Supplementary File 1

Table S1. Summary of Extraction Data

Author/Year/Design	Purpose	Sample/Setting	Intervention	Measures	Results	Strengths/Limitations
Li, He, Yang, Yin, and Xu (2011). Prospective, nonrandomized, observational, cross-sectional	Investigating the medication compliance of infertile patients with polycystic ovarian syndrome (PCOS) Assessing factors which may contribute to noncompliance in order to provide a basis for clinical treatment, specialist consultation, and health education	N= 90 Age: x=28.71± 3.63 years 20-25 (n=17) 26-30 (n=45) ≥31 (n=28) BMI: x=24.0±4.34 kg/m² <23 (n=38, 23-25 (n=18), ≥ 25 (n=34) Race or ethnic group: NR Length of time attempting to conceive: ≥ 1 year No history of pregnancy: 68.9% Previous history of infertility treatment: 34.4% Self-pay for medical expenses: 88.9% Postsecondary education: 57.8%	None	(1) Questionnaire including 3 questions derived from the Morisky-Green test and 1 question addressing weight loss based on the principles of PCOS. (2) Questionnaire with contents assessing (a) Demographic information, (b) Disease diagnostic information, and (c) Self-factors including personal, medical, economic, and social experiences and concerns surrounding PCOS Timing: baseline, study lasted for 6 months	Primary findings: Overall compliance with the treatment: 23 (25.6%) Compliance with the medications: contraceptive drugs- 31 (48.4%), anti-insulin resistance drugs-28 (52.8%), clomiphene-15(60%), traditional Chinese medicine-3(30%). BMI (P=0.040), convenience of medical treatment (*P*=0.012), and concerns about adverse drug effects (*P*=0.043) significantly affected compliance with the treatment. Secondary findings: NR	Strengths: Sample inclusion criteria clearly specified; moderate sample size; adherence measures adequately described; statistical methods appropriate. Limitations: External review of study by an ethics review board not reported; actual compliance rates not reported, only categorized as good compliance rates; convenience sampling and nonrandomized; race/ethnicity of sample not reported; potential for sampling bias and sampling homogeneity limiting generalizability of study results; potential for a lower reliability and accuracy in participant self-reporting on survey questionnaires; potential for inaccuracy in content validity on second questionnaire; no theoretical framework basis was addressed.

		Setting: Reproductive Medical Center in China				
Mc Govern et al. (2008), Retrospective, randomized controlled trial	Examining mediation adherence in the metformin-containing arms of the primary study to determine whether participants within the expected range for similar trials	N=626 Metformin group: n=208, Clomiphene group: n=209, Combination group: n=209 Age (years): Metformin group 28.1± 4.0, Clomiphene group 27.9± 4.0, Combination group 28.3 ± 4.0 BMI: metformin group 35.6 kg/m² ± 8.5, clomiphene group 36.0± 8.9, combination group 34.2 ± 8.4 Race or ethnic group: Metformin group- Caucasian 67.6%, Hispanic 29.3%, Black 19.3%, Asian 2.4%, Native American 13.0%: Clomiphene group- Caucasian 70.7%, Hispanic	Metformin Group: 2000 mg daily plus clomiphene placebo Clomiphene Group: 50, 100, or 150 mg daily for 5 days per cycle plus metformin placebo. Combined Group: metformin 2000mg daily plus clomiphene 50, 100, or 150 mg daily for 5 days per cycle.	Pill counts used to assess adherence by percentage of the recommended tablets not in the returned bottles by counting the remaining tablets. Timing: baseline and monthly for up 6 cycles or 6 months, pill counts from returned bottles performed monthly.	Primary findings: Overall median adherence rate: 81%; median adherence rates: 81.6% in metformin group and 81.7% in combination group. No significant (*P*=0.80) difference in medication adherence between metformin and combined groups. Ovulatory rates were low in the metformin group across all levels of the adherence. Secondary findings: Median adherence for clomiphene group was 100% in clomiphene and combined groups.	Strength: Sample randomization and control, and blinding present, Sample large and heterogeneous. Sampling bias minimized by stratification based on the study site and the presence or absence of previous exposure to either of the study drugs. Instruments were adequately described; appropriate statistical methods. Limitations: Study setting not adequately described; a total of 176 participants dropped out of the study; medication adherence not originally reported in primary study; discrepancy between the sample size of metformin arm (n=195) in secondary study and that of the metformin arm (n=208) reported in primary study; pill counts less reliable in assessing medication adherence if participants removed pills out of bottle which were not taken; primary study was not designed to investigate adherence systematically; no theoretical basis was addressed.

		25.4%, Black 17.8%, Asian 2.4%, Native American 10.1%; Combination group- Caucasian 71.2%, Hispanic 23.9%, Black 15.4%, Asian 3.4%, Native American 11.5% Length of time attempting to conceive (months): metformin group 39.0± 31.9, clomiphene group 41.4± 39.4, combination group 40.7 ± 36.0 Previous history of infertility treatment: metformin group 53.4%, clomiphene group 55.5%, combination group 55.5% Self-pay for medical expenses: NR Postsecondary education: NR Setting: Multi-centers, location unknown.				

Kruse, Eggert-Kruse, Rampmaier, Runnebaum, and Weber (1993). Prospective, cross-sectional, randomized	Evaluating the relationship between adverse reactions and patient compliance in women with primary infertility	N=61 Age: x=30.4±4.4 years (range, 21-39 years) BMI: NR Race or ethnic group: NR Length of time attempting to conceive: 4.5± 2.7 years (range, 9 months- 19 years) No history of pregnancy: NR Self-pay for medical expenses: NR Postsecondary education: NR Setting: Infertility Unit of Women's Hospital, Heidelberg, Germany	Ethinylestra-diol at 40µG twice daily or 20µG four times daily for 7 days	Medication Event Monitoring System used to evaluate the percentage of the prescribed doses taken during the period (administrative compliance- container openings recorded during the period divided by the prescribed number of doses during the period) and adherence to the prescribed dose schedule (regimen compliance- the number of days in which two openings for twice daily regimen or four openings for four times per daily regimen) Timing: daily for 7 days Standardized questionnaire to assess or adverse drug reactions which asked participants to rate symptoms which they attributed to the drug as mild, moderate, or severe	Primary findings: The overall mean compliance was 75.7%; Administration compliance ranged from 7-143% and regimen compliance ranged from 0-100% for twice a day dosing; Administration compliance ranged from 14-136% and regimen compliance ranged from 0-100% for four times per day dosing. Mean administration compliance: 85% for twice a day dosing and 65% for four times per day dosing ($P<0.05$). Mean regimen compliance: 62% for twice a day dosing and 34% for four times per day dosing ($P<0.005$). No significant difference in compliance compared to participants with or without adverse drug reactions. Compliance was significantly lower ($P< 0.05$) when participants reported three or more adverse drug reactions versus one or two: 54% versus 84% in	Strengths: Randomized sampling and moderate sample size; adherence measures adequately described; Key variables were operationalized. Limitations: Limited review of the literature to provide synthesis on the existing evidence of medication adherence; external review of study by an ethics review board not reported; method of sample randomization design not addressed; no report on avenues used to minimize sampling bias; sample inclusion criteria not clearly specified; limited sample demographics; ethnicity not reported-potential of sample homogeneity; potential for a lower reliability and accuracy in participant self-reporting on questionnaires; statistical methods used were not adequately described; no theoretical framework basis addressed.

				Timing: once after completing 7 days of ethinylestradiol	administrative compliance and 31% versus 58% in regimen compliance); Compliance was lower in participants with nausea and vomiting compared to those without these symptoms: 59% versus 91% in administrative compliance and 34% versus 66% in regimen compliance (*P*< 0.005); Compliance was lower with moderate or severe side effects compared to mild side effects: 48% versus 85% in administrative compliance and 25% versus 59% in regimen compliance (*P*< 0.005).	
Kruse, Eggert-Kruse, Rampmaier, Runnebaum, and Weber (1991), Prospective, cross-sectional, randomized	Investigating the patient compliance with two different dosage schedules of ethinyloestradiol 20 µg four times daily versus 40 µg two times daily	N=65 Age: x=29.9 years (range, 21-39 years) BMI: NR Race or ethnic group: NR Length of time attempting to conceive: 4.3 years	Ethinyloestr-adiol at 40 µg twice daily or 20 µg four times daily for 7 days	Medication Event Monitoring System used to evaluate percentage of container openings (recorded pill openings during the period divided by the prescribed number of doses during the period	Mean overall compliance was 75.7% (range 7.1 - 143%); Mean compliance with twice daily dosing was 85% compared to 67% with four times per day dosing (*P*<0.05); Regimen compliance: 63% for twice a day dosing and 36% for four	Strengths: Randomized sampling, moderate sample size; adherence measures adequately described; key variables were operationalized; to minimize attrition, participants who failed to returned MEMs bottles were email reminded; statistical methods appropriate. Limitations: Limited review of the literature to provide synthesis on the

		(range 9 months to 19 years) No history of pregnancy: NR Self-pay for medical expenses: NR Postsecondary education: NR Setting: Infertility Unit of Women's Hospital, Heidelberg, Germany		multiplied by 100) and adherence to the prescribed dose schedule (regimen compliance- the number of days in which two openings for twice daily regimen or four openings for four times per daily regimen) Timing: daily for 7 days	times per day dosing (P<0.005).	existing evidence of medication adherence; external review of study by an ethics review board not reported; method of sample randomization design not addressed; sample inclusion criteria not clearly specified; no discussion of power analysis to estimate sample size; limited sample demographics; potential for limitation of study finding generalizability; ethnicity not reported-potential of sample homogeneity; no theoretical framework basis was addressed.
Kruse, Eggert-Kruse, Rampmaier, Runnebaum, and Weber (1990), Prospective, cross-sectional, randomized	Investigating the patient compliance with ethinyl-oestradiol therapy of 20 µg four times daily.	N=30 Age: x=28.8 years (range, 21-36 years) BMI: NR Race or ethnic group: NR Length of time attempting to conceive: 3.9 years (range 9 months to 8 years) No history of pregnancy: NR Self-pay for medical expenses: NR Postsecondary education: NR	Ethinyl-oestradiol therapy of 20 µg four times daily.	Medication Event Monitoring System; Compliance data was obtained as the listing of the time and date of individual bottle openings and closings, duration of openings, and the hours since previous dose; Compliance was defined as the number of doses taken during period divided by the number of prescribed doses during period multiplied by 100. Timing: daily for 7 days	Mean overall compliance was 64.9% (range 14.3 to 136%). Mean adherence to prescribed QID regimen was 34.3% (range 0-114%); Sixteen of the 30 participants reported adverse drug reaction on response open-question or spontaneously; Twenty-four participants reported side effects on standardized questionnaire; Seventy-nine % of the symptoms were rated as being mild;	Strengths: Randomized sampling, adherence measures adequately described. Strengths: Key variables were operationalized; addressed accuracy of MEMS; to minimize attrition, participants who failed to return MEMs bottles were email reminded; all the participants were interviewed by the same investigator; to reduce cofounding variables, only participants who were not taking other medications during study were included; statistical methods appropriate. Limitations: Small sample size; limited review of the literature to provide synthesis on the existing evidence of medication adherence; external review of the study by an ethics review board not reported; method of sample randomization design not addressed; sample inclusion criteria not clearly

		Setting: Infertility Unit of University Women's Hospital		Interview regarding adverse drug effects which included one open question followed by a standardized questionnaire rating adverse symptoms as mild, moderate, or severe. Timing: Immediately after the 7-day ethinyl oestradiol regimen completed	In participants who had a compliance rate >65%, adherence rate correlated with reported adverse drug reactions inversely (r=0.71; *P*<0.01). The lower the compliance rate, the more adverse drug reactions were reported (no r or *P* value provided). Compliance was positively correlated with duration of infertility (r=0.44; *P*<0.05).	specified; no discussion of power analysis to estimate sample size; limited sample demographics; potential for limitation of study finding generalizability; ethnicity not reported-potential of sample homogeneity; potential for a lower reliability and accuracy in participant self-reporting on questionnaires; no theoretical framework basis addressed.

`

Table S2. Downs and Black Checklist

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Quality Score
MCGovern et al.(2008)	1	1	1	1	1	1	0	1	1	0	0	1	0	1	1	1	1	1	0	0	1	1	1	0	0	1	5	23
Kruse et al. (1993)	0	0	1	1	0	1	1	1	0	0	0	0	1	0	0	1	1	1	1	1	1	0	1	0	0	0	1	14

Key:

Reporting: “Yes=1”, “No=0”

1. Is the hypothesis /aim /objective of the study clearly described?

2. Are the main outcomes to be measured clearly described in the Introduction or Method Section?

3. Are the characteristics of the patients/samples included in the study clearly described?

4. Are the interventions of interest clearly described?

“Yes=2”, “Partially=1”, “No=0”

5. Are the distributions of principal confounders in each group of subjects to be compared clearly described?

“Yes=1”, “No=0”

6. Are the main findings of the study clearly described?

7. Does the study provide estimates of the random variability in the data for the main outcomes?

8. Have all the important adverse events which may be a consequence of the intervention been reported?

9. Have the characteristics of the patients lost to follow-up been described?

10. Have actual probability values been reported (e.g., 0.035 rather than <0.05) for the main outcomes except for where the probability value is less than 0.001?

External Validity: “Yes=1”, “No=0”, “Unable to determine=0”

11. Were the subjects asked to participate in the study represented the entire population from whom they were recruited?

12. Were those subjects who were prepared to participate represented the entire population from whom they were recruited?

13. Were the staff, places, and facilities where the patients were treated, represented the treatment of the majority of patients receive?

Internal Validity-bias: “Yes=1”, “No=0”, “Unable to determine=0”

14. Was an attempt made to blind study subjects to the intervention they received? 15. Was an attempt made to blind those measuring the main outcomes of the intervention? 16. If any of the results of the study were based on “data dredging” was this made clear? 17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?
18. Were the statistical tests used to assess the main outcomes appropriate? 19. Was compliance with the intervention/s reliable? 20.Were the main outcome measures used accurately (valid and reliable)?

Internal Validity-confounding (Selection bias): “Yes=1”, “No=0”, “Unable to determine=0”
21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? 22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?
23.Were the study subjects randomized to intervention groups?
24. Was the randomized intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?
25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?
26. Were losses of patients to follow-up taken into account?

Power
27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%? Sample sizes were calculated to detect a difference of x% and y%.
The size of the smallest intervention group:
1. A 1<n1 0
2. B n1-n2 1
3. C n3-n4 2
4. D n5-n6 3
5. E n7-n8 4
6. F n8+ 5