7. Lemieux J, Goodwin PJ, Bordeleau LJ, Lauzier S, Theberge V. Quality-of-life measurement in randomized clinical trials in breast cancer: an updated systematic review (2001–2009). *J Natl Cancer Inst.* 2011;103(3):178–231.
8. Goodwin PJ, Black JT, Bordeleau LJ, Ganz PA. Health-related quality-of-life measurement in randomized clinical trials in breast cancer—taking stock. *J Natl Cancer Inst.* 2003;95(4):263–81.
9. Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review of the literature from 1974 to 2007. *J Exp Clin Cancer Res.* 2008;27:32.
10. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA StatementThe PRISMA statement. *Ann Intern Med.* 2009;151(4):264–9.
11. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med.* 2012;22(3):276–82.
12. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clin Trials.* 1996;17(1):1–12.
13. Buijs C, Rodenhuis S, Seynaeve CM, van Hoesel QG, van der Wall E, Smit WJ, et al. Prospective study of long-term impact of adjuvant high-dose and conventional-dose chemotherapy on health-related quality of life. *J Clin Oncol.* 2007;25(34):5403–9.
14. Crivellari D, Bonetti M, Castiglione-Gertsch M, Gelber RD, Rudenstam CM, Thurlimann B, et al. Burdens and benefits of adjuvant cyclophosphamide, methotrexate, and fluorouracil and tamoxifen for elderly patients with breast cancer: the International Breast Cancer Study Group Trial VII. *J Clin Oncol.* 2000;18(7):1412–22.
15. Francini G, Petrioli R, Montagnani A, Cadirni A, Campagna S, Francini E, et al. Exemestane after tamoxifen as adjuvant hormonal therapy in postmenopausal women with breast cancer: effects on body composition and lipids. *Br J Cancer.* 2006;95(2):153–8.
16. Land SR, Kopec JA, Yothers G, Anderson S, Day R, Tang G, et al. Health-related quality of life in axillary node-negative, estrogen receptor-negative breast cancer patients undergoing AC versus CMF chemotherapy: findings from the National Surgical Adjuvant Breast and Bowel Project B-23. *Breast Cancer Res Treat.* 2004;86(2):153–64.
17. Ohsumi S, Shimoizuma K, Ohashi Y, Shinji M, Hozumi Y, Mukai H, et al. Health-related quality of life and psychological distress of breast cancer patients after surgery during a phase III randomized trial comparing continuation of tamoxifen with switching to anastrozole after adjuvant tamoxifen for 1–4 years: N-SAS BC 03. *Breast Cancer Res Treat.* 2011;127(1):143–52.

18. Pagani O, Gelber S, Price K, Zahrieh D, Gelber R, Simoncini E, et al. Toremifene and tamoxifen are equally effective for early-stage breast cancer: first results of International Breast Cancer Study Group Trials 12-93 and 14-93. *Ann Oncol*. 2004;15(12):1749–59.
19. Takei H, Ohsumi S, Shimozuma K, Takehara M, Suemasu K, Ohashi Y, et al. Health-related quality of life, psychological distress, and adverse events in postmenopausal women with breast cancer who receive tamoxifen, exemestane, or anastrozole as adjuvant endocrine therapy: National Surgical Adjuvant Study of Breast Cancer 04 (N-SAS BC 04). *Breast Cancer Res Treat*. 2012;133(1):227–36.
20. Mamounas EP, Jeong JH, Wickerham DL, Smith RE, Ganz PA, Land SR, et al. Benefit from exemestane as extended adjuvant therapy after 5 years of adjuvant tamoxifen: intention-to-treat analysis of the National Surgical Adjuvant Breast And Bowel Project B-33 trial. *J Clin Oncol*. 2008;26(12):1965–71.
21. Cella D, Fallowfield L, Barker P, Cuzick J, Locker G, Howell A. Quality of life of postmenopausal women in the ATAC (“Arimidex”, tamoxifen, alone or in combination) trial after completion of 5 years’ adjuvant treatment for early breast cancer. *Breast Cancer Res Treat*. 2006;100(3):273–84.
22. Fallowfield L, Cella D, Cuzick J, Francis S, Locker G, Howell A. Quality of life of postmenopausal women in the arimidex, tamoxifen, alone or in combination (ATAC) adjuvant breast cancer trial. *J Clin Oncol*. 2004;22(21):4261–71.
23. Goss PE, Ingle JN, Pritchard KI, Robert NJ, Muss H, Gralow J, et al. Extending aromatase-inhibitor adjuvant therapy to 10 years. *N Engl J Med*. 2016;375(3):209–19.
24. Muss HB, Tu D, Ingle JN, Martino S, Robert NJ, Pater JL, et al. Efficacy, toxicity, and quality of life in older women with early-stage breast cancer treated with letrozole or placebo after 5 years of tamoxifen: NCIC CTG intergroup trial MA.17. *J Clin Oncol*. 2008;26(12):1956–64.
25. Whelan TJ, Goss PE, Ingle JN, Pater JL, Tu D, Pritchard K, et al. Assessment of quality of life in MA.17: a randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *J Clin Oncol*. 2005;23(28):6931–40.
26. FACIT.org. About Us. <http://www.myendnoteweb.com/EndNoteWeb.html?func=new>.
27. Ware JE Jr. SF-36 health survey update. *Spine*. 2000;25(24):3130–9.
28. Hilditch JR, Lewis J, Peter A, van Maris B, Ross A, Franssen E, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas*. 2008;61(1–2):107–21.
29. Radloff L. Center for epidemiological studies depression scale. *Appl Psychol Meas*. 1977;1:385–401.
30. De Haes J, Olschewski M, Fayers P, Visser M, Cull A, Hopwood P, et al. The Rotterdam symptom checklist (RSCL). Amsterdam: University of Groningen; 1996.
31. Life EQo. EORTC QLQ-C30 EORTC Quality of Life Department. <http://groups.eortc.be/qol/eortc-qlq-c30>.
32. Hürny C, Bernhard J, Bacchi M, van Wegberg B, Tomamichel M, Spek U, et al. The Perceived Adjustment to Chronic Illness Scale (PACIS): a global indicator of coping for operable breast cancer patients in clinical trials. *Support Care Cancer*. 1993;1(4):200–8.
33. Bodian CA, Freedman G, Hossain S, Eisenkraft JB, Beilin Y. The visual analog scale for pain: Clinical significance in postoperative patients. *Anesthesiology*. 2001;95(6):1356–61.
34. Jacobsen PB, Davis K, Cella D. Assessing quality of life in research and clinical practice. *Oncology (Williston Park)*. 2002;16(9 Suppl 10):133–9.