The effectiveness of foot reflexology in inducing ovulation: a sham-controlled randomized trial

Jane Holt, R.N.,[a] *Jonathan Lord, M.D.,*[b,c] *Umesh Acharya, M.D.,*[b] *Adrian White, M.D.,*[d] *Nyree O'Neill, R.N.,*[b,e] *Steve Shaw, Ph.D.,*[f] *and Andy Barton, M.S.*[g]

[a] South Devon School of Reflexology, Morningside, Loddiswell, Devon; [b] South West Centre for Reproductive Medicine, Derriford Hospital, Plymouth, Devon; [c] Royal Cornwall Hospital, Truro, Cornwall; [d] Primary Care Research, Peninsula Medical School, Plymouth, Devon; [e] Bolton Hospice, Bolton, Lancashire; [f] School of Mathematics and Statistics, University of Plymouth, Plymouth, Devon; and [g] Peninsula Research and Development Support Unit (Plymouth), Derriford Hospital, Plymouth, Devon, United Kingdom

Objective: To determine whether foot reflexology, a complementary therapy, has an effect greater than sham reflexology on induction of ovulation.
Design: Sham-controlled randomized trial with patients and statistician blinded.
Setting: Infertility clinic in Plymouth, United Kingdom.
Patient(s): Forty-eight women attending the clinic with anovulation.
Intervention(s): Women were randomized to receive eight sessions of either genuine foot reflexology or sham reflexology with gentle massage over 10 weeks.
Main Outcome Measure(s): The primary outcome was ovulation detected by serum progesterone level of >30 nmol/L during the study period.
Result(s): Twenty-six patients were randomized to genuine reflexology and 22 to sham (one randomized patient was withdrawn). Patients remained blinded throughout the trial. The rate of ovulation during true reflexology was 11 out of 26 (42%), and during sham reflexology it was 10 out of 22 (46%). Pregnancy rates were 4 out of 26 in the true group and 2 out of 22 in the control group. Because of recruitment difficulties, the required sample size of 104 women was not achieved.
Conclusion(s): Patient blinding of reflexology studies is feasible. Although this study was too small to reach a definitive conclusion on the specific effect of foot reflexology, the results suggest that any effect on ovulation would not be clinically relevant. Sham reflexology may have a beneficial general effect, which this study was not designed to detect. (Fertil Steril® 2009;91:2514–9. ©2009 by American Society for Reproductive Medicine.)

Key Words: Infertility, anovulation, reflexology, complementary therapy, randomized controlled trial

Infertility affects around one in seven couples, and a diagnosis of anovulation is made in approximately one third of cases (1). Traditional treatment is ovulation induction using drugs such as clomiphene. Although clomiphene is effective in 70% to 85% of patients (2), it does have side effects. The most important adverse effect is to increase the likelihood of multiple pregnancy from around 1:80 with natural conception to about 1:10 with clomiphene (3). Furthermore, clomiphene is responsible for 17% of all high order multiple births in the United Kingdom (4). There is also a theoretical concern that clomiphene may increase the risk of developing ovarian carcinoma, and for this reason it is only licensed for 6 months' use.

Foot reflexology is a complementary therapy in which the feet are massaged with the intention of gaining specific health benefits. Foot massage itself is recorded from Ancient Egypt, India, and China, but the modern version dates from the early 20th century when Fitzgerald, an ear, nose, and throat surgeon, used massage for its analgesic effects and developed the concept that zones on the feet (and hands) were linked to other areas and organs of the body (5). This theory evolved into a clinical practice in which organs of the body are represented somatotopically on the sole of the foot, reminiscent of the homunculus of cortical sensation. Massage or pressure on the foot is used with the intention of correcting functional disturbances in the remote organ.

Reflexology is used for a range of health problems in the United Kingdom, and accounted for 6% of complementary medicine use in one survey (6). It is available in about 1% of all primary care practices, according to one survey (7). Reflexology is used as an intervention for infertility, and anecdotal reports of success have received publicity through the national press (8). Its reputation has been further enhanced by recommendations from television celebrities (9).

The attractiveness of reflexology as a potential treatment for infertility is understandable as it is pleasant and nonpharmacologic. Couples have been happy to accept anecdotal evidence and pay for a therapy that does not reach the required standard of conventional scientific evidence for it to be an

Received February 25, 2008; revised and accepted April 7, 2008; published online June 19, 2008.
J.H. is a paid tutor of reflexology training courses. J.L. has nothing to disclose. U.A. has nothing to disclose. A.W. has nothing to disclose. N.O'N. has nothing to disclose. S.S. has nothing to disclose. A.B. has nothing to disclose.
Partly funded by the Association of Reflexologists, Taunton, Somerset, UK, whose representatives attended steering committee meetings but had no influence on the analysis.
Reprint requests: Adrian White, M.D., Primary Care Research, Peninsula Medical School, Room 32, ITTC Building, PL6 8BX (FAX: 01752-764234; E-mail: adrian.white@pms.ac.uk).

0015-0282/09/$36.00
doi:10.1016/j.fertnstert.2008.04.016

accepted part of National Health Service care. Such rigorous evidence is urgently required to provide reliable information to patients and to health-care providers on whether reflexology is a cheap and effective intervention for women with anovulatory problems.

Our trial was designed to test the hypothesis that true reflexology is superior to sham reflexology in inducing ovulation. Secondary objectives included determining whether reflexology treatment has any effect on self-reported anxiety, changes to menstrual regularity, and skin greasiness (reported anecdotally by some patients after reflexology).

MATERIALS AND METHODS

Design

The research design was a sham-controlled, randomized trial with patient and statistician blinded to allocation. The study was granted ethics approval by the Plymouth LREC (ref 1404), Torbay LREC (no reference number), and Exeter LREC (204017) local research ethics committees.

Population

All women attending the South West Centre for Reproductive Medicine at Derriford Hospital, Plymouth, with anovulation who would otherwise have been given drugs for ovulation induction were invited to participate. More centers (in Exeter and Torbay) were included later because of poor recruitment; although ethics approval was granted, the trial was terminated before the additional centers had recruited any patients. The inclusion criteria were age 18 to 38 years with anovulation (defined as oligomenorrhea: less than 6 menstrual periods in the preceding 12 months, or with a luteal phase progesterone of $<$28 nmol/L, the standard diagnostic figure of the local clinical chemistry department). All patients had biochemical evidence of raised androgens in keeping with the diagnostic criteria for polycystic ovary syndrome proposed by the U.S. National Institutes of Health and European Society of Human Reproduction and Embryology/American Society for Reproductive Medicine (ESHRE/ASRM).

Exclusion criteria were: coexistence of diabetes mellitus, thyroid disease, late-onset adrenal hyperplasia (raised 17α-hydroxyprogesterone); use of ovulation inducing agents within previous 2 months; previous reflexology treatment; pregnancy (excluded by high sensitivity urine β human chorionic gonadotropin); and contraindications to reflexology (active thrombosis or phlebitis, foot deformity, recent sprain or trauma to ankles or feet, infections or inflammatory condition of the feet, topical steroid use of longer than 6 months, pacemaker). Because of the numbers of very obese women recruited in the early stages of the trial, a group of women who have known treatment resistance (10), body mass index (BMI) $\geq$35 was added as an exclusion criterion in February 2002.

Potential recruits were provided with information about the trial and about reflexology. Inclusion and exclusion criteria were applied by the infertility clinic nurse (N.O.), who then referred the patient to the reflexologist together with the next sequentially numbered, sealed, opaque envelope containing a randomization code generated by computer in blocks of 20. The randomization was conducted and envelopes sealed by another member of the research team (A.B.) in an office unconnected with the infertility clinic.

The reflexologist arranged the first and subsequent appointments at the Centre for Reproductive Medicine or at the School of Reflexology. During the first appointment the reflexologist took a detailed medical history and provided further details about the nature of the treatment, before opening the envelope to determine whether the patient was in treatment or sham control group.

Intervention

Both active and control treatment was provided by the first author, a qualified nurse registered as a reflexologist with the Association of Reflexology (AoR), and a teacher of an AoR accredited course, having 12 years' clinical experience. The true reflexology group were given conventional reflexology (known as the Bayly method) with firm, accurate pressure to tender points in the feet, including the particular points that are believed to represent the following parts or conditions of the body: hypothalamus, pituitary, adrenals, ovaries, fallopian tubes, uterus, and "chronic uterus" (which is reported to reflect the activity of female hormones), thyroid, parathyroids, lumbar/sacral spine, and bladder area, which are considered relevant to fertility; and "solar plexus" as an irrelevant reference point (Fig. 1).

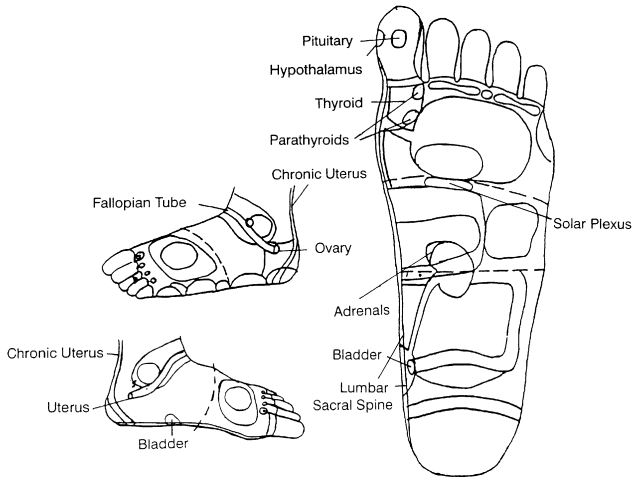
The sham treatment group were given a gentle massage to the foot, avoiding direct pressure on the organ points listed above but giving firm pressure to the "solar plexus" point. In both cases, the duration of the intervention was about 1 hour. Patients received seven sessions of intervention or control over 6 weeks, with a further treatment 1 month later. Patient blinding was maintained throughout the trial. An attempt was made to standardize the interaction between the reflexologist and the patient, to reduce any difference in expectation between the two groups. Neither patient nor reflexologist were made aware of the results of blood tests until 2 months after the end of treatment when the investigations had been completed.

Outcomes

Baseline (no more than 3 months before recruitment) variables included the presence of polycystic ovary syndrome, defined (at that time) as the presence of more than five peripherally located follicles measuring 1 to 5 mm in diameter on transvaginal ultrasound scan together with one biochemical feature of either a raised free androgenic index ($>$5.0) or a raised luteinizing hormone/follicle-stimulating hormone (LH:FSH) ratio ($>$2.5:1). Other endocrine causes of anovulation (diabetes, thyroid disease, raised prolactin, late-onset adrenal hyperplasia, pregnancy) were excluded by applying standard clinical protocols in the management of anovulation in secondary care.

FIGURE 1

Diagram of the foot showing the reflexology areas used in the study, named according to their supposed representation.



Holt. Foot reflexology for ovulation induction. Fertil Steril 2009.

Body mass index was measured at the clinic. The Hospital Anxiety and Depression (HAD) scale was completed by the patient before the first treatment and after the last; at the same time, a symptom checklist was completed, asking participants to rate the regularity and heaviness of their periods, greasiness of skin, and change in skin, using Likert-type categorical scales. These outcome measures were administered by the clinic nurse at enrollment, and by the reflexologist at each patient's last attendance. The reflexologist also recorded the tenderness of the relevant foot points (scale of 0 to 10, called the "reflexology score") in both groups, but only baseline data are presented in this report.

Serum progesterone was measured (for those with regular cycles) 7 days before the expected next onset of menstruation. Oligomenorrheic patients (i.e., with menstruation intervals greater than 6 weeks) were assumed to be not ovulating and categorized as a treatment failure. Those who menstruated during the trial had progesterone measured on days 14, 21, 28, and 35 (unless menstruation occurred on the latter three days). Blood test results were made available to the patients but not to the reflexologist.

Success of participant blinding was assessed after the final reflexology session by the reflexologist asking the single question "Do you think you received the active treatment or the dummy treatment?" Possible responses were active, dummy, and unsure.

All data were entered into Excel spreadsheets by the infertility clinic nurse (N.O.), who remained blinded to patient group allocation.

Sample Size and Power Calculation

Our principal outcome was the number of women in both groups who showed signs of ovulation, indicated by a luteal phase serum progesterone level of >30 nmol/L, during the 10-week course of reflexology treatment. With this group of women, a background rate was likely to be about 20%, taken as a conservative estimate from the mean ovulation rate in the placebo arm of trials included in a Cochrane review investigating women with anovulation (11). An increase to 50% would be detectable with 52 women in each arm of the trial, given a statistical significance level of 5% and 90% power. We therefore originally intended to recruit 52 women to each arm. The effectiveness of 50% was chosen as the minimum acceptable in clinical practice. Data were analyzed blind to allocation at the School of Mathematics and Statistics, University of Plymouth. Groups were compared using chi-square (or Fisher's exact) tests, Mann-Whitney tests, or *t*-tests, as appropriate.

RESULTS

Recruitment to the study commenced in January 2001. In spite of extending the trial to neighboring centers, recruitment

TABLE 1

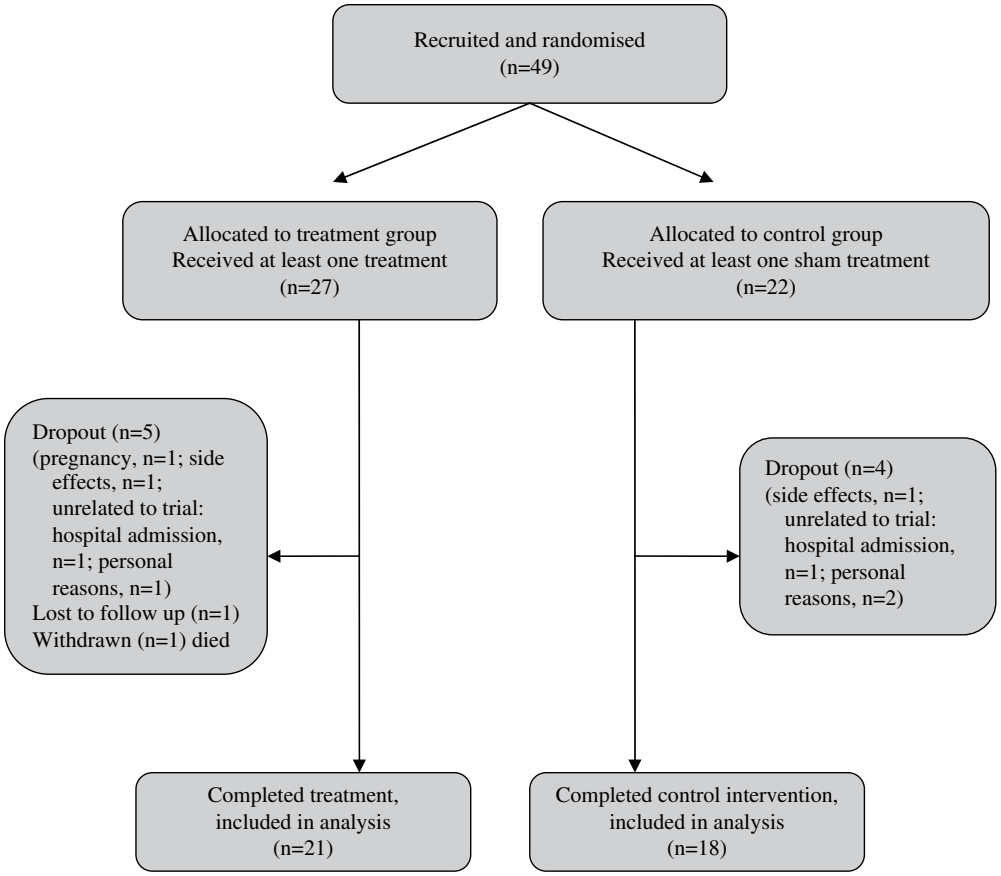
Age, body mass index, and perception of periods at baseline.

Variable	True reflexology (N = 26)	Sham reflexology (N = 22)
Mean age (±SD), years	29.2 (4.5)	28.1 (4.1)
Mean body mass index (±SD)	29.2 (5.5)	29.1 (7.1)
Period flow	(n = 25)	(n = 21)
Heavy (%)	5 (20)	8 (38)
Normal (%)	14 (56)	11 (52)
Absent (%)	6 (24)	2 (10)
Period regularity	(n = 26)	(n = 22)
Regular (%)	6 (23)	5 (23)
Fairly regular (%)	6 (23)	6 (27)
Irregular or absent (%)	14 (54)	11 (50)

Holt. Foot reflexology for ovulation induction. Fertil Steril 2009.

FIGURE 2

Flowchart of participants through the study.



Holt. Foot reflexology for ovulation induction. Fertil Steril 2009.

remained a problem; with no patients recruited after December 2003, the study team reluctantly agreed to abort the trial in the spring of 2004 when 49 patients had been randomized: 27 in the true group, 22 in the sham. With an early termination of recruitment, we acknowledge that our final sample size is unlikely to be large enough to detect anything but a very large difference between the sham and genuine reflexology groups.

Baseline Characteristics

Table 1 demonstrates that baseline characteristics were similar between the two groups (data were omitted from the patient who was withdrawn; Fig. 2). Women in the two groups had similar age and BMI. Their perception of regularity and heaviness of menstruation was also similar.

Generally, the groups are well matched on important prognostic variables. This is true also for initial reflexology scores apart from that for "pituitary right foot": the true group had a mean of 9.69 and the sham of 8.05 (Mann-Whitney, P=.49). All other reflexology scores and the total combined scores are the same for both groups.

There was no evidence of any differences between the groups for either weight (Mann-Whitney, P=.91), BMI (Mann-Whitney, P=.45), or change in BMI (Mann-Whitney, P=.57). Although there was an overall decrease in BMI for both groups, this change was not statistically significant (Wilcoxon test, P=.20).

Retention and Blinding

The flow of patients through the study is shown in Figure 2. Most patients completed the trial: 21 (80.8%) in the true group and 18 (81.8%) in the sham group received eight treatments. There was no statistically significant difference between the dropout rates in the two groups (Fisher's exact test, P=.61). There were no reports of major adverse events, though several patients in the true reflexology group reported discomfort during the treatment.

In the true reflexology group, 11 out of 22 (50%) reported that they had received the true treatment, compared with 4 out of 17 (23.5%) in the sham group. All the others in both groups were unsure which group they were in. Nobody thought they were in the sham group.

Evidence of Ovulation

This was our principal outcome, indicated by a serum progesterone level of over 30 nmol/L. There was no evidence of

TABLE 2

Numbers of patients in each group with evidence of ovulation (serum progesterone level of >30 nmol/L) or pregnancy during the study period.

	True reflexology	Sham reflexology
Ovulation: all patients	11/26 (42%)	10/22 (46%)
Ovulation: excluding participants with body mass index >35	10/22 (55%)	9/17 (53%)
Pregnancy	4/26 (15%)	2/22 (9%)

Holt. Foot reflexology for ovulation induction. Fertil Steril 2009.

a difference between groups: in the true reflexology group, 11 of 26 (42.3%; 95% CI, 23.4–63.1%) showed evidence of ovulation, compared with 10 of 22 (45.5%; 95% CI, 24.4–67.8%) in the sham group (P=.83). The result is very similar if women with a BMI of over 35 at entry are excluded (Table 2).

There were four pregnancies (15.4%; 95% CI, 4.4–34.9%) in the true group and two in the sham group (9.1%; 95% CI, 1.1–29.2%). Again this difference was not statistically significant (Fisher's exact test, P=.68).

Periods

There was no evidence of any difference in either heaviness or regularity of periods between the groups (data not shown; Mann-Whitney, P=.33 for "regular" and P=.27 for "heavy").

Skin Quality

Although there is no difference between the groups at exit with regard to reported greasiness of the skin, there was a statistically significant difference in the assessment of improvement, favoring the sham group. Nine of 17 (52.9%) in the sham felt their skin had improved over the course of treatment and 2 of 21 (9.5%) of the true group (chi-square, P=.01).

HAD Scores

Baseline scores were toward the upper limit of normal (values over 8 indicate mild anxiety or depression). There was no difference between the groups for the preselected measure of anxiety (Table 3). However, an unexpected finding was a statistically significant improvement in depression scores in favor of the true group, though mean baseline scores were within the normal range.

DISCUSSION

Recruitment and Other Practical Considerations

Complementary therapies are notoriously difficult to trial rigorously. They tend to be difficult to blind, and the placebo effect is often strong. In our case, we could not blind the therapist, but the patients were successfully kept blind until the end of the trial. Our recruitment rate was far lower than anticipated before the trial, and, in spite of a great deal of media attention, many of the women who wanted to join the trial had one or more of the exclusion criteria. In particular, it transpired that many were ovulating. However, once randomized, participants in both arms stayed in the trial to the same extent, enjoyed the treatment, and would recommend it to a friend. Our attempt to boost recruitment by approaching other centers proved unduly optimistic: although willing in principle to recruit, fertility experts at other trusts obviously felt less inclined to approach potential participants.

Recruitment was further hampered by a tightening of the BMI entry criterion during the course of the trial. A large proportion of randomized women were very obese at recruitment, which is often associated with polycystic ovary syndrome. The reflexologist found accessing the points successfully to be difficult, and we felt it was fairer to the therapist to exclude these women and believed it also was a fairer test of the therapy. It is also worth noting that

TABLE 3

Anxiety and depression subscale scores Hospital Anxiety and Depression scores (± standard deviation).

	True reflexology		Sham reflexology		
	N	Median (range)	N	Median (range)	*P* value[a]
Anxiety before	25	6 (1–18)	21	7 (1–15)	.18
Anxiety after	22	5 (1–17)	18	6 (0–16)	.23
Anxiety change	21	−1 (−8–+5)	17	0 (−4–+4)	.45
Depression before	23	5 (3–15)	21	4 (3–11)	.15
Depression after	22	4 (2–12)	17	4 (3–12)	.35
Depression change	19	−1 (−3–0)	16	+0.5 (−7–+6)	.009

[a] Mann-Whitney.

Holt. Foot reflexology for ovulation induction. Fertil Steril 2009.

reflexologists, like fertility experts, would recommend diet and lifestyle changes to very obese clients to enhance their chances of becoming pregnant.

Results

Given our relatively small final sample size, it is unsurprising that there is no difference between the groups on any of the main outcome measures. A post hoc power calculation estimates that our sample size would have 80% power to detect an increase in ovulation rate from 20% to 59% (instead of 50% as anticipated). There was no suggestion of any clinically relevant effect; in fact, there was no suggestion of even a small difference between the groups in ovulation induction. The pregnancy rate was so small in both groups that a study would need to be vast (total sample size in excess of 600) to detect any effect of reflexology. On the basis of these data, further research into the specific effects of reflexology (i.e., the effect compared with sham reflexology) for ovulation induction is not justified.

We expected ovulation rates of about 20% in this population (11), and we observed rates of about double that. It remains a possibility that foot massage itself might have a general effect from which both groups benefited.

Blinding is a problem with many therapeutic trials of physical interventions, but complementary therapies such as acupuncture, Reiki, and reflexology present their own problems. We are reasonably confident that participants in this trial were blinded to allocation—none felt that they were in the sham group, for example. However, the blinding of the therapist was not possible.

The assessment of depression scores after treatment was limited because the measures were administered by the reflexologist, who was not blinded. The statistically significant difference between the groups, which was not based on an a priori hypothesis, should be regarded as hypothesis generating.

CONCLUSION

We have conducted a randomized trial of reflexology for induction of ovulation. We have demonstrated that blinding of participants is feasible, but we also encountered a problem recruiting patients from health service infertility clinics. Although the small sample size prevents any conclusive statement about whether reflexology has any specific effects on induction of ovulation, the lack of effect seen here suggests that any specific effects that may exist are unlikely to be clinically useful. However, as the rate of ovulation was about 40% in both groups compared with an expected 20%, it remains possible that sham reflexology in general has an overall positive effect on anovulation, which this trial was not designed to detect. Foot massage may not be an appropriate inactive control procedure for reflexology trials. Future studies should consider exploring the relationship between reflexology and depression.

REFERENCES

1. Templeton A. Infertility—epidemiology, aetiology and effective management. Health Bull 1995;53:294–8.
2. Lord J, Acharya U. The treatment of ovulatory disorders. Obstet Gynaecol 1999;1:33–5.
3. ESHRE Capri Workshop. Female infertility: treatment options for complicated cases. Hum Reprod 1997;12:1191–6.
4. Levene MI, Wild J, Steer P. Higher multiple births and the modern management of infertility in Britain. Br J Obstet Gynaecol 1992;35:607–13.
5. Mackareth P, Tivan D. Clinical reflexology—a guide for professionals. Edinburgh: Churchill Livingstone, 2002.
6. Ernst E, White AR. The BBC survey of complementary use in the United Kingdom. Complement Ther Med 2000;8:32–6.
7. Thomas KJ, Nicholl JP, Fall M. Access to complementary medicine via general practice. Br J Gen Pract 2001;51:25–30.
8. Hardy R. "I only wish I'd tried this sooner." Women with fertility problems are turning to reflexology—and the results are so impressive, now a hospital trial has begun, says Rebecca Hardy. Independent Review. March, 15 2004. Available at: http://www.independent.co.uk/life-style/health-and-wellbeing/health-news/i-only-wish-id-tried-this-sooner-566404.html.
9. Perry A. Is reflexology the new cure for infertility? Daily Mail. April 12, 2001. Available at: http://www.dailymail.co.uk/pages/live/articles/health/womenfamily.html?in_article_id=20980&in_page_id=1799.
10. Pasquali R, Pelusi C, Genghini S, Cacciari M, Gambineri A. Obesity and reproductive disorders in women. Hum Reprod Update 2003;9:359–72.
11. Lord JM, Flight IH, Norman RJ. Insulin-sensitising drugs (metformin, troglitazone, rosiglitazone, pioglitazone, D-chiro-inositol) for polycystic ovary syndrome. Cochrane Database Syst Rev 2003;3:CD003053.