Long-term Postmenopausal Hormone Replacement Therapy Modifies Skeletal Muscle Composition and Function: A Study with Monozygotic Twin Pairs

Paula H.A. Ronkainen1,2, Vuokko Kovanen1,2, Markku Alén1,3, Eija Pöllänen1,2, Eeva-Maija Palonen1,2, Carina Ankarberg-Lindgren4, Esa Hämiäläinen5, Ursula Turpeinen5, Urho M. Kujala1, Jukka Puolakka6, Jaakko Kaprio7,8, Sarianna Sipilä1,2.

1Department of Health Sciences, University of Jyväskylä, Finland
2The Finnish Centre for Interdisciplinary Gerontology, University of Jyväskylä, Jyväskylä, Finland
3Department of Medical Rehabilitation, Oulu University Hospital, and Institute of Health Sciences, University of Oulu, Oulu, Finland
4Department of Pediatrics, The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
5HUSLAB, Helsinki University Central Hospital, Helsinki, Finland
6Central Finland Central Hospital, Jyväskylä, Finland
7Department of Public Health and Institute for Molecular Medicine, University of Helsinki, Helsinki, Finland
8Department of Mental Health and Alcohol Research, National Public Health Institute, Helsinki, Finland

Running title: Hormone replacement therapy, muscle and mobility

Correspondence and reprint requests:
M.Sc. Paula Ronkainen
Department of Health Sciences
University of Jyväskylä, P.O. Box 35 (Viv)
FIN-40014 Jyväskylä, Finland
e-mail: paula.ronkainen@sport.jyu.fi
tel: +358 14 260 2097, fax: +358 14 260 4600

Funding sources: This study was supported by the Academy of Finland, Finnish Ministry of Education, the Lapland Fund of the Finnish Cultural Foundation, Research Foundation of Orion Corporation, “Aging, well-being and technology” graduate school, Petter Silfverskiöld foundation and the Centre of Excellence in Complex Disease Genetics.
Abstract

Our aim was to investigate whether long-term hormone replacement therapy (HRT) is associated with mobility and lower limb muscle performance and muscle composition in early postmenopausal women. A sample of 54-62-yrs-old monozygotic female twin pairs discordant for HRT was recruited from the Finnish Twin Cohort. Habitual (HWS) and maximal walking speeds (MWS) over 10 meters, thigh muscle composition, lower body muscle power assessed as vertical jumping height as well as maximal isometric hand grip and knee extension strengths were measured. Intra-pair differences (IPD%) with 95% confidence intervals were calculated. The mean duration of HRT use was 6.9 ± 4.1 yrs. MWS was 7% (0.9 to 13.1%; p=0.019) and lower body muscle power on average 16% (-0.8 to 32.8%, 0.023) greater in the HRT users compared to their co-twins. Thigh muscle cross-sectional area tended to be larger (IPD%= 6%; 95% CI, -0.07 to 12.1%, p=0.065), relative muscle area greater (8%, 0.8 to 15.0%, 0.047), and relative thigh fat area smaller (-5%, -11.3 to 1.2%, 0.047) in the HRT users compared to their sisters. There were no significant differences in maximal isometric strengths or HWS between the users and the non-users. Further subgroup analyses revealed that estrogen-containing therapies (11 pairs) significantly decreased total body and thigh fat content, whereas tibolone (4 pairs) tended to increase muscle cross-sectional area. This study showed that long-term HRT was associated with better mobility, greater lower body muscle power and favorable body and muscle composition among 54-62-year-old women. The results indicate that HRT is a potential agent in preventing muscle weakness and mobility limitation in older women.
Introduction

Low muscle strength among middle-aged and older people is predictive of adverse health events [1], as well as incident mobility limitation and disability [2]. Among the major reasons for muscle weakness and consequent mobility limitation are the age-induced changes in muscle and body composition such as sarcopenia, the inexorable loss of muscle tissue [3, 4]. Based on the cross-sectional computed tomography (CT) images, sarcopenia is characterized not only by the loss of muscle mass, but also by fat infiltration within the muscle and between the separate muscle groups. Moreover, the loss of muscle mass has been suggested to contribute to fat gain, which again reinforces the decrements in muscle mass [5, 6].

Aging is characterized by changing endocrine activity. In women, the most dramatic hormonal dysfunction is observed during the peri- and the postmenopausal period, when the secretion of ovarian hormones, namely estradiol and progesterone, is drastically reduced [7, 8]. By the mid-sixth decade of life, all women confront the menopause, whereupon women spend on average one-third of their life in a state with a profound sex hormone deprivation. Interestingly, some cross-sectional studies in older women have shown that higher serum estradiol concentration is associated with greater muscle mass [9, 10] and muscle strength [10, 11]. In addition, a few studies propose that an accelerated decline in muscle strength occurs right around the time of menopause [12-14]. These data imply that the female sex steroids may represent a potential mechanism for aging-related muscle weakness and unfavorable changes in the muscle and body composition thus increasing the risk for mobility limitation and disability.
In previous randomized controlled trials (RCT), one-year administration of postmenopausal hormone replacement therapy (HRT) has been documented to improve mobility and increase muscle strength among relatively young postmenopausal women [15-17], whereas among older women similar effects with six months to three year long treatment were not evident [18, 19]. To the best of our knowledge, only two RCTs have investigated the effects of HRT on muscle cross-sectional area (CSA) and muscle composition in postmenopausal women. First, an open trial by Skelton and colleagues [16] showed no change in the CSA of the musculus adductor pollicis after one year of cyclical estrogen and norgestrel treatment in on average 61-year-old women despite the significant strength gain of the same muscle. Second, a double-blinded RCT among early postmenopausal women showed a significant six percent mean increase in the CSA of knee extensor muscle after continuous combined estradiol and progestin treatment compared to the controls [17]. In the same study, thigh muscle density assessed by CT increased after HRT [15] suggesting decreased intramuscular lipid content [20]. In addition, the relative proportion of fat within the knee extensor muscle compartment remained unchanged in women on HRT, whereas it increased in the control subjects [17].

During the last six years the balance between the risks and the benefits concerning the long-term use of postmenopausal HRT has been challenged by large clinical trials [e.g. 21]. The knowledge on the long-term effects of HRT would, however, be of considerable importance to capture the possible preventive impact of HRT on mobility limitation and disability. This is because the clinically relevant aging-induced features such as sarcopenia, muscle weakness and mobility limitation develop over several years. An effective strategy to study the long-term effects of HRT is to utilize a co-twin control design, in which exposed monozygotic (MZ) twins are compared to
their unexposed co-twins. This design holds the incomparable opportunity to examine “clonal controls” and to study the given associations independent of subjects’ genetic makeup and shared past experiences from childhood onwards. Due to the fact that MZ twins are genetically identical at the sequence level, any difference observed between the co-twins has to be founded on acquired factors. The purpose of this study was to investigate whether long-term HRT results in clinically significant alterations in mobility and skeletal muscle mass, composition and function in 54-62-year-old postmenopausal women.

Materials and Methods

Study design

This study is part of a larger research project “Sarcopenia - Skeletal Muscle Adaptation to Postmenopausal Hypogonadism and Effects of Hormone Replacement Therapy and Physical Activity in Older Women: a Genetic and Molecular Biological Study on Estrogen-related Pathways” (SAWEs) investigating the molecular events involved in maintaining proper muscle mass and function after menopause. The participants were recruited from The Finnish Twin Cohort [22, 23], which includes all same-sex twin pairs born in Finland before 1958, and with both co-twins alive in 1967 (n=13 888 pairs). Health- and lifestyle-related factors were assessed in the Finnish Twin Cohort by specific questionnaires mailed to the participants in 1975, 1981 and 1990 [23]. An invitation to the current study along with a short pre-questionnaire concerning subject’s HRT use and history as well as willingness to participate in the study was sent to all the MZ twin pairs born in 1943-1952 (537 pairs, Figure 1.). Only those twin pairs, in which one co-twin was a current HRT user and the other co-twin did not currently use HRT, were asked to respond to the invitation. From all the responders (n=114 pairs), those twin pairs of which one
sister had never used HRT, while the other sister was a current user (n=21 pairs), were contacted by telephone. During a telephone call subject’s gynecological history, possible medication and contraindications for participation the measurements (chronic musculoskeletal disease, type 1 diabetes, type 2 diabetes with medication, diagnosed mental disorder, asthma with oral cortisol treatment, acute cancer, known drug or alcohol abuse/dependence, Crohn’s disease) and for biopsy sampling (haemorrhagic diseases or usage of warfarin) were queried. Subjects undergoing acute diseases or convalescence were excluded or measured after full recovery (inflammatory status confirmed before the measurements). A total of 16 pairs were met the inclusion criteria for the present study and were further invited to the laboratory examinations. The zygosity of the twins participating in the laboratory measurements was verified at the Paternity Testing laboratory (National Public Health Institute, Helsinki, Finland) using DNA extracted from a venous blood sample with a battery of ten highly polymorphic gene markers. Fifteen pairs of the sixteen participating in the measurements were confirmed to be MZ pairs, whereas one twin pair turned out to be dizygotic (DZ) and was excluded from the study.

Of the HRT users five women used estradiol-only preparations (the mount of the estrogenic agent in the preparation 1-2 mg), whereas six were taking a combined treatment including estrogenic (1-2 mg) and progestogenic compounds. Four women used tibolone (2.5 mg), which represents an analogue of the progestin, and is metabolized in the intestine and the liver into metabolites having both estrogenic and progestogenic/androgenic effects on target tissues [24]. Thirteen women were taking preparation as pills, whereas one used hormonal patch and one a gel preparation. Of the non-users, one co-twin had tested HRT for three months seven years prior to
the laboratory measurements, but discontinued after this testing period. The duration of HRT usage was on average 6.9 ± 4.1 years (range, 2-16 years).

Each subject participated in the laboratory measurements during two consecutive days. CT scans were obtained at the local hospital on the first measurement day, while all the other laboratory measurements were carried out on the second day. Both sisters within a given twin pair participated in all the measurements on the same days. All the laboratory measurements and data analysis were carried out blinded with respect to HRT status.

The Ethics Committee of the Central Finland Health Care District approved the study and it was conducted according to the guidelines in The Declaration of Helsinki. Written informed consent was provided by the subjects before participating in the measurements.

**Blood sampling, medical examination and health status**

Fasting blood samples were taken between 0700 and 0900 hours. Sera stored at -70°C after sampling were used for hormone measurements as described later. During a medical examination a physician assessed participant’s general health status and gynecological history and confirmed the presence of chronic conditions. The examination included a detailed review of medication use, including the history of HRT use. Furthermore, the possible presence of contraindications for participating in the study was confirmed. Data on current and past smoking was collected using a standard questionnaire.
Hormone measurements

Serum follicle stimulating hormone (FSH) and sex hormone binding globulin (SHBG) concentrations were measured using solid-phase, chemiluminescent immunometric assays (Immulite 1000, Diagnostic Products Corporation, Los Angeles, CA, USA). The interassay CVs were 5.5% for concentrations of 38.5 IU/l (FSH), and 8.4% for 32.4 nmol/l (SHBG). Limit of quantification (LOQ) was 0.1 IU/l for FSH and 0.2 nmol/l for SHBG. Serum estradiol levels were determined in duplicate by extraction RIA as previously described [25]. The extraction RIA has been validated especially for measuring low serum estradiol concentrations and is an accredited assay by SWEDAC in Sweden, SS-EN ISO 15189 (no 1899). LOQ was 4 pmol/l, while the interassay CV was 19% at 6 pmol/l and below 14% for concentrations of 12 pmol/l and above. Serum testosterone (T) was measured as previously described [26]LOQ being 70 pmol/l and interassay CV 5.2%. for concentration 4.7 pmol/l. E₂ and T levels were utilized together with SHBG in calculating the respective free hormone levels according to previously presented methods [27, 28]. Estrone (E₁) was determined as a dansyl derivative by LC-MS/MS on API 4000 mass spectrometer as previously described [29]. Interassay CV was 7.8% for concentration 200 pmol/l, while the LOQ was 10 pmol/l.

Body anthropometry and composition

Body anthropometry was measured after overnight fasting. Body weight was measured in kilograms using a beam scale with the participant wearing only undergarments and height in centimeters using a stadiometer while the participant was standing in stocking feet. Waist circumference was measured midway between spina iliaca superior and the lower rib margin and hip circumference at the level of the greater trochanters both to the nearest half centimeter [30].
Percentage body fat and lean body mass (LBM) were measured with bioelectrical impedance [InBody (720), Biospace Co. Ltd., Seoul, Korea]. In our laboratory the coefficient of variation (CV) for two consecutive measurements for fat percent is 0.6% [31]. Self-report data on weight and height from prior questionnaires carried out in 1975, 1981 and 1990 of the twin cohort was also accessed, and BMI computed.

Mobility

Mobility was assessed with habitual (HWS) and maximal walking speed (MWS) for 10 meters in a laboratory corridor. Five meters were allowed for acceleration and time used over ten meters was measured using photocells. In HWS test, the subjects were instructed to “walk in your habitual speed, as you were going to a supermarket without a need to hurry”, whereas in the MWS, the subjects were advised to "walk as fast as possible, without compromising your safety". Timing was done twice for both measurements and the faster performance was documented as a result. The participants wore their own walking shoes or sneakers. For the safety, the examiner walked behind the participant during the tests. Walking speed is highly predictive for mobility limitation and activities of daily living related disabilities [32]. The CV for MWS has been 5% in our laboratory [33].

Muscle characteristics

Thigh muscle mass and composition

CT scans (Siemens Somatom Emotion scanner, Siemens AG, Erlangen, Germany) were obtained from the midpoint between the greater trochanter and the lateral joint line of the knee. The scans were analyzed using software developed at the University of Jyväskylä (Jyväskylä, Finland) for
cross-sectional CT image analysis (Geanie 2.1, Commit Ltd, Espoo, Finland), which separates fat and lean tissue based on the given radiological density limits. Total thigh muscle and fat (includes subcutaneous fat and fat infiltrated into the muscle compartment) CSAs were analyzed and relative proportions of muscle and fat within the whole thigh CSA calculated. Fat area within the muscle compartment (infiltrated fat) was determined from manually outlining the muscle compartment by drawing a line along the fascial plane to exclude subcutaneous fat. CSAs and relative proportions of muscle and fat as well as muscle attenuation were analyzed. Skeletal muscle attenuation was defined as the mean attenuation coefficient in Hounsfield units (HU). In our previous studies, the CV between two consecutive measurements was 1-2% for muscle CSA [34] and less than 1% for muscle HU [15].

**Isometric muscle strength and lower body muscle power**

Maximal isometric knee extension strength was measured on both legs in a sitting position using an adjustable dynamometer chair at a knee angle of 60° from full extension (Good Strength, Metitur, Palokka, Finland). In addition, hand grip strength was measured on the dominant side. After familiarization with the test, three to six maximal efforts were conducted. Lower body muscle power, i.e. the ability of the neuromuscular system to produce the greatest possible force as fast as possible, was assessed as the height that a subject is able to elevate the body’s centre of gravity during a vertical jump (vertical jumping height) with counter-movement on a contact mat. The flight time was measured, and the respective height calculated as follows: vertical jumping height (cm) = (g x t^2) : 8 x 100 [35]. Three maximal efforts were conducted. In all measurements, the best performance with the highest value was accepted as the result. CVs between two
consecutive measurements in our laboratory have earlier been 6% for knee extension strength and hand grip [36], and 5% for vertical jumping height [37].

**Physical activity**

Physical activity was assessed using the scale of Grimby [38] with slight modifications. The participants were categorized on the basis of their self-reported physical activity as sedentary (no other activities, but at the most light walking ≤ 2 times/week), moderately active (walking or other light exercise at least 3 times/week, but no other more intensive activities) and active (moderate or vigorous exercise at least 3 times/week) groups. Physical activity at leisure was also assessed in the prior questionnaires in 1975, 1981 and 1990 [39].

**Statistical analyses**

The means were compared using Wilcoxon’s signed rank test due to relatively small number of observations. Data are shown as means and standard deviations unless otherwise stated. Intra-pair differences are expressed as percentages (IPD%) and calculated as follows: (HRT user – non-user) : (non-user) x 100. In addition, the 95% confidence interval (95% CI) was calculated for the IPD%. Data are reported for the whole group (“HRT users”, n=15 pairs) and separately for two subgroups; for the “E users” including the pairs in which the other co-twin used estradiol-only or combined estradiol plus progestogen preparation (n=11 pairs) and for the “tibolone users” using a preparation including tibolone as an effective agent (n=4 pairs). The level of significance was set at p≤0.05. Data analyses were carried out with SPSS (Version 14.0, SPSS Inc., Chicago, IL).
Results

Subjects’ characteristics and hormonal status

The mean age (±SD) of the subjects was 57.2 ± 1.8 years (range, 54-62 years). Four current users and one non-user had undergone perimenopausal combined hyster- and ovariectomy and one user and one non-user a hysterectomy only. Six HRT users and five non-users had medication for hypertension, and three subjects from each group were taking statins for elevated cholesterol levels. One user and one non-user had a history of basal cell carcinoma and one non-user that of melanoma. One co-twin in both groups was rated as sedentary, whereas seven users and five non-users were moderately active and seven and nine active, respectively. Moreover, five users and three non-users were current smokers. According to prior questionnaire information from 1975, 1981 and 1990, there were no differences in physical activity, smoking behavior or alcohol use between the HRT users and their non-using co-twins prior to use of HRT.

The expected differences between the HRT users and non-users in serum hormone concentrations were observed (Table 1). The HRT users had on average five times higher concentrations of E$_2$ and E$_1$ than the non-users. A similar trend was observed for free E$_2$ as well. No difference was observed between the sisters in the levels of SHBG, and total and free T. Subgroup analyses on total and free E$_2$ and E$_1$ showed that the observed differences between the users and the non-users were primarily due to differences between E users and their co-twins. Furthermore, E users had on average 62% (p=0.010) higher SHBG levels compared to their co-twins, whereas the tibolone users had on average 54% (p=0.068) lower SHBG concentrations than their sisters. Of note, the assays utilized do not measure serum tibolone.
Body anthropometry and composition

HRT users were on average 0.8 cm taller than the non users (IPD%=0.5%, p=0.025, Table 2). No statistically significant differences were observed in weight, BMI, waist or hip circumference, LBM and body fat percent between the sisters. In a subgroup analysis, however, the E users had lower body fat percent (p=0.026) and tended to have lower BMI (p=0.091) compared to their co-twins with no HRT history. According to previous self-report, the current HRT sister had been taller than her co-twin already in 1975, 1981 and 1990, although not statistically significantly, the smallest p value being 0.054 in 1981. Moreover, the sisters within a pair did not differ in BMI prior to HRT use.

Mobility

MWS of the HRT users was seven percent greater compared to the non-users (2.2 ± 0.3 vs. 2.0 ± 0.2 m/s, p=0.019, Figure 2). The same tendency towards better performance within both subgroups of users was observed (E users: 2.1 ± 0.3 vs. 2.0 ± 0.2 m/s, p=0.074; tibolone users: 2.3 ± 0.3 vs. 2.1 ± 0.1 m/s, p=0.068). HWS did not differ between the users and the non-users.

Muscle strength and power

Lower body muscle power assessed as vertical jumping height was on average 16% greater in the HRT users compared to their co-twins (14.8 ± 3.7 vs. 13.2 ± 3.7 cm, p=0.023, Figure 2). This difference was clearly due to better muscle power among the E users, who were able to elevate their body on average 21% higher than their twin sisters without HRT history (15.1 ± 4.2 vs. 13.0 ± 4.3 cm, p=0.016). No intra-pair differences were observed within the tibolone group (users vs.
non-users: 14.1 ± 3.2 vs. 13.9 ± 1.7 cm, p=0.72, IPD%=1.8%). There were no significant differences in maximal isometric strength between the HRT users and the non-users.

**Thigh muscle mass and composition**

Total thigh muscle CSA tended to be six percent larger among the HRT users compared to the non-users (p=0.065). Subgroup analysis showed a 15% larger mean thigh muscle CSA (p=0.068) for the tibolone users compared to their co-twins, while a 3% insignificant difference within the E using twin pairs was found.

No significant differences were observed in total thigh and subcutaneous fat areas between the sisters. However, the subgroup analysis revealed that the total thigh fat CSA, including subcutaneous fat and fat infiltrated into the muscle compartment, was on average 12% (p=0.021) lower in E users than in their co-twins. Moreover, E users had less subcutaneous (IPD%=10%, p=0.037) and infiltrated fat (IPD%=10%, p=0.11) than their sisters.

Muscle CSA in relation to total soft tissue CSA of the thigh (relative muscle area of thigh) was an average eight percent (p=0.047) larger among HRT users compared to non-users. When subcutaneous fat was excluded from the analysis and muscle CSA was analyzed in relation to muscle compartment CSA (relative muscle area of muscle compartment), HRT users had an average 1% (p=0.11) larger relative muscle area within the thigh muscle compartment compared to the sisters with no HRT history.
Larger relative muscle areas among the HRT users appeared to be due to higher muscle proportion within the E users (relative muscle area of the total thigh was 11%, p=0.013 and relative muscle area of the muscle compartment 2%, p=0.033 larger in E users than in their sisters). This was due to the clearly smaller amount of subcutaneous and infiltrated fat in the thigh in combination with slightly larger muscles (3%) among the E users compared to their co-twins.

Discussion

Towards the aim of elucidating the effects of postmenopausal HRT on muscle properties, we investigated the association of long-term HRT with mobility and lower limb muscle characteristics in a rare sample consisting of postmenopausal monozygotic twin pairs. Our major finding was that HRT was associated with better mobility and greater lower body muscle power among 54-62-year-old women. In addition, the sisters using HRT had lower relative proportion of fat within the thigh and tended to have larger muscles than their sisters with no HRT history. Subgroup analysis concerning muscle composition revealed that estrogen-containing therapies had influence primarily on the fat tissue and consequently also on relative proportion of muscle. Tibolone had no effects on either total body fat or relative proportion of fat within the thigh, but it may have had a hypertrophic effect on skeletal muscle.

The greater maximal walking speed in HRT users compared to their sisters irrespective of the preparation used is a finding of notable importance. Mobility and good functional capacity are the key factors for independent living and good quality of life in older people [32]. Difficulty in walking is associated with adverse health events [40, 41] and also increases dependency on social
and health care services [42]. Supporting our finding, a previous RCT with one-year double-blinded HRT showed a significant 2% net increase in 20-m running speed in 50-57-year-old women [15]. Earlier cross-sectional studies involving substantially older postmenopausal women, however, failed to show significant association between estrogen containing therapy and walking speed.

The mechanism for greater walking speed of HRT users compared to their sisters observed in this study may include larger thigh muscles, greater lower body muscle power, as well as lower fat percentage observed in the HRT users. Fast and safe walking is a multifactorial task requiring sufficient neuromuscular performance as well as control for posture [43]. In addition, body composition has been documented to be associated with mobility among the elderly. More specifically, earlier studies have shown that high body fat mass and BMI are significantly associated with mobility limitation [44-46]. Especially among those older persons who are overweight and have low muscle strength or muscle mass [46, 47] the prevalence and the risk for physical disability is substantial.

In this study, the use of HRT was associated with greater muscle power of the lower extremities. Those women with estradiol containing therapy had over 20% greater muscle power compared to their sisters. The intra-pair difference between the tibolone users and their sisters was, however, minor and insignificant. These results are supported by our previous study among early postmenopausal women, in which one-year administration of combined estradiol and progestogen treatment improved jumping height an average by 7% compared to the mean 5% decrease in the placebo group[17]. Muscle power is the product of force and speed of contraction,
and thus represents the ability of the neuromuscular system to produce the greatest possible force as fast as possible. Earlier studies have shown that muscle power is more sensitive to aging process and that it may be more important in mobility than muscle force alone ([48-50]. Therefore, the difference between the sisters in muscle power, but not in strength, may be at least partly due to the deteriorated muscle power of the non-HRT co-twin.

The current study showed that HRT tended to be associated with larger thigh muscles. A 15% intra-pair difference was observed between tibolone users and their sisters, whereas the difference between estradiol containing HRT users and their sisters was minor (3%). Earlier case-control [10] and RCT [17] studies showed a three to six percent significant difference in thigh muscle mass between women using estrogen containing HRT and the non-users. In the latter one, the treatment included continuous combined estradiol (2 mg, high dose) and norethisterone acetate (1 mg). On the other hand, the RCT by Kenny et al. [18] showed no effect of three year ultra low-dose estrogen (0.25 mg) therapy on appendicular muscle mass in over 65-year-old women. In our case-control study, the most of the sisters were using products including 1 mg of the estrogenic effective agent. The discrepancy between the earlier RCTs and the present study in muscle cross-sectional area may be due to the differences in the preparations and doses used in these studies, as well as the age of the subjects.

We were unable to identify any previous studies, in which the effects of tibolone would have been studied on muscle mass specifically. Nonetheless, supporting our findings several studies have reported positive effects of tibolone on total lean body mass [51-55]. In addition to the
association of tibolone with muscle size, our results suggest that the tibolone users have higher LBM compared to the non-users, although the result did not reach statistical significance.

The sisters using estrogen-containing preparations had significantly smaller thigh fat area, subcutaneous fat area and relative proportion of fat in the whole thigh area, as well as within the muscle compartment than their co-twins with no HRT history. In addition, they also had lower body fat percent compared to their sisters. A previous study reported that abdominal fat percent increases upon placebo compared to a combined estrogen and progestogen treatment, although no difference in total body fat was observed [56]. Moreover, in another study the placebo group gained more weight over a five-year period than women taking estrogen-containing preparation, a notion especially due to changes in body fat [57]. Similar results were evident elsewhere as well [58], although also reports documenting contradictory results [59] or no effects [60-62] of HRT on body fat have been published. Presuming that HRT could, however, prevent the accumulation of excess body fat, may hold clinical relevance, since the amount of adipose tissue especially within the fascia surrounding skeletal muscle has been documented to be related with insulin resistance [63-65] and, at the level of the whole body, adiposity has further been strongly associated with adverse health outcomes such as cardiovascular disease, diabetes, and cancer [66]. Our study suggests that both estrogen-related and tibolone therapies exert positive effects on muscle composition, albeit the target tissue may be different.

The exact molecular level mechanisms responsible for the changes found in muscle phenotypes due to HRT use remain to be clarified and are further challenged by the diversity of the HRT regimens. Data on human skeletal muscle demonstrate that both known ERs, ERα and ERβ, are
expressed at mRNA level [67, 68] and ERβ protein has also been detected [67]. Intriguingly, in our previous study with postmenopausal women a menu of changes in the expression levels of specific genes in muscle samples was observed after one year without HRT during early postmenopausal years, whereas HRT use was shown to largely counteract these changes [69]. Considering these data, skeletal muscle can be considered to potentially be responsive to estradiol signaling. Furthermore, experimental animal studies by D’Eon and colleagues support a role for estrogen itself in reducing overall adiposity by downregulating the expression of lipogenic genes not only in the adipocytes but also in skeletal muscle [70]. Moreover, an alternative HRT, tibolone, along with estradiol have been reported to effectively restore the episodic release of growth hormone [71], which, again, has been suggested to promote lipolysis resulting in a potential anabolic effect on muscle tissue and a decrement in the amount of fat in the muscle compartment [15, 72]. Consequently, estradiol should be recognized when the intricate pathways leading to sarcopenia are constructed.

Our co-twin analysis, in which the drug-exposed monozygotic co-twins are compared to their unexposed sisters, holds several advantages over the traditional case-control and experimental designs. Co-twin analysis intrinsically adjusts for genetic factors and in addition for a number of shared environmental factors starting from childhood. Thereby, our data set and the analysis strategy together offer a powerful tool to evaluate the association of long-term HRT with skeletal muscle characteristics. Our analyses are based on fifteen pairs. This number of subjects in a study design with discordant MZ twins, however, is comparable to previous reports with similar number of participants [73-76] and has sufficient power to detect clinically relevant differences.
We are, however, well aware that in the subgroup analysis, the number of pairs using tibolone is small.

This genetically controlled case-control study showed that long-term HRT was associated with better mobility, greater lower body muscle power and favorable body and muscle composition among 54-62-year-old women. The results indicate that HRT is a potential agent in preventing muscle weakness and mobility limitation in older women.

Acknowledgements
We are grateful for all the twins for participating in the study. Kaisa-Leena Tulla, Erkki Helkala, Tuovi Nykänen and Mervi Matero are acknowledged for excellent technical assistance in the laboratory. In addition, we wish to thank Eila Voipio for assistance in the recruitment of twins and Kirsti Salo for taking good care of our study subjects during the whole period of data collection. In addition, the personnel at the Paternity Testing laboratory, National Public Health Institute, Helsinki, Finland is acknowledged for carrying out the determination of the zygosity of the twins.
References


Table 1. Hormone status of MZ twin pairs discordant for the long-term use of HRT.

<table>
<thead>
<tr>
<th>Variable</th>
<th>non-user</th>
<th>HRT user</th>
<th>Intra-pair difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSH (IU/l), n=15</td>
<td>84.8 ± 20.5</td>
<td>56.4 ± 23.3</td>
<td>-33.4 (-45.0 to -21.9)</td>
<td>0.001</td>
</tr>
<tr>
<td>E users, n=11</td>
<td>93.2 ± 16.7</td>
<td>62.8 ± 23.6</td>
<td>-31.8 (-47.7 to -15.9)</td>
<td>0.006</td>
</tr>
<tr>
<td>Tibolone users, n=4</td>
<td>61.8 ± 7.6</td>
<td>38.8 ± 9.9</td>
<td>-23.8 (-55.9 to -1.6)</td>
<td>0.068</td>
</tr>
<tr>
<td>SHBG (nmol/l)</td>
<td>50.3 ± 21.5</td>
<td>58.0 ± 33.3</td>
<td>7.7 (-13.8 to 28.2)</td>
<td>0.43</td>
</tr>
<tr>
<td>E users</td>
<td>42.6 ± 14.6</td>
<td>68.2 ± 33.1</td>
<td>25.6 (14.2 to 109.3)</td>
<td>0.010</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>71.4 ± 25.3</td>
<td>30.2 ± 10.7</td>
<td>-41.2 (-84.7 to -17.3)</td>
<td>0.068</td>
</tr>
<tr>
<td>E2 (pmol/l)</td>
<td>30.2 ± 23.8</td>
<td>132.8 ± 185.1</td>
<td>104.6 (-66.7 to 1073.8)</td>
<td>0.002</td>
</tr>
<tr>
<td>E users</td>
<td>33.3 ± 27.4</td>
<td>172.9 ± 203.2</td>
<td>139.6 (-146.7 to 1539.5)</td>
<td>0.003</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>21.8 ± 2.1</td>
<td>22.5 ± 5.9</td>
<td>0.7 (-30.5 to 35.2)</td>
<td>0.58</td>
</tr>
<tr>
<td>Free E2 (pmol/l)</td>
<td>0.71 ± 0.52</td>
<td>2.59 ± 3.0</td>
<td>1.87 (-11.7 to 768.3)</td>
<td>0.03</td>
</tr>
<tr>
<td>E users</td>
<td>0.81 ± 0.58</td>
<td>3.31 ± 3.27</td>
<td>2.5 (33.3 to 1034.6)</td>
<td>0.006</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>0.45 ± 0.08</td>
<td>0.63 ± 0.20</td>
<td>0.18 (-37.5 to 120.8)</td>
<td>0.14</td>
</tr>
<tr>
<td>E1 (pmol/l)</td>
<td>97.1 ± 26.1</td>
<td>690.5 ± 1281.0</td>
<td>615.4 (-100.8 to 1225.5)</td>
<td>0.001</td>
</tr>
<tr>
<td>E users</td>
<td>97.7 ± 26.8</td>
<td>899.6 ± 1454.7</td>
<td>801.9 (-153.0 to 1672.8)</td>
<td>0.003</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>95.5 ± 27.6</td>
<td>115.5 ± 46.0</td>
<td>20.0 (-11.4 to 49.5)</td>
<td>0.14</td>
</tr>
<tr>
<td>T (pmol/l)</td>
<td>763 ± 360</td>
<td>701 ± 273</td>
<td>42 (-15.5 to 15.4)</td>
<td>0.87</td>
</tr>
<tr>
<td>E users</td>
<td>639 ± 269</td>
<td>715 ± 306</td>
<td>76 (2.1 to 13.4)</td>
<td>0.061</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>1105 ± 390</td>
<td>663 ± 184</td>
<td>-442 (-54.3 to -20.7)</td>
<td>0.068</td>
</tr>
<tr>
<td>Free T (pmol/l)</td>
<td>10.6 ± 5.0</td>
<td>9.7 ± 5.3</td>
<td>-0.9 (-21.2 to 9.3)</td>
<td>0.23</td>
</tr>
<tr>
<td>E users</td>
<td>9.9 ± 4.7</td>
<td>8.4 ± 4.7</td>
<td>-1.5 (-28.7 to 4.7)</td>
<td>0.075</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>12.5 ± 6.3</td>
<td>13.3 ± 5.7</td>
<td>0.8 (-39.2 to 60.5)</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Data are expressed as mean±SD.

FSH = follicle stimulating hormone, SHBG = sex hormone binding globulin, E2 = estradiol, T = testosterone, E users=pairs in which the HRT user is on estradiol-only or combined HRT.
Table 2. Body anthropometry of MZ twin pairs discordant for the long-term use of HRT.

<table>
<thead>
<tr>
<th>Variable</th>
<th>non-user</th>
<th>HRT user</th>
<th>Intra-pair difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body height (cm), n=15</td>
<td>162.8 ± 4.4</td>
<td>163.6 ± 4.9</td>
<td>0.5 (0.1 to 0.9)</td>
<td>0.025</td>
</tr>
<tr>
<td>E users, n=11</td>
<td>161.8 ± 4.4</td>
<td>162.4 ± 4.7</td>
<td>0.36 (-0.09 to 0.81)</td>
<td>0.013</td>
</tr>
<tr>
<td>Tibolone users, n=4</td>
<td>165.5 ± 3.1</td>
<td>167.0 ± 3.9</td>
<td>0.9 (-0.5 to 2.3)</td>
<td>0.14</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>71.3 ± 13.6</td>
<td>69.1 ± 8.6</td>
<td>-1.1 (-10.2 to 8.1)</td>
<td>0.50</td>
</tr>
<tr>
<td>E users</td>
<td>73.6 ± 15.3</td>
<td>67.6 ± 8.7</td>
<td>-5.9 (-16.1 to 4.2)</td>
<td>0.13</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>65.2 ± 4.1</td>
<td>73.0 ± 8.2</td>
<td>12.3 (-9.7 to 34.2)</td>
<td>0.14</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.1 ± 5.9</td>
<td>25.8 ± 3.3</td>
<td>-2.2 (-10.8 to 6.5)</td>
<td>0.39</td>
</tr>
<tr>
<td>E users</td>
<td>28.2 ± 6.5</td>
<td>25.7 ± 3.8</td>
<td>-6.7 (-16.4 to 3.1)</td>
<td>0.091</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>23.8 ± 1.8</td>
<td>26.1 ± 1.9</td>
<td>10.2 (-9.5 to 29.8)</td>
<td>0.14</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>90.5 ± 11.8</td>
<td>87.1 ± 7.7</td>
<td>-2.7 (-9.3 to 4.0)</td>
<td>0.31</td>
</tr>
<tr>
<td>E users</td>
<td>92.8 ± 12.8</td>
<td>86.3 ± 8.2</td>
<td>-6.0 (-13.2 to 1.3)</td>
<td>0.11</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>84.1 ± 5.2</td>
<td>89.2 ± 6.6</td>
<td>6.4 (-12.2 to 25.1)</td>
<td>0.27</td>
</tr>
<tr>
<td>Hip circumference (cm)</td>
<td>102.5 ± 8.5</td>
<td>100.5 ± 5.7</td>
<td>-1.6 (-4.6 to 1.3)</td>
<td>0.17</td>
</tr>
<tr>
<td>E users</td>
<td>104.4 ± 9.1</td>
<td>101.0 ± 6.1</td>
<td>-2.9 (-6.2 to 0.3)</td>
<td>0.075</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>97.2 ± 3.0</td>
<td>99.0 ± 4.7</td>
<td>2.0 (-7.2 to 11.1)</td>
<td>1.0</td>
</tr>
<tr>
<td>LBM (kg)</td>
<td>46.4 ± 4.2</td>
<td>47.3 ± 3.7</td>
<td>2.2 (-2.2 to 6.7)</td>
<td>0.69</td>
</tr>
<tr>
<td>E users</td>
<td>46.3 ± 4.6</td>
<td>46.2 ± 3.4</td>
<td>0.1 (-5.0 to 5.3)</td>
<td>0.37</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>46.8 ± 3.5</td>
<td>50.4 ± 2.9</td>
<td>8.0 (-2.8 to 18.9)</td>
<td>0.14</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>33.3 ± 8.2</td>
<td>30.6 ± 6.4</td>
<td>-5.5 (-16.8 to 5.8)</td>
<td>0.100</td>
</tr>
<tr>
<td>E users</td>
<td>35.2 ± 8.9</td>
<td>30.1 ± 7.1</td>
<td>-10.6 (-23.4 to 2.1)</td>
<td>0.026</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>28.0 ± 1.2</td>
<td>30.3 ± 4.8</td>
<td>8.7 (-22.0 to 39.3)</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Data are expressed as mean±SD.
BMI = body mass index, LBM = lean body mass, E users=pairs in which the HRT user is on estradiol-only or combined HRT.
Table 3. Muscle composition of MZ twin pairs discordant for the long-term use of HRT.

<table>
<thead>
<tr>
<th>Variable</th>
<th>non-user</th>
<th>HRT user</th>
<th>Intra-pair difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL THIGH MUSCLE CROSS-SECTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thigh muscle area (mm$^2$), n=15</td>
<td>9153.3 ± 1301.6</td>
<td>9626.7 ± 1171.6</td>
<td>6.0 (-0.07 to 12.1)</td>
<td>0.065</td>
</tr>
<tr>
<td>E users. n=11</td>
<td>9410.9 ± 1157.7</td>
<td>9627.3 ± 1169.0</td>
<td>2.8 (-4.2 to 9.9)</td>
<td>0.51</td>
</tr>
<tr>
<td>Tibolone users. n=4</td>
<td>8445.0 ± 1589.3</td>
<td>9625.0 ± 1360.4</td>
<td>14.8 (3.1 to 26.5)</td>
<td>0.068</td>
</tr>
<tr>
<td>Relative muscle area (%)</td>
<td>50.4 ± 11.2</td>
<td>53.5 ± 9.5</td>
<td>7.9 (0.8 to 15.0)</td>
<td>0.047</td>
</tr>
<tr>
<td>E users</td>
<td>49.8 ± 12.7</td>
<td>54.2 ± 10.0</td>
<td>11.1 (3.1 to 19.1)</td>
<td>0.013</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>52.2 ± 6.9</td>
<td>51.7 ± 9.1</td>
<td>-0.9 (-20.4 to 18.7)</td>
<td>0.72</td>
</tr>
<tr>
<td>Muscle attenuation (HU)</td>
<td>53.3 ± 5.3</td>
<td>53.4 ± 4.4</td>
<td>0.9 (-5.0 to 6.8)</td>
<td>0.26</td>
</tr>
<tr>
<td>E users</td>
<td>53.4 ± 6.0</td>
<td>53.3 ± 4.7</td>
<td>0.6 (-7.8 to 9.0)</td>
<td>0.53</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>52.8 ± 3.2</td>
<td>53.6 ± 3.9</td>
<td>1.6 (-1.0 to 4.1)</td>
<td>0.14</td>
</tr>
<tr>
<td>Thigh fat area (mm$^2$)</td>
<td>9992.0 ± 5458.1</td>
<td>8734.0 ± 3038.4</td>
<td>-3.0 (-20.5 to 14.5)</td>
<td>0.14</td>
</tr>
<tr>
<td>E users</td>
<td>10786.4 ± 6170.8</td>
<td>8598.2 ± 3399.6</td>
<td>-11.9 (-28.5 to 4.7)</td>
<td>0.021</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>7807.5 ± 1848.7</td>
<td>9107.5 ± 2074.9</td>
<td>21.5 (-41.1 to 84.1)</td>
<td>0.27</td>
</tr>
<tr>
<td>Relative fat area (%)</td>
<td>49.6 ± 11.2</td>
<td>46.5 ± 9.5</td>
<td>-5.0 (-11.3 to 1.2)</td>
<td>0.047</td>
</tr>
<tr>
<td>E users</td>
<td>50.2 ± 12.7</td>
<td>45.8 ± 10.0</td>
<td>-7.4 (-13.3 to -1.5)</td>
<td>0.013</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>47.8 ± 6.9</td>
<td>48.3 ± 9.1</td>
<td>1.4 (-24.4 to 27.2)</td>
<td>0.72</td>
</tr>
<tr>
<td>Subcutaneous fat area (mm$^2$)</td>
<td>8542.9 ± 4979.4</td>
<td>7391.3 ± 2645.0</td>
<td>-1.3 (-20.8 to 18.2)</td>
<td>0.20</td>
</tr>
<tr>
<td>E users</td>
<td>9294.0 ± 5605.7</td>
<td>7340.9 ± 3000.6</td>
<td>-9.9 (-31.2 to 11.3)</td>
<td>0.037</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>6477.5 ± 1790.6</td>
<td>7530.0 ± 1613.0</td>
<td>22.5 (-37.3 to 82.2)</td>
<td>0.27</td>
</tr>
<tr>
<td>MUSCLE COMPARTMENT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative muscle area (%)</td>
<td>84.5 ± 5.3</td>
<td>85.5 ± 5.4</td>
<td>1.3 (-0.3 to 2.9)</td>
<td>0.11</td>
</tr>
<tr>
<td>E users</td>
<td>84.7 ± 5.6</td>
<td>86.4 ± 5.1</td>
<td>2.0 (0.4 to 3.7)</td>
<td>0.033</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>83.9 ± 4.7</td>
<td>83.4 ± 6.3</td>
<td>-0.7 (-5.9 to 4.4)</td>
<td>0.47</td>
</tr>
<tr>
<td>Fat area (mm$^2$)</td>
<td>1663.7 ± 689.3</td>
<td>1592.7 ± 641.0</td>
<td>-2.1 (-16.1 to 11.8)</td>
<td>0.53</td>
</tr>
<tr>
<td>E users</td>
<td>1711.5 ± 795.1</td>
<td>1500.9 ± 638.0</td>
<td>-9.6 (-21.9 to 2.6)</td>
<td>0.11</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>1532.5 ± 280.2</td>
<td>1845.0 ± 667.3</td>
<td>18.4 (-33.8 to 70.7)</td>
<td>0.27</td>
</tr>
<tr>
<td>Relative fat area (%)</td>
<td>15.5 ± 5.3</td>
<td>14.4 ± 5.4</td>
<td>-7.3 (-15.5 to 0.8)</td>
<td>0.11</td>
</tr>
<tr>
<td>E users</td>
<td>15.3 ± 5.6</td>
<td>13.6 ± 5.1</td>
<td>-10.6 (-18.8 to -2.3)</td>
<td>0.033</td>
</tr>
<tr>
<td>Tibolone users</td>
<td>16.1 ± 4.7</td>
<td>16.6 ± 6.3</td>
<td>1.6 (-28.4 to 31.7)</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Data are expressed as mean±SD. E users=pairs in which the HRT user is on estradiol-only or combined HRT.
Figure legends

Figure 1. Flow chart of the study.

Figure 2. Muscle performance and mobility of MZ twins discordant for the long-term use of HRT. * p<0.05.