**Human Reproduction, Vol.27, No.4 pp. 1130–1138, 2012**
Advanced Access publication on February 16, 2012 doi:10.1093/humrep/des004

human reproduction

ORIGINAL ARTICLE *Reproductive endocrinology*

# Physiological sex steroid replacement in premature ovarian failure: randomized crossover trial of effect on uterine volume, endometrial thickness and blood flow, compared with a standard regimen

**R.L. O'Donnell[1], P. Warner[2], R.J. Lee[2], J. Walker[1], L.E. Bath[3], C.J. Kelnar[3], W.H.B. Wallace[3], and H.O.D. Critchley[1,*]**

[1]MRC Centre for Reproductive Health, The Queen's Medical Research Institute, University of Edinburgh, 47 Little France Crescent, Edinburgh EH16 4TJ, UK [2]Centre for Population Health Sciences, University of Edinburgh, Medical School, Teviot Place, Edinburgh EH8 9AG, UK [3]Child Life and Health Section, University of Edinburgh, 20 Sylvan Place, Edinburgh EH9 1UW, UK

*Correspondence address. E-mail: hilary.critchley@ed.ac.uk

*Submitted on June 13, 2011; resubmitted on October 3, 2011; accepted on December 21, 2011*

**BACKGROUND:** Premature ovarian failure (POF) is currently managed by non-physiological sex steroid regimens which are inadequate at optimizing uterine characteristics. Previous short-term studies have demonstrated some benefits of a sex steroid replacement (SSR) regimen devised to replicate the physiological cycle. This study aimed to directly compare the effects of longer-term administration of physiological SSR (pSSR) and standard SSR (sSSR) regimens on the uterine volume, blood flow and endometrial thickness (ET) in women with POF.

**METHODS:** In a controlled crossover trial, 34 women with POF were randomized to receive 12 months of 4-week cycles of transdermal estradiol and vaginal progesterone (pSSR) followed by 12 months of 4-week cycles of oral ethinylestradiol and norethisterone (sSSR), or *vice versa*. Each treatment period was preceded by a 2-month washout period. At 0, 3, 6 and 12 months of each treatment period, transvaginal ultrasound examined the uterine volume and ET, as primary end-points, and uterine artery resistance (UARI) and pulsatility indices (UAPI), as secondary end-points. Serum estradiol, progesterone and gonadotrophins were also measured.

**RESULTS:** Of the 29 women eligible for the uterine analysis, 17 completed the entire study protocol, but 25 women contributed data to statistical analysis of treatment effect. There was a greater estimated mean ET with the use of pSSR (4.8 mm) compared to that with standard therapy (3.0 mm), with an estimated difference of 1.8 mm [95% confidence interval (CI), 0.7 to 2.8, $P = 0.002$]. The estimated mean uterine volume was also greater during physiological treatment (24.8 $cm^3$) than during standard treatment (20.6 $cm^3$), but the estimated difference of 4.2 $cm^3$ (95% CI −0.4 to 8.7) was not statitsically significant, $P = 0.070$. The small differences between the two treatments in the mean UARI and mean UAPI were not statistically significant. The estimated treatment differences were fairly constant across the treatment periods, suggesting that prolonged treatment does not increase response.

**CONCLUSIONS:** pSSR has a greater beneficial effect upon ET in women with POF in comparison with standard therapy. A similar trend was seen for uterine volume. Further studies are required to optimize treatment and to assess pregnancy rate and outcome.

**TRIAL REGISTRATION:** www.ClinicalTrials.gov, NCR00732693.

**Key words:** premature ovarian failure / hormone replacement therapy / uterine size / endometrial thickness / uterine artery blood flow

# Introduction

Cessation of menstruation before the age of 40 years is termed premature ovarian failure (POF), a disorder with a prevalence of ~1% (Cooper and Sandler, 1998). While idiopathic POF remains the most common presentation for women who have had a spontaneous onset of puberty (Rebar and Connolly, 1990), the most common reasons for ovarian failure pre-pubertally are Turner syndrome and gonadotoxic treatment of childhood cancer, specifically direct irradiation to the ovary and high-dose alkylating agents (Ossewaarde *et al.*, 2005). Turner syndrome is associated with POF in over 90% of cases, and the majority fail to enter puberty spontaneously (Hovatta, 1999). The idiopathic form of POF can show sporadic and familial forms and despite the description of several candidate genes (Aittomaki *et al.*, 1995; Beau *et al.*, 1998; De Baere *et al.*, 2003; Di Pasquale *et al.*, 2004), the cause of POF still remains undetermined in the majority of cases.

Young women with pre-pubertal POF require pubertal induction to induce secondary sexual characteristics and facilitate the pubertal growth spurt. Once they have completed pubertal induction, they are currently managed with sex steroid replacement (SSR) often in the convenient form of the combined oral contraceptive pill (COCP) or hormone replacement therapy (HRT) formulations designed for older women following the menopause. With these non-physiological preparations, secondary sexual development is reasonable, but uterine morphology is often unsatisfactory with failure to achieve normal volume and configuration, and a low success rate on oocyte donation programmes (Paterson *et al.*, 2002). Bath *et al.* (2001) reported a wide variety of hormone regimens used by post-pubertal young women, but there is no clear evidence to dictate best practice (Conway, 2001). Current non-physiological hormone regimens (COCP and HRT) are sufficient to suppress menopausal symptoms but are not designed to achieve physiological replacement of estrogen or progesterone, and these regimens have also been shown to be inadequate at achieving normal uterine volume, endometrial thickness (ET) and uterine blood flow (Critchley and Wallace, 2005). It has been shown, however, that physiological SSR (pSSR) achieves sex steroid serum concentrations similar to levels in women with normal ovarian function (Critchley *et al.*, 1990) and endometrial sampling has also demonstrated functional response of the endometrium in women treated with pSSR therapy (Critchley *et al.*, 1990; Critchley and Wallace, 2005).

We thus hypothesized that the use of a pSSR regimen among women with POF would offer therapeutic advantage in terms of improving uterine parameters and that prolonged use would result in sustained improvement during the treatment period. A clinical trial (registration number: NCR00732693) was conducted to investigate the effects of pSSR in comparison with a standard non-physiological regimen of hormone replacement (where each regimen was given for a period of 12 months), upon cardiovascular (Langrish *et al.*, 2009) and skeletal health (Crofton *et al.*, 2010) as well as uterine parameters (uterine volume, ET and uterine blood flow). We here report the findings for the uterine parameters.

# Materials and Methods

The trial received approval from the local research ethics committee and was given Clinical Trial Authorization by the Medicines and Healthcare products Regulatory Agency (UK).

## Subjects

Women with POF secondary to chemotherapy or radiotherapy for the treatment of cancer [leukaemia with 14.4 Gy radiotherapy total body irradiation (TBI), lymphoma with no pelvic radiotherapy or Wilms' tumour with 30 Gy radiotherapy to the abdomen], to surgical oophorectomy with hysterectomy, Turner syndrome or idiopathic POF were recruited from outpatient oncology follow-up and gynaecological clinics in tertiary centres in the South-East of Scotland. Patients with intercurrent illness were ineligible for the trial and women who had undergone hysterectomy and so could not contribute uterine data were also excluded from this analysis. Participants were between 19 and 40 years of age. Recruitment took place between February 2002 and September 2004 and the last follow-up was in November 2006. Written informed consent was provided by participants.

## Study design

This was an open-label randomized controlled crossover trial as previously reported, with each participant receiving both treatments (Langrish *et al.*, 2009; Crofton *et al.*, 2010). Equal 1:1 randomization (to order of treatments) was performed separately for the three aetiological groups (idiopathic/surgical, following cancer treatment, Turner syndrome) in balanced blocks of 10 by opaque standard assignment 'envelopes' produced at the Medical Statistics Unit (University of Edinburgh) using custom-written computer software. Patients and study staff were aware of their treatment regimens throughout the study period. At entry, all patients were receiving, or had previously received, a standard non-physiological hormone replacement regimen. Following a 2-month washout period of no therapy, patients were randomized to first receive either Loestrin 30 (Galen Ltd, UK) (referred to throughout as 'sSSR') or a combination of Estraderm TTS patches (Novartis Pharmaceuticals UK Ltd) and Cyclogest vaginal pessaries (Actavis UK Ltd) (referred to throughout as 'pSSR'). The active ingredients and 4-week regimens are detailed in Table I. After 12 months of the first treatment regimen, a further 2-month washout was completed before changing to the alternative treatment regimen for a further 12 months. The overall design and timing of assessments is summarized in Fig. 1. One woman receiving pSSR during her second treatment period changed after 6 months to using oral progesterone (dydrogesterone 10 mg twice daily; Duphaston, Solvay Healthcare Ltd, UK) in preference to vaginal progesterone pessaries.

Assessments, including a nurse-led consultation to monitor for complications, were made at the start of the first washout period, at the end of each washout and at 3, 6 and 12 months into each treatment period. Primary outcomes were uterine volume and ET and secondary outcomes were uterine blood flow [uterine artery resistance index (UARI) and pulsatility index (UAPI)].

The study also had non-uterine end-points [e.g. cardiovascular (Langrish *et al.*, 2009) and bone (Crofton *et al.*, 2010)], and it was decided *a priori* that the analyses for the three aspects would be undertaken independently, without adjustment across parts for the total number of hypotheses tested. Study size was calculated in relation to bone markers, with a target of 46–56 patients. In total, 42 patients were recruited.

## Measurement of the uterine volume, ET and uterine artery blood flow

At each visit, every patient underwent a detailed transvaginal ultrasound scan using a Siemens Elegra 6.5 MHz endovaginal probe. If transvaginal scanning was not appropriate, a Siemens Elegra 3.5 MHz transabdominal probe was used. One consultant radiologist (J.W.) performed all scans, to minimize inter-observer variability. The radiologist was informed of the aetiology of POF of each patient but was blind to the treatment

**Table I** 'Physiological' and 'standard' sex steroid regimens.

| | Product | Component | Dose | | | | Dosing freq. |
|---|---|---|---|---|---|---|---|
| | | | Week 1 | Week 2 | Week 3 | Week 4 | |
| pSSR | Patches<br>Vaginal pessaries | Estradiol<br>Progesterone | 100 μg | 150 μg | 150 μg<br>200 mg | 150 μg<br>200 mg | /24 h<br>/12 h |
| sSSR | Pill | Ethinylestradiol 30 μg<br>Norethisterone 1.5 mg | Pill-free | 1 pill | 1 pill | 1 pill | /24 h |

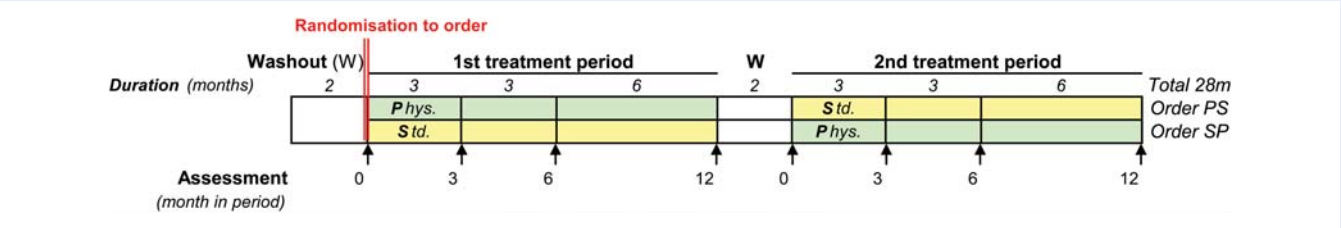


**Figure 1** Time-course of study showing wash-out and treatment periods, plus timing of randomization and assessments.

received. The measurements recorded were uterine volume, ET and uterine artery blood flow as assessed by uterine artery PI and RIs.

If 'cycle' day 1 is taken as the first day of the pill-free week (on the sSSR regimen) or the first day after discontinuing progesterone (on the pSSR regimen), then cycle day 21 can be considered to be the 'mid-luteal phase' of the menstrual cycle. The target for uterine parameter measurements was therefore set as Days 21–28, so that the endometrium would be measured when it was at its maximum thickness. In order to avoid a reduction in the reliability of estimation of mean thickness, measurements made on cycle days 1–7 have been excluded from analyses of ET (six measurements on four patients on sSSR and two measurements on two patients on pSSR). Blood for measurement of gonadotrophins was taken at the same visit and samples taken on cycle days 1–7 have been excluded from analyses.

Uterine body length ($L$), from the fundus to the internal os, the transverse ($T$) and antero-posterior (AP) diameters were measured (cm). Assuming the uterine form to be ellipsoid, the uterine volume was calculated using the formula for a prolate ellipsoid: uterine volume ($cm^3$) $= L \times T \times AP \times 0.523$.

ET was measured at the thickest section of the endometrium on a sagittal section (mm), and is the double thickness of both endometrial layers, excluding any intrauterine fluid.

Uterine artery blood flow was evaluated using colour and pulsed flow Doppler. The uterine artery was identified lateral to the cervix with colour Doppler and then blood waveforms were obtained by placing the Doppler gate over the appropriate area. The PI and RI were calculated as a measure of impedance to uterine blood flow distal to the point of sampling. PI was derived from the flow velocity waveform, using the formula: $PI = (V_{max} - V_{min})/\text{mean}$, where $V_{max}$ is the peak systolic shift frequency, $V_{min}$ the end diastolic shift frequency and mean the mean maximum shift frequency over the cardiac cycle (Taylor *et al.*, 1985). RI was derived using the formula: $RI = (V_{max} - V_{min})/V_{max}$ (Taylor *et al.*, 1985).

## Measurement of circulating estradiol and gonadotrophin concentrations

Venous blood was collected during visits at 0, 3 and 12 months of each hormone treatment regimen targeting Days 21–28 of each 28 day treatment cycle, and circulating concentrations of FSH, LH and estradiol levels were measured. Serum LH and FSH concentrations were determined by fluoroimmunoassay and serum estradiol concentrations were measured by radioimmunoassay.

## Statistical analysis

The uterine end-points (uterine volume, ET, UARI and UAPIs) were each analysed using a mixed model approach, which can take into account the crossover design and repeated measurements within each treatment period (Brown and Prescott, 2006). The model has both a random effect (patient) and fixed effects [treatment, treatment period (first 12 months or second 12 months), time within the treatment period (3, 6 or 12 months assessment) and treatment by time interaction]. A general covariance structure was specified for the repeated measurements on the same patient within each treatment period. For all models, checks were made to ensure that data satisfied the assumptions for this method of analysis (including normality of residuals and homogeneity of variances). There was no evidence that the assumptions were unreasonable. The mixed model approach has the important advantage that it enables patients to be included even if they did not provide data at all three time-points in both treatment periods. For each end-point, there was consideration of the inclusion of baseline covariates, the corresponding baseline (pretreatment) measurements for the two treatment periods and the mean of the two baseline (pretreatment) measurements, which is in accordance with current guidance (Kenward and Roger, 2010). Baseline covariates proved to be valuable for the uterine volume model but not for any of the other end-points. As a precaution, sensitivity analyses were undertaken to check for alterations to estimates of treatment effect if patients with incomplete data were excluded.

# Results

Details for patient flow through the uterine study are documented in Fig. 2. Patient flow for the entire study has previously been reported (Langrish *et al.*, 2009). Forty-two women were recruited and 8 of these withdrew prior to randomization, while a further 5, who had

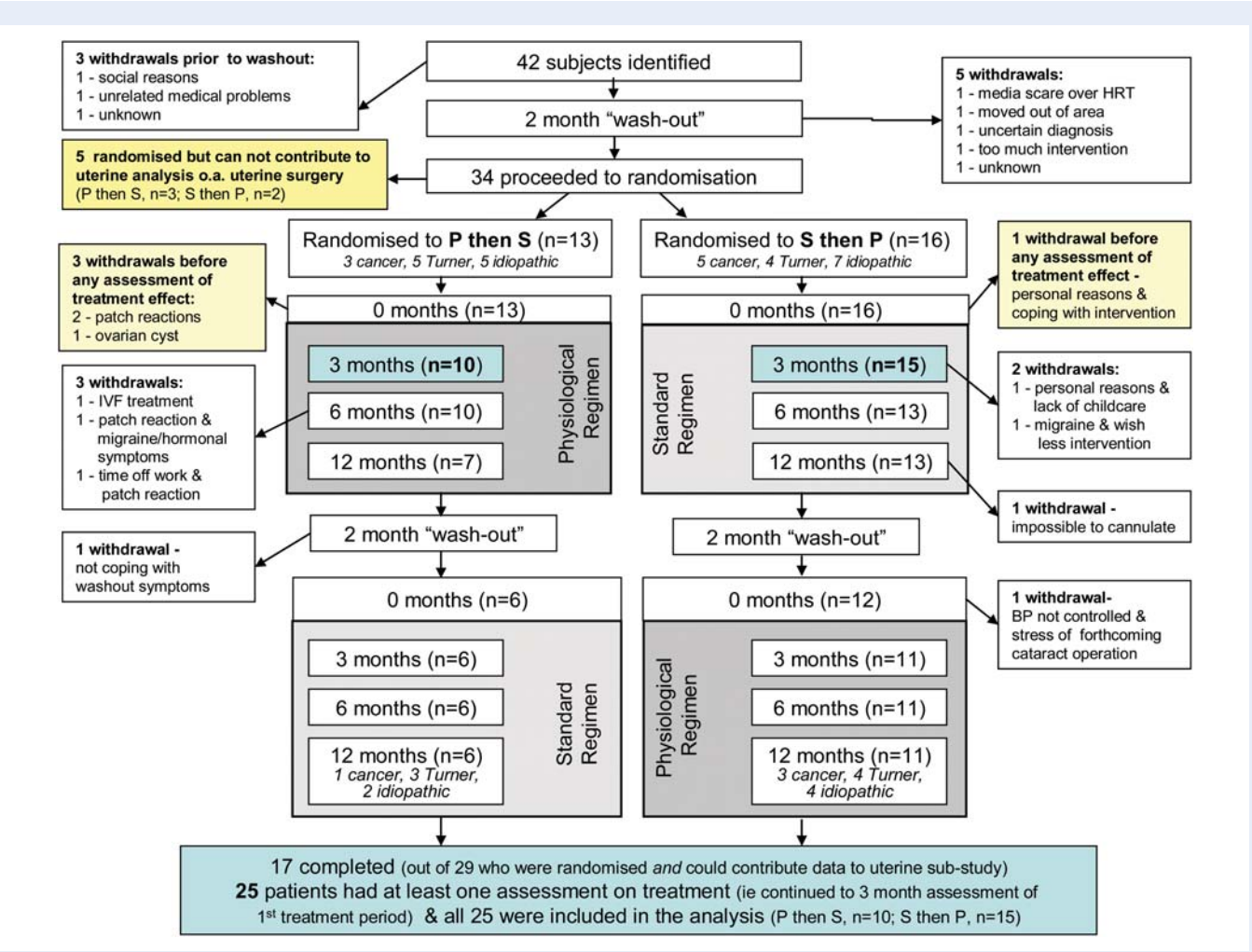


**Figure 2** Flow of participants through uterine study. Adapted from Langrish *et al.* (2009); P, physiological SSR and S, standard SSR.

undergone hysterectomy and oophorectomy and could therefore not contribute uterine data, were ineligible for the uterine study (Fig. 2).

Of the 29 randomized patients able to provide uterine data, 17 (59%) completed the full 28-month study period, although 25 (86%) continued with treatment until at least the 3-month assessment, and so could contribute data to the statistical analysis of treatment effect.

Five patients reported irritant reactions with the Estraderm patches and four of these patients withdrew on account of this side effect. One patient who completed the study changed to Evorel patches (50 and 100 μg estradiol matrix patch, Janssen-Cilag Ltd, UK) due to a local patch reaction.

Baseline data (after first washout, before first treatment commenced) are summarized in Table II by aetiological subgroup. It can be seen that there is a variation in the baseline characteristics between the aetiological groups, with the patients with Turner syndrome being younger and those with idiopathic POF having greater baseline ET and uterine volume. Five of the eight cancer subgroup patients received radiotherapy to or including the pelvis. Two patients received 30 Gy to the abdomen for treatment of Wilms' tumours (at 3 and 4 years of age), while three patients received 14.4 Gy of TBI for treatment of acute lymphoblastic leukaemia (at ages 16, 23 and 27 years).

Figure 3 plots baseline levels for aetiological subgroups, with the idiopathic and cancer subgroups, subdivided according to radiation exposure (cancer subgroup) and parity (idiopathic POF subgroup). Numbers are very small, but the data suggest that the cancer patients with no radiotherapy exposure have greater mean baseline uterine volume than those who received radiotherapy to or including the pelvis (by a factor of 4–7 times). Similarly, within the idiopathic subgroup, parous women had the mean baseline uterine volume nearly four times that of nulliparous women.

## Endometrial thickness

Overall, there was a greater increase in ET in women when receiving physiological replacement, in comparison to when receiving standard therapy (Table III)—a model estimated mean difference of 1.8 mm [95% confidence interval (CI), 0.7–2.8 mm, $P = 0.002$]. This treatment effect was consistent over the three assessment points (3–12 months of treatment), as seen in Fig. 4a.

## Uterine volume

Although there was an overall greater uterine volume among women when receiving pSSR, in comparison to when receiving sSSR (Table III),

**Table II Characteristics of study participants after initial washout, before first treatment, by aetiological subgroup ($n = 25$).**

| | Turner ($n = 7$)[a] | Cancer ($n = 8$)[a] | Idiopathic ($n = 10$)[a] | All ($n = 25$)[a] |
|---|---|---|---|---|
| Age (years) at recruitment to study | | | | |
| Mean (range) | 23 (20–25) | 32 (23–40) | 28 (19–34) | 27 (19–40) |
| Uterine volume ($cm^3$) | | | | |
| Mean (range) | 10.2 (2.8–18.0) | 10.1 (0.6–33.1) | 23.1 (2.8–49.6) | 15.6 (0.6–49.6) |
| Endometrial thickness (mm) | | | | |
| Mean ($n = 7, 7, 10, 24$) (range) | 1.7 (0.8–2.7) | 1.8 (0.6–4.1) | 2.6 (1.1–7.0) | 2.0 (0.6–7.0) |
| Uterine RI | | | | |
| Mean ($n = 5, 7, 10, 22$) (range) | 0.97 (0.91–1.00) | 0.92 (0.84–1.00) | 0.94 (0.90–1.00) | 0.93 (0.84–1.00) |
| Uterine PI | | | | |
| Mean ($n = 5, 7, 10, 22$) (range) | 4.84 (3.27–8.67) | 3.72 (2.39–6.73) | 4.32 (2.68–6.57) | 4.21 (2.68–8.67) |

[a]Where there are missing data for a variable, numbers for analysis of that variable are noted in the first column.

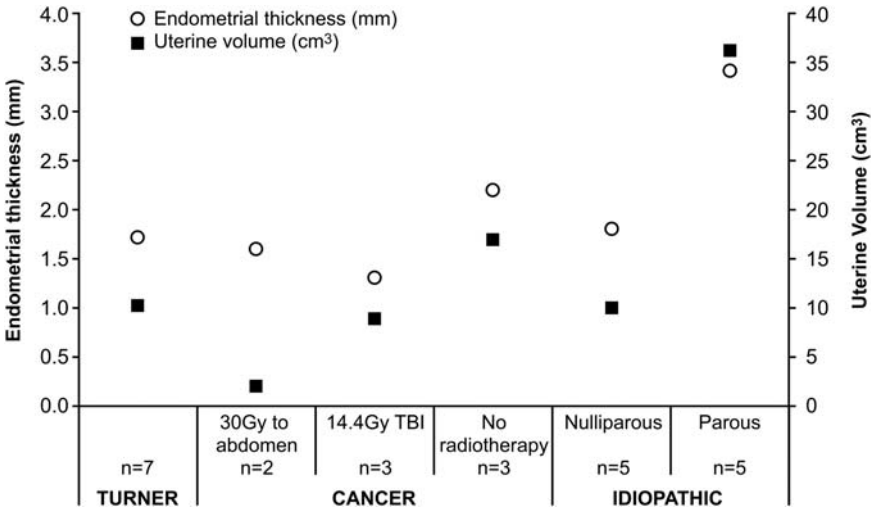


**Figure 3** Mean after washout, before treatment regimens commenced, of ET and uterine volume, by aetiological subgroup.

this difference was not statistically significant, with a model estimated mean difference of 4.2 $cm^3$ (95% CI −0.4 to 8.7 $cm^3$), $P = 0.070$. This small difference was also consistent over the three assessment points, as seen in Fig. 4b.

### Uterine blood flow

Overall, the mean UARI and mean UAPI were only slightly lower when receiving pSSR compared with sSSR, and the differences between the treatments were not statistically significant (Table III).

### Circulating estradiol, and gonadotrophin (FSH and LH) concentrations during washout and treatment

Following washout, the mean hormone assay levels demonstrated elevated gonadotrophins and low estradiol levels, consistent with ovarian failure (Table IV). Serum estradiol levels were substantially higher with pSSR than sSSR, by a factor of six or more. However, mean gonadotrophins were very similar on the two regimens.

## Discussion

This study demonstrates that 12 months of treatment with pSSR therapy results overall in a greater ET (1.8 mm difference, $P = 0.002$) in comparison to a 12-month standard regimen of hormone replacement. There was also a borderline increase in the uterine volume (difference 4.2 $cm^3$, $P = 0.070$). No significant difference in uterine blood flow was demonstrated between the two regimens. Although the numbers in this study appear small, the study comprises the largest randomized trial comparing physiological with standard SSR in women with POF. Furthermore, appropriate statistical

**Table III Overall mean with SE for each treatment regimen and difference between treatment regimens with 95% CI ($n = 25$).**

| Uterine end-points | Model-adjusted mean[a] (SE) | | Treatment effect | |
|---|---|---|---|---|
| | Physiological SSR | Standard SSR | Difference[b] (95% CI) | *P*-value |
| Endometrial thickness (mm) | 4.8 (0.38) | 3.0 (0.41) | 1.8 (0.7–2.8) | 0.002 |
| Uterine volume ($cm^3$) | 24.8 (1.9) | 20.6 (2.0) | 4.2 (−0.4–8.7) | 0.070 |
| Uterine artery resistance index | 0.89 (0.013) | 0.90 (0.014) | −0.01 (−0.03–0.01) | 0.39 |
| Uterine artery pulsatility index | 3.04 (0.20) | 3.24 (0.21) | −0.20 (−0.56–0.17) | 0.27 |

[a]Overall mean using assessments at 3, 6 and 12 months.
[b]Difference is physiological SSR minus standard SSR.

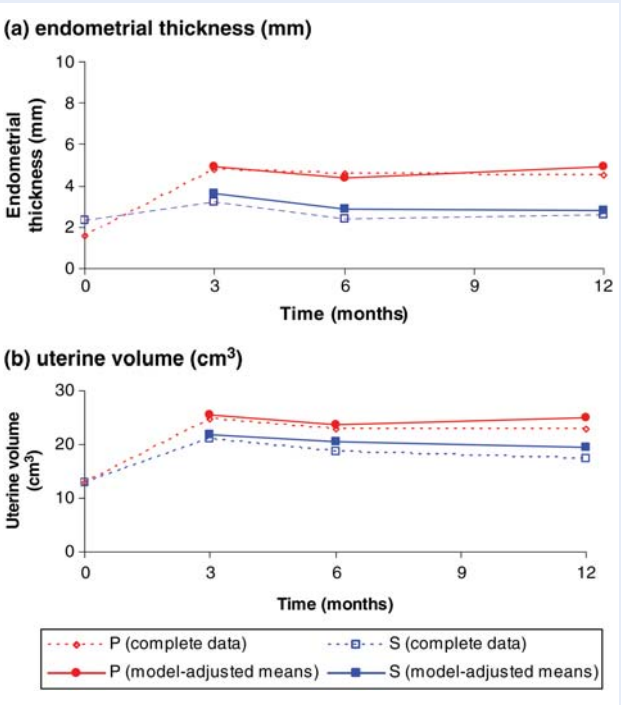


**Figure 4** Model-adjusted means over time and means of observed values over time for patients with complete data from all time points for each treatment regimen. P, physiological; S, standard. Solid lines represent model-adjusted means ($n = 25$). Dashed lines represent means of observed values for patients with complete data from all time points ($n = 14$).

methods have been used to ensure valid analysis of treatment effect and to utilize as much as possible of the data collected.

No previous study has directly compared pSSR and sSSR regimens, regarding uterine parameters in women with POF, but our study findings are consistent with results of multiple smaller studies, using similar physiological regimens in POF patient subgroups (Critchley *et al.*, 1990; Biljan *et al.*, 1995; Bath *et al.*, 1999; Holm *et al.*, 1999).

One of the most important issues for patients with POF is its impact upon their reproductive function and the implication for the health of their offspring (British Fertility Society Multidisciplinary Working Group, 2003; Hudson, 2010). The potential for preservation of fertility for young women affected by childhood cancer is an area where assisted reproductive technologies are rapidly advancing (Wallace *et al.*, 2005a). This is important for those women in whom a premature menopause may be predicted (Wallace *et al.*, 2005b). In most cases, however, the diagnosis of POF is made in retrospect and these women are therefore dependent upon donor oocytes and embryo transfer for pregnancy. A successful pregnancy, however, not only depends upon a functioning hypothalamic–pituitary–ovarian axis but also on a uterine cavity that is receptive to implantation, with the capacity to accommodate growth of the fetus to term. pSSR improves ET significantly in comparison with standard SSR with the mean ET after 12 months of pSSR being 5.0 mm. This might be considered suboptimal for embryo transfer as part of traditional assisted reproduction in women with ovarian function, but it might be sufficient to achieve a pregnancy. Al-Ghamdi *et al.* (2008) demonstrated a steady and gradual increase in the IVF pregnancy rate in women with ovarian function with an increasing ET, but as ET was not a reliable predictor of pregnancy rate, they recommended that IVF treatment should not be restricted based upon this alone. There have also been several reports of successful pregnancies resulting from cycles with ET of <4 mm, indicating that a thinner endometrium does not necessarily preclude the possibility of implantation (Remohi *et al.*, 1997; Sundstrom, 1998). It is not yet clear whether higher doses of pSSR would improve ET further. However, since pSSR has previously been shown to establish a physiological endometrium in women with POF (Critchley *et al.*, 1990), the combination of a functional change and an improvement in thickness may be sufficient to achieve a pregnancy.

Achieving and maintaining a pregnancy to term may not be possible for all women with POF even with pSSR. The uterine volume in women with POF is often reduced to 40% of the normal adult size (Bath *et al.*, 1999; Critchley and Wallace, 2005), with poor blood flow and a thin endometrium. This provides a suboptimal environment for pregnancy. Further damage may also be induced by radiotherapy during treatment of cancer, resulting in a reduced uterine volume and decreased elasticity of uterine musculature (Critchley *et al.*, 1992; Holmes and Shalet, 1996; Bath *et al.*, 1999). The extent of damage caused by gonadotoxic treatments of cancer is related to age at treatment, the chemotherapeutic regimen used and the dose of abdominal radiation (Meirow *et al.*, 2010).

**Table IV Mean serum gonadotrophin (FSH, LH) and estradiol levels for each treatment regimen ($n = 25$).**

| | Assessment (months) | Physiological SSR | | | Standard SSR | | |
|---|---|---|---|---|---|---|---|
| | | *n* | Mean | SD | *n* | Mean | SD |
| FSH (U/l) | 0 | 20 | 80.4 | 35.8 | 20 | 78.5 | 41.5 |
| | 3 | 20 | 18.3 | 15.4 | 19 | 11.1 | 10.9 |
| | 12 | 17 | 26.8 | 25.7 | 15 | 22.4 | 24.2 |
| LH (U/l) | 0 | 20 | 34.5 | 15.5 | 20 | 38.3 | 19.9 |
| | 3 | 20 | 10.6 | 9.7 | 19 | 9.3 | 11.9 |
| | 12 | 17 | 19.7 | 24.5 | 15 | 9.5 | 10.5 |
| Estradiol (pmol/l) | 0 | 20 | 114 | 179 | 20 | 94 | 102 |
| | 3 | 20 | 381* | 230 | 19 | 67 | 41 |
| | 12 | 17 | 525* | 536 | 15 | 63 | 25 |

'0 months' is assessment post-washout, before treatment commenced.
*Normalized estradiol levels.

Evidence for pregnancy outcome in women treated for childhood cancer demonstrates that there is risk of early pregnancy loss, premature labour and low birthweight, and that these complications are more prevalent in patients exposed to abdominal radiation, with the risk being greatest for those treated prepubertally (Sanders *et al.*, 1996; Green *et al.*, 2009; Hudson, 2010; Lie Fong *et al.*, 2010). Women with ovarian failure, who achieve a pregnancy, should therefore be considered high-risk obstetric patients with regular assessment of fetal growth. Reassuringly, available evidence demonstrates no excess of congenital anomaly in the offspring of survivors of childhood cancer (Hudson, 2010; Meirow *et al.*, 2010).

The main limitations of this study are the withdrawal rate and the heterogeneity of the study population. The study protocol was intensive and more women withdrew during the first treatment phase compared with the second phase, suggesting a lack of tolerance of the research protocol as much as the intervention. Of those who withdrew, only four women clearly withdrew due to intolerance of the treatment, and this was regarding adverse reactions to the patch. The open nature of the trial with participant knowledge of the cyclical pattern on the hormone replacement within the physiological regimen may have contributed to the psychological intolerance of the patch. However, the use of statistical methods suited to such data ensured that of the 29 patients who were able to provide uterine data and were randomized, 25 contributed data to the statistical analysis (86%).

A further difficulty with research into POF is the heterogeneous aetiology of the cases. There may be a variation in response to SSR between different aetiological groups. Several earlier studies have suggested a difference in response to pSSR with respect to aetiology of POF and the age of diagnosis. Biljan *et al.* (1995) demonstrated an increased response in UAPI with pSSR therapy (2 mg oral estradiol valerate with 500 μg norgesterel for 10 days of a 28-day cycle) in patients with Turner syndrome ($n = 5$) in comparison with those patients with surgical removal of ovaries ($n = 6$) . There is also evidence that young women ($n = 12$) treated with TBI, for bone marrow transplantation for childhood leukaemia or lymphoma, have impaired uterine growth and blood flow in comparison with normal controls (Holm *et al.*, 1999). These patients have a reduced uterine volume and sSSR has been reported to be inadequate to achieve normal uterine growth (Holm *et al.*, 1999). The timing of the exposure to irradiation appears to be important. Bath *et al.* (1999) described that following 3 months of the same pSSR used in this study, a smaller increase in the uterine volume from baseline was observed in one young woman, who had undergone TBI pre-pubertally ($n = 1$), compared with the average response for three young women exposed to radiation post-puberty. This study also reported a positive correlation between the uterine volume and age at TBI (Bath *et al.*, 1999). It has also been reported that the endometrium of women exposed to abdominal irradiation for malignant disease in childhood ($n = 3$) were less responsive to pSSR (Estraderm 50–150 μg/24 h with cyclical 300–400 mg of vaginal progesterone in divided doses/24 h) compared with that of women who had absent ovarian function for other reasons ($n = 22$) (Critchley *et al.*, 1992).

Observations from small studies (Critchley *et al.*, 1990; Biljan *et al.*, 1995; Bath *et al.*, 1999; Holm *et al.*, 1999) using pSSR short term, alongside the data from the present study, indicate potential for improvement in uterine characteristics in at least selected POF patients. Evidence for the potential for successful pregnancy following optimization of uterine characteristics is, however, still unclear and the initial observations of pregnancy following irradiation suggest that women exposed to radiotherapy pre-pubertally are less likely to achieve a successful pregnancy than those exposed later. Realistic counselling must be provided for this subgroup of women. Long-term studies of SSR regimens are required to further assess uterine response and pregnancy outcome according to aetiological subgroup and exposure to radiotherapy.

Women with early menopause are at increased risk of mortality compared with the normal population (Jacobsen *et al.*, 1999; Ossewaarde *et al.*, 2005). The high morbidity and mortality observed in women with Turner syndrome (Schoemaker *et al.*, 2008) and as a consequence of cancer treatment (Reulen *et al.*, 2010) is believed to be multi-factorial, and the extent to which the POF is a contributing factor is not known. Our study group has recently reported that pSSR therapy results in lower mean 24-h systolic and diastolic blood pressure, improves renal function and results in less activation of the

renin–angiotensin system in comparison with sSSR, thereby improving cardiovascular health (Langrish *et al.*, 2009). pSSR also has a beneficial effect upon bone mass acquisition on the lumbar spine, mediated by increased bone formation and decreased bone resorption (Crofton *et al.*, 2010). However, while there is evidence demonstrating that SSR has a protective effect against fracture in women with a normally timed menopause, there is a lack of evidence regarding the possibility that SSR provides substantial protection against coronary heart disease in this group (Beral *et al.*, 2005). Furthermore, balanced against this skeletal benefit, there are well-documented risks associated with SSR in post-menopausal women, such as increased risks of breast cancer, cerebral vascular accidents, venous thromboembolism and gallbladder disease (Beral *et al.*, 2005). However, for women with POF, there is good evidence demonstrating an excess cardiovascular and skeletal morbidity without steroid replacement (Jacobsen *et al.*, 1999; Ossewaarde *et al.*, 2005), but as yet, there is no robust evidence regarding the relative safety of long-term sex steroid hormone replacement. Further studies with long-term follow-up of POF cases are required.

In order to achieve the cardiovascular and skeletal health benefits of pSSR, long-term treatment is needed. In contrast, for reproductive purposes, pSSR treatment may not be needed long term, being used by patients only for a limited time for the purpose of fertility treatment. This study demonstrates that, in general, treatment response is apparent after only 3 months of physiological treatment (Fig. 3). Although the mean absolute levels of the uterine volume and ET achieved in this study are less than those seen in a normal population of women of reproductive age, such a response could still be worthwhile for women with POF who seek infertility treatment, since it demonstrates physical improvements in uterine characteristics with the use of pSSR. It may be possible for women to adopt a regimen of pSSR for short periods of time prior to fertility treatment in order to optimize the uterine parameters for embryo transfer. Further studies addressing the impact of pSSR on the success of assisted reproduction and pregnancy outcome will be required.

## Conclusions

In women with POF, pSSR results in significantly greater ET than standard non-physiological regimens, with a trend towards increased uterine volume. Further studies are required to evaluate the impact of pSSR on uterine characteristics, fertility and pregnancy outcomes, within the various POF aetiological subgroups.

## Acknowledgements

We thank the women who participated, our research nurses (Caroline Valentine and Morag Charles), Angela Smith for her administrative support and Dr Jeremy Langrish for his assistance with data collaboration. We acknowledge the secretarial assistance of Mrs Sheila Milne with manuscript preparation.

## Authors' roles

R.L.O'D.: data collation, analysis and interpretation, manuscript preparation and critical discussion. P.W.: study design, data analysis and interpretation, manuscript preparation and critical discussion. R.J.L.: data analysis and interpretation. J.W.: ultrasound examination of patients, manuscript preparation. L.E.B.: study design, execution of study and critical discussion. C.J.K.: study design, execution of study. W.H.B.W. and H.O.D.C.: study design, execution of study, manuscript preparation and critical discussion.

## Funding

This research was supported by a project grant from CLIC Sargent (R35464) and conducted with the assistance of the Wellcome Trust Clinical Research Facility, Edinburgh.

## Conflict of interest

There are no author conflicts of interest to declare and no drug company sponsorship.

## References

Aittomaki K, Lucena JL, Pakarinen P, Sistonen P, Tapanainen J, Gromoll J, Kaskikari R, Sankila EM, Lehvaslaiho H, Engel AR *et al.* Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell* 1995;**82**:959–968.

Al-Ghamdi A, Coskun S, Al-Hassan S, Al-Rejjal R, Awartani K. The correlation between endometrial thickness and outcome of *in vitro* fertilization and embryo transfer (IVF-ET) outcome. *Reprod Biol Endocrinol* 2008;**6**:37.

Bath LE, Critchley HO, Chambers SE, Anderson RA, Kelnar CJ, Wallace WH. Ovarian and uterine characteristics after total body irradiation in childhood and adolescence: response to sex steroid replacement. *Br J Obstet Gynaecol* 1999;**106**:1265–1272.

Bath LE, Critchley HO, Kelnar CJ, Wallace WH. Choice of hormone replacement therapy in young women with ovarian failure. *Clin Endocrinol (Oxf)* 2001;**55**:697–698.

Beau I, Touraine P, Meduri G, Gougeon A, Desroches A, Matuchansky C, Milgrom E, Kuttenn F, Misrahi M. A novel phenotype related to partial loss of function mutations of the follicle stimulating hormone receptor. *J Clin Invest* 1998;**102**:1352–1359.

Beral V, Reeves G, Banks E. Current evidence about the effect of hormone replacement therapy on the incidence of major conditions in postmenopausal women. *Br J Obstet Gynaecol* 2005;**112**:692–695.

Biljan MM, Taylor CT, Matijevic R, Jones SV, Garden AS, Fraser WD, Diver MJ, Kingsland CR. Exaggerated effects of progestogen on uterine artery pulsatility index in Turner's syndrome patients receiving hormone replacement therapy. *Fertil Steril* 1995;**64**:1104–1108.

British Fertility Society Multidisciplinary Working Group. A strategy for fertility services for survivors of childhood cancer. *Hum Fertil (Camb)* 2003;**6**:A1–A39.

Brown H, Prescott RJ. *Applied Mixed Models in Medicine,* 2nd edn. Chichester: Wiley, 2006.

Conway GS. Oestrogen replacement in young women with Turner's syndrome. *Clin Endocrinol (Oxf)* 2001;**54**:157–158.

Cooper GS, Sandler DP. Age at natural menopause and mortality. *Ann Epidemiol* 1998;**8**:229–235.

Critchley HO, Wallace WH. Impact of cancer treatment on uterine function. *J Natl Cancer Inst Monogr* 2005;**2005**:64–68.

Critchley HO, Buckley CH, Anderson DC. Experience with a 'physiological' steroid replacement regimen for the establishment of a receptive endometrium in women with premature ovarian failure. *Br J Obstet Gynaecol* 1990;**97**:804–810.

Critchley HO, Wallace WH, Shalet SM, Mamtora H, Higginson J, Anderson DC. Abdominal irradiation in childhood; the potential for pregnancy. *Br J Obstet Gynaecol* 1992;**99**:392–394.

Crofton PM, Evans N, Bath LE, Warner P, Whitehead TJ, Critchley HO, Kelnar CJ, Wallace WH. Physiological versus standard sex steroid replacement in young women with premature ovarian failure: effects on bone mass acquisition and turnover. *Clin Endocrinol (Oxf)* 2010;**73**:707–714.

De Baere E, Beysen D, Oley C, Lorenz B, Cocquet J, De Sutter P, Devriendt K, Dixon M, Fellous M, Fryns JP *et al.* FOXL2 and BPES: mutational hotspots, phenotypic variability, and revision of the genotype-phenotype correlation. *Am J Hum Genet* 2003;**72**:478–487.

Di Pasquale E, Beck-Peccoz P, Persani L. Hypergonadotropic ovarian failure associated with an inherited mutation of human bone morphogenetic protein-15 (BMP15) gene. *Am J Hum Genet* 2004;**75**:106–111.

Green DM, Sklar CA, Boice JD Jr, Mulvihill JJ, Whitton JA, Stovall M, Yasui Y. Ovarian failure and reproductive outcomes after childhood cancer treatment: results from the Childhood Cancer Survivor Study. *J Clin Oncol* 2009;**27**:2374–2381.

Holm K, Nysom K, Brocks V, Hertz H, Jacobsen N, Muller J. Ultrasound B-mode changes in the uterus and ovaries and Doppler changes in the uterus after total body irradiation and allogeneic bone marrow transplantation in childhood. *Bone Marrow Transplant* 1999;**23**:259–263.

Holmes SJ, Shalet SM. Role of growth hormone and sex steroids in achieving and maintaining normal bone mass. *Horm Res* 1996;**45**:86–93.

Hovatta O. Pregnancies in women with Turner's syndrome. *Ann Med* 1999;**31**:106–110.

Hudson MM. Reproductive outcomes for survivors of childhood cancer. *Obstet Gynecol* 2010;**116**:1171–1183.

Jacobsen BK, Knutsen SF, Fraser GE. Age at natural menopause and total mortality and mortality from ischemic heart disease: the Adventist Health Study. *J Clin Epidemiol* 1999;**52**:303–307.

Kenward MG, Roger JH. The use of baseline covariates in crossover studies. *Biostatistics* 2010;**11**:1–17.

Langrish JP, Mills NL, Bath LE, Warner P, Webb DJ, Kelnar CJ, Critchley HO, Newby DE, Wallace WH. Cardiovascular effects of physiological and standard sex steroid replacement regimens in premature ovarian failure. *Hypertension* 2009;**53**:805–811.

Lie Fong S, van den Heuvel-Eibrink MM, Eijkemans MJ, Schipper I, Hukkelhoven CW, Laven JS. Pregnancy outcome in female childhood cancer survivors. *Hum Reprod* 2010;**25**:1206–1212.

Meirow D, Biederman H, Anderson RA, Wallace WH. Toxicity of chemotherapy and radiation on female reproduction. *Clin Obstet Gynecol* 2010;**53**:727–739.

Ossewaarde ME, Bots ML, Verbeek AL, Peeters PH, van der Graaf Y, Grobbee DE, van der Schouw YT. Age at menopause, cause-specific mortality and total life expectancy. *Epidemiology* 2005;**16**:556–562.

Paterson WF, Hollman AS, Donaldson MD. Poor uterine development in Turner syndrome with oral oestrogen therapy. *Clin Endocrinol (Oxf)* 2002;**56**:359–365.

Rebar RW, Connolly HV. Clinical features of young women with hypergonadotropic amenorrhea. *Fertil Steril* 1990;**53**:804–810.

Remohi J, Ardiles G, Garcia-Velasco JA, Gaitan P, Simon C, Pellicer A. Endometrial thickness and serum oestradiol concentrations as predictors of outcome in oocyte donation. *Hum Reprod* 1997;**12**:2271–2276.

Reulen RC, Winter DL, Frobisher C, Lancashire ER, Stiller CA, Jenney ME, Skinner R, Stevens MC, Hawkins MM. Long-term cause-specific mortality among survivors of childhood cancer. *J Am Med Assoc* 2010;**304**:172–179.

Sanders JE, Hawley J, Levy W, Gooley T, Buckner CD, Deeg HJ, Doney K, Storb R, Sullivan K, Witherspoon R *et al.* Pregnancies following high-dose cyclophosphamide with or without high-dose busulfan or total-body irradiation and bone marrow transplantation. *Blood* 1996;**87**:3045–3052.

Schoemaker MJ, Swerdlow AJ, Higgins CD, Wright AF, Jacobs PA. Cancer incidence in women with Turner syndrome in Great Britain: a national cohort study. *Lancet Oncol* 2008;**9**:239–246.

Sundstrom P. Establishment of a successful pregnancy following in-vitro fertilization with an endometrial thickness of no more than 4 mm. *Hum Reprod* 1998;**13**:1550–1552.

Taylor KJ, Burns PN, Wells PN, Conway DI, Hull MG. Ultrasound Doppler flow studies of the ovarian and uterine arteries. *Br J Obstet Gynaecol* 1985;**92**:240–246.

Wallace WH, Anderson RA, Irvine DS. Fertility preservation for young patients with cancer: who is at risk and what can be offered? *Lancet Oncol* 2005a;**6**:209–218.

Wallace WH, Thomson AB, Saran F, Kelsey TW. Predicting age of ovarian failure after radiation to a field that includes the ovaries. *Int J Radiat Oncol Biol Phys* 2005b;**62**:738–744.