European Journal of Obstetrics & Gynecology and Reproductive Biology 249 (2020) 72–83

Contents lists available at ScienceDirect

European Journal of Obstetrics & Gynecology and Reproductive Biology

journal homepage: www.elsevier.com/locate/ejogrb

Editorial

Data integrity of 35 randomised controlled trials in women' health

ARTICLE INFO

Article history:
Received 16 March 2020

Keywords:
Research integrity
Data integrity
Fabricated data
Randomization
Ovulation induction

ABSTRACT

While updating a systematic review on the topic of ovulation of induction, we observed unusual similarities in a number of randomised controlled trials (RCTs) published by two authors from the same institute in the same disease spectrum in a short period of time. We therefore undertook a focused analysis of the data integrity of all RCTs published by the two authors. We made pairwise comparisons to find identical or similar values in baseline characteristics and outcome tables between trials. We also assessed whether baseline characteristics were compatible with chance, using Monte Carlo simulations and Kolmogorov-Smirnov test.

For 35 trials published between September 2006 and January 2016, we found a large number of similarities in both the baseline characteristics and outcomes of 26. Analysis of the baseline characteristics of the trials indicated that their distribution was unlikely to be the result of proper randomisation. The procedures demonstrated in this paper may help to assess data integrity in future attempts to verify the authenticity of published RCTs.

Introduction

We recently performed an update of a systematic review for the Cochrane Database for systematic reviews on the topic of ovulation induction [1]. While performing data extraction, we noticed similarities between values in the baseline characteristics (i.e. age, Body Mass Index (BMI), parity) and outcome tables in different randomised controlled trials (RCTs) authored by Ahmed Badawy and Hatem Abu Hashim (Badawy and Abu Hashim for short in the following), from the University of Mansoura in Egypt. Following these concerns, we undertook a systematic assessment of the RCTs published by these two authors.

Materials and methods

Inclusion of RCTs

We searched the PubMed database for authors 'Badawy' or 'Abu Hashim' using the affiliation 'Mansoura' restricting to RCTs. We excluded trials not related to the University of Mansoura in Egypt and not published in full-text.

Data extraction

We extracted information regarding year of publication, journal, number of authors, study centers, baseline characteristics, number and type of participants, interventions, outcome data, conclusions, study start and end dates, and date of submission to the journal. We calculated the average number of randomised participants per month for each study by dividing the total number of randomised participants by the number of months over which recruitment was said to have occurred. We obtained the trial registration numbers by searching the World Health Organization (WHO) and International Standard randomised Controlled Trial Number (ISRCTN) registers.

Comparison of baseline characteristics and outcomes

For both Badawy and Abu Hashim, we performed pairwise comparisons with respect to the values in the baseline characteristics and outcome tables to find identical or similar values across trials. We only compared studies that reported at least five of the same parameters in either their baseline or outcome tables. We compared values of mean, standard deviation (SD), percentage, t-value, p-value, and confidence intervals (CIs) whenever available. We reported values that were identical or highly similar in the pairwise comparisons, with high similarity being defined as plus or minus one for any corresponding digit in the decimal (i.e. base 10) display of a number.

Last digits analyses

We analysed the frequency of last digits for the mean and standard deviations of some baseline characteristics that are reported in all or most randomised trials of the two authors. We plotted the observed frequencies with their expected frequencies determined by the number of numerals to the right of the last zero and Newcomb-Benford's [2].

http://dx.doi.org/10.1016/j.ejogrb.2020.04.016
0301-2115/

Probability of random sampling for baseline characteristics

Baseline characteristics between groups should, in general, be balanced because the allocation of participants to these groups is random. For any given baseline characteristic, a statistical comparison of the group means or proportions generates a p-value addressing the null hypothesis that the group means or proportions are the same. If randomisation has been conducted properly, the collection of p-values across a series of independent baseline characteristics will have a uniform distribution between 0 and 1, that is, they should look like a series of randomly drawn values between 0 and 1 [3]. This statistical property originates from the nature of randomisation and is unrelated to the population distribution of any of the baseline characteristics. We used Monte Carlo simulations (computational algorithms that use random sampling to generate numerical results) to generate a p-value for differences between means (proportions) for each baseline continuous (categorical) variable [4]. If randomisation and data recording are done correctly in all or at least most trials, the set of simulation-generated p-values from baseline variables (especially continuous ones) should not deviate much from a uniform [0,1] distribution. A preponderance of simulation-generated p-values close to either 0 or 1 is evidence of systematic baseline imbalance or extreme similarity, which violates the assumption that group allocation was properly randomised [5]. In addition to building cumulative distribution plots for the simulation-generated p-values to visualize their deviations from the null hypothesis (uniform distribution), we also used the Kolmogorov–Smirnov test, against a uniform distribution [0,1], to assess quantitatively the evidence against the null hypothesis that the simulation-generated p-values follow a uniform distribution. We planned to perform the analyses for different groups: all trials of Badawy, the trials of which Badawy was first author, the trials of which Badawy was co-author, all the trials of Abu Hashim, the trials of which Abu Hashim was first author, and the trials of which Abu Hashim was co-author.

The statistical analyses were performed using Stata (v16.0) and the R statistical software (v3.5.1).

Results

Inclusion of RCTs

We identified 35 (24 by Badawy, 11 by Abu Hashim of which 2 co-authored by Badawy) RCTs that included a total of 8770 participants, published between September 2006 and January 2016 (Table A1 in Appendix A). We refrain from citing these papers in view of the integrity concerns we are examining. Badawy was first author in 19 trials and co-authored seven. Abu Hashim was first author in 10 trials and co-authored one. In two trials where Abu Hashim was first author, Badawy co-authored.

Characteristics of trials

All 35 RCTs state they were performed in Mansoura University Hospitals, Mansoura University, Egypt, and associated private practices, either in the fertility outpatient clinic, gynaecology outpatient clinic or labour ward.

Ten Badawy trials studied women with polycystic ovary syndrome (PCOS), and seven couples with unexplained infertility. Interventions under study were ovulation induction, laparoscopy and laparoscopic ovarian drilling, and endometrial scratching. In the other seven Badawy trials different types of participants undergoing different interventions were included (e.g. healthy primiparous women or women with pre-eclampsia or eclampsia). The median number of trial participants in the 24 Badawy trials was 235 (range 32–804 per study), with a median number of participants per month of 8 (range 1–42). Of the 24 Badawy trials, five were retrospectively registered and 19 not registered. None were prospectively registered (Table 1, Supp. Table 1).

Eight Abu Hashim trials included women with PCOS who underwent interventions including ovulation induction, intrauterine insemination, or laparoscopic ovarian drilling. In three other Abu Hashim trials, a variety of interventions in different types of participants (e.g. women with heavy menstrual bleeding) were evaluated. The median number of participants in the 11 Abu Hashim trials was 176 (range 48–282), with a median number of randomised participants per month of 6 (range 3–8). Seven of the 11 Abu Hashim trials were retrospectively registered, four not registered and none prospectively registered (Table 2, Supp Table 2).

Comparison of baseline characteristics and outcomes

RCTs published by Badawy

Out of the 24 published trials, we did not include seven in our pairwise comparisons, since they presented variables in their baseline and outcome tables that were not reported in the other trials. The 17 included trials presented at least five (range 5–11) of the same baseline characteristics in the tables, including age, duration of infertility, parity, oligo/anovulation, hyperandrogenism, polycystic ovaries, height, weight, BMI, serum follicle stimulating hormone (FSH), serum luteinizing hormone (LH), serum estradiol (E_2), serum progesterone (P), FSH/LH ratio, and insulin resistance. The median number of identical or highly similar values of baseline characteristics between these 17 trials was 5 (range 0–31) (Fig. 1). Badawy 2008c and Badawy 2009f had the highest number (31) of identical or similar baseline values (Fig. 2).

For outcomes, the 17 trials presented at least 2 (range 2–10) of the same outcomes in their results tables, including number of follicles, number of follicles $\geq$ 14 mm, number of follicles $\geq$18 mm, number of days of stimulation, serum FSH, serum LH, serum E_2, serum P, pre-treatment endometrial thickness (EMT), EMT at human chorionic gonadotropin (hCG), ovulation, pregnancy, and miscarriage. The median number of identical or highly similar values for outcomes between these 17 trials was 0 (range 0–29) (Fig. 1). Badawy 2008b and Badawy 2009e had the highest number (29) of identical or highly similar outcome values (Fig. 3). Apart from studies of interventions for PCOS showing highly similar values, highly similar values were also found in the comparisons between studies reporting on women with PCOS and unexplained infertility. Each of these 17 studies had at least five highly similar values in the baseline characteristics or outcome variables with one other study.

RCTs published by Abu Hashim

Out of 11 trials published by Abu Hashim, we were not able to compare two, since they presented variables in their baseline and outcome tables that were not reported in the other trials. Nine trials presented at least three (range 3–18) of the same baseline characteristics in the tables, including age, duration of infertility, BMI, nulliparous, multiparous, waist-to-hip ratio, oligomenorrhoea, amenorrhoea, hyperandrogenism, serum FSH, serum LH, LH/FSH ratio, serum sex hormone binding globulin (SHBG), serum testosterone, fasting glucose, fasting insulin, fasting glucose/insulin ratio, and ovarian volume. The median number of identical or highly similar values of baseline characteristics between these nine trials was 21 (range 0–49) (Fig. 4). Abu Hashim 2010a and Abu Hashim 2011c had the highest number (41) of identical or highly similar baseline values (Fig. 5).

Table 1
Characteristics of RCTs authored by Badawy.

No.	Study	No. of authors	Journal	No. of patients	Interventions	Population	Calculated RR and 95 % CI for PR
1	Badawy 2006	3	Fertil Steril	804	CC + N-acetyl. CC	Unexplained infertility	1.21 (0.95–1.55)
2	Badawy 2007a	3	Acta Obstet Gynecol Scand	573	CC + N-acetyl. CC	PCOS	NA
3	Badawy 2007b	3	Reprod Biomed Online	179	2.5 mg Letrozole. 5 mg Letrozole. 7.5 mg Letrozole	Unexplained infertility	1.21 (0.51–1.53) (2.5 mg vs 7.5 mg)
4	Badawy 2007c	3	Reprod Biomed Online	58	GnRH agonist + FSH + corticosteroids. GnRH agonist + FSH	Idiopathic karyotypically normal premature ovarian failure	5.00 (0.25–99.82)
5	Badawy 2008a	6	J Obstet Gynaecol	340	LMWH + folic acid. folic acid	Unexplained spontaneous recurrent miscarriages	–
6	Badawy 2008b	3	Fertil Steril	220	Letrozole. Anastrozole	CC-resistant PCOS	1.09 (0.76–1.56)
7	Badawy 2008c	3	Reprod Biomed Online	318	CC. FSH	CC-resistant PCOS	1.76 (1.32–2.35) in favour of FSH
8	Badawy 2009a	4	Fertil Steril	80	GnRHa + chemotherapy. chemotherapy	Unilateral adenocarcinoma of the breast and with no metastasis who had undergone modified radical mastectomy or breast-conserving surgery plus full axillary lymph node dissection	–
9	Badawy 2009b	4	Fertil Steril	212	Early CC, late CC	PCOS	0.75 (0.42–1.34)
10	Badawy 2009c	5	Fertil Steril	163	UTND. LEOD	CC-resistant PCOS	1.12 (0.64–1.97)
11	Badawy 2009d	4	Fertil Steril	218	Long letrozole, short letrozole	CC-resistant PCOS	1.39 (0.89–2.19)
12	Badawy 2009e	3	Fertil Steril	438	Letrozole. CC	PCOS	1.14 (0.90–1.43)
13	Badawy 2009f	3	Fertil Steril	216	Anastrozole. CC	PCOS	1.22 (0.71–2.10)
14	Badawy 2009g	3	Fertil Steril	412	Letrozole. CC	Unexplained infertility	1.04 (0.80–1.34)
15	Badawy 2009h	4	Acta Obstet Gynecol Scand	769	Letrozole. CC. Anastrozole	Unexplained infertility	1.05 (0.60–1.83) (L vs A) 1.37 (0.95, 1.97) (CC vs L) 1.31 (0.78–2.18) (CC vs A)
16	Badawy 2010a	3	J Obstet Gynaecol	280	FSH + CC. FSH + letrozole	Unexplained infertility	1.10 (0.72–1.68)
17	Badawy 2010b	5	J Obstet Gynaecol	312	Laparoscopy + OI. LOD + OI	Unexplained infertility	1.07 (0.88–1.31)
18	Badawy 2011	2	Eur J Obstet Gynecol Reprod Biol	317	CC. Tamoxifen	PCOS	1.72 (1.03–2.87) in favour of CC
19	Badawy 2012	3	Acta Obstet Gynecol Scand	32	Letrozole. GnRH agonist	Premenopausal women with uterine adenomyosis	–
20	Gibreel 2013	4	J Obstet Gynaecol Res	105	Endometrial scratching. placebo	Unexplained infertility	2.64 (1.03–6.82) in favour of endometrial scratching
21	Ragab 2013	6	Arch Gynecol Obstet	420	Immediate postpartum curettage. None	Pre-eclampsia or eclampsia	–
22	El Refaeey 2014	3	Reprod Biomed Online	101	CoQ10 + CC. CC	CC-resistant PCOS	6.21 (1.96–19.68) in favour of CoQ10 + CC
23	Marzouk 2015	5	J Obstet Gynaecol	60	Lavender-thymol. Placebo	Primiparous women	–
24	El-Refaie 2016	3	Arch Gynecol Obstet	250	Vaginal progesterone suppositories. none	Dichorionic twin pregnancy, asymptomatic with cervical length of 20–25 mm at 20–24 weeks of gestation	–

No. = number of. CC = clomiphene citrate. FSH = follicle stimulating hormone. LMWH = Low-molecular weight heparin. GnRH = gonadotropin-releasing hormone agonists. UTND = ultrasound-guided transvaginal ovarian needle drilling LEOD = laparoscopic electrosurgery ovarian drilling LOD = laparoscopic ovarian drilling CoQ10 = vombined coenzyme Q10. PCOS = polycystic ovary syndrome. RR = risk ratio. CI = confidence interval. PR = pregnancy rate. NA = not applicable.

Regarding outcomes, the nine trials presented at least 5 (range 5–16) of the same outcomes in the table, including number of follicles, number of follicles > 14 mm, number of follicles ≥ 18 mm, number of days of stimulation, serum E_2, serum P, pre-treatment EMT, EMT at hCG, ovulation, (clinical) pregnancy, live birth, twin/multiple pregnancies and miscarriage. The median number of identical or highly similar values of outcomes between these nine trials was 6 (range 0–35) (Fig. 4). Abu Hashim 2011b and Abu Hashim 2012a had the highest number (33) of identical or similar outcome values (Fig. 6). Highly similar values were not confined to studies including women with PCOS. The comparisons between studies including women who underwent surgery for endometriosis and those including women with PCOS also showed highly similar values. Each of the nine studies had at least five highly similar values in the baseline characteristics or outcome variables with one other study.

Table 2
Characteristics of RCTs authored by Abu Hashim.

No.	Study	No. of authors	Journal	No. of patients	Interventions	Population	Calculated RR and 95 % CI for PR
1	Abu Hashim 2010a	3	Arch Gynecol Obstet	260	Letrozole. LOD	CC-resistant PCOS	1.00 (0.68–1.48)
2	Abu Hashim 2010b	3	Fertil Steril	250	Letrozole. CC + metformin	CC-resistant PCOS	1.01 (0.71–1.42)
4	Abu Hashim 2010c	3	J. Women's Health	192	CC + NAC. CC + metformin	CC-resistant PCOS	4.31 (1.70–10.91) in favour of CC + metformin
3	Abu Hashim 2011a	3	Gynecol Endocrinol	153	CC + metfomin. FSH	CC-resistant PCOS	1.71 (1.05–2.77) in favour of FSH
5	Abu Hashim 2011b	3	Acta Obstet Gynecol Scand	188	CC + IUI. CC timed intercourse	PCOS	1.07 (0.63–1.81)
6	Abu Hashim 2011c	3	J. Obstet. Gynaecol. Res.	282	CC + metformin. LOD	CC-resistant PCOS	1.02 (0.86–1.21)
7	Shokeir 2011	5	J Minim Invasive Gynecol	48	Hysteroscopic Endometrial Resection with either oxytocin infusion or saline	Abnormal uterine bleeding that was unresponsive to conservative medical management	–
8	Abu Hashim 2011d	4	Arch Gynecol Obstet	176	LOD. CC	CC-failure PCOS	1.16 (0.78–1.72)
9	Abu Hashim 2012a	3	Gynecol Endocrinol	113	CC + FSH. CC	PCOS	1.11 (0.56–2.18)
10	Abu Hashim 2012b	3	Contraception	95	CVR. norethisterone	Heavy menstrual bleeding	–
11	Abu Hashim 2012c	3	Acta Obstet Gynecol Scand	136	Letrozole. CC	Recently surgically treated minimal to mild endometriosis	1.10 (0.77–1.55)

No. = number of. CC = clomiphene citrate. FSH = follicle stimulating hormone. IUI = intra-uterine insemination. LOD = laparoscopic ovarian drilling. PCOS = polycystic ovary syndrome. RR = risk ratio. CI = confidence interval. PR = pregnancy rate. NA = not applicable.

		Baseline / Study No.																No. of studies with similarities			
																		Baseline		Outcome	
		1	**2**	**3**	**4**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**≥5**	**≥10**	**≥5**	**≥10**
	1		9	2	6	6	1	3	2	1	3	7	5	7	11	3	4	1	6	1	1	1
	2	11		2	3	5	7	0	5	1	7	12	6	7	5	3	3	3	8	1	1	1
Outcome	**3**	1	2		2	5	3	1	4	2	6	4	4	4	8	8	3	1	4	0	2	5
	4	5	7	2		7	6	4	4	1	7	6	4	6	4	6	1	3	7	0	3	0
	6	0	0	11	0		12	6	13	4	10	17	17	11	11	13	3	4	5	8	2	7
	7	1	0	14	0	14		6	14	8	17	19	31	9	11	7	6	5	9	6	2	6
	8	0	0	0	1	2	1		4	0	5	5	2	4	6	4	3	2	5	0	0	0
	9	0	1	13	0	15	21	0		2	16	16	17	8	14	5	7	1	4	6	4	5
Study No.	**10**	0	0	0	0	0	0	0	1		6	8	4	2	5	6	4	3	5	0	0	0
	11	0	0	13	1	17	18	1	9	0		15	12	8	12	7	7	4	8	6	5	4
	12	0	0	10	0	29	16	1	14	0	9		15	21	16	19	5	2	5	9	1	6
	13	0	0	8	0	16	15	0	12	0	16	10		10	20	7	4	2	3	7	2	5
	14	1	1	5	0	10	9	3	8	0	9	18	7		15	19	4	2	6	5	5	3
	15	0	0	0	0	0	0	0	2	0	0	0	0	0		15	8	3	5	9	0	0
	16	0	0	4	0	9	5	2	9	0	5	10	5	21	0		3	1	7	4	5	2
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		2	5	0	0	0
	18	2	0	2	0	3	4	0	2	0	2	1	4	2	0	0	0		1	0	0	0

Fig. 1. League table of the baseline characteristics and outcomes of 17 Badawy studies.
The number of similarities (same number, or +1 or -1 for any digit) between studies mean values, standard deviations (SDs) and, if available, percentages, t-values, p-values, and confidence intervals (CI). For example, as shown in Fig. 1 the baseline characteristics of study number 7 (Badawy 2008c) and 13 (Badawy 2009f) are compared and we found 31 similarities. Between these two studies we found 15 similarities in the outcome table. Upper triangular block = amount of similarities in baseline characteristics. Lower triangular block = amount of similarities in outcomes. Green = amount of similarities <5. Orange = amount of similarities ≥ 5 but < 10. Red = amount of similarities ≥ 10.

Table 1. Characteristics of polycystic ovary syndrome patients in the extended clomiphene citrate and gonadotrophin treatment groups.

Parameter	*Clomiphene citrate group*	*Gonadotrophin group (n = 158)*	*95% confidence interval*
No. of cycles (mean no. per patient)	405 (2.53)	397 (2.51)	–
Age (years)	24.1 ± 3.1	26.3 ± 3.0	−0.12 to 0.16
Parity	0.3 ± 0.2	0.3 ± 0.3	−0.30 to 0.05
Height (cm)	160.3 ± 6.2	158.1 ± 5.8	−0.002 to 0.15
Weight (kg)	78.3 ± 6.4	81.1 ± 4.2	−0.26 to 0.45
Clinical presentation			
Oligo/anovulation (%)	136 (85.0)	140 (88.6)	0.36 to 1.3
Hyperandrogenism (%)	76 (47.5)	70 (44.3)	0.63 to 1.99
Polycystic ovaries (%)	111 (69.4)	103 (65.2)	1.05 to 1.44
BMI (kg/m²)	30.5 ± 3.1	32.5 ± 2.9	−0.02 to 5.4
FSH (IU/ml)	4.1 ± 2.7	5.1 ± 2.1	−0.07 to 2.3
LH (IU/ml)	10.9 ± 1.8[a]	13.1 ± 2.2[a]	−0.07 to 2.5

Values are mean ± SEM unless otherwise stated; BMI = body mass index.
[a] $P = 0.04$. There were no other statistically significant differences.

TABLE 1
Patients' characteristics.

Parameter	Anastrozole group (n = 115)	CC group (n = 101)	Values of χ^2 or t[a]	*P* value	CI
No. of cycles	243	226			
Age (y)	23.8 ± 3.1	25.3 ± 2.9	1.04	.67	−0.12–0.15
Parity	0.3 ± 0.12	0.3 ± 0.16	0.98	.71	−0.30–0.06
Height (cm)	158.3 ± 5.12	155.1 ± 4.20	1.65	.08	−0.002–0.15
Weight (kg)	80.3 ± 5.42	79.1 ± 4.22	0.22	.95	−0.26–0.45
Clinical presentation, n (%)					
Oligoovulation or anovulation	110 (95.6)	92 (91.0)	1.84[b]	1.75	0.36–1.31
Hyperandrogenism	51 (44.3)	42 (41.5)	0.17[b]	.68	0.63–1.99
Polycystic ovaries	98 (85.2)	71 (70.2)	7[b]	.008[c]	1.05–1.41
BMI (kg/m²)	31.1 ± 2.91	29.1 ± 3.12	1.4	.31	−0.02–5.4
FSH (IU/mL)	6.1 ± 2.92	6.3± 2.22	2.43	.06	−0.07–2.1
LH (IU/mL)	13.2 ± 1.82	12.1 ± 3.11	2.55	.052	−0.06–3.2

[a] Data are *t* values unless otherwise indicated.
[b] Data are χ^2.
[c] Statistically significant difference at $P < .01$.
Badawy. Clomiphene citrate or anastrozole for ovulation induction in women with PCOS. Fertil Steril 2009.

Fig. 2. Similarities in baseline characteristics between Badawy 2008c and Badawy 2009f (number 7 and 13 in Table 1).
Badawy 2008c (7) was published in *Reproductive BioMedicine Online* in January 2008 and includes 318 women with clomiphene citrate-resistant polycystic ovarian syndrome comparing ovulation induction with clomiphene citrate versus follicle-stimulating hormone.
Badawy 2009f (13) was published in *Fertility and Sterility* in September 2009 and includes 216 women with polycystic ovarian syndrome comparing ovulation induction with anastrozole versus clomiphene citrate.
The green dots represent the exact same number and the red dots represents -1 or +1 for any digit. Two means, five SDs, and fourteen values in 95 % confidence interval are identical and two means, five SDs, and three values in the 95 % confidence interval are highly similar. A total of 31 similarities.

TABLE 2
Outcome in letrozole and anastrozole groups.

	Letrozole group (n = 111)	Anastrozole Group (n = 119)	*t*-test	*P* value
Total number of follicles	5.4 ± 0.4	5.8 ± 0.4	5.21	.01[a]
Number of follicles >14 mm	3.1± 0.3	2.7 ± 0.2	5.33	.004[a]
Number of follicles >18 mm	2.3 ± 0.1	3.1 ± 0.2	8.62	.001[a]
Pretreatment endometrial thickness (mm)	5.5 ± 0.5	5.3 ± 0.6	1.31	.22
Endometrial thickness at hCG (mm)	9.1 ± 0.2	10.2 ± 0.7	4.45	.04[a]
Serum E_2 (pg/mL)	455.1 ± 64.2	484 ± 91.3	2.39	.08
Serum P (ng/mL)	9.2 ± 0.9	10.1 ± 1.2	2.81	.06
Duration of stimulation (days)	11.9 ± 1.3	10.8 ± 2.2	2.30	.21
Pregnancy/cycle	36/295 (12.2%)	42/279 (15.1%)	0.99	.31
Miscarriage/patient	4 (11.1%)	4(9.5%)	0.01	.92

[a] Statistically significant differences as $P < .05$.
Badawy. Letrozole versus anastrozole. Fertil Steril 2008.

TABLE 2
Outcome in letrozole and clomiphene citrate (CC) groups.

	Letrozole group (n = 218)	CC group (n = 220)	*t*	*P* value
Total number of follicles	4.4 ± 0.4	6.8 ± 0.3	4.3	.042[a]
Number of follicles >14 mm	2.1 ± 0.3	3.7 ± 0.5	6.13	.008[a]
Number of follicles >18 mm	2.3 ± 0.1	3.1 ± 0.8	5.03	.03[a]
Pretreatment endometrial thickness (mm)	4.5 ± 0.4	4.3 ± 0.5	1.41	.52
Endometrial thickness at hCG (mm)	8.1 ± 0.2	9.2 ± 0.7	5.44	.021[a]
Serum E_2 (pg/mL)	255.1 ± 64.2	384 ± 91.3	4.12	.022[a]
Serum progesterone (ng/mL)	7.1 ± 0.9	11.1 ± 1.2	6.33	.024[a]
Duration of stimulation (days)	12.1 ± 1.38	8 ± 2.9	4.91	.036[a]
Pregnancy/cycle	82/540 (15.1%)	94/523 (17.9%)	1.33	.72
Miscarriage/patient	4 (12.1%)	4 (9.7%)	1.73	.43

[a] Statistically significant difference: $P < .05$.
Badawy. Clomiphene citrate or letrozole. Fertil Steril 2009.

Fig. 3. Similarities in outcomes between Badawy 2008b and Badawy 2009e (number 6 and 12 in Table 1).
Badawy 2008b (6) was published in *Fertility and Sterility* in May 2008 and includes 220 women with clomiphene citrate-resistant polycystic ovarian syndrome comparing ovulation induction with letrozole versus anastrozole.
Badawy 2009e (12) was published in *Fertility and Sterility* in September 2009 and includes 438 women with polycystic ovarian syndrome comparing ovulation induction with letrozole versus clomiphene citrate.
The green dots represent the exact same number and the red dots represents -1 or +1 for any digit. Two means, two numbers, and ten SDs are identical and ten means, three SDs, one percentage, and one *t*-test value are highly similar. A total of 29 similarities.

Study No. (Outcome)	Study No. (Baseline) 1	2	3	4	5	6	8	9	11	No. of studies with similarities: Baseline ≥5	Baseline ≥10	Outcome ≥5	Outcome ≥10
1		5	25	33	20	41	27	21	3	1	6	0	0
2	5		12	2	1	3	4	4	3	1	1	1	3
3	3	19		29	25	24	19	31	4	0	7	3	3
4	3	7	10		26	36	26	28	3	0	6	4	1
5	4	11	7	5		21	25	49	3	0	6	3	2
6	7	5	3	2	2		23	19	0	0	6	4	0
8	7	5	5	3	2	7		29	6	1	6	5	0
9	3	8	6	7	33	5	4		5	1	6	4	1
11	7	15	10	8	9	4	8	0		2	0	4	2

Fig. 4. League table of the baseline characteristics and outcomes of nine Abu Hashim studies.
The number of similarities (same number, or +1 or -1 for any digit) between studies mean values, standard deviations (SDs), and if available, percentages, t-values, p-values, and confidence intervals (CI). For example, as shown in Fig. 3, the baseline characteristics of study number 1 (Abu Hashim 2010a) and 6 (Abu Hashim 2011c) were compared and we found 41 similarities. Between these two studies we found seven similarities in the outcome table. Upper triangular block = amount of similarities in baseline characteristics. Lower triangular block = amount of similarities in outcomes. Green = amount of similarities <5. Orange = amount of similarities ≥5 but <10. Red = amount of similarities ≥ 10.

Table 1 Patients' characteristics

Variable	Group A (letrozole) (n = 128)	Group B (LOD) (n = 132)
Age (years)	27.3 ± 2.6	26.4 ± 2.4
Parity		
Nulliparous	114 (89%)	116 (87.8%)
Multiparous	14 (11%)	16 (12.2%)
Duration of infertility (years)	4.3 ± 1.11	4.5 ± 1.24
BMI (kg/m^2)	26.4 ± 3.3	26.6 ± 3.6
Waist-to-hip ratio	0.8 ± 0.03	0.8 ± 0.1
Menstrual cycle		
Oligomenorrhoea	118 (92.2%)	120 (91%)
Amenorrhoea	10 (7.8%)	12 (9%)
Hyperandrogenism	59 (46%)	56 (42.4%)
LH (mIU/mL)	12.4 ± 2.1	13.2 ± 2.4
FSH (mIU/mL)	5.2 ± 1.3	5.4 ± 1.2
LH/FSH ratio	2.4 ± 1.3	2.4 ± 1.2
Fasting glucose (mg/dL)	89.5 ± 1.8	91.5 ± 1.7
Fasting insulin (μU/mL)	11.3 ± 3.1	10.8 ± 2.8
Fasting glucose/insulin ratio	7.2 ± 3.3	7.8 ± 4.3
Ovarian volume (mL)	11.3 ± 2.4	11.4 ± 2.5

Table 1 Patients' characteristics

	Group A (combined metformin–CC) (n = 138)	Group B (LOD) (n = 144)
Age (years)	27.2 ± 2.5	26.5 ± 2.3
Parity		
Nulliparous	123 (89%)	126 (87.5%)
Multiparous	15 (11%)	18 (12.5%)
Duration of infertility (years)	4.4 ± 1.2	4.6 ± 1.3
BMI (kg/m^2)	26.2 ± 3.4	26.1 ± 3.5
Waist-to-hip ratio	0.8 ± 0.04	0.8 ± 0.1
Menstrual cycle		
Oligomenorrhea	127 (92%)	131 (91%)
Amenorrhea	11 (8%)	13 (9%)
Hyperandrogenism	64 (46.3%)	61 (42.3%)
LH (mIU/mL)	12.5 ± 2.2	13.4 ± 2.3
FSH (mIU/mL)	5.3 ± 1.2	5.4 ± 1.1
LH/FSH ratio	2.4 ± 1.1	2.4 ± 1.2
Fasting glucose (mg/dL)	90.5 ± 1.7	88.5 ± 1.8
Fasting insulin (μU/mL)	10.9 ± 2.8	11.4 ± 3.1
Fasting glucose/insulin ratio	7.7 ± 4.3	7.1 ± 3.3
Ovarian volume (mL)	11.4 ± 2.6	11.3 ± 2.5

Values are mean ± SD or numbers (percentages) of women. None of the differences were statistically significant (P > 0.05). BMI, body mass index; CC, clomiphene citrate; FSH, follicle stimulating hormone; LH, luteinizing hormone; LOD, laparoscopic ovarian diathermy.

Fig. 5. Similarities in baseline characteristics between Abu Hashim 2010a and Abu Hashim 2011c (number 1 and 6 in Table 2).
Abu Hashim 2010a (1) was published in *Archives of Gynecology and Obstetrics* in November 2010 and includes 260 women with clomiphene citrate-resistant polycystic ovarian syndrome comparing ovulation induction with letrozole versus laparoscopic ovarian drilling.
Abu Hashim 2011c (6) was published in *Journal of Obstetrics and Gynaecology Research* in March 2011 and includes 282 women with clomiphene citrate-resistant polycystic ovarian syndrome comparing ovulation induction with clomiphene citrate plus metformin versus laparoscopic ovarian drilling.
The green dots represent the exact same number and the red dots represents -1 or +1 for any digit. Five means, three SDs, and seven percentages are identical and nine means, two numbers, and fifteen SDs are highly similar. A total of 41 similarities.

Table 2. Outcome in CC/IUI and CC/TI groups.

	Group A (CC/IUI) (n=93)	Group B (CC/TI) (n=95)	p-value
Number of cycles	259	266	0.72
Total number of follicles	4.2 ± 0.7	3.9 ± 0.6	0.57
Number of follicles ≥14 mm	2.4 ± 0.3	2.3 ± 0.4	0.43
Number of follicles ≥18 mm	1.8 ± 0.4	1.6 ± 0.2	0.64
Pretreatment endometrial thickness (mm)	5.6 ± 0.4	5.4 ± 0.6	0.45
Endometrial thickness at hCG day	7.7 ± 0.4	7.5 ± 0.6	0.54
Serum E2 at hCG day (pg/ml)	290.4 ± 91.2	282.6 ± 86.2	0.23
Serum progesterone (ng/ml)	8.7 ± 0.5	8.8 ± 0.3	0.34
Duration of stimulation (days)	12.7 ± 2.2	12.6 ± 2.1	0.36
Ovulation/cycle	135/259 (52.1%)	137/266 (51.5%)	0.64
Ovulation/woman	72/93 (77.4%)	73/95 (76.8%)	0.57
Clinical pregnancy/cycle	22/259 (8.49%)	21/266 (7.89%)	0.26
Clinical pregnancy/ovulatory cycle	22/135 (16.3%)	21/137 (15.3%)	0.28
Clinical pregnancy/woman	22/93 (23.6%)	21/95 (22.1%)	0.33
No. of twin pregnancies (%)	2/22 (9%)	2/21 (9.5%)	0.46
Miscarriage/pregnancy	4/22 (18.1%)	4/21 (19%)	0.31
Live birth rate	18/93 (19.35%)	17/95 (17.89%)	0.33

CC, clomiphene citrate; IUI, intrauterine insemination; TI, timed vaginal intercourse.
Values are mean ±SD or numbers (percentages).
None of the differences was statistically significant (p>0.05).

Table II. Outcome in minimal stimulation and CC groups.

	Group A (minimal stimulation) (n = 58)	Group B (CC) (n = 55)	p Value
Number of cycles	159	153	0.82
Total number of follicles	4.1 ± 0.7	3.8 ± 0.6	0.56
Number of follicles ≥14 mm	2.3 ± 0.3	2.2 ± 0.4	0.41
Number of follicles ≥18 mm	1.8 ± 0.4	1.6 ± 0.2	0.62
Pretreatment endometrial thickness (mm)	5.6 ± 0.4	5.4 ± 0.6	0.41
Endometrial thickness at hCG day	7.8 ± 0.5	7.6 ± 0.6	0.52
Serum E2 at hCG day (pg/ml)	321.4 ± 92.2	308 ± 88.2	0.31
Serum P (ng/ml)	8.6 ± 0.4	8.5 ± 0.3	0.29
Duration of stimulation (days)	12.6 ± 2.3	12.5 ± 2.2	0.33
Ovulation/cycle	93/159 (58.5%)	88/153 (57.2%)	0.78
Ovulation/patient	46/58 (79.3%)	42/55 (76.3%)	0.58
Clinical pregnancy/cycle	14/159 (8.8%)	12/153 (7.8%)	0.23
Clinical pregnancy/ovulatory cycle	14/93 (15%)	12/88 (13.6%)	0.29
Clinical pregnancy/patient	14/58 (24.1%)	12/55 (21.8%)	0.36
No. of twin pregnancies (%)	1/14 (7.1%)	1/12 (8.3%)	0.21
Miscarriage/pregnancy	2/14 (14.3%)	2/12 (16.7)	0.38

CC, Clomiphene citrate. Values are mean ± SD or numbers (percentages). None of the differences were statistically significant (p>0.05).

Fig. 6. Similarities in outcomes between Abu Hashim 2011b and Abu Hashim 2012a (number 5 and 9 in Table 2).
Abu Hashim 2011b (5) in *Acta Obstetricia et Gynecologica Scandinavica* in April 2011 and includes 188 women with polycystic ovarian syndrome comparing ovulation induction with clomiphene citrate plus intrauterine insemination versus clomiphene citrate plus timed intercourse.
Abu Hashim 2012a (9) in *Gynecological Endocrinology* in February 2012 and includes 113 women with polycystic ovarian syndrome comparing ovulation induction with clomiphene citrate plus follicle-stimulating hormone versus clomiphene citrate.
The green dots represent the exact same number and the red dots represents -1 or +1 for any digit. Four means, and ten SDs are identical and nine means, one number, five SDs, and four p-values are highly similar. A total of 33 similarities.

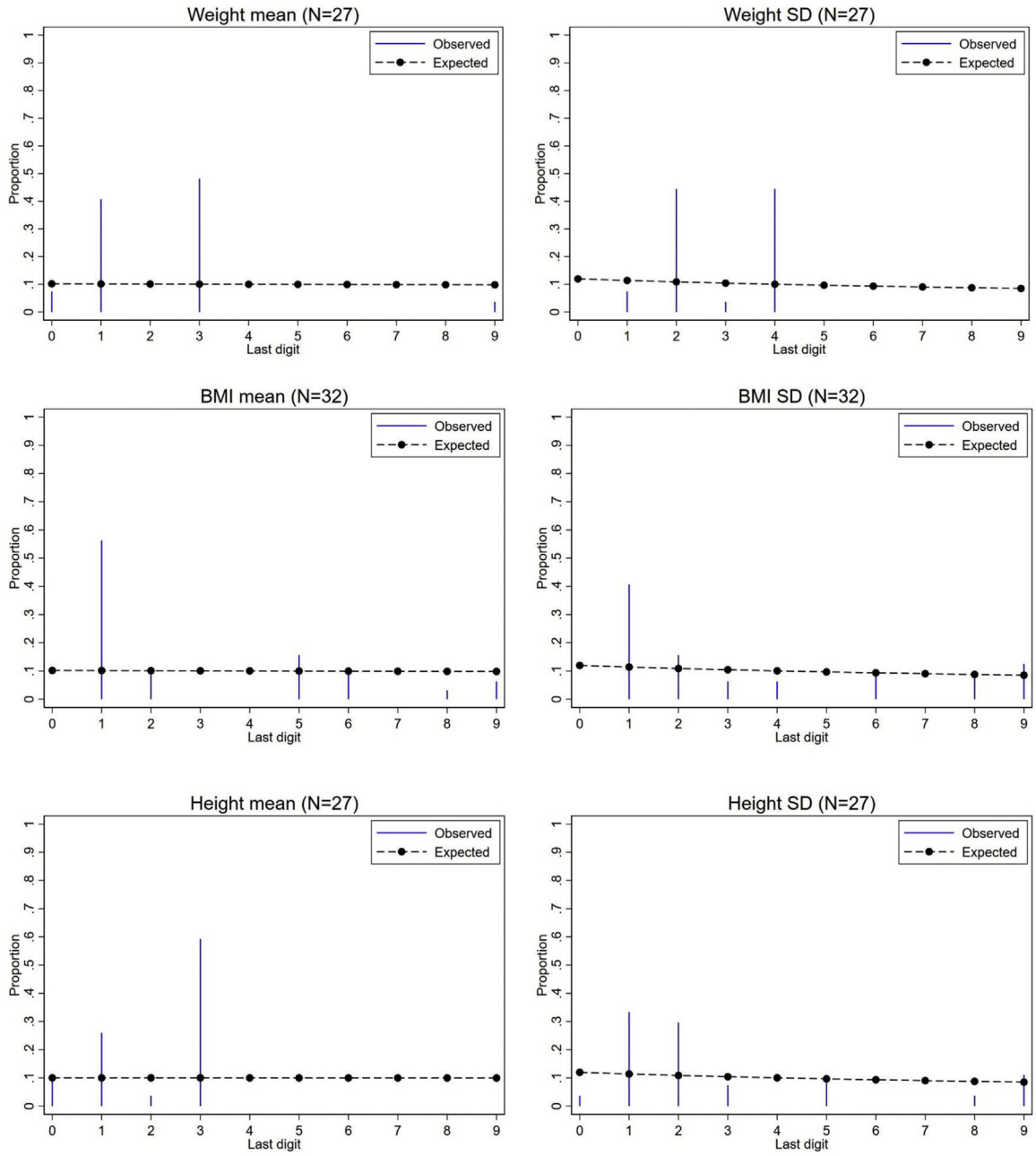

Fig. 7. Distribution of last digits of the mean and SD of height, weight, and body mass index of 17 Badawy studies.

Last digits distribution

By analyzing the data we saw a pattern in the last digits for height, weight, and BMI in the trials of Badawy. In the Abu Hashim trials, we noticed a pattern for the outcomes pre-treatment endometrial thickness and number of miscarriages.

The baseline characteristics of 17 of the 24 Badawy trials showed the same values after the decimal separator for height, weight and BMI. The distributions of the last digits deviated from their expected distributions. (Fig. 7, Supp. Fig. 1). Conspicuously, the digit seven never presented in these variables.

In 9 of 11 Abu Hashim studies and one Badawy study, the intervention arms had similar values of mean pre-treatment endometrial thickness with identical SDs. The distributions of the last digits also deviated from their expected distributions. (Fig. 8, Supp. Fig. 2). The digits zero, one and two were absent. Furthermore, in six of these studies the number of miscarriages was precisely four in either one or both intervention arms (Supp. Fig. 2).

Probability of random sampling for baseline characteristics

We used mean and SD of baseline characteristics to perform the Monte Carlo simulations. There were a number of trials in which the authors claim to have reported standard error (SE) rather than SD for baseline characteristics. After performing conversions from SE to SD by multiplying by the square root of the sample size, we are convinced that the authors reported SE as SD in these publications in view of extremely large converted SDs. An example is given in Table B1 in Appendix B. We assumed, therefore, that the reported SE was meant to have been reported as SD.

For the 24 Badawy trials, the distribution of p-values from a two-sample comparison of summary statistics of baseline variables

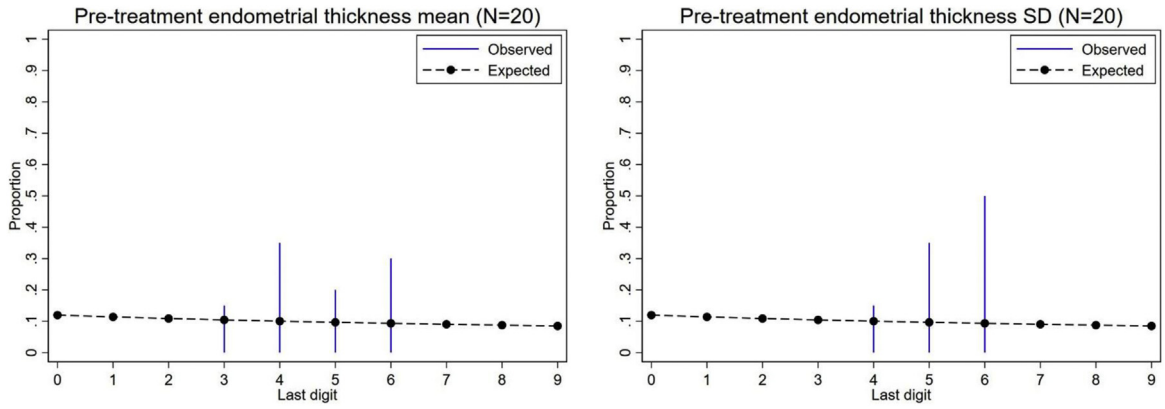

Fig. 8. Distribution of last digits of the mean and SD of pre-treatment endometrial thickness (SD of 9 Abu Hashim studies and one Badawy study.

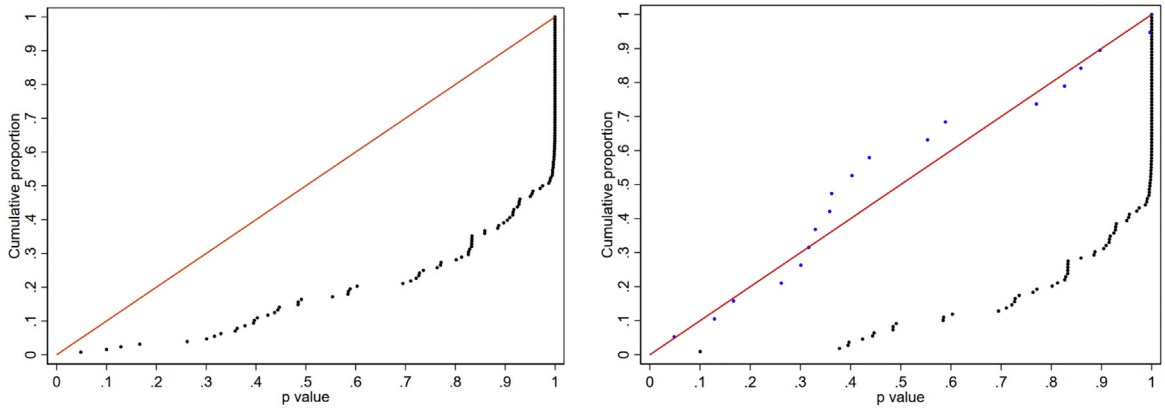

Fig. 9. Cumulative distribution of Monte Carlo simulations generated p-values for baseline characteristics from Badawy studies.
The null hypothesis is that the baseline characteristics in intervention and controls groups in these RCTs are the results of a properly conducted randomisation process. 9A All Badawy studies. The distribution was inconsistent with the null hypothesis (black dots, $p < 2.2*10^{-16}$). 9B First-author studies. The distribution was inconsistent with the null hypothesis (black dots, $p < 2.2*10^{-16}$). 9B Co-author studies. The distribution was consistent with the null hypothesis (blue dots, $p = 0.9192$).

resulted in a p-value from the Monte Carlo simulations of less than $2.2*10^{-16}$, representing very strong evidence against the null hypothesis that these summaries of baseline characteristics are the result of a properly conducted randomisation process (Fig. 9A, black dots). For 19 of 24 trials in which Badawy was first-author, the corresponding p-value for comparison with the uniform distribution was also less than $2.2*10^{-16}$ (Fig. 9B, black dots). For five of 24 trials in which Badawy was not first-author, the observed distribution of two-group comparison p-values was consistent with the expected uniform distribution, with a p-value of 0.9192 (Fig. 9B, blue dots).

For the 11 trials of Abu Hashim, comparing the observed values of the baseline p-values to the distribution of simulation generated

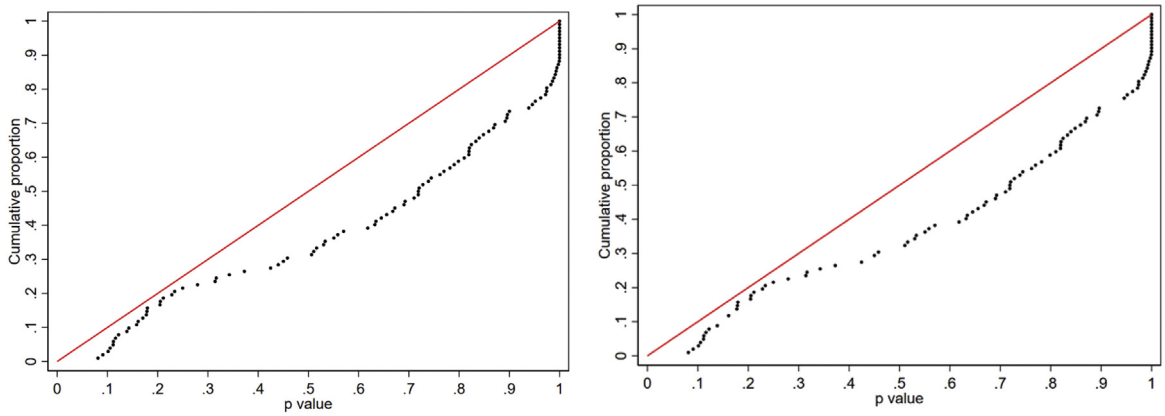

Fig. 10. Cumulative distribution of Monte Carlo simulations generated p-values for baseline characteristics from all studies of Abu Hashim.
The null hypothesis is that the baseline characteristics in intervention and controls groups in these RCTs are the results of a properly conducted randomisation process. 10A All Abu Hashim studies. The distribution was inconsistent with the null hypothesis (black dots, $p = 7.364*10^{-5}$). 10B First-author studies. The distribution was inconsistent with the null hypothesis (black dots, $p = 6.426*10^{-6}$).

p-values from the baseline variables resulted in a p-value of $7.364*10^{-5}$ (Fig. 10A, black dots). For 10 of 11 trials in which Abu Hashim was first-author, the corresponding p-value was $6.426*10^{-6}$ (Fig. 10B, black dots). There was only one trial in which Abu Hashim was not first-author, so it was not possible to perform a comparative analysis.

Discussion

In our analysis on 35 RCTs from the authors A. Badawy and H. Abu Hashim, we found among 26 individual trials (17 Badawy and 9 Abu Hashim) at least five highly similar values from the baseline characteristics or outcome tables between the studies. For 21 of the 26 evaluated trials, more than 10 similarities were observed between at least two trials. These similarities were not only found between studies including women with PCOS, but also when we compared studies which included women with PCOS with studies including women with unexplained infertility. Furthermore, our analyses suggest that the baseline characteristics of 19 trials in which Badawy was first-author and the 11 trials in which Abu Hashim was first-author are unlikely to be the result of properly conducted randomisation. Analysis of the distribution of last digits suggests that the data did not originate from the recording of measurements of empirical observations of real biological processes. These results question whether the reported results are based on RCTs that recruited real patients.

There are limitations of the method we used to assess the patterns of randomisation. The assumed null hypothesis about the distribution of simulation generated p-values under proper randomisation is not precisely correct if baseline variables are correlated. The extent of the problem caused by the violation of the null hypothesis remains the subject of current research [6,7]. The extent to which this procedure under-performs with realistic correlation structures is unknown, but it remains the best method currently available. This method has identified at least three authors from two different groups who are now known to have fabricated a large number of RCTs [4,8,9]. We have shown that the numeric results presented are very unlikely to either represent true biological values or to have been derived using the methods described by the authors. Apart from the findings presented, we also draw the reader's attention to the number of recruits, and the number of distinct trials recruiting similar or even identical patient groups in the same hospital during overlapping periods of time (Supp. Table 1, Supp. Table 2).

The findings presented here are important for several reasons. First, results of medical research in general, and from RCTs in particular, are used to inform clinical practice. The studies included in our analyses have been used, either directly or via meta-analyses, in clinical decision making. They have been cited frequently (1041 times on Scopus and 2021 on Google Scholar), including 13 citations in meta-analyses within Cochrane reviews, and seven citations in the recent international PCOS Guideline [10]. As a consequence, inappropriate practice recommendations could have been drawn, and inappropriate treatments may have been offered to patients. Second, fabricated data could misinform future research, as investments made in new large clinical research projects are usually based on meta-analyses of existing evidence.

We hope the evidence presented here leads to further investigation of all publications by Badawy and Abu Hashim, by all journals involved, including non-randomised studies, according to Committee on Publication Ethics (COPE) guidelines. We advise that the following additional information should be collected: original datasets, protocols, ethics approval and documentation of study medication. A collaborative approach by all journals involved will be needed to assess any overall patterns.

Once integrity problems in RCTs have been identified, there is often considerable delay in retractions by journals. The average retraction time for an article is two years for papers in obstetrics and gynaecology and close to three years in all disciplines [11,12]. Editors and publishers are committed to the guidelines of the Committee on Publication Ethics (https://publicationethics.org/). Thorough investigation of suspected research fraud is currently a difficult, time consuming and labour-intensive process. We urgently need simpler tools that can be used to make routine assessment straightforward, and for the availability of these tools to be promoted widely.

While the first checklists to detect research integrity have already been presented [13], we question whether a uniform "checklist" approach is the best way to address this complex problem. The type and pattern of misconduct in several RCTs identified over the last few years varies from fabricated datasets, to tables that are falsified or copied from other articles [14–16].

In some areas such as genomics, artificial intelligence techniques using pattern recognition have been developed to detect likely fraudulent data and/or publications [17]. Some have suggested that study size and/or prospective trial registration can be used as markers of trial "quality" [18], but empirical work has failed to confirm this and calls to exclude trials from meta-analyses based solely on these criteria have been rejected [19–21]. Similarly, relying on journal peer reviewers, working in isolation and under confidentiality conditions, to detect data integrity patterns and problems without access to the underlying raw individual participant data (IPD) is also not ideal. The recent ICMJE [22] calls for greater public sharing of de-identified IPD may alleviate this problem somewhat, but it will not help a lone reviewer to detect widespread data integrity issues without more open communication across journals and their editorial staff.

Conclusions

Our analyses suggest serious data integrity issues in published RCTs from these authors. The similarities between baseline and outcome tables, Monte Carlo simulations and "last digit" analyses indicate that the presence of data integrity problems is beyond that expected by chance. Based on our findings, we suggest further investigation of all publications by Badawy, and Abu Hashim, by all journals involved, including non-randomised studies. We also suggest further urgent development of tools that can be used to reliably detect, test and measure suspected data integrity issues within and across RCTs to help mitigate this increasingly prevalent problem.

Authors' roles

EMB designed the study, extracted and analysed data, and critically revised the manuscript; RW assisted in the study design, oversaw data interpretation, and critically revised the manuscript; LMA, LCG, JGT and MvW oversaw data interpretation, and critically revised the manuscript; WL designed the study, extracted and analysed data, and critically revised the manuscript; BWM designed the study, oversaw data interpretation, and critically revised the manuscript. All authors approved the version to be published.

Funding

No external funding was sought to support this work.

Declaration of Competing Interest

LMA is a co-convener of the Cochrane Prospective Meta-analysis Methods Group and manager of the Australian New Zealand Clinical Trials Registry. MvW is co-oordinating editor of the Netherlands Satellite of the Cochrane Gynaecology and Fertiliy Group. BWM reports grants from NHMRC, personal fees from ObsEva, personal fees from Merck Merck KGaA, personal fees from Guerbet, personal fees from iGenomix, outside the submitted work. All other authors have no conflicts of interest.

Appendix A

Table A1
Investigated studies.

No.	Study	Journal	Title
1	Badawy 2006	Fertil Steril [23]	Clomiphene citrate plus N-acetyl cysteine versus clomiphene citrate for augmenting ovulation in the management of unexplained infertility: a randomized double-blind controlled trial.
2	Badawy 2007a	Acta Obstet Gynecol Scand [24]	N-Acetyl cysteine and clomiphene citrate for induction of ovulation in polycystic ovary syndrome: a cross-over trial.
3	Badawy 2007b	Reprod Biomed Online [25]	Randomized controlled trial of three doses of letrozole for ovulation induction in patients with unexplained infertility.
4	Badawy 2007c	Reprod Biomed Online [26]	Induction of ovulation in idiopathic premature ovarian failure: a randomized double-blind trial.
5	Badawy 2008a	J Obstet Gynaecol [27]	Low-molecular weight heparin in patients with recurrent early miscarriages of unknown aetiology.
6	Badawy 2008b	Fertil Steril [28]	Anastrozole or letrozole for ovulation induction in clomiphene-resistant women with polycystic ovarian syndrome: a prospective randomized trial.
7	Badawy 2008c	Reprod Biomed Online [29]	Extending clomiphene treatment in clomiphene-resistant women with PCOS: a randomized controlled trial.
8	Badawy 2009a	Fertil Steril [30]	Gonadotropin-releasing hormone agonists for prevention of chemotherapy-induced ovarian damage: prospective randomized study.
9	Badawy 2009b	Fertil Steril [31]	Luteal phase clomiphene citrate for ovulation induction in women with polycystic ovary syndrome: a novel protocol.
10	Badawy 2009c	Fertil Steril [32]	Ultrasound-guided transvaginal ovarian needle drilling (UTND) for treatment of polycystic ovary syndrome: a randomized controlled trial.
11	Badawy 2009d	Fertil Steril [33]	Extended letrozole therapy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a novel protocol.
12	Badawy 2009e	Fertil Steril [34]	Clomiphene citrate or letrozole for ovulation induction in women with polycystic ovarian syndrome: a prospective randomized trial.
13	Badawy 2009f	Fertil Steril [35]	Clomiphene citrate or anastrozole for ovulation induction in women with polycystic ovary syndrome? A prospective controlled trial.
14	Badawy 2009g	Fertil Steril [36]	Clomiphene citrate or aromatase inhibitors for superovulation in women with unexplained infertility undergoing intrauterine insemination: a prospective randomized trial.
15	Badawy 2009h	Acta Obstet Gynecol Scand [37]	Pregnancy outcome after ovulation induction with aromatase inhibitors or clomiphene citrate in unexplained infertility.
16	Badawy 2010a	J Obstet Gynaecol [38]	Clomiphene citrate or aromatase inhibitors combined with gonadotropins for superovulation in women undergoing intrauterine insemination: a prospective randomised trial.
17	Badawy 2010b	J Obstet Gynaecol [39]	Laparoscopy or not for management of unexplained infertility.
18	Badawy 2011	Eur J Obstet Gynecol Reprod Biol [40]	Clomiphene citrate versus tamoxifen for ovulation induction in women with PCOS: a prospective randomized trial.
19	Badawy 2012	Acta Obstet Gynecol Scand [41]	Aromatase inhibitors or gonadotropin-releasing hormone agonists for the management of uterine adenomyosis: a randomized controlled trial.
20	Gibreel 2013	J Obstet Gynaecol Res [42]	Endometrial scratching to improve pregnancy rate in couples with unexplained subfertility: a randomized controlled trial.
21	Ragab 2013	Arch Gynecol Obstet [43]	Does immediate postpartum curettage of the endometrium accelerate recovery from preeclampsia-eclampsia? A randomized controlled trial.
22	El Refaeey 2014	Reprod Biomed Online [44]	Combined coenzyme Q10 and clomiphene citrate for ovulation induction in clomiphene-citrate-resistant polycystic ovary syndrome.
23	Marzouk 2015	J Obstet Gynaecol [45]	Lavender-thymol as a new topical aromatherapy preparation for episiotomy: A randomised clinical trial.
24	El-Refaie 2016	Arch Gynecol Obstet [46]	Vaginal progesterone for prevention of preterm labor in asymptomatic twin pregnancies with sonographic short cervix: a randomized clinical trial of efficacy and safety.

No.	Study	Journal	Title
1	Abu Hashim 2010a	Arch Gynecol Obstet [47]	Letrozole versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial.
2	Abu Hashim 2010b	Fertil Steril [48]	Letrozole versus combined metformin and clomiphene citrate for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial.
3	Abu Hashim 2010c	J. Women's Health [49]	N-Acetyl Cysteine Plus Clomiphene Citrate Versus Metformin and Clomiphene Citrate in Treatment of Clomiphene-Resistant Polycystic Ovary Syndrome: A Randomized Controlled Trial
4	Abu Hashim 2011a	Gynecol Endocrinol [50]	Combined metformin and clomiphene citrate versus highly purified FSH for ovulation induction in clomiphene-resistant PCOS women: a randomised controlled trial
5	Abu Hashim 2011b	Acta Obstet Gynecol Scand [51]	Intrauterine insemination versus timed intercourse with clomiphene citrate in polycystic ovary syndrome: a randomized controlled trial
6	Abu Hashim 2011c	J. Obstet. Gynaecol. Res [52]	Combined metformin and clomiphene citrate versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: A randomized controlled trialjog_1383 169..177
7	Shokeir 2011	J Minim Invasive Gynecol [53]	An RCT: Use of Oxytocin Drip during Hysteroscopic Endometrial Resection and Its Effect on Operative Blood Loss and Glycine Deficit
8	Abu Hashim 2011d	Arch Gynecol Obstet [54]	Laparoscopic ovarian diathermy after clomiphene failure in polycystic ovary syndrome: is it worthwhile? A randomized controlled trial
9	Abu Hashim 2012a	Gynecol Endocrinol [55]	Minimal stimulation or clomiphene citrate as first-line therapy in women with polycystic ovary syndrome: a randomized controlled trial
10	Abu Hashim 2012b	Contraception [56]	Contraceptive vaginal ring treatment of heavy menstrual bleeding: a randomized controlled trial with norethisterone
11	Abu Hashim 2012c	Acta Obstet Gynecol Scand [57]	Randomized comparison of superovulation with letrozole vs. clomiphene citrate in an IUI program for women with recently surgically treated minimal to mild endometriosis

Appendix B

Table B1
The Mean, SD, so called SE and SD after conversion of the baseline characteristics from Badawy 2009e.

variable	group	participants	mean	reported SE	sd after conversion
Age (years)	1	218	27.1	3.2	47.2
Age (years)	2	220	29.3	2.9	43
Parity	1	218	0.3	0.1	1.5
Parity	2	220	0.4	0.2	3
Height (cm)	1	218	163.3	6.1	90.1
Height (cm)	2	220	158.1	5.2	77.1
Weight (kg)	1	218	98.3	6.4	94.5
Weight (kg)	2	220	91.1	4.2	62.3
BMI (kg/m^2)	1	218	28.1	3.2	47.2
BMI (kg/m^2)	2	220	27.1	3.1	46
FSH (IU/mL)	1	218	4.1	3.1	45.8
FSH (IU/mL)	2	220	5.1	2.2	32.6
LH (IU/mL)	1	218	11.2	1.8	26.6
LH (IU/mL)	2	220	14.1	2.2	32.6

Appendix C. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejogrb.2020.04.016.

References

[1] Bordewijk EM, Ng KYB, Rakic L, Mol BWJ, Brown J, Crawford TJ, et al. Laparoscopic ovarian drilling for ovulation induction in women with anovulatory polycystic ovary syndrome. Cochrane Database Syst Rev 2020.
[2] Hullemann S, Schupfer G, Mauch J. Application of Benford's law: a valuable tool for detecting scientific papers with fabricated data? : a case study using proven falsified articles against a comparison group. Anaesthesist 2017;66: 795–802.
[3] Bolland MJ, Gamble GD, Avenell A, Grey A, Lumley T. Baseline P value distributions in randomized trials were uniform for continuous but not categorical variables. J Clin Epidemiol 2019;112:67–76.
[4] Carlisle JB, Loadsman JA. Evidence for non-random sampling in randomised, controlled trials by Yuhji Saitoh. Anaesthesia 2017;72:17–27.
[5] Li W, Suke S, Wertaschnigg D, Lensen S, Wang R, Gurrin L, et al. Randomised controlled trials evaluating endometrial scratching: assessment of methodological issues. Hum Reprod 2019;34:2372–80.
[6] Betensky RA, Chiou SH. Correlation among baseline variables yields non-uniformity of p-values. PLoS One 2017;12:e0184531.
[7] Bolland MJ, Gamble GD, Avenell A, Grey A. Rounding, but not randomization method, non-normality, or correlation, affected baseline P-value distributions in randomized trials. J Clin Epidemiol 2019;110:50–62.
[8] Carlisle JB. The analysis of 168 randomised controlled trials to test data integrity. Anaesthesia 2012;67:521–37.
[9] Bolland MJ, Avenell A, Gamble GD, Grey A. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. Neurology 2016;87:2391–402.
[10] Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Fertil Steril 2018;110:364–79.
[11] Chambers LM, Michener CM, Falcone T. Plagiarism and data falsification are the most common reasons for retracted publications in obstetrics and gynaecology. BJOG 2019;126:1134–40.
[12] Steen RG, Casadevall A, Fang FC. Why has the number of scientific retractions increased? PLoS One 2013;8: e68397-e.
[13] Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CK. Check for publication integrity before misconduct. Nature 2020;577:167–9.
[14] Shehata NAA. Retraction notice to: "Calcium versus oral contraceptive pills containing drospirenone for the treatment of mild to moderate premenstrual syndrome: A double blind randomized placebo controlled trial" [Eur J Obstet Gynecol Reprod Biol (2016) 100-104]. Eur J Obstet Gynecol Reprod Biol 2019;240:387.
[15] Statement of Retraction: peri-conceptional progesterone treatment in women with unexplained recurrent miscarriage: a randomized double-blind placebo-controlled trial. J Maternal-Fetal Neonatal Med 2020;33:1073.
[16] Notice of retraction: “Sildenafil citrate therapy for oligohydramnios: a randomized controlled trial” (Maher MA, Sayyed TM, and elkhouly N). Obstet Gynecol 2020 9000; Publish Ahead of Print.
[17] Byrne JA, Grima N, Capes-Davis A, Labbe C. The possibility of systematic research fraud targeting under-studied human genes: causes, consequences, and potential solutions. Biomark Insights 2019;14:1177271919829162.
[18] Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. BMJ 2015;350:h2463.
[19] Tan AC, Jiang I, Askie L, Hunter K, Simes RJ, Seidler AL. Prevalence of trial registration varies by study characteristics and risk of bias. J Clin Epidemiol 2019;113:64–74.
[20] Farquhar CM, Showell MG, Showell EAE, Beetham P, Baak N, Mourad S, et al. Clinical trial registration was not an indicator for low risk of bias. J Clin Epidemiol 2017;84:47–53.
[21] Tovey D, Bero L, Farquhar C, Lasserson T, MacLehose H, Macdonald G, et al. A response to Ian Roberts and his colleagues [Letter]. BMJ 2015;2015(350):h2463.
[22] Taichman D, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the international committee of medical journal editors. Natl Med J India 2017;30:121–4.
[23] Badawy A, Baker El Nashar A, El Totongy M. Clomiphene citrate plus N-acetyl cysteine versus clomiphene citrate for augmenting ovulation in the management of unexplained infertility: a randomized double-blind controlled trial. Fertil Steril 2006;86:647–50.
[24] Badawy A, State O, Abdelgawad S. N-Acetyl cysteine and clomiphene citrate for induction of ovulation in polycystic ovary syndrome: a cross-over trial. Acta Obstet Gynecol Scand 2007;86:218–22.
[25] Badawy A, Metwally M, Fawzy M. Randomized controlled trial of three doses of letrozole for ovulation induction in patients with unexplained infertility. Reprod Biomed Online 2007;14:559–62.
[26] Badawy A, Goda H, Ragab A. Induction of ovulation in idiopathic premature ovarian failure: a randomized double-blind trial. Reprod Biomed Online 2007;15:215–9.
[27] Badawy AM, Khiary M, Sherif LS, Hassan M, Ragab A, Abdelall I. Low-molecular weight heparin in patients with recurrent early miscarriages of unknown aetiology. J Obstet Gynaecol 2008;28:280–4.
[28] Badawy A, Mosbah A, Shady M. Anastrozole or letrozole for ovulation induction in clomiphene-resistant women with polycystic ovarian syndrome: a prospective randomized trial. Fertil Steril 2008;89:1209–12.
[29] Badawy A, Allam A, Abulatta M. Extending clomiphene treatment in clomiphene-resistant women with PCOS: a randomized controlled trial. Reprod Biomed Online 2008;16:825–9.
[30] Badawy A, Elnashar A, El-Ashry M, Shahat M. Gonadotropin-releasing hormone agonists for prevention of chemotherapy-induced ovarian damage: prospective randomized study. Fertil Steril 2009;91:694–7.
[31] Badawy A, Inany H, Mosbah A, Abulatta M. Luteal phase clomiphene citrate for ovulation induction in women with polycystic ovary syndrome: a novel protocol. Fertil Steril 2009;91:838–41.
[32] Badawy A, Khiary M, Ragab A, Hassan M, Sherief L. Ultrasound-guided transvaginal ovarian needle drilling (UTND) for treatment of polycystic ovary syndrome: a randomized controlled trial. Fertil Steril 2009;91:1164–7.
[33] Badawy A, Mosbah A, Tharwat A, Eid M. Extended letrozole therapy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a novel protocol. Fertil Steril 2009;92:236–9.
[34] Badawy A, Abdel Aal I, Abulatta M. Clomiphene citrate or letrozole for ovulation induction in women with polycystic ovarian syndrome: a prospective randomized trial. Fertil Steril 2009;92:849–52.
[35] Badawy A, Abdel Aal I, Abulatta M. Clomiphene citrate or anastrozole for ovulation induction in women with polycystic ovary syndrome? A prospective controlled trial. Fertil Steril 2009;92:860–3.
[36] Badawy A, Elnashar A, Totongy M. Clomiphene citrate or aromatase inhibitors for superovulation in women with unexplained infertility undergoing intrauterine insemination: a prospective randomized trial. Fertil Steril 2009;92:1355–9.
[37] Badawy A, Shokeir T, Allam AF, Abdelhady H. Pregnancy outcome after ovulation induction with aromatase inhibitors or clomiphene citrate in unexplained infertility. Acta Obstet Gynecol Scand 2009;88:187–91.
[38] Badawy A, Elnashar A, Totongy M. Clomiphene citrate or aromatase inhibitors combined with gonadotropins for superovulation in women undergoing intrauterine insemination: a prospective randomised trial. J Obstet Gynaecol 2010;30:617–21.
[39] Badawy A, Khiary M, Ragab A, Hassan M, Sherif L. Laparoscopy–or not–for management of unexplained infertility. J Obstet Gynaecol 2010;30:712–5.
[40] Badawy A, Gibreal A. Clomiphene citrate versus tamoxifen for ovulation induction in women with PCOS: a prospective randomized trial. Eur J Obstet Gynecol Reprod Biol 2011;159:151–4.
[41] Badawy AM, Elnashar AM, Mosbah AA. Aromatase inhibitors or gonadotropin-releasing hormone agonists for the management of uterine adenomyosis: a randomized controlled trial. Acta Obstet Gynecol Scand 2012;91:489–95.
[42] Gibreel A, Badawy A, El-Refai W, El-Adawi N. Endometrial scratching to improve pregnancy rate in couples with unexplained subfertility: a randomized controlled trial. J Obstet Gynaecol Res 2013;39:680–4.
[43] Ragab A, Goda H, Raghib M, Barakat R, El-Samanoudy A, Badawy A. Does immediate postpartum curettage of the endometrium accelerate recovery from preeclampsia-eclampsia? A randomized controlled trial. Arch Gynecol Obstet 2013;288:1035–8.
[44] El Refaeey A, Selem A, Badawy A. Combined coenzyme Q10 and clomiphene citrate for ovulation induction in clomiphene-citrate-resistant polycystic ovary syndrome. Reprod Biomed Online 2014;29:119–24.

[45] Marzouk T, Barakat R, Ragab A, Badria F, Badawy A. Lavender-thymol as a new topical aromatherapy preparation for episiotomy: a randomised clinical trial. J Obstet Gynaecol 2015;35:472–5.
[46] El-Refaie W, Abdelhafez MS, Badawy A. Vaginal progesterone for prevention of preterm labor in asymptomatic twin pregnancies with sonographic short cervix: a randomized clinical trial of efficacy and safety. Arch Gynecol Obstet 2016;293:61–7.
[47] Abu Hashim H, Mashaly AM, Badawy A. Letrozole versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial. Arch Gynecol Obstet 2010;282:567–71.
[48] Abu Hashim H, Shokeir T, Badawy A. Letrozole versus combined metformin and clomiphene citrate for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial. Fertil Steril 2010;94:1405–9.
[49] Abu Hashim H, Anwar K, El-Fatah RA. N-acetyl cysteine plus clomiphene citrate versus metformin and clomiphene citrate in treatment of clomiphene-resistant polycystic ovary syndrome: a randomized controlled trial. J Womens Health 2002;2010(19):2043–8.
[50] Abu Hashim H, Wafa A, El Rakhawy M. Combined metformin and clomiphene citrate versus highly purified FSH for ovulation induction in clomiphene-resistant PCOS women: a randomised controlled trial. Gynecol Endocrinol 2011;27:190–6.
[51] Abu Hashim H, Ombar O, Abd Elaal I. Intrauterine insemination versus timed intercourse with clomiphene citrate in polycystic ovary syndrome: a randomized controlled trial. Acta Obstet Gynecol Scand 2011;90:344–50.
[52] Abu Hashim H, El Lakany N, Sherief L. Combined metformin and clomiphene citrate versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial. J Obstet Gynaecol Res 2011;37:169–77.
[53] Shokeir T, El-Lakkany N, Sadek E, El-Shamy M, Abu Hashim H. An RCT: use of oxytocin drip during hysteroscopic endometrial resection and its effect on operative blood loss and glycine deficit. J Minim Invasive Gynecol 2011;18:489–93.
[54] Abu Hashim H, Foda O, Ghayaty E, Elawa A. Laparoscopic ovarian diathermy after clomiphene failure in polycystic ovary syndrome: is it worthwhile? A randomized controlled trial. Arch Gynecol Obstet 2011;284:1303–9.
[55] Abu Hashim H, Bazeed M, Abd Elaal I. Minimal stimulation or clomiphene citrate as first-line therapy in women with polycystic ovary syndrome: a randomized controlled trial. Gynecol Endocrinol 2012;28:87–90.
[56] Abu Hashim H, Alsherbini W, Bazeed M. Contraceptive vaginal ring treatment of heavy menstrual bleeding: a randomized controlled trial with norethisterone. Contraception 2012;85:246–52.
[57] Abu Hashim H, El Rakhawy M, Abd Elaal I. Randomized comparison of superovulation with letrozole vs. Clomiphene citrate in an IUI program for women with recently surgically treated minimal to mild endometriosis. Acta Obstet Gynecol Scand 2012;91:338–45.

Esmée M Bordewijk[a,b,*]
[a]*Department of Obstetrics and Gynecology, Monash University, Clayton, Australia*

[b]*Centre for Reproductive Medicine, Amsterdam UMC, Amsterdam, the Netherlands*

Rui Wang
Department of Obstetrics and Gynecology, Monash University, Clayton, Australia

Lisa M. Askie
NHMRC Clinical Trials Centre, University of Sydney, Australia

Lyle C. Gurrin
Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, Parkville, Australia

Jim G. Thornton
Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, United Kingdom

Madelon van Wely
Centre for Reproductive Medicine, Amsterdam UMC, Amsterdam, the Netherlands

Wentao Li
Ben W. Mol
Department of Obstetrics and Gynecology, Monash University, Clayton, Australia

* Corresponding author at: Department of Obstetrics and Gynecology, Monash University, Clayton, Australia.
E-mail address: e.m.bordewijk@amsterdamumc.nl (E. Bordewijk).

Received 16 March 2020