Cochrane Library

Cochrane Database of Systematic Reviews

Metformin versus the combined oral contraceptive pill for hirsutism, acne, and menstrual pattern in polycystic ovary syndrome (Review)

Fraison E, Kostova E, Moran LJ, Bilal S, Ee CC, Venetis C, Costello MF

Fraison E, Kostova E, Moran LJ, Bilal S, Ee CC, Venetis C, Costello MF.
Metformin versus the combined oral contraceptive pill for hirsutism, acne, and menstrual pattern in polycystic ovary syndrome.
Cochrane Database of Systematic Reviews 2020, Issue 8. Art. No.: CD005552.
DOI: 10.1002/14651858.CD005552.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Metformin versus the combined oral contraceptive pill for hirsutism, acne, and menstrual pattern in polycystic ovary syndrome

Eloise Fraison[1], Elena Kostova[2], Lisa J Moran[3], Sophia Bilal[4], Carolyn C Ee[5], Christos Venetis[6], Michael F Costello[7]

[1]Department of Gynecology and Obstetrics, Hôpital Femme Mère Enfant, Bron, France. [2]Center for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands. [3]Monash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University, Clayton, Australia. [4]Department of Obstetrics and Gynaecology, Westmead Hospital, Westmead, Australia. [5]NICM Health Research Institute, Western Sydney University, Penrith, Australia. [6]Centre for Big Data Research in Health & School of Women's & Children's Health, UNSW Medicine, Sydney, Australia. [7]Division of Obstetrics & Gynaecology, School of Women's and Children's Health, UNSW and Royal Hospital for Women and IVF Australia, Sydney, Australia

Contact: Michael F Costello, mfcostello@unsw.edu.au.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 8, 2020.

Citation: Fraison E, Kostova E, Moran LJ, Bilal S, Ee CC, Venetis C, Costello MF. Metformin versus the combined oral contraceptive pill for hirsutism, acne, and menstrual pattern in polycystic ovary syndrome. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No.: CD005552. DOI: 10.1002/14651858.CD005552.pub3.

ABSTRACT

Background

Metformin has been proposed as possibly a safer and more effective long-term treatment than the oral contraceptive pill (OCP) in women with polycystic ovary syndrome (PCOS). It is important to directly compare the efficacy and safety of metformin versus OCP in the long-term treatment of women with PCOS. This is an update of a Cochrane Review comparing insulin sensitising agents with the OCP and only includes studies on metformin.

Objectives

To assess the effectiveness and safety of metformin versus the OCP (alone or in combination) in improving clinical, hormonal, and metabolic features of PCOS.

Search methods

In August 2019 we searched the Cochrane Gynaecology and Fertility Group Trials Register, Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase and CINAHL, the trial registers, handsearched references of the identified articles, and contacted experts in the field to identify additional studies.

Selection criteria

We included randomised controlled trials (RCTs) of the use of metformin versus the OCP (alone or in combination) for women with PCOS.

Data collection and analysis

We used standard methods recommended by Cochrane. The primary review outcomes were the clinical parameters of hirsutism and adverse events, both severe (requiring stopping of medication), and minor. In the presence of substantial heterogeneity (I^2 statistic > 50), which could be explained by pre-specified subgroup analyses on the basis of BMI, we reported the subgroups separately.

Main results

This is a substantive update. We identified 38 additional studies. We included 44 RCTs (2253 women), which comprised 39 RCTs on adult women (2047 women) and five RCTs on adolescent women (206 women). Evidence quality ranged from very low to low. The main limitations were risk of bias, imprecision and inconsistency.

Metformin versus the OCP

In adult women, we are uncertain of the effect of metformin compared to the OCP on hirsutism in subgroup body mass index (BMI) < 25 kg/m^2 (mean difference (MD) 0.38, 95% confidence interval (CI) -0.44 to 1.19, 3 RCTs, n = 134, I^2 = 50%, very low-quality evidence) and subgroup BMI > 30 kg/m^2 (MD -0.38, 95% CI -1.93 to 1.17; 2 RCTs, n = 85, I^2 = 34%, low-quality evidence). Metformin may be less effective in improving hirsutism compared to the OCP in the subgroup BMI 25 kg/m^2 to 30 kg/m^2 (MD 1.92, 95% CI 1.21 to 2.64, 5 RCTs, n = 254, I^2 = 0%, low-quality evidence). Metformin may increase severe gastro-intestinal adverse events rate compared to the OCP (Peto odds ratio (OR) 6.42, 95% CI 2.98 to 13.84, 11 RCTs, n = 602, I^2 = 0%, low-quality evidence). Metformin may decrease the incidence of severe other adverse events compared to the OCP (Peto OR 0.20, 95% CI 0.09 to 0.44, 8 RCTs, n = 363, I^2 = 0%, low-quality evidence). There were no trials reporting on minor adverse events.
In adolescents, we are uncertain whether there is a difference between Metformin and the OCP, on hirsutism and adverse events.

Metformin versus metformin combined with the OCP

In adult women, metformin may be less effective in improving hirsutism compared to Metformin combined with the OCP (MD 1.36, 95% CI 0.62 to 2.11, 3 RCTs, n = 135, I^2= 9%, low-quality evidence). We are uncertain if there was a difference between metformin and metformin combined with the OCP for severe gastro-intestinal adverse events (OR 0.74, 95% CI 0.21 to 2.53, 3 RCTs, n = 171, I^2 = 0%, low-quality evidence), or for severe other adverse events (OR 0.56, 95% CI 0.11 to 2.82, 2 RCTs, n = 109, I^2 = 44%, low-quality evidence). There were no trials reporting on minor adverse events. In adolescents, there were no trials for this comparison.

The OCP versus metformin combined with the OCP

In adult women, the OCP may be less effective in improving hirsutism compared to metformin combined with the OCP (MD 0.54, 95% CI 0.20 to 0.89, 6 RCTs, n = 389, I^2= 1%, low-quality evidence). The OCP may decrease the incidence of severe gastro-intestinal adverse events compared to metformin combined with the OCP (OR 0.20, 95% CI 0.06 to 0.72, 5 RCTs, n = 228, I^2 = 0%, low-quality evidence). We are uncertain if there is a difference between the OCP and metformin combined with the OCP for severe other adverse events (OR 1.61, 95% CI 0.49 to 5.37, 4 RCTs, n = 159, I^2 = 12%, low-quality evidence). The OCP may decrease the incidence of minor (gastro-intestinal) adverse events compared to metformin combined with the OCP (OR 0.06, 95% CI 0.01 to 0.44, 2 RCTs, n = 98, I^2 = 0%, low-quality evidence). In adolescents, we are uncertain whether there is a difference between the OCP, compared to metformin combined with the OCP, on hirsutism or adverse events.

Authors' conclusions

In adult women with PCOS, metformin may be less effective in improving hirsutism compared to the OCP in the subgroup BMI 25 kg/m^2 to 30 kg/m^2 but we are uncertain if there was a difference between metformin and the OCP in subgroups BMI < 25 kg/m^2 and BMI > 30kg/m^2. Compared to the OCP, metformin may increase the incidence of severe gastro-intestinal adverse events and decrease the incidence of severe other adverse events with no trials reporting on minor adverse events. Either metformin alone or the OCP alone may be less effective in improving hirsutism compared to metformin combined with the OCP. We are uncertain whether there is a difference between the OCP alone and metformin alone compared to metformin combined with the OCP for severe or minor adverse events except for the OCP versus metformin combined with the OCP where the OCP may decrease the incidence of severe and minor gastro-intestinal adverse events. In adolescent women with PCOS, we are uncertain whether there is a difference between any of the comparisons for hirsutism and adverse events due to either no evidence or very low-quality evidence.

Further large well-designed RCTs that stratify for BMI are needed to evaluate metformin versus the OCP and combinations in women with PCOS, in particular adolescent women.

PLAIN LANGUAGE SUMMARY

Metformin versus the combined oral contraceptive pill for excessive facial/body hair, acne, and menstrual disorders in polycystic ovary syndrome

Review question

Is Metformin more effective and safer than the oral contraceptive pill (OCP) (alone or in combination) in improving clinical, hormonal, and metabolic features (irregular/prolonged menstrual cycles, excessive facial and body hair, acne, obesity) in women with polycystic ovary syndrome (PCOS)?

Background

PCOS is a common hormonal and metabolic problem affecting approximately 1 in 10 women of childbearing age, often resulting in infrequent menstrual periods, excess body and facial hair, acne and polycystic ovaries (enlarged ovaries due to numerous small collections of fluid (follicles)). The OCP has long been a proven effective treatment for women with PCOS who are not trying to fall pregnant. More recently, metformin (a medication that lowers insulin and blood sugar levels and often used to treat type 2 diabetes) has been advocated as possibly a more effective and safer long-term treatment than the OCP in women with PCOS. Therefore, it is important to directly compare the benefits and risks of these two treatments in women with PCOS.

Study characteristics

We found 44 randomised controlled trials (RCTs) comparing metformin versus the OCP (alone or in combination) in a total of 2253 women with PCOS which comprised 39 RCTs on adult women (2047 women) and five RCTs on adolescent women (206 women). We combined results from the studies and assessed the quality of the studies to judge how confident we could be in their results. The evidence is current to August 2019.

Key results

In adult women, when we compared metformin to the OCP in terms of improving excessive facial and body hair, metformin may be less effective in women with PCOS with a body mass index (BMI) between 25 kg/m^2 to 30 kg/m^2, but we are uncertain of the effect with BMI less than 25 kg/m^2 or greater than 30 kg/m^2. In terms of severe adverse events (requiring stopping of medication), metformin may result in a higher incidence of gastro-intestinal (i.e. nausea, vomiting, diarrhoea), but a lower incidence of other adverse events. Evidence suggests that if the severe gastro-intestinal adverse event rate following the OCP is 0.3%, then the severe gastro-intestinal adverse event rate after metformin would be between 1% and 4.5%. Evidence also suggests that if the severe other adverse event rate following the OCP is 12%, the severe other adverse event rate after metformin would between 1% and 6%.

Either metformin alone or the OCP alone may be less effective in improving excessive facial and body hair compared to the combination of the OCP with metformin. In terms of severe adverse events, we are uncertain if there was a difference between metformin and metformin combined with the OCP for gastro-intestinal or other adverse events. If the severe gastro-intestinal adverse event rate following metformin combined with the OCP is 7%, then the corresponding rate after metformin would be between 2% and 17%, and if the severe other adverse event rate following metformin combined with the OCP is 6%, the corresponding rate after metformin would be between 0.7% and 15%.

When comparing the OCP to metformin combined with the OCP in terms of severe adverse events, there may be a lower incidence of gastro-intestinal adverse events with the OCP, but we are uncertain if there is a difference in other adverse events. If the severe gastro-intestinal adverse event rate is 10% following metformin combined with the OCP, the corresponding rate following the OCP would be between 1% and 7%. If the severe other adverse event rate is 4% following Metformin combined with the OCP, the corresponding rate following the OCP would be between 2% and 18%.

In adolescent women, we are uncertain as to whether there is a difference between any of the three comparisons in this review in terms of hirsutism and adverse events (both severe requiring stopping medication and minor) due to either a lack of evidence or very low-quality evidence based on one trial.

Quality of the evidence

The evidence was of very low to low quality. The main limitations in the evidence were poor reporting of study methods and a lack of both precision and consistency in the results.

Summary of findings 1. Metformin compared to oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS)

Metformin compared to OCP for hirsutism, acne, and menstrual pattern in adult women with PCOS

Patient or population: adult women with PCOS
Setting: Hospital or University Clinics
Intervention: metformin
Comparison: OCP

Outcomes		Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
		Risk with OCP	Risk with metformin				
Hirsutism - Clinical F-G score	BMI ≤ 25kg/m^2	The mean hirsutism - Clinical F-G score was 7.5	MD 0.38 higher (0.44 lower to 1.19 higher)	-	134 (3 RCTs)	⊕⊝⊝⊝ VERY LOW[1,2,3]	
	BMI > 25 kg/m^2 < 30 kg/m^2	The mean hirsutism - Clinical F-G score was 6.44	MD 1.92 higher (1.21 higher to 2.64 higher)	-	254 (5 RCTs)	⊕⊕⊝⊝ LOW[1,4]	
	BMI ≥ 30 kg/m^2	The mean hirsutism - Clinical F-G score was 6.05	MD 0.38 lower (1.93 lower to 1.17 higher)	-	85 (2 RCTs)	⊕⊕⊝⊝ LOW[1,5]	
Adverse events - Severe	Gastro-intestinal	3 per 1 000	21 per 1 000 (10 to 45)	OR 6.42 (2.98 to 13.84)	602 (11 RCTs)	⊕⊕⊝⊝ LOW [6,7]	
	Others	122 per 1000	27 per 1000 (12 to 57)	OR 0.20 (0.09 to 0.44)	363 (8 RCTs)	⊕⊕⊝⊝ LOW [6,7]	
Adverse events - Minor	Gastro-intestinal	No trials reported on outcome "Adverse events - Minor - Gastro-intestinal"					
	Others	No trials reported on outcome "Adverse events - Minor - Others"					
Improved menstrual pattern	Shortening of intermenstrual days	The mean improved menstrual pattern (ie. shortening of intermenstrual days) was 32.4	MD 6.05 higher (2.37 higher to 9.74 higher)	-	153 (2 RCTs)	⊕⊕⊝⊝⊝ LOW [4,8]	

	An initiation of menses or cycle regularity) - ≤ 25 kg/m^2	1000 per 1 000	1000 per 1 000 (1000 to 1000)	OR 0.07 (0.01 to 0.65)	17 (1 RCT)	⊕⊕⊖⊖ LOW [6,7]
	An initiation of menses or cycle regularity)- BMI > 25 kg/m^2 < 30 kg/m^2	931 per 1000	669 per 1000 (486 to 817)	OR 0.15 (0.07 to 0.33)	129 (3 RCTs)	⊕⊖⊖⊖ VERY LOW [7,8,9]
	An initiation of menses or cycle regularity) - BMI ≥ 30 kg/m^2	1000 per 1 000	1000 per 1 000 (1000 to 1000)	OR 0.09 (0.01 to 1.62)	18 (1 RCT)	⊕⊖⊖⊖ VERY LOW [6,10]
	An initiation of menses or cycle regularity) - BMI not stated	500 per 1000	661 per 1000 (281 to 906)	OR 1.95 (0.39 to 9.65)	25 (1 RCT)	⊕⊖⊖⊖ VERY LOW [8,10]
Acne - Visual analogue scale		The mean acne - Visual analogue scale was 1	MD 0.90 higher (0.40 lower to 2.20 higher)	-	34 (1 RCT)	⊕⊕⊖⊖ LOW [11]
BMI (kg/m^2)	BMI ≤ 25 kg/m^2	The mean BMI (kg/m^2) was 22.7	MD 0.59 lower (1.02 lower to 0.17 lower)	-	451 (9 RCTs)	⊕⊖⊖⊖ VERY LOW [1,12,13]
	BMI > 25 kg/m^2 < 30 kg/m^2	The mean BMI (kg/m^2) was 27.4	MD 0.11 higher (0.48 lower to 0.7 higher)	-	353 (8 RCTs)	⊕⊖⊖⊖ VERY LOW [1,14,15]
	BMI ≥ 30 kg/m^2	The mean BMI (kg/m^2) was 35.1	MD 2.31 lower (4.40 lower to 0.21 lower)	-	119 (3 RCTs)	⊕⊖⊖⊖ VERY LOW [1, 15,16]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

BMI: body mass index; **CI:** Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio; **RCT:** Randomised controlled trial.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[1] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have unclear or high risk of bias
[2] Evidence downgraded by one level for serious inconsistency (I^2 = 50%) as unexplained heterogeneity (i.e. heterogeneity not explained by subgrouping of data according to mean study BMI)
[3] Evidence downgraded by one level for serious imprecision – 95% CI includes both appreciable effect and little or no effect and low number of participants (total number of participants < 400)
[4] Evidence downgraded by one level for serious imprecision – low number of participants (total number of participants < 400)
[5] Evidence downgraded by one level for serious imprecision - low number of participants (total number of participants < 400) and 95% CI includes both appreciable benefit and harm
[6] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have unclear risk of bias
[7] Evidence downgraded by one level for serious imprecision – low number of events (total number of events < 300)
[8] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have high risk of bias
[9] Evidence downgraded by one level for serious inconsistency (I^2 = 51%) as unexplained heterogeneity (i.e. heterogeneity not explained by subgrouping of data according to mean study BMI)
[10] Evidence downgraded by two levels for very serious imprecision – 95% CI includes both appreciable benefit and harm or no effect and very low number of events (total number of events < 300)
[11] Evidence downgraded by two levels for serious imprecision – 95% CI includes both appreciable benefit and harm or no effect and low number of participants (total number of participants < 400)
[12] Evidence downgraded by one level for serious inconsistency (I^2 = 76%) as unexplained heterogeneity (i.e. heterogeneity not explained by subgrouping of data according to mean study BMI)
[13] Evidence downgraded by one level for serious imprecision – 95% CI includes both appreciable effect and little or no effect
[14] Evidence downgraded by one level for serious inconsistency (I^2 = 72%) as unexplained heterogeneity (i.e. heterogeneity not explained by subgrouping of data according to mean study BMI)
[15] Evidence downgraded by one level for serious imprecision - low number of participants (total number of participants < 400) and 95% CI includes both appreciable effect and little or no effect
[16] Evidence downgraded by one level for serious inconsistency (I^2 = 52%) as unexplained heterogeneity (i.e. heterogeneity not explained by subgrouping of data according to mean study BMI)

Summary of findings 2. Metformin compared to metformin combined with oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS)

Metformin compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adult women with PCOS

Patient or population: adult women with PCOS
Setting: Hospital or University Clinics
Intervention: metformin
Comparison: metformin combined with OCP

Outcomes	Anticipated absolute effects* (95% CI)	Relative effect	№ of participants	Quality of the evidence	Comments

		Risk with Metformin combined with OCP	Risk with metformin	(95% CI)	(studies)	(GRADE)
Hirsutism - Clinical F-G score		The mean hirsutism - Clinical F-G score was 5.6	MD 1.36 higher (0.62 higher to 2.11 higher)	-	135 (3 RCTs)	⊕⊕⊖⊖ LOW [1,2]
Adverse events - Severe	Gastro-intestinal	74 per 1 000	56 per 1 000 (17 to 168)	OR 0.74 (0.21 to 2.53)	171 (3 RCTs)	⊕⊕⊖⊖ LOW [1,3]
	Others	60 per 1 000	35 per 1 000 (7 to 153)	OR 0.56 (0.11 to 2.82)	109 (2 RCTs)	⊕⊕⊖⊖ LOW [1,3]
Adverse events - Minor	Gastro-intestinal	No trials reported on outcome "Adverse events - Minor - Gastro-intestinal"				
	Others	No trials reported on outcome "Adverse events - Minor - Others"				
Improved menstrual pattern	Shortening of intermenstrual days	No trials reported on outcome "Improved menstrual pattern (i.e. shortening of intermenstrual days)"				
	An initiation of menses or cycle regularity	No trials reported on outcome "Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)"				
Acne - Visual analogue scale/Clinical acne score		No trials reported either on outcome "Acne - Visual analogue scale" or "Acne - Clinical acne score"				
BMI (kg/m^2)		The mean Body Mass Index (kg/m^2) was 25.49	MD 1.47 lower (2.27 lower to 0.66 lower)	-	199 (5 RCTs)	⊕⊕⊖⊖ LOW [1,2]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

BMI: Body mass index; CI: Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio; **RCT:** Randomised controlled trial.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[1] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have unclear risk of bias
[2] Evidence downgraded by one level for serious imprecision – low number of participants (total number of participants < 400)
[3] Evidence downgraded by one level for serious imprecision – low number of events (total number of events < 300) and 95% CI includes both appreciable benefit and harm

Summary of findings 3. Oral contraceptive pill (OCP) compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS)

OCP compared to Metformin combined with OCP for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome

Patient or population:aAdult women with PCOS
Setting: Hospital or University Clinics
Intervention: OCP
Comparison: metformin combined with OCP

Outcomes		Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
		Risk with metformin combined with OCP	Risk with OCP				
Hirsutism - Clinical F-G score		The mean hirsutism - Clinical F-G score was 5.57	MD 0.54 higher (0.20 higher to 0.89 higher)	-	389 (6 RCTs)	⊕⊕⊖⊖ LOW [1,2]	
Adverse events - Severe	Gastro-intestinal	98 per 1000	21 per 1000 (6 to 73)	OR 0.20 (0.06 to 0.72)	228 (5 RCTs)	⊕⊕⊖⊖ LOW [1,3]	
	Others	39 per 1000	61 per 1000 (19 to 179)	OR 1.61 (0.49 to 5.37)	159 (4 RCTs)	⊕⊕⊖⊖ LOW [1,4]	
Adverse events - Minor	Gastro-intestinal	260 per 1000	21 per 1000 (4 to 134)	OR 0.06 (0.01 to 0.44)	98 (2 RCTs)	⊕⊕⊖⊖ LOW [5,6]	
	Others	No trials reported on outcome "Adverse events - Minor - Others"					
Improved menstrual pattern	Shortening of inter menstrual days	No trials reported on outcome "Improved menstrual pattern (i.e. shortening of inter menstrual days)"					
	An initiation of menses or cycle regularity	No trials reported on outcome "Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)"					

Acne - Clinical acne score	The mean acne - Clinical acne score was 0.54	MD 0.09 lower (0.10 lower to 0.08 lower)	-	82 (1 RCT)	⊕⊕⊖⊖ LOW [7]
BMI (kg/m^2)	The mean BMI (kg/m^2) was 28.6	MD 0.21 lower (0.53 lower to 0.12 higher)	-	661 (13 RCTs)	⊕⊖⊖⊖ VERY LOW [1,8,9]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

BMI: body mass index; **CI:** Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio; **RCT:** Randomised controlled trial.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[1] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have unclear risk of bias
[2] Evidence downgraded by one level for serious imprecision – low number of participants (total number of participants < 400) and 95% CI includes both appreciable effect and little or no effect
[3] Evidence downgraded by one level for serious imprecision - low number of events (total number of events < 300)
[4] Evidence downgraded by one level for serious imprecision – low number of events (total number of events < 300) and 95% CI includes both appreciable benefit and harm
[5] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have high risk of bias
[6] Evidence downgraded by one level for serious imprecision – low number of events (total number of events < 300)
[7]Evidence downgraded by one level for serious imprecision – very low number of participants (total number of participants < 400; i.e n = 82; single RCT) and/or 95% CI includes both appreciable effect and little or no effect
[8]Evidence downgraded by one level for serious inconsistency - substantial heterogeneity was detected which was not explained by the difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.64, I^2 = 0%)
[9] Evidence downgraded by one level for serious imprecision - 95% CI includes both appreciable effect and little or no effect

Summary of findings 4. Metformin compared to oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adolescent women with polycystic ovary syndrome (PCOS)

Metformin compared to OCP for hirsutism, acne, and menstrual pattern in adolescent women with PCOS

Patient or population: adolescent women with PCOS
Setting: Hospital or University Clinics
Intervention: metformin
Comparison: OCP

Outcomes	Anticipated absolute effects* (95% CI)	Relative effect	№ of participants	Quality of the evidence	Comments

		Risk with OCP	Risk with metformin	(95% CI)	(studies)	(GRADE)	
Hirsutism - Clinical F-G score		The mean hirsutism - Clinical F-G score was 8.6	MD 0.40 lower (3.42 lower to 2.62 higher)	-	16 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2]	
Adverse event - Severe	Gastro-intestinal	No trials reported on outcome "Adverse event - Severe - Gastro-intestinal"					
	Others	150 per 1 000	100 per 1 000 (27 to 300)	OR 0.63 (0.16 to 2.43)	80 (1 RCT)	⊕⊖⊖⊖ VERY LOW [3,4]	
Adverse event - Minor	Gastro-intestinal	0 per 1 000	3 per 1 000 (0 to 0)	OR 11.67 (0.53 to 258.56)	22 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,5]	There were only 3 events in the arm metformin and 0 in the arm OCP
	Others	No trials reported on outcome "Adverse event - Minor - Others"					
Improved menstrual pattern	Shortening of inter menstrual days	No trials reported on outcome "Improved menstrual pattern (i.e. shortening of inter menstrual days)"					
	An initiation of menses or cycle regularity	1 000 per 1 000	1000 per 1 000 (1 000 to 1 000)	OR 0.10 (0.01 to 1.92)	80 (1 RCT)	⊕⊖⊖⊖ VERY LOW [3,6]	40 out of 40 participants had improved menstrual patter in the OCP group compared to 36 out of 40 in the metformin group
Acne - Visual analogue scale or Clinical acne score		No trials reported either on outcome "Acne - Visual analogue scale" or "Acne - Clinical acne score"					
BMI (kg/m^2)		The mean BMI (kg/m^2) was 36	MD 1.45 lower (5.08 lower to 2.17 higher)	-	69 (3 RCTs)	⊕⊖⊖⊖ VERY LOW [7,8]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

BMI: body mass index; **CI:** Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio; **RCT:** Randomised controlled trial.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[1] Evidence downgraded by one level for serious risk of bias – a single RCT which has unclear risk of bias
[2] Evidence downgraded by two levels for very serious imprecision – very low number of participants (total number of participants < 400 i.e. n = 16 participants) and 95% CI includes both appreciable benefit and appreciable harm
[3] Evidence downgraded by one level for serious risk of bias – a single RCT which has high risk of bias
[4] Evidence downgraded by two levels for very serious imprecision – very low number of events (total number of events < 300 i.e. n = 10 events) and 95% CI includes both appreciable benefit and appreciable harm
[5] Evidence downgraded by two levels for very serious imprecision – very low number of events (total number of events < 300 i.e. n = 3 events) and 95% CI includes both appreciable benefit and appreciable harm
[6] Evidence downgraded by two levels for very serious imprecision – very low number of events (total number of events < 300 i.e. n = 76 events) and 95% CI includes both appreciable benefit and appreciable harm
[7] Evidence downgraded by one level for serious risk of bias - the majority of the RCTs have unclear risk of bias
[8] Evidence downgraded by two levels for very serious imprecision – very low number of participants (total number of participants < 400 i.e. n = 69 participants) and 95% CI includes both appreciable benefit and appreciable harm

Summary of findings 5. Metformin compared to metformin combined with oral contraceptive pill (OCP) for hirsutism, acne and menstrual pattern in adolescent women with polycystic ovary syndrome (PCOS)

Metformin compared to metformin combined with OCP for hirsutism, acne and menstrual pattern in adolescent women with PCOS

Patient or population: adolescent women with PCOS
Setting: Hospital or University Clinics
Intervention: metformin
Comparison: metformin combined with OCP

Outcomes		Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
		Risk with metformin combined with OCP	Risk with metformin				
Hirsutism - Clinical F-G score		No trials reported on outcome "Hirsutism - Clinical F-G score"					
Adverse event - Severe	Gastro-intestinal	No trials reported on outcome "Adverse event - Severe - Gastro-intestinal"					
	Others	No trials reported on outcome "Adverse event - Severe - Others"					

Adverse event - Minor	Gastro-intestinal	No trials reported on outcome "Adverse event - Minor - Gastro-intestinal"
	Others	No trials reported on outcome "Adverse event - Minor - Others"
Improved menstrual pattern	Shortening of inter menstrual day	No trials reported on outcome "Improved menstrual pattern (i.e. shortening of intermenstrual days)"
	An initiation of menses or cycle regularity	No trials reported on outcome "Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)"
Acne - Visual analogue scale or Clinical acne score		No trials reported either on outcome "Acne - Visual analogue scale" or "Acne - Clinical acne score"
BMI (kg/m^2)		No trials reported on outcome "BMI"

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

BMI: body mass index; **CI:** Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

Summary of findings 6. Oral contraceptive pill (OCP) compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adolescent women with polycystic ovary syndrome PCOS

OCP compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adolescent women with PCOS

Patient or population: adolescent women with PCOS
Setting: Hospital or University Clinics
Intervention: OCP
Comparison: metformin combined with OCP

Outcomes	Anticipated absolute effects* (95% CI)	Relative effect	№ of participants	Quality of the evidence	Comments

		Risk with metformin combined with OCP	Risk with OCP	(95% CI)	(studies)	(GRADE)
Hirsutism - Clinical F-G score		The mean hirsutism - Clinical F-G score was 6.2	MD 0.80 higher (1.19 lower to 2.79 higher)	-	32 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2]
Adverse events - Severe	Gastro-intestinal	56 per 1 000	56 per 1 000 (4 to 505)	OR 1.00 (0.06 to 17.33)	36 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,3]
	Others	No trials reported on outcome "Adverse events - Severe - Others"				
Adverse events - Minor	Gastro-intestinal	No trials reported on outcome "Adverse events - Minor - Gastro-intestinal"				
	Others	No trials reported on outcome "Adverse events - Minor - Others"				
Improved menstrual pattern	Shortening of intermenstrual days	No trials reported on outcome "Improved menstrual pattern (i.e. shortening of intermenstrual days)"				
	An initiation of menses or cycle regularity	No trials reported on outcome "Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)"				
Acne - Visual analogue scale or Clinical acne score		No trials reported either on outcome "Acne - Visual analogue scale" or "Acne - Clinical acne score"				
BMI (kg/m^2)		The mean BMI (kg/m^2) was 32.4	MD 1.5 higher (1.63 lower to 4.63 higher)	-	32 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

BMI: body mass index; **CI:** Confidence interval; **F-G:** Ferriman-Gallwey score; **MD:** Mean difference; **OR:** Odds ratio; **RCT:** Randomised controlled trial.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: we are very uncertain about the estimate.

[1] Evidence downgraded by one level for serious risk of bias – a single RCT which has unclear risk of bias
[2] Evidence downgraded by two levels for very serious imprecision – very low number of participants (total number of participants < 400 i.e. n = 32 participants) and 95% CI includes both appreciable benefit and appreciable harm
[3] Evidence downgraded by two levels for very serious imprecision – very low number of events (total number of events < 300 i.e. n = 2 events) and 95% CI includes both appreciable benefit and appreciable harm

BACKGROUND

Description of the condition

Polycystic ovary syndrome (PCOS) is characterised by chronic anovulation (failure or absence of ovulation) and hyperandrogenism (excessive production of male hormones in women) with clinical manifestations of irregular menstrual cycles (periods), infertility (failure to conceive), hirsutism (excessive hairiness) and acne (pimples). This condition is the most common endocrinopathy in women, affecting approximately 8% to 18% of women of reproductive age (Teede 2010a; Teede 2018). PCOS is a heterogenous condition, both clinically and biochemically. Women with PCOS are at increased risk of a number of metabolic disturbances, including gestational diabetes, impaired glucose tolerance (IGT), type 2 diabetes mellitus (T2DM) and metabolic syndrome. However, it remains unclear whether women with PCOS have a higher risk of cardiovascular disease (CVD) (Fauser 2012; Ovalle 2002; Teede 2010a; Teede 2018; Wild 2002a; Wild 2002b).

Recent international evidence-based guidelines for the assessment, diagnosis and management of polycystic ovary syndrome have been published (Teede 2018). The guidelines recommend that all clinicians and investigators now use an internationally agreed definition of PCOS according to Rotterdam criteria (Rotterdam ESHRE 2004). Therefore, the diagnosis of PCOS in adult women requires that at least two of the following three criteria are met: (1) oligo- or anovulation (infrequent or no ovulation); (2) clinical and/or biochemical signs of hyperandrogenism; (3) polycystic ovaries on ultrasound. However, ultrasound is not indicated for adolescent patients due to overlap with normal reproductive physiology. Other causes for hyperandrogenism which mimic PCOS (such as congenital adrenal hyperplasia, Cushing's syndrome, or androgen-secreting tumours) and amenorrhoea (such as thyroid disease or hyperprolactinaemia) should be excluded. These guidelines also recommend that standardised visual scales are preferred when assessing hirsutism, and one such scale is the Ferriman-Gallwey (F-G) score (Martin 2018; Teede 2018).

The exact pathophysiological mechanism (body characteristics) leading to the characteristic PCOS phenotype remains unclear. Some investigators explain it as primarily an intrinsic ovarian problem (excess ovarian production of androgens), others as adrenal (excess adrenal gland production of androgens), and again others as hypothalamic-pituitary dysfunction (exaggerated gonadotropin-releasing hormone pulsatility that results in hypersecretion of luteinising hormone). Insulin resistance (IR) (defined as a reduced glucose response to a given amount of insulin) seems to be one of the key pathophysiological feature of PCOS leading to both reproductive and metabolic disorders. Evidence of decreased insulin sensitivity is seen in both lean (30% incidence) and obese women (75% incidence) with PCOS; but IR accompanied by compensatory hyperinsulinaemia is most marked when there is an interaction between obesity and the syndrome (Conway 1990; Diamanti-Kandarakis 2012; Dunaif 1989; Dunaif 1994). Hyperinsulinaemia directly stimulates both ovarian and adrenal androgen secretion and suppresses liver sex hormone-binding globulin (SHBG) synthesis, resulting in an increase in free, biologically-active androgens. This excess in local ovarian androgen production, augmented by hyperinsulinaemia, causes premature follicular atresia (the breakdown of the ovarian follicles) and anovulation along with the other clinical manifestations of hyperandrogenism such as hirsutism and acne (Costello 2003; Utiger 1996).

Description of the intervention

Metformin, an insulin-sensitising drug (ISD), has been advocated as a long-term treatment given the importance of hyperinsulinaemia in the development of hyperandrogenism and disrupted folliculogenesis in PCOS. Metformin may be useful in the restoration of normal endocrinological and clinical parameters of PCOS by lowering insulin secretion (Hasegawa 1999). The most common side effects of metformin include gastrointestinal complaints such as nausea, diarrhoea, and abdominal cramping. These occur in up to 50% of treated patients, usually improving or completely subsiding with continued treatment (Hundal 2003).

Oral contraceptive pills (OCP) have been the traditional therapy for the long-term treatment of PCOS to regularise and lighten menses, improve hirsutism and acne by reducing ovarian androgen production and to provide endometrial protection. It has been advocated that the OCP may reduce insulin sensitivity, glucose tolerance in women with PCOS, and increase triglycerides in women with PCOS (Diamanti-Kandarakis 2003; Freitas de Medeiros 2017). The most common side effects in women taking the OCP include headache, mood changes, gastrointestinal disturbances, and breast pain (Gallo 2013).

Oral contraceptive pills have been demonstrated to be effective therapy for hirsutism and acne, whilst the evidence for such efficacy with metformin for these outcomes is less so and inconsistent (Buzney 2014; Martin 2018; Teede 2018).

How the intervention might work

Metformin, improves insulin sensitivity, reduces insulin and consequently androgen levels, and therefore could improve menstrual cyclicity, acne and hirsutism (Katsiki 2010). OCPs contain oestrogen and progestin components allowing a regularisation of the menstrual cycle. OCP therapy reduces hyperandrogenism via a number of mechanisms, including the following: suppression of luteinising hormone secretion (and therefore ovarian androgen secretion), stimulation of hepatic production of sex hormone binding globulin (thereby increasing androgen binding in serum and reducing serum free androgen concentrations), and a slight reduction in both adrenal androgen secretion and binding of androgens to their receptor. Consequently, there is a reduction of androgen production and action yielding an improvement in hirsutism and acne (Martin 2018).

Why it is important to do this review

The OCP has been the traditional therapy for the long-term treatment of women with PCOS not seeking fertility treatment in terms of the reproductive features of menstrual dysfunction and hyperandrogenic symptoms such as hirsutism and acne. However, metformin has more recently been proposed and used as an alternative therapy to the OCP for these reproductive manifestations with perhaps more favourable effect on the metabolic features of PCOS. It is therefore important to directly compare these two interventions in terms of both efficacy and adverse events to help guide clinical practice in the management of women with PCOS. Therefore, the overall aim of this review was to compare the efficacy and safety of metformin versus the OCP (alone

or in combination) in improving clinical, hormonal, and metabolic features of PCOS.

This review is a substantive update of a previous Cochrane Review (Costello 2007).

OBJECTIVES

To assess the effectiveness and safety of metformin versus the oral contraceptive pill (OCP) (alone or in combination) for the long-term treatment of women with polycystic ovary syndrome (PCOS).

METHODS

Criteria for considering studies for this review

Types of studies

We considered randomised controlled trials (RCTs) for inclusion in the review. Cross-over trials were not eligible for inclusion unless phase-one data (i.e. pre-cross-over) were available.

Types of participants

Women (adults or adolescents) with PCOS based on clinical (ovulatory dysfunction, hirsutism, acne, androgen dependent alopecia), biochemical (hyperandrogaenemia), or ultrasound (polycystic ovaries) evidence as defined by included studies.

Note was taken as to whether the participants of the included studies met the internationally agreed definitions of PCOS (ESHRE/ASRM 2004), which were endorsed as the diagnostic criteria for PCOS in the recently published international guidelines on PCOS (Teede 2018).

Note was also taken of whether any of the participants had diabetes mellitus or were taking any other medications which might alter insulin sensitivity.

The RCTs of adult women with PCOS were analysed separately to those involving adolescent women with PCOS.

Types of interventions

(a) Metformin versus the combined oral contraceptive pill (MET versus OCP).

(b) Metformin versus metformin in combination with the combined oral contraceptive pill (MET versus MET + the OCP).

(c) Combined oral contraceptive pill versus metformin in combination with the combined oral contraceptive pill (the OCP versus MET + the OCP).

Types of outcome measures

Outcomes measures were defined as primary (clinical) outcomes and secondary (clinical, hormonal, and metabolic) outcomes.

This review considered trials with a minimum length of follow-up of three months.

Primary outcomes

(a) Clinical parameters

1. Hirsutism as assessed clinically by a trained observer (using a scoring system such as the Ferriman and Gallwey (F-G) score or a Visual Analogue Scale (VAS)), or participant self-scoring of subjective improvement or not.
2. Adverse events (gastro-intestinal and other): severe (requiring stopping of medication), and minor.

Secondary outcomes

(a) Clinical parameters

3. Improved menstrual pattern (i.e. an initiation of menses or cycle regularity or significant shortening of intermenstrual days).

4. Acne as assessed clinically by a trained observer (using a scoring system such as a VAS, or participant self-scoring of subjective improvement or not.

5. Diagnosis of type 2 diabetes mellitus
6. Body weight (kg)
7. Body mass index (BMI) (kg/m^2)
8. Blood pressure systolic (mmHg)
9 Blood pressure diastolic (mmHg)

(b) Hormonal parameters

10. Serum total testosterone (nmol/L)
11. Free androgen index (FAI) (%)

(c) Metabolic parameters

12. Fasting insulin (mLU/L)
13. Fasting glucose (mmol/L)
14. Fasting total cholesterol (mmol/L)
15. Fasting high-density lipoprotein (HDL) cholesterol (mmol/L)
16. Fasting low-density lipoprotein (LDL) cholesterol (mmol/L)
17. Fasting triglycerides (mmol/L)

Search methods for identification of studies

We searched for all published and unpublished RCTs on the use of metformin and the OCP (alone or in combination) in women with PCOS, without language restriction and in consultation with the Gynaecology and Fertility Group Information Specialist.

Electronic searches

We searched the following electronic databases for relevant trials:

(1) The Cochrane Gynaecology and Fertility Group (CGF) Specialised Register of Controlled Trials; PROCITE platform (searched 15 August 2019) (Appendix 1)

(2) The Cochrane Central Register of Controlled Trials; Ovid platform (searched 15 August 2019, Issue July 2019) (Appendix 2)

(3) MEDLINE (Epub Ahead of Print, In-Process & Other Non-Indexed Citations); Ovid platform (searched from 1946 to 15 August 2019) (Appendix 3)

(4) Embase; Ovid platform (searched from 1980 to 15 August 2019) (Appendix 4)

(5) PsycINFO; Ovid platform (searched from 1806 to 15 August 2019) (Appendix 5)

(6) CINAHL (Cumulative Index to Nursing and Allied Health Literature); EBSCO platform (searched from 1961 to 15 August 2019) (Appendix 6)

Searching other resources

(7) Reference lists of included studies, other relevant review articles and textbooks were handsearched.

(8) Trial registers for ongoing and registered trials were checked and authors contacted if required.

(9) We contacted experts in the field to identify additional studies.

Data collection and analysis

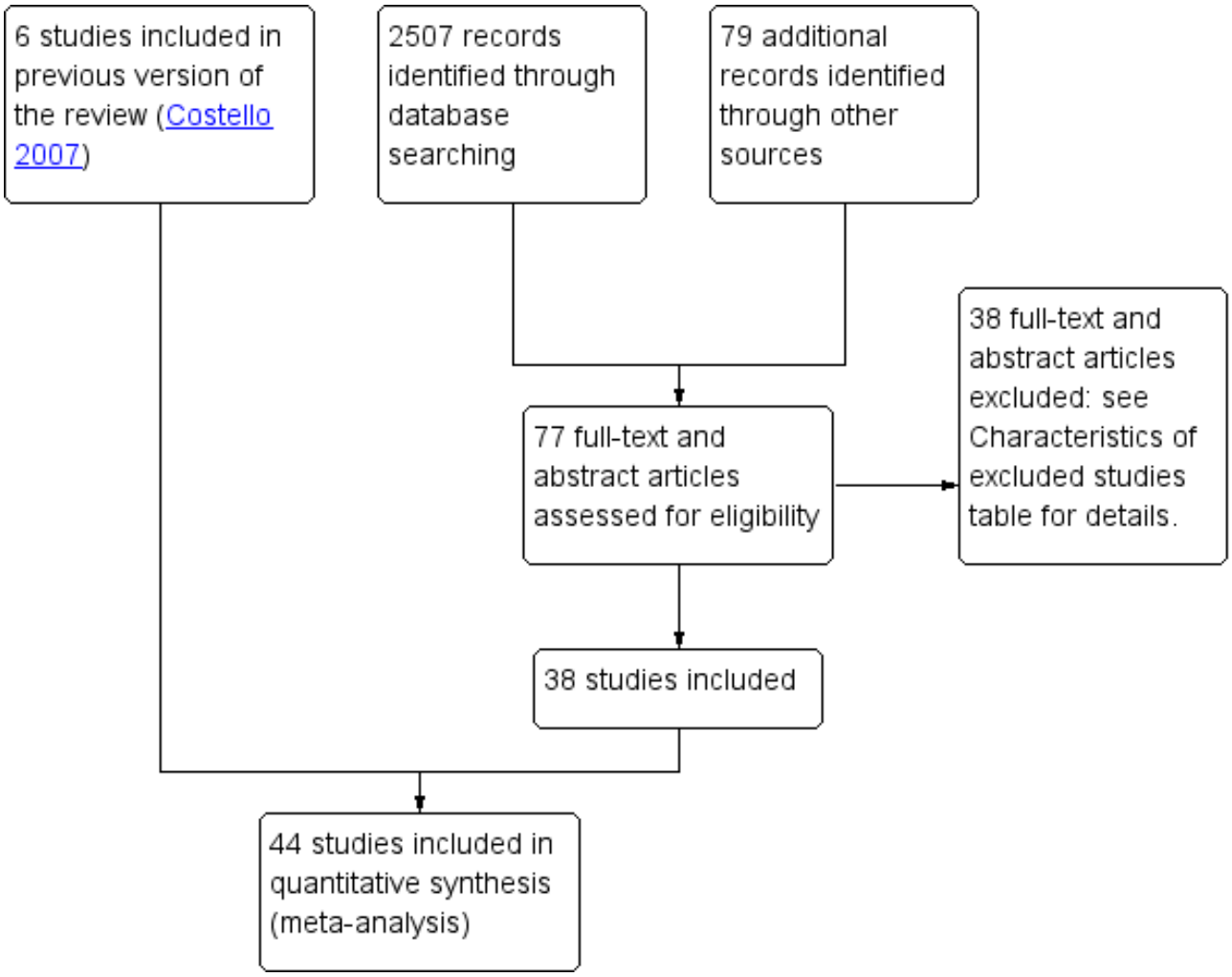
We conducted data collection and analysis in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2017).

Selection of studies

After an initial screen of titles and abstracts retrieved by the search, conducted by review authors EF, LM, CV, and MC, the full texts of all potentially eligible studies were retrieved. Three review authors (LM, EF and MC) independently examined these full-text articles for compliance with the inclusion criteria and selected eligible studies. If papers contained insufficient information to make a decision about eligibility, we contacted the authors of those papers in order to seek further information to clarify study eligibility. Disagreements were resolved by discussion.

Studies from non-English language journals were translated if necessary. The selection process is documented in the PRISMA flow chart (Figure 1). We provide a list of excluded studies, showing the reasons for exclusion in the Characteristics of excluded studies table.

Figure 1. Study flow diagram.



Data extraction and management

Four review authors (EF, LM, CE and SB) independently extracted data from eligible studies using a data extraction form designed and pilot-tested by the authors. Any disagreements were resolved by discussion. Data extracted included study characteristics and outcome data. Unit conversion factors are shown in Table 1. Where studies had multiple publications, we collated multiple reports of the same under a single study ID with multiple references. We corresponded with study investigators for further data on methods

or results, or both, when required. In multiple-arm studies; data from arms that did not meet the inclusion criteria were not used.

Assessment of risk of bias in included studies

Four review authors (EF, LM, CE and MC) independently assessed the included studies for risk of bias using the Cochrane 'Risk of bias' assessment tool (Higgins 2017). We assessed selection (random sequence generation and allocation concealment); performance (blinding of participants and personnel); detection (blinding of outcome assessors); attrition (incomplete outcome data); reporting (selective reporting) bias; and other bias (other potential bias), and summarised our judgements in the 'Risk of bias' tables, Figure 2 and Figure 3. Judgements were assigned as low, high or unclear risk using the criteria from the *Cochrane Handbook Table* 8.5.d: 'Criteria for judging risk bias' in the 'Risk of bias' assessment tool (Higgins 2017). Disagreements were resolved by discussion. All judgements were fully described and the conclusions are presented in the 'Risk of Bias' table, and incorporated into the interpretation of review findings by means of sensitivity analyses where indicated.

Figure 2. 'Risk of bias' graph: review authors' judgements about each 'Risk of bias' item presented as percentages across all included studies

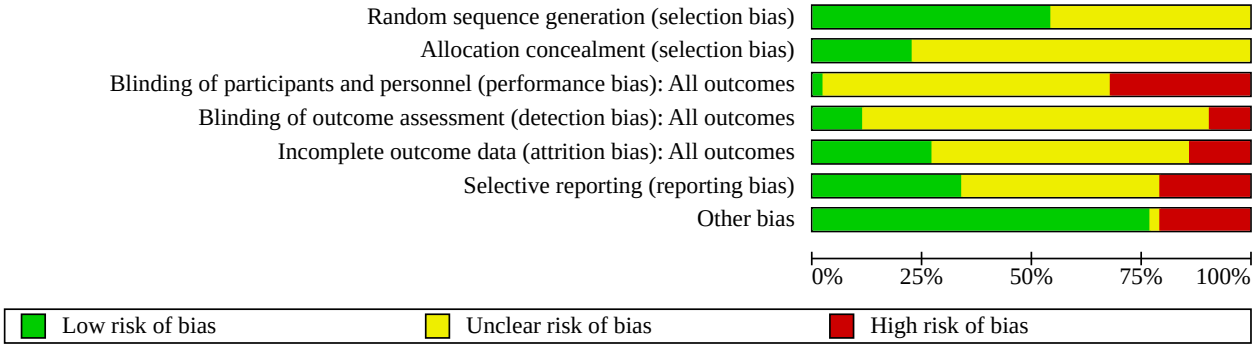


Figure 3. 'Risk of bias' summary: review authors' judgements about each 'Risk of bias' item for each included study .

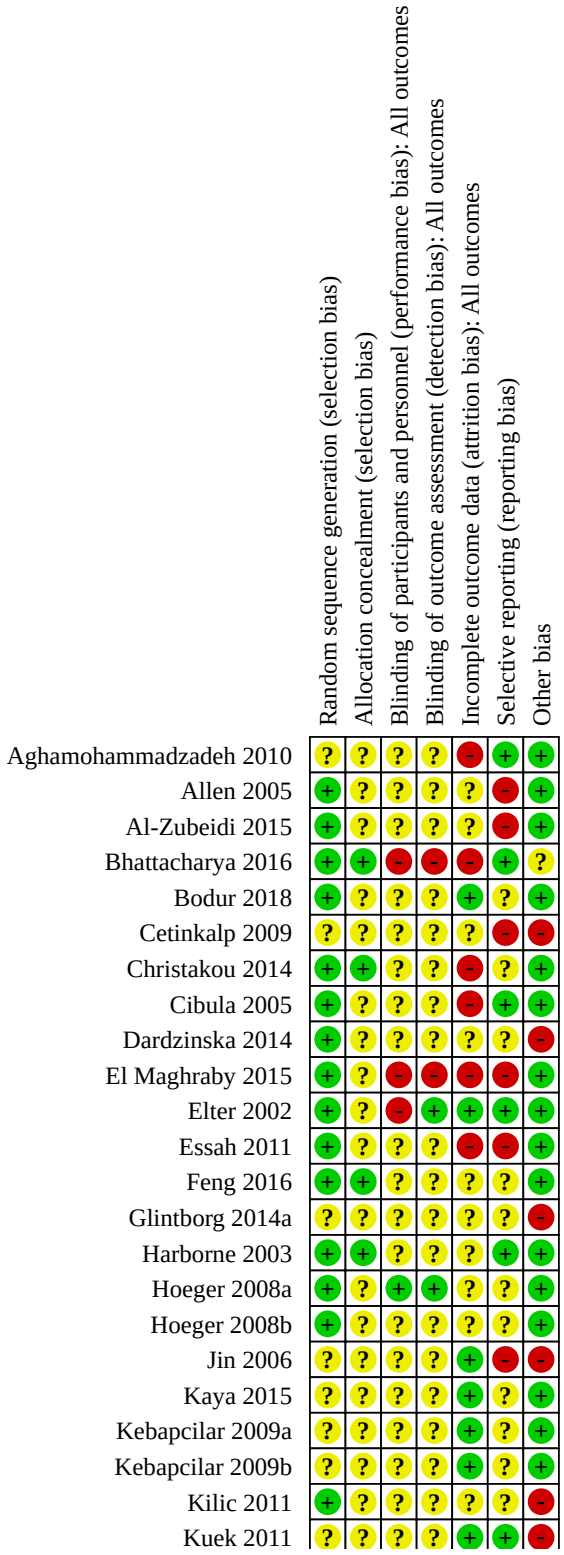
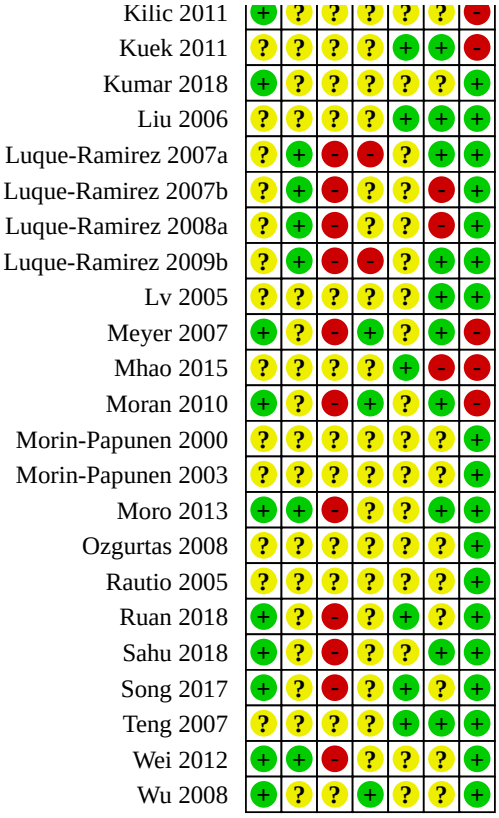


Figure 3. (Continued)



Measures of treatment effect

For dichotomous data, results for each study were expressed as odds ratios (ORs) with 95% confidence intervals (CIs) and combined for meta-analysis with Review Manager software using the fixed-effect model (Peto method). The goal was to calculate a pooled estimate of treatment effect for each outcome across studies.

For continuous data, we measured the mean post-treatment or intervention values and standard deviations for each group and calculated the weighted mean differences (MDs) with 95% CIs. If different scales measured the same continuous data outcome, we planned to measure the mean post-treatment or intervention values and standard deviations for each group and to calculate the standardised mean difference (SMD) with 95% CI.

Unit of analysis issues

We took into account the level at which the randomisation occurred in each trials.

We considered whether in each study:

1. groups of individuals were randomised together to the same intervention (i.e. cluster- randomised trials);
2. or each individual was individually randomised to one of the intervention groups;
3. or if individuals underwent more than one intervention (e.g. in cross-over trial). Only first-phase data from cross-over trials were included.

The analysis was by woman randomised.

Dealing with missing data

If data were missing from included studies, we contacted the investigators to request the relevant missing data.

If this was not possible, we analysed the data on an intention-to-treat basis as far as possible (i.e. including all randomised participants in analysis, in the groups to which they were randomised). Otherwise, available data were analysed**.**

If studies reported sufficient detail to calculate mean differences (MD), but no information on associated standard deviation, we assumed the outcome to have a standard deviation equal to the highest standard deviation from other studies within the same analysis.

Assessment of heterogeneity

Heterogeneity reflects any type of variability among the studies in a systematic review. The clinical and methodological characteristics of the included studies were considered in order to check if they were sufficiently similar for meta-analysis to provide a clinically-meaningful summary. A consistent treatment effect among the included studies suggests there is sufficient homogeneity for pooled analysis. Heterogeneity (inconsistency) between the results of different studies was examined by inspecting the scatter in the data points on the graph and the overlap in their confidence intervals on the forest plot and, more formally, by checking the

results of the Chi^2 tests and the measure of the I^2 statistic (Higgins 2003). An I^2 statistic greater than 50% was taken to indicate substantial heterogeneity (Deeks 2017).

Substantial heterogeneity for the review outcomes was explored (investigated) with subgroup and sensitivity analyses by consideration of factors such as study quality, differences in population, interventions and outcomes (see below section 'Subgroup analysis and investigation of heterogeneity'.

Assessment of reporting biases

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, we aimed to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert for duplication of data. If there were 10 or more studies in an analysis, we used a funnel plot for the main review outcomes to explore the possibility of small-study effects (a tendency for estimates of the intervention effect to be more beneficial in smaller studies).

Data synthesis

Studies were sufficiently similar in order to combine the data using a fixed-effect model in the following comparisons.

(a) Metformin versus the combined oral contraceptive pill (MET versus the OCP).

(b) Metformin versus the combined oral contraceptive pill in combination with metformin (MET versus the OCP + MET).

(c) Combined oral contraceptive pill versus the combined oral contraceptive pill in combination with metformin (the OCP versus the OCP + MET).

Statistical analysis was performed using Review manager 5.3 in accordance with the guidelines for statistical analysis developed by Cochrane (Review Manager 2014).

For all outcomes, where data were available, we stratified comparisons by subgroups of studies of women with different mean body mass index (BMI) (e.g. BMI ≤ 25 kg/m^2, BMI > 25 kg/m^2 but < 30 kg/m^2, BMI ≥ 30 kg/m^2) with an additional stratum for studies in which BMI was not reported, apart from the outcome of adverse events (which were divided into gastro-intestinal and other adverse events).

Subgroup analysis and investigation of heterogeneity

Where data were available, we conducted pre-specified subgroup analyses to determine the separate evidence within the following subgroups for all outcomes.

All outcomes (apart from the outcome of adverse events) were divided or subgrouped according to studies of women with different mean BMI (e.g. BMI ≤ 25 kg/m^2, BMI > 25 kg/m^2 but < 30 kg/m^2, BMI ≥ 30 kg/m^2) in order to assess any differences in intervention effect between these subgroups and to assess whether any substantial heterogeneity, if detected, could be explained by such subgroup analysis according to BMI (i.e. whether BMI is an important effect modifier). The rationale for pre-specifying BMI for subgroup analysis was that metformin is an insulin-sensitising agent and insulin resistance in PCOS is exacerbated by obesity (see Background section), and therefore it is clinically plausible for metformin to possibly have a larger (or different) relative effect with increasing BMI.

The assessment of whether there was a statistically significant difference between the subgroups was performed by comparing the different subgroups directly with each other using the "statistical test for subgroup differences" (I^2 statistic > 50 and/or P value < 0.05) in the forest plot graph.

If substantial heterogeneity could be explained by such pre-specified subgroup analyses (i.e. explained substantial heterogeneity), we analysed/reported the subgroups separately (and did not analyse/report the pooled results).

If substantial heterogeneity was detected, which could not be explained by such pre-specified subgroup analyses (i.e. unexplained substantial heterogeneity), we conducted sensitivity analyses to determine whether the conclusions are robust to arbitrary decisions made regarding the eligibility and analysis. These analyses included consideration of whether the review conclusions would have differed if (i) eligibility was restricted to studies at low risk of bias (defined as low risk of selection bias (both random sequence generation and allocation concealment) and not at high risk of bias in any domain) for the main review outcomes and if (ii) a random-effects model had been adopted for all outcomes. If the random-effects model and the fixed-effect model produced substantially different pooled estimates, then this is an excellent indication of heterogeneity and the random-effects model is the preferred model and as such was used. If the two models yielded similar pooled estimates then the fixed-effect model is preferred and as such was used, because usually it will have a narrower confidence interval; that is, it is more precise than the random-effects model (Ryan 2016). Unexplained substantial heterogeneity was taken into account when assessing the quality of the evidence using GRADE in terms of inconsistency and interpreting the results, especially when there was variation in the direction of effect.

Sensitivity analysis

We conducted sensitivity analyses to determine whether the conclusions are robust to arbitrary decisions made regarding the eligibility and analysis. These analyses included consideration of whether the review conclusions would have differed if eligibility was restricted to studies at low risk of bias (defined as low risk of selection bias (both random sequence generation and allocation concealment) and not at high risk of bias in any domain) for the main review outcomes and if a random-effects model had been adopted in the presence of unexplained substantial heterogeneity for all outcomes.

Overall quality of the body of the evidence: 'Summary of findings' table

We generated 'Summary of findings' tables using GRADEpro software and Cochrane methods (GRADEpro GDT). This tables evaluated the overall quality of the body of the evidence for the main review outcomes (hirsutism, improvement in menstrual pattern, acne assessed clinically (as opposed to subjectively) by either visual analogue scale (VAS) or clinical acne score, BMI and adverse events (severe and minor) (subgrouped according to type of adverse event: gastro-intestinal or other)) for the main review comparisons (metformin versus the OCP, metformin versus metformin in combination with the OCP, and the OCP versus

metformin in combination with the OCP). We assessed the quality of the evidence using GRADE criteria: risk of bias, inconsistency, imprecision, indirectness and publication bias). Judgements about evidence quality (high, moderate, low or very low) were made by three review authors (EF, EK and MC) working independently, with disagreements resolved by discussion. Judgements were justified, documented and incorporated into reporting of results for each main review outcome. Unexplained substantial heterogeneity was taken into account when assessing the quality of the evidence using GRADE in terms of inconsistency and interpreting the results, especially when there was variation in the direction of effect.

RESULTS

Description of studies

Results of the search

The previous version of this review included six trials (Costello 2007). The search for the current review update resulted in the retrieval of 77 full-text papers and abstracts (Figure 1). We included 38 new studies, 37 full-text papers and one abstract (Characteristics of included studies). Three included studies from non-English language journals required translation (Jin 2006; Liu 2006;Teng 2007). We excluded 38 new studies (Characteristics of excluded studies). Four new studies are awaiting classification (Characteristics of studies awaiting classification); we have contacted the authors and still await a response. We one study from awaiting classification to included studies (Sahu 2018), and another study from awaiting classification to excluded studies (NCT02866786). We classified three new studies as ongoing (NCT02744131; NCT03229057;NCT03905941), (Characteristics of ongoing studies).

Included studies

Study design and setting

In this 2020 updated review, we included 44 randomised controlled trials (RCTs) consisting of 43 parallel-designed RCTs and 1 cross-over RCT (in which only pre-cross over data were used). The updated review includes 37 full-text articles, one abstract and six full-text references from the former review.

The studies were performed in different locations around the world.

- Australia (Meyer 2007; Moran 2010)
- Czech Republic (Cibula 2005)
- China (Feng 2016; Jin 2006; Kuek 2011; Liu 2006; Lv 2005; Song 2017; Ruan 2018; Teng 2007; Wei 2012; Wu 2008)
- Denmark (Glintborg 2014a)
- Egypt (El Maghraby 2015)
- Finland (Morin-Papunen 2000; Morin-Papunen 2003; Rautio 2005)
- Greece (Christakou 2014)
- Italy (Moro 2013)
- India (Bhattacharya 2016 (abstract); Kumar 2018; Sahu 2018)
- Iran (Aghamohammadzadeh 2010)
- Iraq (Mhao 2015)
- Poland (Dardzinska 2014)
- Scotland (Harborne 2003)
- Spain (Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b)
- Turkey (Bodur 2018; Cetinkalp 2009; Elter 2002; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kilic 2011; Ozgurtas 2008)
- USA (Allen 2005; Al-Zubeidi 2015; Essah 2011; Hoeger 2008a; Hoeger 2008b).

Participants

The studies included 2253 women with polycystic ovary syndrome (PCOS).

- 43/44 studies fulfilled the Rotterdam PCOS criteria (Bhattacharya 2016; Bodur 2018; Cetinkalp 2009; Dardzinska 2014; El Maghraby 2015; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Harborne 2003; Jin 2006; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kilic 2011; Kuek 2011; Kumar 2018; Lv 2005; Mhao 2015; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Rautio 2005; Song 2017; Sahu 2018; Ruan 2018; Teng 2007; Wei 2012; Wu 2008); including 12/44 studies which fulfilled the stricter National Institutes of Health (NIH) PCOS criteria (Aghamohammadzadeh 2010; Allen 2005; Al-Zubeidi 2015; Christakou 2014; Cibula 2005; Hoeger 2008a; Hoeger 2008b; Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b; Meyer 2007); and including 1/44 studies fulfilled either NIH or Rotterdam PCOS criteria (Moran 2010).
- 1/44 studies did not described any diagnostic criteria (Liu 2006).

The mean age of the women ranged across studies from 12 to 40 years.

- 5/44 studies (206 women analysed) recruited adolescent women (Allen 2005; Al-Zubeidi 2015; El Maghraby 2015; Hoeger 2008a; Hoeger 2008b).
- 39/44 studies (2047 women analysed) recruited adult women (Aghamohammadzadeh 2010; Bhattacharya 2016; Bodur 2018; Cetinkalp 2009; Christakou 2014; Cibula 2005; Dardzinska 2014; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Harborne 2003; Jin 2006; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kilic 2011; Kuek 2011; Kumar 2018; Liu 2006; Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b; Lv 2005; Meyer 2007; Mhao 2015; Moran 2010; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Rautio 2005; Song 2017; Sahu 2018; Ruan 2018; Teng 2007; Wei 2012; Wu 2008).

Interventions

- 26/44 studies compared metformin (MET) versus the combined oral contraceptive (OCP) (Aghamohammadzadeh 2010; Allen 2005; Al-Zubeidi 2015; Cetinkalp 2009; Christakou 2014; Cibula 2005; Dardzinska 2014; El Maghraby 2015; Harborne 2003; Hoeger 2008a; Jin 2006; Kilic 2011; Kuek 2011; Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b; Meyer 2007; Mhao 2015; Moran 2010; Morin-Papunen 2000; Morin-Papunen 2003; Ozgurtas 2008; Rautio 2005; Sahu 2018; Teng 2007).
- 11/44 studies compared MET versus MET + the OCP (Bhattacharya 2016; Elter 2002; Essah 2011; Feng 2016; Hoeger 2008b; Kaya 2015; Kebapcilar 2009b; Lv 2005; Song 2017; Ruan 2018; Wei 2012).

- 7/44 studies compared MET versus the OCP versus MET + the OCP (Bodur 2018; Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Moro 2013; Wu 2008).

(See Characteristics of included studies table).

Outcomes

- 18/44 studies reported hirsutism by woman randomised (Bhattacharya 2016; Cetinkalp 2009; Dardzinska 2014; El Maghraby 2015; Elter 2002; Feng 2016; Glintborg 2014a; Harborne 2003; Hoeger 2008a; Hoeger 2008b; Jin 2006; Kumar 2018; Luque-Ramirez 2007a; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Sahu 2018; Wu 2008).
- 17/44 studies reported adverse events (severe or minor) by woman randomised (Aghamohammadzadeh 2010; Bodur 2018; Christakou 2014; Cibula 2005; Dardzinska 2014; El Maghraby 2015; Elter 2002; Essah 2011; Glintborg 2014a; Harborne 2003; Hoeger 2008b; Kilic 2011; Luque-Ramirez 2007a; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Wei 2012).
- 7/44 studies reported improved menstrual pattern by woman randomised (El Maghraby 2015; Jin 2006; Luque-Ramirez 2007a; Mhao 2015; Morin-Papunen 2000; Morin-Papunen 2003; Sahu 2018).
- 6/44 studies reported acne (Bhattacharya 2016; Cetinkalp 2009; Feng 2016; Harborne 2003; Jin 2006; Mhao 2015).
- 39/44 studies reported body mass index (BMI) by woman randomised (Aghamohammadzadeh 2010; Allen 2005; Al-Zubeidi 2015; Bhattacharya 2016; Bodur 2018; Cetinkalp 2009; Christakou 2014; Cibula 2005; Dardzinska 2014; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Harborne 2003; Hoeger 2008a; Hoeger 2008b; Jin 2006; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kilic 2011; Kuek 2011; Kumar 2018; Liu 2006; Luque-Ramirez 2007a; Lv 2005; Meyer 2007; Mhao 2015; Moran 2010; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Song 2017; Sahu 2018; Ruan 2018; Teng 2007; Wei 2012; Wu 2008).

Excluded studies

We excluded 38 trials from this review, 36 full texts (Alpanes 2017; Altinok 2018; Bachani 2016; Bhattacharya 2012; Bredella 2013; Burchall 2015; Cakiroglu 2013; Diaz 2016; Glintborg 2014b; Glintborg 2015; Glintborg 2017; Hadziomerovic-Pekic 2010; Harris-Glocker 2009; Hu 2010; Hutchison 2008; Ibanez 2010; Ibanez 2017; Kebapcilar 2010; Kim 2010; Ladson 2011; Lazaro 2011; Lemay 2006; Luque-Ramirez 2008c; Luque-Ramirez 2009; Luque-Ramirez 2010a; Luque-Ramirez 2011; Mehrabian 2016; Mitkov 2005; NCT02866786; Orbetzova 2011; Panidis 2011; Pedersen 2018; Romualdi 2010; Suvarna 2016; Teede 2010b; Wang 2016) and two abstract articles (Moghtadaei 2009; Moretti 2016). The primary reasons for exclusion of the studies were no randomisation, other interventions, or other outcome of interest .

(See Characteristics of excluded studies table).

Risk of bias in included studies

Allocation

Twenty-four studies were at low risk of selection bias in relation to random sequence generation. These 24 studies used a computer random number generator or a random number table. In the remaining 20 studies, insufficient information about the random sequence generation was given and therefore were rated at unclear risk of bias for this domain (Figure 3).

Ten studies were at low risk of selection bias in relation to allocation concealment. These 10 studies used central allocation, sequentially numbered drug containers of identical appearance or sequentially numbered, opaque, sealed envelopes. The other 34 studies did not describe allocation concealment sufficiently and therefore were rated at unclear risk of bias for this domain (Figure 3).

Blinding

In relation to the blinding of all outcomes:

- participants (performance bias): three out of 44 studies described the blinding of participants and were thus rated at low risk of performance bias. Thirty-one out of 44 studies did not report on blinding of participants and were rated at unclear risk of performance bias. Ten out of 44 studies described no blinding for participants and were rated at high risk of performance bias (Figure 3)
- personnel (performance bias): one out of 44 studies described the blinding of personnel and were thus rated at low risk of performance bias. Thirty-two out of 44 studies did not mention blinding of personnel and were rated at unclear risk of performance bias. Eleven out of 44 studies described no blinding for personnel and were rated at high risk of performance bias (Figure 3)
- outcome assessor (detection bias): Five out of 44 studies described the blinding of outcome assessor and were thus rated at low risk of detection bias. Thirty-five out of 44 studies did not mention blinding of outcome assessor and were rated at unclear risk of detection bias. Four out of 44 studies described no blinding for outcome assessor and were rated at high risk of detection bias (Figure 3)

Incomplete outcome data

Thirteen out of 44 studies had no missing outcome data and were rated at low risk of attrition bias.

Twenty six out of 44 studies were rated at unclear risk of attrition bias. In 20 of these studies the reasons for missing data were judged as unclear: no show, unreachable, declined treatment, regrets, loss of follow up, personal reason, not stated, discontinued treatment, protocol violation, voluntary drop off, wanted the other intervention, and incomplete data (Allen 2005; Al-Zubeidi 2015; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Hoeger 2008a; Hoeger 2008b; Kilic 2011; Kumar 2018; Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b; Meyer 2007; Moran 2010; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Sahu 2018; Wei 2012). In one study the number or patient randomised was unclear (states n = 100, n = 99, n=94) (Cetinkalp 2009). In two studies no information was provided after randomisation and the authors were contacted without success (Feng 2016; Lv 2005). In two studies the number of dropouts were considered quite important (50% for MET group and 37.5% for the OCP group), and it was unclear whether this could have had a clinically relevant impact on the intervention effect estimate (Morin-Papunen 2000; Rautio 2005). In one study, there was a discrepancy between the figure and the text regarding number and reason for withdrawal (Wu 2008).

Five out of 44 studies were rated at high risk of attrition bias. In four of these studies, the reason for missing data were imbalanced and related to intervention (Aghamohammadzadeh 2010; Christakou 2014; Cibula 2005; Essah 2011). In one of these studies, the manner in which the use of imputation was performed was not appropriate (Bhattacharya 2016) (Figure 3).

Selective reporting

Fifteen out of 44 studies reported the outcomes that were stated in the methods section and thus were judged as low risk of reporting bias.

In 20 out of 44 studies, insufficient information was available to permit a judgement of 'low risk' or 'high risk' and therefore we rated them at unclear risk of reporting bias. In 19 of these studies, some secondary outcomes (BMI, BP, fasting glucose, weight, menstrual pattern, lipid profile, testosterone level, acne) were described in the methods section but not in the results and the study protocol was not available (Bodur 2018; Cetinkalp 2009; Dardzinska 2014; Feng 2016; Glintborg 2014a; Hoeger 2008a; Hoeger 2008b; Kebapcilar 2009b; Kilic 2011; Kumar 2018; Mhao 2015; Morin-Papunen 2000; Morin-Papunen 2003; Ozgurtas 2008; Rautio 2005; Song 2017; Ruan 2018; Wei 2012; Wu 2008). In two of these studies, the outcome measurements were not described in the methods section and the protocol was not available (Kaya 2015; Kebapcilar 2009b).

Nine out of 44 studies were rated at high risk of reporting bias. In four of these studies, hirsutism (which was a primary outcome) was described in the methods section but not reported in the results (Allen 2005; Al-Zubeidi 2015; El Maghraby 2015; Jin 2006). In the other three of these studies, some of the primary outcomes were reported in the results without being prespecified in the methods section (Figure 3).

Other potential sources of bias

In six studies there were substantial baseline imbalances between the two intervention groups and thus we deemed the risk of other bias to be high (Cetinkalp 2009; Dardzinska 2014; Jin 2006; Kilic 2011; Kuek 2011; Mhao 2015). In one study the use of laser and waxing was considered a source of bias in terms of the evaluation of hirsutism (Glintborg 2014a), and therefore rated at high risk of other bias. In two studies there were substantial baseline differences in clinical and biochemical hyperandrogenism between the two intervention groups and we judged them at high risk of other bias (Meyer 2007; Moran 2010). We found no potential sources of other bias in the other 35 studies (Figure 3).

Effects of interventions

See: **Summary of findings 1** Metformin compared to oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS); **Summary of findings 2** Metformin compared to metformin combined with oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS); **Summary of findings 3** Oral contraceptive pill (OCP) compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adult women with polycystic ovary syndrome (PCOS); **Summary of findings 4** Metformin compared to oral contraceptive pill (OCP) for hirsutism, acne, and menstrual pattern in adolescent women with polycystic ovary syndrome (PCOS); **Summary of findings 5** Metformin compared to metformin combined with oral contraceptive pill (OCP) for hirsutism, acne and menstrual pattern in adolescent women with polycystic ovary syndrome (PCOS); **Summary of findings 6** Oral contraceptive pill (OCP) compared to metformin combined with OCP for hirsutism, acne, and menstrual pattern in adolescent women with polycystic ovary syndrome PCOS

1. Metformin compared to the combined oral contraceptive pill (OCP) in adult women (clinical parameters)

Summary of findings 1

Twenty-eight randomised controlled trials (RCTs) including 1403 adult women compared metformin with the OCP (Aghamohammadzadeh 2010; Bodur 2018; Cetinkalp 2009; Christakou 2014; Cibula 2005; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Jin 2006; Kebapcilar 2009a; Kilic 2011; Kuek 2011; Kumar 2018; Luque-Ramirez 2007a; Luque-Ramirez 2007b; Luque-Ramirez 2008a; Luque-Ramirez 2009b; Meyer 2007; Mhao 2015; Moran 2010; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Rautio 2005; Sahu 2018; Teng 2007, Wu 2008).

Primary outcomes

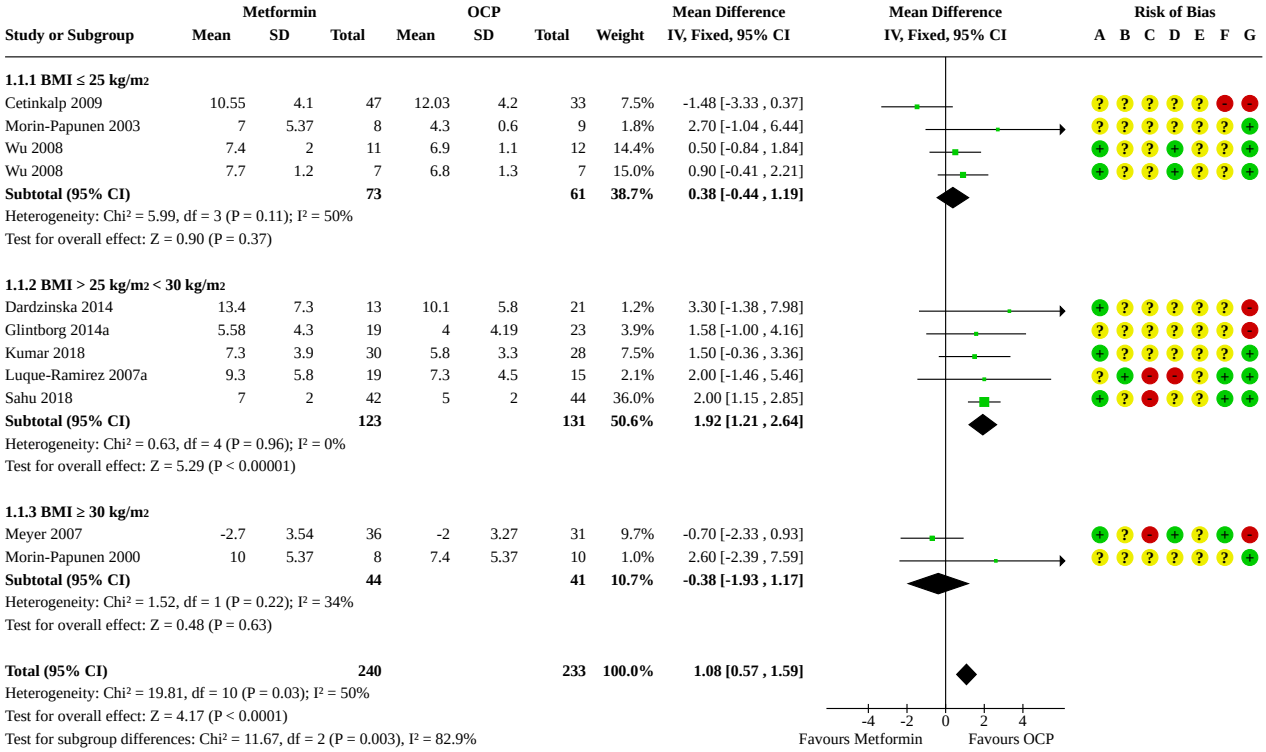
1.1 Hirsutism - Clinical Ferriman-Gallwey (F-G) score

Ten trials including 473 women compared metformin versus the OCP and reported hirsutism clinically using the Ferriman-Gallwey (F-G) score (Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Kumar 2018; Luque-Ramirez 2007a; Sahu 2018; Wu 2008).

Substantial heterogeneity was detected, which may be explained by the difference in effect of the interventions between the mean study body mass index (BMI) subgroups (test for subgroup difference: P = 0.003, I^2 = 82.9%). Therefore, we analysed the results per mean BMI subgroup.

We are uncertain if there was a difference between metformin and the OCP on F-G score in subgroup BMI < 25 kg/m^2 (mean difference (MD) 0.38, 95% confidence interval (CI) -0.44 to 1.19; 3 RCTs; n = 134; I^2 = 50%; very low-quality evidence, Figure 4, Analysis 1.1). This suggests that for a mean F-G score of 7.5 following the OCP, the mean F-G score following metformin would be between 0.44 lower to 1.19 higher.

Figure 4. Forest plot of comparison: 1 Adult - metformin versus OCP (Clinical parameters), outcome: 1.1 Hirsutism - Clinical F-G score.



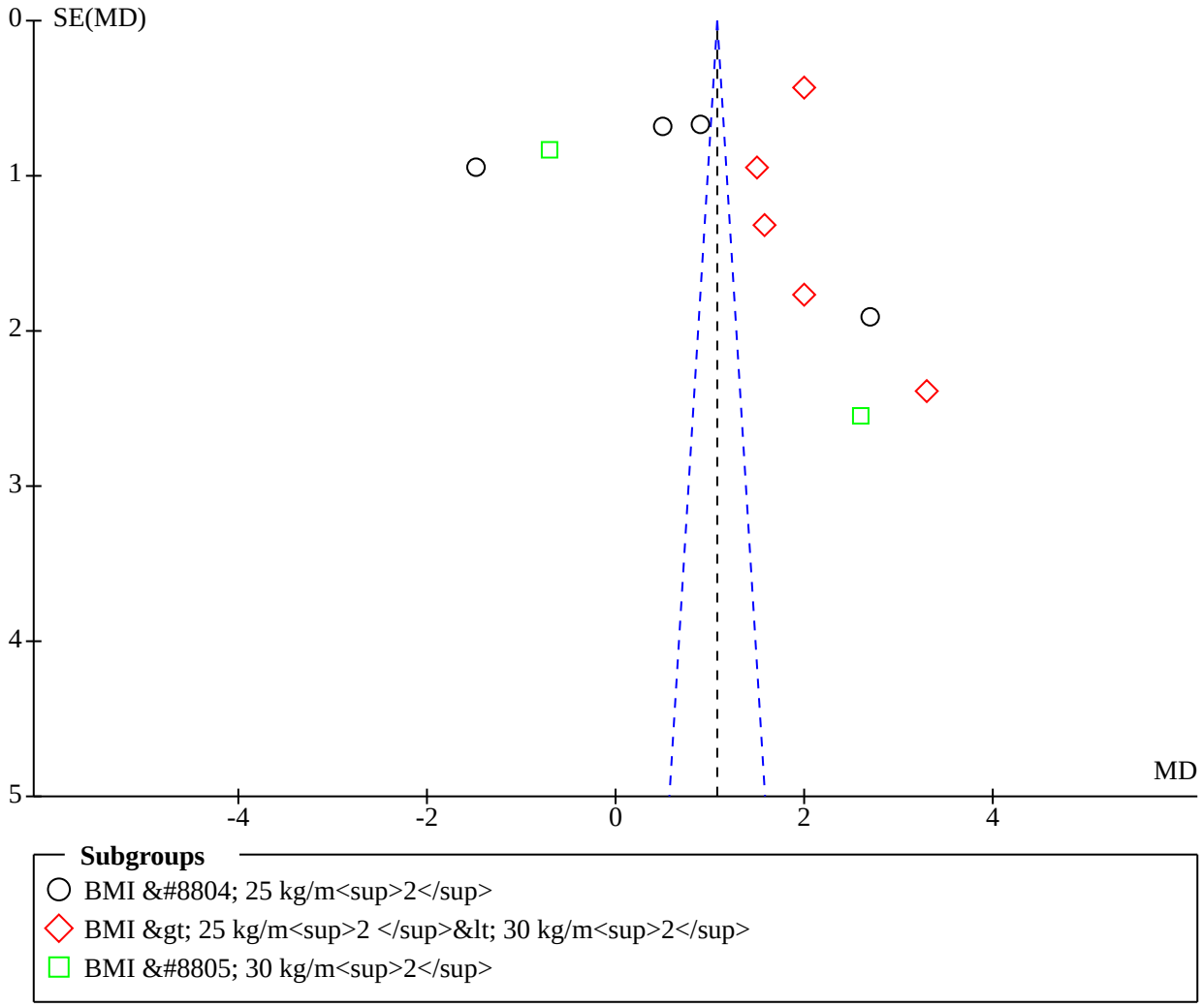
Metformin may be less effective in improving F-G score compared to the OCP in the subgroup BMI > 25 /kg2 < BMI 30 kg/m^2 (MD 1.92, 95% CI 1.21 to 2.64; 5 RCTs; n = 254; I^2 = 0%; low-quality evidence, Figure 4, Analysis 1.1). This suggests that for a mean F-G score of 6.44 following the OCP, the mean F-G score following metformin would be between 1.21 higher to 2.64 higher.

We are uncertain if there was a difference between metformin and the OCP on F-G score in subgroup BMI > 30 kg/m^2 (MD -0.38, 95% CI -1.93 to 1.17; 2 RCTs; n = 85; I^2 = 34%; low-quality evidence, Figure 4, Analysis 1.1). This suggests that for a mean F-G score of 6.05 following the OCP, the mean F-G score following metformin would be between 1.93 lower to 1.17 higher.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

The funnel plot (n = 10 studies) was asymmetrical indicating that our findings might be influenced by publication bias although asymmetrical funnel plots can also be due to true heterogeneity (Figure 5, Analysis 1.1).

Figure 5. Funnel plot of comparison: 1 Adult - metformin versus OCP (Clinical parameters), outcome: 1.1 Hirsutism - Clinical F-G score.



1.2 Hirsutism - Subjective visual analogue scale (VAS)

One trial including 34 women compared metformin versus the OCP and reported hirsutism subjectively (patient self-assessed) using a VAS ranging from 0 to 10. All the participants in this trial had hirsutism (F-G score > 8) (Harborne 2003).

Metformin resulted in an improvement of hirsutism compared to the OCP (MD -2.70, 95% CI -4.41 to -0.99, 1 RCT, n = 34, Analysis 1.2).

1.3 Hirsutism - Subjective improvement

One trial including 25 women compared metformin versus the OCP and reported on hirsutism subjective improvement (not reported if patient self-assessed or clinician assessed) (Jin 2006).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for subjective improvement of hirsutism (odds ratio (OR) 0.64, 95% CI 0.04 to 11.63, 1 RCT, n = 25, Analysis 1.3).

One trial including 28 women compared metformin versus the OCP and reported an improvement of acne and hirsutism with both metformin and the OCP in the Discussion, but did not report supporting data and therefore could not be included in any meta-analysis for this review (Kuek 2011).

1.4 Adverse events: severe (requiring stopping of medication) (gastro-intestinal and others)

Twelve trials including 965 women compared metformin versus the OCP and reported severe adverse events (Aghamohammadzadeh 2010; Bodur 2018; Christakou 2014; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kilic 2011; Luque-Ramirez 2007a; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Wu 2008).

Findings were not influenced by sensitivity analyses restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis').

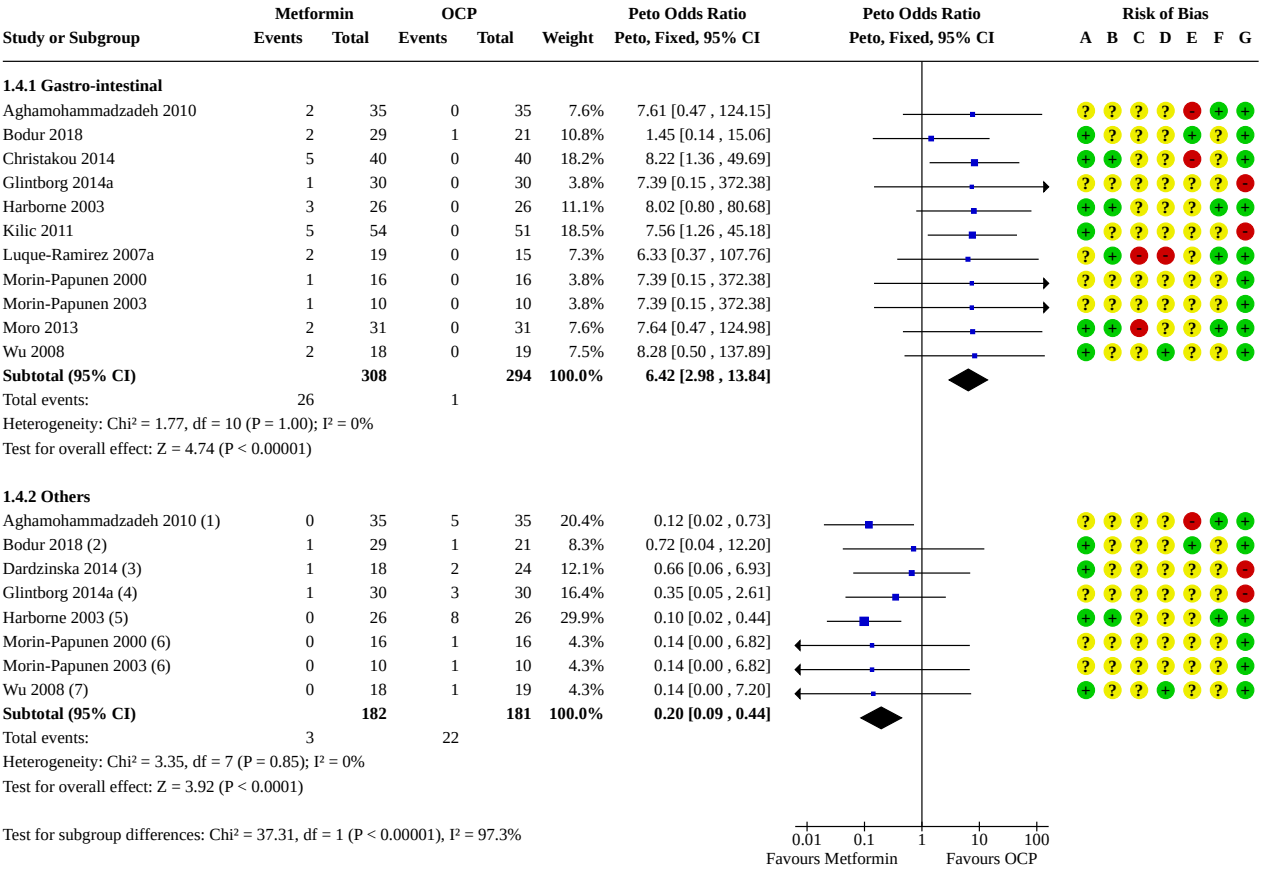
The funnel plot (n = 12 studies) was symmetrical indicating that our findings might not be influenced by publication bias (funnel plot not shown).

1.4.1 Gastro-intestinal

Eleven trials including 602 women compared metformin versus the OCP and reported severe gastro-intestinal adverse events (Aghamohammadzadeh 2010; Bodur 2018; Christakou 2014; Glintborg 2014a; Harborne 2003; Kilic 2011; Luque-Ramirez 2007a; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Wu 2008).

Metformin resulted in a higher incidence of severe gastro-intestinal adverse events compared to the OCP (Peto OR 6.42, 95% CI 2.98 to 13.84, 11 RCTs, n = 602, I^2 = 0%; low-quality evidence, Figure 6, Analysis 1.4) This suggests that if the severe gastro-intestinal adverse event rate following the OCP is 0.3%, then the severe gastro-intestinal adverse event rate after metformin would be between 1% and 4.5%.

Figure 6. Forest plot of comparison: 1 Adult - metformin versus OCP (Clinical parameters), outcome: 1.4 Adverse events - severe.



Footnotes

(1) OCP: combined: nausea, increasing BW, drug intolerance
(2) MET: dizziness/ OCP: sexual reluctance
(3) MET: depression/ OCP: intolerance
(4) Met: depression/ OCP: side effects
(5) Weight gain, depression, HBP, chest pain
(6) Headache, HBP
(7) Weight gain

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Findings were not influenced by sensitivity analysis restricting to the one study with low risk of bias (defined in methods section 'Sensitivity analysis') (Harborne 2003).

1.4.2 Others

Eight trials including 363 women compared metformin versus the OCP and reported other severe adverse events (Aghamohammadzadeh 2010; Bodur 2018; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Morin-Papunen 2000; Morin-Papunen 2003; Wu 2008.

Metformin may result in a lower incidence of severe other adverse events compared to the OCP (Peto OR 0.20, 95% CI 0.09 to 0.44, 8 RCTs, n = 363, I^2 = 0%, low-quality evidence, Figure 6;Analysis 1.4). This suggests that if the severe other adverse event rate following the OCP is 12%, then the severe other adverse event rate after metformin would between 1% and 6%.

Findings were not influenced by sensitivity analysis restricting to the one study with low risk of bias (defined in methods section 'Sensitivity analysis') (Harborne 2003).

One trial including 101 women compared metformin versus the OCP and reported severe adverse events with both metformin and the OCP, but did not report supporting data and therefore could not be included in any meta-analysis for this review (Sahu 2018).

One trial including 24 women compared metformin versus the OCP and reported minor adverse events with metformin only, but did not report supporting data and therefore could not be included in any meta-analysis for this review (Kebapcilar 2009a).

Adverse events - minor (gastro-intestinal and others)

No RCT reported on this outcome.

Secondary outcomes

1.5 Improved menstrual pattern: i.e. significant shortening of intermenstrual days

Two trials including 153 women compared metformin versus the OCP and reported improvement in menstrual pattern (Meyer 2007; Sahu 2018).

Overall, metformin may be less effective in improving menstrual pattern compared to the OCP (MD 6.05, 95% CI 2.37 to 9.74, 2 RCTs, n = 153, I^2= 0%; low-quality evidence, Analysis 1.5). This suggests that for a mean Improved menstrual pattern (i.e. significant shortening of intermenstrual days) of 32.4 following the OCP, the mean Improved menstrual pattern following metformin would be between 2.37 higher to 9.74 higher.

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.73, I^2 = 0%).

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

1.6 Improved menstrual pattern: i.e. an initiation of menses or cycle regularity

Six trials including 189 women compared metformin versus the OCP and reported improvement in menstrual pattern (Cetinkalp 2009; Jin 2006, Luque-Ramirez 2007b; Mhao 2015; Morin-Papunen 2000; Morin-Papunen 2003).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.03, I^2 = 67.8%) (Analysis 1.6). Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to the OCP may be less effective in improving menstrual pattern in subgroup BMI ≤ 25 kg/m^2 (Peto OR 0.07, 95% CI 0.01 to 0.65; 1 RCT; n = 17; low-quality evidence).

We are uncertain if metformin is less effective in improving menstrual pattern compared to the OCP in subgroup BMI > 25 kg/m^2 < 30kg/m^2 (Peto OR 0.15, 95% CI 0.07 to 0.33; 3 RCTs; n = 129; I^2 = 51%; verylLow-quality evidence).

We are uncertain of the effect of metformin compared to the OCP on menstrual pattern in subgroup BMI ≥30 kg/m^2 (Peto OR 0.09, 95% CI 0.01 to 1.62; 1 RCT; n = 18; very low-quality evidence).

We are uncertain of the effect of metformin compared to the OCP on menstrual pattern in subgroup BMI not stated (Peto OR 1.95, 95% CI 0.39 to 9.65; 1 RCT; n = 25; very low-quality evidence).

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

One trial including 60 women compared metformin versus the OCP and reported improvement in menstrual pattern in both group but did not report supporting and therefore could not be included in any meta-analysis in this review (Wu 2008).

1.7 Acne - visual analogue scale (VAS)

One trial including 34 women compared metformin versus the OCP and reported acne subjectively (patient self assessed) using a VAS ranging from 0 to 10. (Harborne 2003).

There was uncertainty as to whether there was a difference between metformin and the OCP for acne (patient self-assessed) (MD 0.90, 95% CI -0.40 to 2.20, 1 RCT, n = 34, low-quality evidence; Analysis 1.7). This suggests that for a mean acne - VAS of 1 following the OCP, the mean acne - VAS following metformin would be between 0.40 lower to 2.2 higher.

1.8 Acne - Subjective improvement

Three trials including 131 women compared metformin versus the OCP and reported acne subjectively (Cetinkalp 2009; Jin 2006, Mhao 2015).

Metformin was less effective in improving subjective improvement of acne compared to the OCP (OR 0.30, 95% CI 0.11 to 0.79; 3 RCTs; n = 131; I^2 = 18%) (Analysis 1.8).

One trial including 28 women compared metformin versus the OCP and reported an improvement of acne and hirsutism with both metformin and the OCP in the Discussion but did not report

supporting data and therefore could not be included in any meta-analysis in this review (Kuek 2011).

1.9 Diagnosis of Type II diabetes mellitus

One trial including 18 women compared metformin versus the OCP and reported on diagnosis of T2DM (Morin-Papunen 2000).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for the diagnosis of T2DM. (Peto OR 0.17, 95% CI 0.00 to 8.54, 1 RCT, n = 18, Analysis 1.9).

1.10 Body weight (kg)

Seven trials including 358 women compared metformin versus the OCP and reported on body weight (Aghamohammadzadeh 2010; Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Kuek 2011; Kumar 2018; Moran 2010).

Overall, there was insufficient evidence to determine whether there was a difference between metformin and the OCP for body weight (MD -0.93, 95% CI -2.93 to 1.08, 7 RCTs, n = 358, I^2 = 30%, Analysis 1.10).

Subgroup analysis showed sufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.03, I^2 = 70.5%).

1.11 Body Mass Index (BMI) (kg/m^2)

Nineteen trials including 923 women compared metformin versus the OCP and reported on BMI (Aghamohammadzadeh 2010; Cetinkalp 2009; Christakou 2014; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kilic 2011; Kuek 2011; Kumar 2018; Luque-Ramirez 2007a; Mhao 2015; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Ozgurtas 2008; Teng 2007; Sahu 2018; Wu 2008).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.03, I^2 = 70.9%) (Analysis 1.11). Therefore, we analysed the results per mean BMI subgroup.

We are uncertain if metformin decreases BMI compared to the OCP in subgroup BMI < 25 kg/m^2 (MD -0.59, 95% CI -1.02 to -0.17; 9 RCTs; n = 451; I^2 = 76%, very low-quality evidence). This suggests that for a mean BMI of 22.7 following the OCP, the mean BMI following metformin would be between 1.02 lower to 0.17 lower.

We are uncertain of the effect of metformin compared to the OCP on BMI in subgroup BMI > 25 kg/m^2 < 30 kg/m^2.(MD 0.11, 95% CI -0.48 to 0.70; 8 RCTs; n = 353; I^2 = 72%, very low-quality evidence). This suggests that for a mean BMI of 27.4 following the OCP, the mean BMI following metformin would be between 0.48 lower to 0.7 higher.

We are uncertain if metformin decreases BMI compared to the OCP in subgroup BMI > 30 kg/m^2 (MD -2.31, 95% CI -4.40 to -0.21; 3 RCTs; n = 119; I^2 = 52%), very low-quality evidence). This suggests that for a mean BMI of 35.1 following the OCP, the mean BMI following metformin would be between 4.4 lower to 0.21 lower.

Findings were not influenced by sensitivity analysis restricting to the one study with low risk of bias (defined in methods section 'Sensitivity analysis') (Harborne 2003).

The funnel plot (n = 19 studies) was asymmetrical indicating that our findings might be influenced by publication bias although asymmetrical funnel plots can also be due to true heterogeneity (funnel plot not shown).

1.12 Blood pressure - systolic (mmHg)

Five trials including 209 women compared metformin versus the OCP and reported on systolic blood pressure (Dardzinska 2014; Glintborg 2014a; Harborne 2003; Luque-Ramirez 2009b; Meyer 2007).

Overall, metformin resulted in an improvement in systolic blood pressure compared to the OCP (MD -4.81, 95% CI -8.55 to -1.06, 5 RCTs, n = 209, I^2 = 0%, Analysis 1.12).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.41, I^2 = 0%).

1.13 Blood pressure - diastolic (mmHg)

Four trials including 142 women compared metformin versus the OCP and reported on diastolic blood pressure (Dardzinska 2014; Glintborg 2014a; Harborne 2003; Luque-Ramirez 2009b).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.0009, I^2 = 90.9%) (Analysis 1.13). Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to the OCP showed appreciable benefit in the " BMI > 25 kg/m^2< 30 kg/m^2 " subgroup (MD -4.25, 95% CI -7.30 to -1.20; 3 RCTs; n = 108; I^2 = 0%) and appreciable harm in the BMI ≥ 30kg/m^2 " subgroup (MD 7.50, 95% CI 1.27 to 13.73; 1 RCT; n = 34; I^2 = 0%)

2. Metformin compared to the combined oral contraceptive pill (OCP) in adult women (hormonal parameters)

2.1 Serum total testosterone (nmol/L)

Seventeen trials including 818 women compared metformin versus the OCP and reported on serum total testosterone (Aghamohammadzadeh 2010; Cetinkalp 2009; Christakou 2014; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Jin 2006; Kuek 2011; Kumar 2018; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Ozgurtas 2008; Sahu 2018; Teng 2007; Wu 2008).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.006, I^2 = 76.1%). Therefore, we analysed the results per mean BMI subgroup.

The effect of metformin compared to the OCP showed a slight benefit for the OCP in all three subgroups: BMI ≤ 25 kg/m^2 (MD 0.48, 95% CI 0.38 to 0.57; 9 RCTs; n = 454; I^2 = 57%); BMI > 25kg/m^2 < 30kg/m^2 (MD 0.21, 95% CI 0.07 to 0.35; 4 RCTs; n = 220; I^2 = 0%); BMI

≥30 kg/m^2(MD 0.42, 95% CI 0.11 to 0.74; 3 RCTs; n = 119; I^2 = 0%) (Analysis 2.1)

2.2 Free androgen index (FAI) (%)

Ten trials including 433 women compared metformin versus the OCP and reported on FAI (Christakou 2014; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kuek 2011; Luque-Ramirez 2008a; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.006, I^2 = 78%). Therefore, we analysed the results per mean BMI subgroup.

The effect of metformin compared to the OCP showed an important benefit for the OCP in all three subgroups: BMI ≤ 25 kg/m^2(MD 4.48, 95% CI 3.56 to 5.40; 4 RCTs; n = 204; I^2 = 84%); BMI > 25 kg/m^2 < 30 kg/m^2(MD 3.06, 95% CI 2.14 to 3.97; 3 RCTs; n = 110; I^2 = 78%); BMI ≥30 kg/m^2 (MD 7.12, 95% CI 4.46 to 9.79; 3 RCTs; n = 119; I^2 = 27%) (Analysis 2.2).

3. Metformin compared to the combined oral contraceptive pill (OCP) in adult women (metabolic parameters)

3.1 Fasting insulin (mLU/L)

Twelve trials including 474 women compared metformin versus the OCP and reported on fasting insulin (Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kumar 2018; Luque-Ramirez 2007a; Meyer 2007; Morin-Papunen 2000; Morin-Papunen 2003; Moro 2013; Sahu 2018; Teng 2007; Wu 2008).

Overall, metformin resulted in an improvement of fasting insulin compared to the OCP (MD - 3.85, 95% CI -4.73 to -2.97, 12 RCTs, n = 474, I^2= 23%, Analysis 3.1).

Subgroup analysis showed sufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.04, I^2 = 65.1%).

3.2 Fasting glucose (mmol/L)

Twelve trials including 519 women compared metformin versus the OCP and reported on fasting glucose (Bodur 2018; Cetinkalp 2009; Glintborg 2014a; Harborne 2003; Kumar 2018; Kuek 2011; Luque-Ramirez 2007a; Moran 2010; Morin-Papunen 2000; Morin-Papunen 2003; Sahu 2018; Teng 2007).
Overall, metformin resulted in an improvement of fasting glucose compared to the OCP (MD -0.15, 95% CI -0.22 to -0.07, 12 RCTs, n = 519; I^2= 46%, Analysis 3.2).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.16, I^2 = 45%).

3.3 Total cholesterol (mmol/L)

Thirteen trials including 610 women compared metformin versus the OCP and reported on total cholesterol (Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kumar 2018; Luque-Ramirez 2007a; Mhao 2015; Meyer 2007; Moro 2013; Ozgurtas 2008; Rautio 2005; Sahu 2018).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P < 0.0001, I^2 = 91.2%). Therefore, we analysed the result per mean BMI subgroup (Analysis 3.3). Metformin decreased total cholesterol compared to the OCP in the subgroup BMI < 25 kg/m^2 (MD -0.77, 95% CI -1.00 to -0.53; 4 RCTs; n = 206; I^2 = 61%). There was insufficient evidence to determine whether there was a difference between metformin and the OCP for total cholesterol in the subgroup BMI > 25 kg/m^2 < 30 kg/m^2 (MD -0.14, 95% CI -0.30 to 0.01; 7 RCTs; n = 303; I^2 = 0%), and in the subgroup BMI > 30 kg/m^2 (MD -0.02, 95% CI -0.32 to 0.28; 2 RCTs; n = 101; I^2 = 0%).

3.4 High-density lipoprotein (HDL) cholesterol (mmol/L)

Thirteen trials including 610 women compared metformin versus the OCP and reported on HDL cholesterol (Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kumar 2018; Luque-Ramirez 2007a; Mhao 2015; Meyer 2007; Moro 2013; Ozgurtas 2008; Rautio 2005; Sahu 2018).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.04, I^2 = 69.5%) (Analysis 3.4).Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to the OCP showed no benefit in subgroup BMI ≤ 25 kg/m^2 (MD -0.02, 95% CI -0.01 to 0.07; 4 RCTs; n = 206; I^2 = 90%) and BMI > 25 kg/m^2 but < 30 kg/m^2 (MD 0.02, 95% CI -0.02 to 0.06; 7 RCTs; n = 303; I^2 = 75%), or slight harm in subgroup BMI ≥ 30kg/m^2 (MD 0.20, 95% CI 0.05 to 0.35; 2 RCTs; n = 101; I^2 = 0%) for the metformin in the three subgroups.

3.5 Low-density lipoprotein (LDL) cholesterol (mmol/L)

Thirteen trials including 610 women compared metformin versus the OCP and reported on LDL cholesterol (Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kumar 2018; Luque-Ramirez 2007a; Mhao 2015; Meyer 2007; Moro 2013; Ozgurtas 2008; Rautio 2005; Sahu 2018).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P < 0.00001, I^2 = 92.7%). Therefore, we analysed the results per mean BMI subgroup.

The effect of the metformin compared to the OCP showed a slight harm in subgroup BMI ≥30 kg/m^2 (MD 0.35, 95% CI 0.02 to 0.67; 2 RCTs; n = 101; I^2 = 0%), or benefit in subgroup BMI ≤ 25kg/m^2 (MD -0.39, 95% CI -0.54 to -0.23; 4 RCTs; n = 206; I^2 = 50%), or no important benefit/harm in subgroup BMI > 25 kg/m^2 but < 30 kg/m^2 (MD 0.02, 95% CI -0.06 to 0.10; 7 RCTs; n = 303; I^2 = 12%) for metformin in the three subgroups (Analysis 3.5).

3.6 Triglycerides (mmol/L)

Thirteen trials including 610 women compared metformin versus the OCP and reported on triglycerides (Cetinkalp 2009; Dardzinska 2014; Glintborg 2014a; Harborne 2003; Kebapcilar 2009a; Kumar 2018; Luque-Ramirez 2007a; Mhao 2015; Meyer 2007; Moro 2013; Ozgurtas 2008; Rautio 2005; Sahu 2018).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P < 0.00001, I^2 = 93.1%). (Analysis 3.6) Therefore, we analysed the results per BMI subgroup.

The effect of metformin compared to the OCP showed slight benefit in subgroup BMI ≤ 25 kg/m^2 (MD -0.45, 95% CI -0.61 to -0.30; 4 RCTs; n = 206; I^2 = 67%), and no improvement in groups BMI > 25k g/m^2 < 30kg/m^2 (MD -0.01, 95% CI -0.07 to 0.05; 7 RCTs; n = 303; I^2 = 20%), and BMI ≥30 kg/m^2 (MD -0.31, 95% CI -0.64 to 0.01; 2 RCTs; n = 101; I^2 = 0%).

4. Metformin compared to metformin combined with the OCP in adult women (clinical parameters)

Summary of findings 2

Six trials including 295 adult women compared metformin combined with the OCP compared to metformin (Bodur 2018; Glintborg 2014a; Harborne 2003; Jin 2006; Kebapcilar 2009a; Kumar 2018; Liu 2006; Moro 2013; Wu 2008).

Primary outcomes

4.1 Hirsutism - Clinical Ferriman-Gallwey (F-G) score

Three trials including 135 women compared metformin versus metformin combined with the OCP and reported hirsutism clinically using the F-G score (Glintborg 2014a; Kumar 2018; Wu 2008).

Overall, metformin may be less effective in improving hirsutism compared to metformin combined with the OCP (MD 1.36, 95% CI 0.62 to 2.11, 3 RCTs, n = 135, I^2= 9%, low-quality evidence, Analysis 4.1).

This suggests that for a mean F-G score of 5.6 following the metformin combined with the OCP, the mean F-G score following metformin would be between 0.62 higher to 2.11 higher.

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.18, I^2 = 43.5%).

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

Hirsutism - Subjective visual analogue scale (VAS)

No RCT reported on this outcome

Hirsutism - subjective improvement

No RCT reported on this outcome

4.2 Adverse events: severe (requiring stopping of medication) (gastro-intestinal and others)

Three trials compared metformin versus metformin combined with the OCP and reported severe adverse events (Bodur 2018; Glintborg 2014a; Moro 2013).

4.2.1 Gastro-intestinal

Three trials including 171 women compared metformin versus metformin combined with the OCP and reported severe gastro-intestinal adverse events (Bodur 2018; Glintborg 2014a; Moro 2013).

We are uncertain if there was a difference between metformin and metformin combined with the OCP for severe gastro-intestinal adverse events (OR 0.74, 95% CI 0.21 to 2.53, 3 RCTs, n = 171, I^2 = 0%, low-quality evidence Analysis 4.2). This suggests that if the severe gastrointestinal adverse event rate following metformin combined with the OCP is 7%, then the severe gastrointestinal adverse event rate after metformin would be between 2% and 17%.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

4.2.2 Others

Two trials including 109 women compared metformin versus metformin combined with the OCP and reported severe other adverse events (Bodur 2018; Glintborg 2014a).

We are uncertain if there was a difference between metformin and metformin combined with the OCP for severe other adverse events (OR 0.56, 95% CI 0.11 to 2.82, 2 RCTs, n = 109, I^2 = 44%, low-quality evidence Analysis 4.2). This suggests that if the severe other adverse event rate following metformin combined with the OCP is 6%, then the severe other adverse event rate after metformin would between 1% and 15%.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

One trial including 24 women compared metformin versus metformin combined with the OCP and reported minor adverse events with metformin combined with the OCP and metformin only, but did not report supporting data and therefore could not be included in any meta-analysis in this review (Kebapcilar 2009a).

Adverse events - minor (gastro-intestinal and others)

No RCT reported on this outcome.

Secondary outcomes

Improved menstrual pattern (i.e. shortening of intermenstrual days)

No RCT reported on this outcome.

Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)

No RCT reported on this outcome.

Acne - visual analogue scale (VAS)

No RCT reported on this outcome.

Acne - subjective improvement

No RCT reported on this outcome.

Diagnosis of Type II diabetes mellitus

No RCT reported on this outcome.

4.3 Body weight (kg)

Two trials including 101 women compared metformin versus metformin combined with the OCP and reported on body weight (Glintborg 2014a; Kumar 2018).

Overall, metformin resulted in an improvement of body weight compared to metformin combined with the OCP (MD -5.39, 95% CI -10.70 to -0.08, 2 RCTs, n = 101, I^2= 0%, Analysis 4.3).

Subgroup analysis was not applicable as both RCTs were in the same mean study BMI subgroup.

4.4 Body mass index (BMI) (kg/m^2)

Five trials including 199 women compared metformin versus metformin combined with the OCP and reported on BMI (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Wu 2008).

Overall, metformin may improve BMI compared to metformin combined with the OCP (MD -1.47, 95% CI -2.27 to -0.66, 5 RCTs, n = 199, I^2= 25%, low-quality evidence, Analysis 4.4). This suggests that for a mean BMI of 25.49 following metformin combined with the OCP, the mean BMI following metformin would be between 2.27 lower to 0.66 lower.

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.90, I^2 = 0%).

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

4.5 Blood pressure - systolic (mmHg)

One trial including 42 women compared metformin versus metformin combined with the OCP and reported on systolic blood pressure (Glintborg 2014a).

Metformin resulted in an improvement of systolic blood pressure compared to metformin combined with the OCP (MD -10.59, 95% CI -15.76 to -5.42, 1 RCT, n = 42, Analysis 4.5).

4.6 Blood pressure - diastolic (mmHg)

One trial including 42 women compared metformin versus metformin combined with the OCP and reported on diastolic blood pressure (Glintborg 2014a).

Metformin resulted in an improvement of diastolic blood pressure compared to metformin combined with the OCP (MD -7.93, 95% CI -14.01 to -1.85, 1 RCT, n = 42, Analysis 4.6).

5. Metformin compared to metformin combined with the OCP in adult women (hormonal parameters)

5.1 Serum total testosterone (nmol/L)

Five trials including 226 women compared metformin versus metformin combined with the OCP and reported on serum total testosterone (Glintborg 2014a; Kumar 2018; Liu 2006; Moro 2013; Wu 2008).

Substantial heterogeneity was detected, which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P < 0.00001, I^2 = 95.9%). (Analysis 5.1). Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to metformin combined with the OCP showed no important benefit/harm for metformin in the subgroup BMI ≤ 25 kg/m^2(MD 0.09, 95% CI -0.08 to 0.26; 2 RCTs; n = 74; I^2 = 0%), and important harm in the other subgroup BMI > 25 kg/m^2 < 30kg/m^2 (MD 0.79, 95% CI 0.57 to 1.00; 3 RCTs; n = 152; I^2 = 65%).

5.2 Free androgen index (FAI) (%)

Three trials including 133 women compared metformin versus metformin combined with the OCP and reported on FAI (Glintborg 2014a; Liu 2006; Moro 2013).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P < 0.0001, I^2 = 93.7%) (Analysis 5.2). Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to metformin combined with the OCP showed appreciable benefit for the OCP combined with metformin in the " BMI > 25 kg/m^2 < 30 kg/m^2" subgroup (MD 3.80, 95% CI 2.91 to 4.69; 2 RCTs; n = 93; I^2 = 97%), and no important benefit/harm in the other subgroup (BMI ≤ 25 kg/m^2) (MD 0.35, 95% CI -1.09 to 1.79; 1 RCT; n = 40; I^2 = 0%).

6. Metformin compared to metformin combined with the OCP in adult women (metabolic parameters)

6.1 Fasting insulin (mLU/L)

Five trials including 199 women compared metformin versus metformin combined with the OCP and reported on fasting insulin (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Wu 2008).

Overall, metformin resulted in an improvement of fasting insulin compared to metformin combined with the OCP (MD -1.32, 95% CI -2.63 to -0.01, 5 RCTs, n = 199, I^2= 33%, Analysis 6.1).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.72, I^2 = 0%).

6.2 Fasting glucose (mmol/L)

Four trials including 170 women compared metformin versus metformin combined with the OCP and reported on fasting glucose (Bodur 2018; Glintborg 2014a; Kumar 2018; Liu 2006).

Overall, metformin resulted in an improvement of fasting glucose compared to metformin combined with the OCP (MD -0.21, 95% CI -0.37 to -0.06, 4 RCTs, n = 170, I^2= 49%, Analysis 6.2).
Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.76, I^2 = 0%).

6.3 Total cholesterol (mmol/L)

Five trials including 216 women compared metformin versus metformin combined with the OCP and reported on total cholesterol (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Moro 2013).
Overall, metformin resulted in an improvement of total cholesterol compared to metformin combined with the OCP (MD -0.54, 95% CI -0.97 to -0.11, 5 RCTs, n = 216, I^2= 65%, Analysis 6.3).

Substantial heterogeneity was detected, which was not explained by a difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.81, I^2 = 0%).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

6.4 High-density lipoprotein (HDL) cholesterol (mmol/L)

Five trials including 216 women compared metformin versus metformin combined with the OCP and reported on HDL cholesterol (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Moro 2013).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.0002, I^2 = 93.0%) (Analysis 6.4). Therefore, we analysed the results per mean BMI subgroup.

Metformin compared to metformin combined with the OCP showed appreciable benefit for metformin in the "BMI ≤ 2 5kg/m^2 " subgroup (MD -0.64, 95% CI -0.99 to -0.29; n = 40; 1 RCT; I^2 = 0%), and no important benefit/harm in the other subgroup (BMI > 25 kg/m^2 < 30 kg/m^2) (MD 0.05, 95% CI -0.04 to 0.14; n = 176; 4 RCTs; I^2 = 0%)

6.5 Low-density lipoprotein (LDL) cholesterol (mmol/L)

Four trials including 176 women compared metformin versus metformin combined with the OCP and reported on LDL cholesterol (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Moro 2013).
Overall, there was insufficient evidence to determine whether there was a difference between metformin combined with the OCP and metformin for LDL cholesterol (MD -0.13, 95% CI -0.32 to 0.06, 4 RCTs, n = 176, I^2 = 40%, Analysis 6.5).

Subgroup analysis was not applicable as all four RCTs were in the same mean study BMI subgroup.

6.6 Triglycerides (mmol/L)

Five trials including 216 women compared metformin combined with the OCP versus metformin and reported on triglycerides (Glintborg 2014a; Kebapcilar 2009a; Kumar 2018; Liu 2006; Moro 2013).
Overall, metformin resulted in an improvement of triglycerides compared to metformin combined with the OCP (MD -0.22, 95% CI -0.37 to -0.07, 5 RCTs, n = 216, I^2= 85%, Analysis 6.6).

Substantial heterogeneity was detected which could not be explained by the difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.03, I^2 = 78.6%) but could be explained by the effect of one of the included studies which showed benefit for metformin combined with the OCP (Kebapcilar 2009a).

7. Oral contraceptive pill (OCP) compared with metformin combined with the OCP in adult women (clinical parameters)

Summary of findings 3

Sixteen trials including 1043 adult women compared metformin combined with the OCP to the combined OCP (Bhattacharya 2016; Bodur 2018; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Ruan 2018; Wei 2012; Wu 2008).

Primary outcomes

7.1 Hirsutism - Clinical Ferriman-Gallwey (F-G) score

Six trials including 389 women compared the OCP versus metformin combined with the OCP and reported hirsutism clinically using the F-G score (Bhattacharya 2016; Elter 2002; Feng 2016; Glintborg 2014a; Kumar 2018; Wu 2008).

Overall, the OCP may be less effective in improving hirsutism compared to metformin combined with the OCP (MD 0.54, 95% CI 0.20 to 0.89, 6 RCTs, n = 389, I^2= 1%, low-quality evidence, Analysis 7.1). This suggests that for a mean F-G score of 5.57 following metformin combined with the OCP, the mean F-G score following the OCP would be between 0.20 higher to 0.89 higher.

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.36, I^2 = 0%).

Findings were not influenced by sensitivity analysis restricting to the one study with low risk of bias (defined in methods section 'Sensitivity analysis') (Feng 2016).

Hirsutism - subjective visual analogue scale (VAS)

No RCT reported on this outcome.

Hirsutism - Subjective improvement

No RCT reported on this outcome.

7.2 Adverse events: severe (requiring stopping of medication) (gastro-intestinal and others)

Six trials including 387 women compared the OCP versus metformin combined with the OCP and reported severe adverse events (Bodur 2018; Cibula 2005; Glintborg 2014a; Essah 2011; Moro 2013; Wu 2008).

7.2.1 Gastro-intestinal

Five trials including 228 women compared the OCP versus metformin combined with the OCP and reported severe gastro-intestinal adverse events (Bodur 2018; Cibula 2005; Glintborg 2014a; Moro 2013; Wu 2008).

The OCP may result in a lower incidence of severe gastro-intestinal adverse events compared to metformin combined with the OCP (OR 0.20, 95% CI 0.06 to 0.72, 5 RCTs, n = 228, I^2 = 0%, low-quality evidence Analysis 7.2). This suggests that for severe gastro-

intestinal adverse event rate of 10% following metformin combined with the OCP, the severe gastro-intestinal adverse event rate following the OCP would be between 1% and 7%.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

7.2.2 Others

Four trials including 159 women compared the OCP versus metformin combined with the OCP and reported other severe adverse events (Bodur 2018; Essah 2011; Glintborg 2014a; Wu 2008).

We are uncertain if there is a difference between the OCP and metformin combined with the OCP for other severe adverse events (OR 1.61, 95% CI 0.49 to 5.37, 4 RCTs, n = 159, I^2 = 12%, low-quality evidence, Analysis 7.2). This suggests that for severe other adverse event rate of 4% following metformin combined with the OCP, the severe other adverse event rate following the OCP would be between 1.9% and 17.9%.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

7.3 Adverse events: minor (gastro-intestinal and others)

Two trials including 98 women compared the OCP versus metformin combined with the OCP and reported minor adverse events (Elter 2002; Wei 2012).

7.3.1 Gastro-intestinal

Overall, the OCP may have a lower incidence of minor (gastro-intestinal) adverse events compared to metformin combined with the OCP (OR 0.06, 95% CI 0.01 to 0.44, 2 RCTs, n = 98, I^2 = 0%,low-quality evidence; Analysis 7.3). This suggests that for an overall minor (gastro-intestinal) adverse event rate of 26% following metformin combined with the OCP, the minor (gastro-intestinal) adverse event rate following the OCP would be between 0.4% and 13%.

Subgroup analysis was not applicable as all minor adverse events were gastro-intestinal in nature.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

One trial including 24 women compared the OCP versus metformin combined with the OCP and reported minor adverse events with metformin combined with the OCP, but did not report supporting data and therefore could not be included in any meta-analysis in this review (Kebapcilar 2009a).

Secondary outcomes

Improved menstrual pattern (i.e. shortening of intermenstrual days)

No RCT reported on this outcome.

Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)

No RCT reported on this outcome

7.4 Acne - clinical acne score

One trial including 82 women compared the OCP versus metformin combined with the OCP and reported acne using a clinical acne score. (Feng 2016).

The OCP may improve slightly acne compared to metformin combined with the OCP (MD -0.09, 95% CI -0.10 to - 0.08, 1 RCT, n = 82, low-quality evidence, Analysis 7.4).

This one RCT had a low risk of bias (defined in methods section 'Sensitivity analysis').

7.5 Acne - subjective improvement

One trial including 129 women compared the OCP versus metformin combined with the OCP and reported acne subjectively. (Bhattacharya 2016).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for acne (OR 0.67, 95% CI 0.33 to 1.35, 1 RCT, n = 129, Analysis 7.5).

Diagnosis of Type II diabetes mellitus

No RCT reported on this outcome.

7.6 Body weight (kg)

Seven trials including 387 women compared the OCP versus metformin combined with the OCP and reported on body weight (Bhattacharya 2016; Cibula 2005; Essah 2011; Glintborg 2014a; Kaya 2015; Kumar 2018; Wei 2012).

Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for body weight (MD -0.63, 95% CI -1.58 to 0.33, 7 RCTs, n = 387, I^2 = 18%, Analysis 7.6).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.80, I^2 = 0%).

7.7 Body Mass Index (BMI) (kg/m^2)

Thirteen trials including 661 women compared the OCP versus metformin combined with the OCP and reported on BMI (Bhattacharya 2016; Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Lv 2005; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Wei 2012; Wu 2008).

Overall, there was uncertainty as to whether there was a difference between the OCP and metformin combined with the OCP for BMI (MD -0.21, 95% CI -0.53 to 0.12, 13 RCTs, n = 661, I^2 = 50%, very low-quality evidence, Analysis 7.7).

Substantial heterogeneity was detected, which was not explained by the difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.64, I^2 = 0%).

Findings were not influenced by sensitivity analysis restricting to the one study with low risk of bias (defined in methods section'Sensitivity analysis') (Feng 2016).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

The funnel plot (n = 13 studies) was symmetrical indicating that our findings might not be influenced by publication bias (funnel plot not shown)

7.8 Blood pressure - systolic (mmHg)

Five trials including 326 women compared the OCP versus metformin combined with the OCP and reported on systolic blood pressure (Bhattacharya 2016; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015).

Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for systolic blood pressure (MD -1.75, 95% CI -4.03 to 0.53, 5 RCTs, n = 326, I^2 = 76%, Analysis 7.8).

Substantial heterogeneity was detected which was not explained by a difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.16, I^2 = 45.8%).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

7.9 Blood pressure - diastolic (mmHg)

Five trials including 326 women compared the OCP versus metformin combined with the OCP and reported on diastolic blood pressure (Bhattacharya 2016; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015).

Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for diastolic blood pressure (MD -1.05, 95% CI -2.79 to 0.68, 5 RCTs, n = 326, I^2 = 76%, Analysis 7.9).

Substantial heterogeneity was detected which was not explained by a difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.28, I^2 = 20.8%).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

8. The OCP compared to metformin combined with the OCP in adult women (hormonal parameters)

8.1 Serum total testosterone (nmol/L)

Twelve trials including 715 women compared the OCP versus metformin combined with the OCP and reported on serum total testosterone (Bhattacharya 2016; Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Wei 2012; Wu 2008).

Overall, the OCP resulted in an increase of serum total testosterone compared to metformin combined with the OCP (MD 0.08, 95% CI 0.01 to 0.16, 12 RCTs, n = 715, I^2 = 0%, Analysis 8.1).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.91, I^2 = 0%).

8.2 Free androgen index (FAI) (%)

Seven trials including 482 women compared the OCP versus metformin combined with the OCP and reported on FAI (Bhattacharya 2016; Cibula 2005; Glintborg 2014a; Kaya 2015; Moro 2013; Song 2017; Wei 2012).

Overall, the OCP resulted in an increase of FAI compared to metformin combined with the OCP (MD 0.51, 95% CI 0.30 to 0.71, 7 RCTs, n = 482, I^2= 28%, Analysis 8.2).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.90, I^2 = 0%).

9. The oral contraceptive pill (OCP) compared to metformin combined with the OCP in adult women (metabolic parameters)

9.1 Fasting insulin (mLU/L)

Twelve trials including 602 women compared the OCP versus metformin combined with the OCP and reported on fasting insulin (Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Song 2017; Wei 2012; Wu 2008).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.0006, I^2 = 86.6%). Therefore, we analysed the results per mean BMI subgroup.

The OCP compared to metformin combined with the OCP showed appreciable benefit for metformin combined with the OCP in the BMI < 25 kg/m^2 subgroup (MD 4.77, 95% CI 3.26 to 6.28; 5 RCTs; n = 198; I^2 = 0%), and in the BMI > 25 kg/m^2 < 30 kg/m^2 subgroup (MD 0.25, 95% CI 0.14 to 0.36; 4 RCTs; = 305; n = 4; I^2 = 39%). There was no important effect between the OCP and metformin combined with the OCP in the BMI > 30 kg/m^2 subgroup (MD -0.30, 95% CI -0.81 to 0.21; 1 RCT; n = 19; I^2 = 0%).

9.2 Fasting glucose (mmol/L)

Ten trials including 529 women compared the OCP versus metformin combined with the OCP and reported on fasting glucose (Bodur 2018; Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kumar 2018; Lv 2005; Song 2017; Wei 2012).

Subgroup analysis showed sufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.06, I^2 = 64.1%), and therefore we did not pool the results and performed the analysis per BMI subgroup. (Analysis 9.2). The OCP compared to metformin combined with the OCP showed insufficient evidence of a difference in effect in the BMI ≤ 25 kg/m^2 subgroup (MD 0.11, 95% CI -0.07 to 0.29; 5 RCTs; n = 205; I^2 = 42%), and in the BMI ≥30 kg/m^2 subgroup (MD -0.30, 95% CI -0.81 to 0.21; participants = 19; studies = 1; I^2 = 0%). The OCP compared to metformin combined with the OCP showed an increase in fasting glucose levels in subgroup BMI > 25 kg/m^2 < 30 kg/m^2 (MD 1.37, 95% CI 0.42 to 2.32; participants = 385; studies = 7; I^2 = 51%).

9.3 Total cholesterol (mmol/L)

Thirteen trials including 668 women compared the OCP versus metformin combined with the OCP and reported on total cholesterol (Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Wei 2012).
Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for total cholesterol (MD 0.07, 95% CI -0.05 to 0.17, 13 RCTs, n = 668, I^2 = 55%, Analysis 9.3).

Substantial heterogeneity was detected which was not explained by a difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.43, I^2 = 0%).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

9.4 High-density lipoprotein (HDL) cholesterol (mmol/L)

Thirteen trials including 668 women compared the OCP versus metformin combined with the OCP and reported on HDL cholesterol (Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Wei 2012).

Overall, the OCP resulted in a slight worsening of HDL cholesterol compared to metformin combined with the OCP (MD 0.05, 95% CI 0.01 to 0.09, 13 RCTs, n = 668, I^2= 70%, Analysis 9.4).

Substantial heterogeneity was detected which was not explained by a difference in effect of the interventions between the BMI subgroups (test for subgroup difference: P = 0.40, I^2 = 0%).

Findings were not influenced by sensitivity analyses using a random-effects model (in the presence of unexplained substantial heterogeneity).

9.5 Low-density lipoprotein (LDL) cholesterol (mmol/L)

Thirteen trials including 668 women compared the OCP versus metformin combined with the OCP and reported on LDL cholesterol (Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Wei 2012).
Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for LDL cholesterol (MD 0.04, 95% CI -0.05 to 0.14, 13 RCTs, n = 668, I^2 = 25%, Analysis 9.5).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.65, I^2 = 0%).

9.6 Triglycerides (mmol/L)

Thirteen trials including 668 women compared the OCP versus metformin combined with the OCP and reported on triglycerides (Cibula 2005; Elter 2002; Essah 2011; Feng 2016; Glintborg 2014a; Kaya 2015; Kebapcilar 2009a; Kebapcilar 2009b; Kumar 2018; Lv 2005; Moro 2013; Song 2017; Wei 2012).

Overall, there was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for total cholesterol (MD -0.02, 95% CI -0.09 to 0.04, 13 RCTs, n = 668, I^2 = 45%, Analysis 9.6).

Subgroup analysis showed insufficient evidence to suggest a difference in effect by mean study BMI (test for subgroup difference: P = 0.43, I^2 = 0%).

10. Metformin compared to the combined oral contraceptive pill (OCP) in adolescent women (clinical parameters)

Summary of findings 4

Four trials including 42 adolescent women compared metformin to the OCP (Al-Zubeidi 2015; Allen 2005; El Maghraby 2015; Hoeger 2008a).

Primary outcomes

10.1 Hirsutism - Clinical Ferriman-Gallwey (F-G) score

One trial including 16 adolescent women compared metformin versus the OCP and reported hirsutism clinically using the F-G score (Hoeger 2008a).

There was uncertainty as to whether there was a difference between metformin and the OCP for hirsutism (MD -0.40, 95% CI -3.42 to 2.62, 1 RCT, n = 16, very low-quality evidence, Analysis 10.1). This suggests that for a mean F-G score of 8.6 following the OCP, the mean F-G score following metformin would be between 3.42 lower to 2.62 higher.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

Hirsutism - Subjective visual analogue scale (VAS)

No RCT reported on this outcome.

10.2 Hirsutism - Subjective improvement

One trial including 80 adolescent women compared metformin versus the OCP and reported hirsutism subjectively (not reported if patient self-assessed or clinician assessed) (El Maghraby 2015).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for hirsutism (Peto OR 0.50, 95% CI 0.19 to 1.30, 1 RCT, n = 80, Analysis 10.2).

10.3 Adverse events: severe (requiring stopping of medication) (gastro-intestinal and others)

10.3.1 Others

One trial including 80 adolescent women compared metformin versus the OCP and reported severe other adverse events (El Maghraby 2015).

There was uncertainty as to whether there was a difference between metformin and the OCP for severe other adverse events (OR 0.63, 95% CI 0.16 to 2.43, 1 RCT, n = 80, very low-quality evidence, Analysis 10.3).

This suggests that for a severe other adverse event rate of 1.5% following the OCP, the severe other adverse event rate following metformin would be between 2.7% and 30%.

10.4 Adverse events: minor (gastro-intestinal and others)

10.4.1 Gastro-intestinal

One trial including 22 adolescent women compared metformin versus the OCP and reported minor gastro-intestinal adverse events (Al-Zubeidi 2015).

There was uncertainty as to whether there was a difference between metformin and the OCP for minor gastro-intestinal adverse events (OR 11.67, 95% CI 0.53 to 258.56, 1 RCT, n = 12, very low-quality evidence, Analysis 10.4). This suggests that for a minor (gastro-intestinal) adverse event rate of 0% following the OCP, the minor (gastro-intestinal) adverse event rate following metformin would be between 0% and 0%.

Secondary outcomes

Improved menstrual pattern (i.e. shortening of intermenstrual days)

No RCT reported on this outcome.

10.5 Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)

One trial including 80 adolescent women compared metformin versus the OCP and reported improvement in menstrual pattern (El Maghraby 2015).

There was uncertainty as to whether there was a difference between metformin and the OCP for improvement of menstrual pattern (OR 0.10, 95% CI 0.01 to 1.92, 1 RCT, n = 80, very low-quality evidence, Analysis 10.5).

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

One trial including 34 adolescent women compared metformin versus the OCP and reported improvement in menstrual pattern but did not report supporting data, and therefore could not be included in any meta-analysis in this review (Al-Zubeidi 2015).

Acne - visual analogue scale (VAS)

No RCT reported on this outcome.

Acne - Subjective improvement

No RCT reported on this outcome.

Diagnosis of type II diabetes mellitus

No RCT reported on this outcome.

10.6 Body weight (kg)

Two trials including 111 adolescent women compared metformin versus the OCP and reported on body weight (Allen 2005; El Maghraby 2015).

Substantial heterogeneity was detected which may be explained by the difference in effect of the interventions between the mean BMI subgroups (test for subgroup difference: P = 0.04, I^2 = 77.1%) (Analysis 10.6) . Therefore, we analysed the results per mean BMI subgroup. The effect of metformin compared to the OCP showed an appreciable benefit for metformin in subgroup, mean BMI not stated (MD -19.00, 95% CI -20.81 to -17.19; 1 RCT; n = 80; I^2 = 0%), and no appreciable benefit/harm in subgroup BMI ≥30 kg/m^2 (MD -2.60, 95% CI -17.87 to 12.67; 1 RCT; n = 31; I^2 = 0%).

10.7 Body Mass Index (kg/m^2)

Three trials including 69 adolescent women compared metformin versus the OCP and reported on BMI (Al-Zubeidi 2015; Allen 2005; Hoeger 2008a).

Overall, there was uncertainty as to whether there was a difference between metformin and the OCP for BMI (MD -1.45, 95% CI -5.08 to 2.17, 3 RCTs, n = 69, I^2 = 0%, very low-quality evidence, Analysis 10.7). This suggests that for a mean BMI of 36 following the OCP, the mean BMI following metformin would be between 5.08 lower to 2.17 higher.

Subgroup analysis was not applicable as all three RCTs were in the same mean study BMI subgroup.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

10.8 Blood pressure - systolic (mmHg)

One trial including 16 adolescent women compared metformin versus the OCP and reported on systolic blood pressure (Hoeger 2008a).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for systolic blood pressure (MD -1.40, 95% CI -12.87 to 10.07, 1 RCT, n = 16, Analysis 10.8).

10.9 Blood pressure - diastolic (mmHg)

One trial including 16 adolescent women compared metformin versus the OCP and reported on diastolic blood pressure (Hoeger 2008a).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for diastolic blood pressure (MD -5.10, 95% CI -13.69 to 3.49, 1 RCT, n = 16, Analysis 10.9).

11. Metformin compared to the combined oral contraceptive pill (OCP) in adolescent women (hormonal parameters)

11.1 Serum total testosterone (nmol/L)

Three trials including 69 adolescent women compared metformin versus the OCP and reported on serum total testosterone (Al-Zubeidi 2015; Allen 2005; Hoeger 2008a).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for serum total

testosterone (MD 0.23, 95% CI -0.21 to 0.68, 3 RCTs, n = 69, I^2= 0%, Analysis 11.1).

Subgroup analysis was not applicable as all three RCTs were in the same mean study BMI subgroup.

11.2 Free androgen index (FAI) (%)

One trial including 16 adolescent women compared metformin versus the OCP and reported on FAI (Hoeger 2008a).

Metformin resulted in an increase of FAI compared to the OCP (MD 8.50, 95% CI 1.99 to 15.01, 1 RCT, n = 16, Analysis 11.2).

12. Metformin compared to the combined oral contraceptive pill (OCP) in adolescent women (metabolic parameters)

12.1 Fasting insulin (mLU/L)

Two trials including 53 adolescent women compared metformin versus the OCP and reported on fasting insulin (Al-Zubeidi 2015; Allen 2005).

Overall, there was insufficient evidence to determine whether there was a difference between metformin and the OCP for fasting insulin (MD 4.55, 95% CI -4.82 to 13.92, 2 RCTs, n = 53, I^2= 21%, Analysis 12.1).

Subgroup analysis was not applicable as both RCTs were in the same mean study BMI subgroup.

12.2 Fasting glucose (mmol/L)

One trial including 16 adolescent women compared metformin versus the OCP and reported on fasting glucose (Hoeger 2008a).

There was insufficient evidence to determine whether there was a difference between metformin and the OCP for fasting glucose (MD 0.11, 95% CI -0.55 to 0.77, 1 RCT, n = 16, Analysis 12.2).

12.3 Total cholesterol (mmol/L)

Two trials including 47 adolescent women compared metformin versus the OCP and reported on total cholesterol (Allen 2005; Hoeger 2008a).

Overall, metformin resulted in an improvement of total cholesterol compared to the OCP (MD -1.12, 95% CI -1.66 to -0.58, 2 RCTs, n = 47, I^2= 0%, Analysis 12.3).

Subgroup analysis was not applicable as both RCTs were in the same mean study BMI subgroup.

12.4 High-density lipoprotein (HDL) cholesterol (mmol/L)

Two trials including 47 adolescent women compared metformin versus the OCP and reported on HDL cholesterol (Allen 2005; Hoeger 2008a).

Overall, there was insufficient evidence to determine whether there was a difference between metformin and the OCP for HDL cholesterol (MD 0.12, 95% CI -0.10 to 0.34, 2 RCTs, n = 47, I^2= 6%, Analysis 12.4).

Subgroup analysis was not applicable as both RCTs were in the same mean study BMI subgroup.

12.5 Low-density lipoprotein (LDL) cholesterol (mmol/L)

Two trials including 47 adolescent women compared metformin versus the OCP and reported on LDL cholesterol (Allen 2005; Hoeger 2008a).

Overall, metformin resulted in an improvement of LDL cholesterol compared to the OCP (MD -0.92, 95% CI -1.49 to -0.35, 2 RCTs, n = 47, I^2= 0%, Analysis 12.5).

Subgroup analysis was not applicable as both RCTs were in the same mean study BMI subgroup.

12.6 Triglycerides (mmol/L)

Three trials including 69 adolescent women compared metformin versus the OCP and reported on triglycerides (Bhattacharya 2016; Allen 2005; Hoeger 2008a).

Overall, there was insufficient evidence to determine whether there was a difference between metformin and the OCP for triglycerides (MD -0.13, 95% CI -0.37 to 0.10, 3 RCTs, n = 69, I^2= 9%, Analysis 12.6).

Subgroup analysis was not applicable as all three RCTs were in the same mean study BMI subgroup.

Metformin compared to metformin combined with the oral contraceptive pill (OCP) in adolescent women

There were no trials identified comparing metformin with metformin combined with the OCP in adolescent women on the selected outcomes for this review.

13. The oral contraceptive pill (OCP) compared to metformin combined with the OCP in adolescent women (clinical parameters)

Summary of findings 6

One trial including 36 adolescent women compared metformin combined with the OCP compared to the OCP (Hoeger 2008b).

Primary outcomes

13.1 Hirsutism - Clinical Ferriman-Gallwey (F-G) score

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported hirsutism clinically using the F-G score (Hoeger 2008b).

There was uncertainty as to whether there was a difference between the OCP and metformin combined with the OCP for hirsutism (MD 0.80, 95% CI -1.19 to 2.79, 1 RCT, n = 32, very low-quality evidence, Analysis 13.1). This suggests that for a mean F-G score of 6.2 following metformin combined with the OCP, the mean F-G score following the OCP would be between 1.19 lower to 2.79 higher.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

Hirsutism - subjective visual analogue scale (VAS)

No RCT reported on this outcome.

Hirsutism - subjective improvement

No RCT reported on this outcome.

13.2 Adverse events: severe (requiring stopping of medication) (gastro-intestinal and others)

13.2.1 Gastro-intestinal

One trial including 36 women compared the OCP versus metformin combined with the OCP and reported severe gastro-intestinal adverse events (Hoeger 2008b).

There was uncertainty as to whether there was a difference between the OCP and metformin combined with the OCP for severe gastro-intestinal adverse event (OR 1.00, 95% CI 0.06 to 17.33, 1 RCT, n = 36, very low-quality evidence, Analysis 13.2). This suggests that for a severe gastro-intestinal adverse event rate of 5.6% following metformin combined with the OCP, the severe gastro-intestinal adverse event rate following the OCP would be between 0.4% and 50.5%.

Adverse events - minor (gastro-intestinal and others)

No RCT reported on this outcome.

Secondary outcomes

Improved menstrual pattern (i.e. shortening of intermenstrual days)

No RCT reported on this outcome.

Improved menstrual pattern (i.e. an initiation of menses or cycle regularity)

No RCT reported on this outcome.

Acne - visual analogue scale (VAS)

No RCT reported on this outcome.

Acne - subjective improvement

No RCT reported on this outcome.

Diagnosis of Type II diabetes mellitus

No RCT reported on this outcome.

Body weight (kg)

No RCT reported on this outcome.

13.3 Body mass index (kg/m^2)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported BMI (Hoeger 2008b).

There was uncertainty as to whether there was a difference between the OCP and metformin combined with the OCP for BMI (MD 1.50, 95% CI -1.63 to 4.63, 1 RCT, n = 32, very low-quality evidence, Analysis 13.3). This suggests that for a mean BMI of 32.4 following metformin combined with the OCP, the mean BMI following the OCP would be between 1.63 lower to 4.63 higher.

Sensitivity analysis restricting studies to low risk of bias (defined in methods section 'Sensitivity analysis') was unable to be performed due to no RCTs having a low risk of bias.

13.4 Blood pressure - systolic (mmHg)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported systolic blood pressure (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for systolic blood pressure (MD -3.50, 95% CI -12.65 to 5.65, 1 RCT, n = 32, Analysis 13.4).

13.5 Blood pressure - diastolic (mmHg)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported diastolic blood pressure (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for diastolic blood pressure (MD -1.00, 95% CI -6.94 to 4.94, 1 RCT, n = 32, Analysis 13.5).

14 The oral contraceptive pill (OCP) compared to metformin combined with the OCP in adolescent women (hormonal parameters)

14.1 Serum total testosterone (nmol/L)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported serum total testosterone (Hoeger 2008b)

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for serum total testosterone (MD 0.37, 95% CI -0.29 to 1.03, 1 RCT, n = 32, Analysis 14.1).

14.2 Free androgen index (FAI) (%)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported FAI (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for FAI (MD 0.70, 95% CI -0.32 to 1.72, 1 RCT, n = 32, Analysis 14.2).

15 The oral contraceptive pill (OCP) compared to metformin combined with the OCP in adolescent women (metabolic parameters)

Fasting insulin (mLU/L)

No RCT reported on this outcome.

15.1 Fasting glucose (mmol/L)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported fasting glucose (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for fasting glucose (MD 0.03, 95% CI -0.23 to 0.29, 1 RCT, n = 32, Analysis 15.1).

15.2 Total cholesterol (mmol/L)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported total cholesterol (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for total cholesterol (MD -0.50, 95% CI -1.32 to 0.32, 1 RCT, n = 32, Analysis 15.2).

15.3 High-density lipoprotein (HDL) cholesterol (mmol/L)

One trial including 32 women compared the OCP versus metformin combined with the OCP and reported HDL cholesterol (Hoeger 2008b)

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for HDL cholesterol (MD 0.23, 95% CI -0.00 to 0.46, 1 RCT, n = 32, Analysis 15.3).

15.4 Low-density lipoprotein (LDL) cholesterol (mmol/L)

One trial including 32 women compared metformin combined with the OCP versus the OCP and reported LDL cholesterol (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for LDL cholesterol (MD -0.29, 95% CI -0.99 to 0.41, 1 RCT, n = 32, Analysis 15.4).

15.5 Triglycerides (mmol/L)

One trial including 32 women compared metformin combined with the OCP versus the OCP and reported triglycerides (Hoeger 2008b).

There was insufficient evidence to determine whether there was a difference between the OCP and metformin combined with the OCP for triglycerides (MD 0.02, 95% CI -0.68 to 0.72, 1 RCT, n = 32, Analysis 15.5).

DISCUSSION

Summary of main results

Metformin compared to the oral contraceptive pill (OCP) in adult women

Main outcomes

Metformin, when compared with the OCP, may be less effective in improving hirsutism in the subgroup BMI > 25/kg2 < BMI 30kg/m^2 but we are uncertain of the effect of metformin compared to the OCP on hirsutism in subgroups body mass index (BMI) < 25 kg/m^2 and BMI > 30 kg/m^2. Metformin may result in a higher incidence of severe gastro-intestinal side-effects and a lower incidence of severe other side effects compared to the OCP. There were no trials reporting on minor adverse events. Metformin may be less effective in improving menstrual pattern compared to the OCP by lengthening of intermenstrual days. In terms of an initiation of menses or cycle regularity, Metformin compared to the OCP may be less effective in improving menstrual pattern in subgroup BMI ≤ 25 kg/m^2 but we are uncertain if metformin is less effective in improving menstrual pattern in subgroup BMI > 25 kg/m^2 < 30kg/m^2 and uncertain of the effect of metformin compared to the OCP in subgroup BMI ≥ 30 kg/m^2 and subgroup BMI not stated. There was uncertainty as to whether there was a difference between metformin and the OCP for acne. We are uncertain if metformin decreases BMI compared to the OCP in subgroup BMI < 25 kg/m^2 and subgroup BMI > 30 kg/m^2,whist there is uncertainty in the effect of metformin compared to the OCP on BMI in subgroup BMI > 25 kg/m^2 < 30 kg/m^2. The quality of the evidence for all reported main outcomes between these two interventions were either low (indicating that our confidence in the effect estimate is limited and that the true effect may be substantially different from the estimate of the effect), or very low (indicating that we have very little confidence in the effect estimate and that the true effect is likely to be substantially different from the estimate of effect).

Other outcomes

Metformin, when compared to the OCP, resulted in an improvement of the clinical parameters systolic blood pressure with insufficient evidence to determine whether there was a difference in diagnosis of type 2 diabetes mellitus, diastolic blood pressure and body weight. Metformin was less effective in improving hormonal parameters serum total testosterone and free androgen index (FAI) compared to the OCP. In terms of metabolic parameters, metformin resulted in an improvement in fasting insulin, glucose, total cholesterol and triglycerides with insufficient evidence to determine whether there was a difference in fasting high-density lipoprotein (HDL) or low-density lipoprotein (LDL) cholesterol when compared to the OCP.

Metformin compared to metformin combined with the OCP in adult women

Main outcomes

Metformin alone when compared with *m*etformin combined with the OCP, may be less effective in improving hirsutism, but we are uncertain as to whether there is a difference between these two interventions for severe gastro-intestinal adverse events and severe other adverse events requiring stopping medication. There were no trials reporting on minor adverse events, menstrual pattern or acne. Metformin may improve BMI compared to metformin combined with the OCP. The quality of the evidence for all reported main outcomes between these two interventions was low.

Other outcomes

Metformin, when compared to metformin combined with the OCP, resulted in an improvement of the clinical parameters body weight, systolic blood pressure, and diastolic blood pressure with no randomised controlled trials (RCTs) reporting on diagnosis of type 2 diabetes mellitus. Metformin alone was less effective in improving hormonal parameters serum total testosterone and FAI compared to metformin combined with the OCP. In terms of metabolic parameters, metformin resulted in an improvement in fasting insulin, glucose, total cholesterol (in the presence of

unexplained substantial heterogeneity), and triglycerides with insufficient evidence to determine whether there was a difference in fasting HDL or LDL cholesterol when compared to metformin combined with the OCP.

The OCP compared to metformin combined with the OCP in adult women

Main outcomes

The OCP alone when compared with metformin combined with the OCP, may be less effective in improving hirsutism. The OCP may result in a lower incidence of severe gastro-intestinal side effects, but there is uncertainty as to whether there is a difference for other severe adverse events when compared to metformin combined with the OCP. The OCP may have a lower incidence of minor (gastro-intestinal) adverse events compared to metformin combined with the OCP. There were no trials reporting on menstrual pattern. The OCP alone may slightly improve acne compared to metformin combined with the OCP. We are uncertain as to whether there is a difference between these two interventions for BMI. The quality of the evidence for all reported main outcomes between these two interventions were low except for the outcome of BMI which had very low-quality evidence .

Other outcomes

There was insufficient evidence to determine whether there was a difference in the clinical parameters body weight, systolic blood pressure, and diastolic blood pressure when comparing the OCP to metformin combined with the OCP, with no RCTs reporting on diagnosis of type 2 diabetes mellitus. The OCP was less effective in improving hormonal parameters serum total testosterone and FAI compared to metformin combined with the OCP. In terms of metabolic parameters, the OCP was less effective in improving fasting insulin and glucose with insufficient evidence to determine whether there was a difference in fasting lipids (total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) when compared to metformin combined with the OCP.

Metformin compared to the OCP in adolescent women

Main outcomes

We are uncertain as to whether there is a difference between metformin, when compared to the OCP, in the outcomes hirsutism, adverse events (severe other requiring stopping medication and minor gastro-intestinal), menstrual pattern, acne (no trials reporting this outcome) and BMI. The quality of the evidence for all reported main outcomes between these two interventions were very low due primarily to very serious imprecision as the evidence for all reported main outcomes derived from a single RCT.

Other outcomes

Metformin, when compared to the OCP, resulted in an improvement of the clinical parameter body weight with insufficient evidence to determine whether there was a difference in systolic or diastolic blood pressure and no RCTs reporting on diagnosis of type 2 diabetes mellitus. Metformin was less effective in improving the hormonal parameter FAI when compared to the OCP, with insufficient evidence to determine whether there was a difference in serum total testosterone. In terms of metabolic parameters, metformin resulted in an improvement in fasting total and LDL cholesterol with insufficient evidence to determine whether there was a difference in fasting insulin, glucose, HDL cholesterol and triglycerides when compared to the OCP.

Metformin compared to metformin combined with the OCP in adolescent women

There were no trials identified comparing metformin versus metformin combined with the OCP in adolescent women reporting on the outcomes for this review.

The OCP compared to metformin combined with the OCP in adolescent women

Main outcomes

We are uncertain as to whether there is a difference between the OCP alone when compared with metformin combined with the OCP, in the outcomes hirsutism, adverse events (both severe requiring stopping medication and minor (no trials reporting this latter outcome)), menstrual pattern (no trials reporting this outcome), acne (no trials reporting this outcome), and BMI. The quality of the evidence for all reported main outcomes between these two interventions were very low due primarily to very serious imprecision as the evidence for all reported main outcomes derived from a single RCT.

Other outcomes

There was insufficient evidence to determine whether there was a difference in the clinical parameters systolic and diastolic blood pressure with no RCTs reporting on diagnosis of type 2 diabetes mellitus and body weight. There was insufficient evidence to determine whether there was a difference in the hormonal parameters serum total testosterone and FAI when comparing the OCP with metformin combined with the OCP. In terms of metabolic parameters, there was also insufficient evidence to determine whether there was a difference in fasting glucose and lipids (total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) when comparing the OCP with metformin combined with the OCP, with no RCTs reporting fasting insulin.

Overall completeness and applicability of evidence

In adult women with PCOS, we found sufficient studies reporting on our main clinical outcomes of hirsutism, menstrual pattern and adverse events to answer our main review questions for our main comparison of metformin compared to the OCP. However, there were only two RCTs reporting on the other main clinical outcome of acne for this comparison. For the other comparisons for this review (metformin versus metformin in combination with the OCP; and the OCP versus metformin in combination with the OCP), there was only a small number of RCTs reporting on our main clinical outcomes. In adolescent women with PCOS, there was only a small number of RCTs reporting on our main clinical outcomes for all three comparisons. Unfortuately, the predominantly low or very low quality of the evidence provided by the included studies limited the confidence we have in the effect estimate for our main clinical outcomes for all three comparisons in both adult and adolescent women with PCOS.

In terms of the applicability of the evidence with respect to the participants, this review includes 44 RCTs with 2253 PCOS women who met the Rotterdam diagnostic criteria for PCOS (Rotterdam

ESHRE 2004) in 43/44 studies as 1/44 studies did not described any diagnostic criteria. 2047 adult women were recruited in 39/44 studies and 206 adolescent women were recruited in 5/44 studies. The women with PCOS were recruited in and thus representative of a number of continents around the world including Europe (see section on description of studies for list of European countries), North America (USA), Africa (Egypt), Asia (Iran, Iraq and Turkey) and Australia.

In terms of the applicability of the evidence with respect to reporting results per pre-specified population mean BMI subgroups (e.g. BMI ≤ 25kg/m^2, BMI > 25kg/m^2 but < 30kg/m^2, BMI ≥30kg/m^2) in the presence of explained (by the BMI subgroup) significant heterogeneity, it is important to note that when interpreting these subgroup findings that the subgroups were defined on this basis of mean BMI, rather than on the basis of a population restricted to a certain BMI range according to strict study inclusion criteria. Participants in the particular subgroup may be very heterogeneous with some having a BMI much lower or higher than the mean BMI and thus does not imply that participants in the study generally have a BMI close to that value. Therefore, caution is required in that the translation of findings into treatment recommendations from a population with a particular mean BMI is different from saying that this is the effect in a population restricted to a given BMI range and thus should not be misinterpreted as showing the latter.

In terms of the applicability of the evidence with respect to the interventions, all the studies compared the relevant interventions of metformin versus the OCP (alone or in combination). There were no trials identified comparing metformin with metformin combined with the OCP in adolescent women on the selected outcomes for this review. The daily dose of metformin ranged from 500 mg to 2000 mg in the included studies. This review compared OCPs as a group, however, different types and doses of oestrogen and progesterone were used. Indeed, 31 trials used ethinyl estradiol (EE) 35 µg combined with cyproterone acetate (CPA) 2 mg (EE 35/CPA2) as the OCP, four trials used EE 35 µg combined with drospirenone 3 mg, three trials used EE 35 µg combined with desogestrel 150, three trials used EE 35 µg combined with norgestimate 0.25, one trial used EE 30 µg with norethisterone 1mg, one trial used EE 20 µg with drospirenone 3 µg, and one trial used EE 30 µg with progestin 15 mg. Therefore, the results of this review are generally applicable to these specific types of the OCP.

In terms of the applicability of the evidence with respect to the outcomes, the review investigated the main clinically important outcomes of hirsutism, acne, improvement of menstrual pattern, BMI and adverse events for the three comparisons (the OCP versus metformin, metformin versus metformin combined with the OCP, and the OCP versus metformin combined with the OCP) in both adolescent and adult women with PCOS separately. The majority of the outcomes reported in the studies were hirsutism, improvement of menses (the OCP versus metformin), BMI and adverse outcomes with only two studies reporting on acne. There was only one study identified reporting on the diagnosis of type 2 diabetes mellitus (the OCP versus metformin). The review also investigated the important surrogate outcomes related to the clinically important outcomes including hormonal (serum total testosterone and FAI) and metabolic (fasting insulin, glucose and lipids [total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) parameters.

There were four studies, all recruiting adult women with PCOS, that measured outcomes but had no 'usable' data. Two of these studies were for the comparison of metformin versus the OCP, one study was for the outcome of 'Hirsutism - Subjective improvement' and the other study was for the outcome of 'Acne - Subjective improvement'. One study was for the comparison of metformin versus metformin combined with the OCP for the outcome of severe adverse events. The last such study was for the comparison of the OCP versus metformin combined with the OCP for the comparison of minor adverse events.

The review findings both support and will help guide the current common practice of using the OCP and metformin, either alone or in combination, in the long-term treatment of both adult and adolescent women with PCOS.

Quality of the evidence

Evidence quality for the main outcomes of the review ranged from very low to low based on GRADE assessment. The main limitations were risk of bias (very few studies at low risk of bias defined as low risk of selection bias (both random sequence generation and allocation concealment) and not at high risk of bias in any domain), imprecision and inconsistency. Only two of 44 studies in total were judged to be at low risk of bias (Feng 2016; Harborne 2003). This review had six comparisons as follows.

1. Metformin compared to the OCP in adult women.

The quality of the evidence for all reported main outcomes between these two interventions was either low (indicating that our confidence in the effect estimate is limited and that the true effect may be substantially different from the estimate of the effect) or very low (indicating that we have very little confidence in the effect estimate and that the true effect is likely to be substantially different from the estimate of effect).

2. Metformin compared to metformin combined with the OCP in adult women.

The quality of the evidence for all reported main outcomes between these two interventions was low .

3. The OCP compared to metformin combined with the OCP in adult women.

The quality of the evidence for all reported main outcomes between these two interventions was low except for the outcome of BMI which had very low-quality evidence.

4. Metformin compared to the OCP in adolescent women.

The quality of the evidence for all reported main outcomes between these two interventions was very low due primarily to very serious imprecision as the evidence for all reported main outcomes derived from a single RCT.

5. Metformin compared to metformin combined with the OCP in adolescent women.

There were no trials identified comparing metformin versus metformin combined with the OCP in adolescent women reporting on the outcomes for this review.

6. The OCP compared to metformin combined with the OCP in adolescent women.

The quality of the evidence for all reported main outcomes between these two interventions was very low due primarily to very serious imprecision as the evidence for all reported main outcomes derived from a single RCT.

We observed significant (substantial) heterogeneity in many of the analyses, which was explained in most cases by the pre-defined subgroup analysis of mean study BMI. Unexplained substantial heterogeneity was observed for the outcomes of BMI (the OCP versus metformin combined with the OCP in adults), systolic and diastolic blood pressure (the OCP versus metformin combined with the OCP in adults), fasting total cholesterol (metformin versus metformin combined with the OCP in adults), fasting total cholesterol (the OCP versus metformin combined with the OCP in adults) and fasting HDL cholesterol (the OCP versus metformin combined with the OCP in adults). Of these outcomes, all showed insufficient evidence to determine whether there was a difference between the interventions except for fasting total cholesterol (metformin versus metformin combined with the OCP in adults), which favoured the intervention of metformin and the result being unchanged with random-effects meta-analysis. Unexplained substantial heterogeneity resulted in the downgrading of evidence quality for the main review outcome of BMI (the OCP versus metformin combined with the OCP in adults).

Potential biases in the review process

We conducted a comprehensive search with the help of an experienced Information Specialist, and ran extensive manual searches in order to identify all relevant studies and in an effort to minimise the risk of publication bias. At the review level, we performed funnel plots for the main review outcomes if there were 10 or more studies in an analysis in order to explore the possibility of small-study effects that may indicate incomplete identification of studies (publication bias). Four funnel plots were performed in total (metformin versus the OCP in adult women for the outcomes of hirsutism - Clinical Ferriman-Gallwey (F-G) score, severe adverse events and BMI; the OCP versus metformin combined with the OCP for the outcome of BMI). Two of the funnel plots were symmetrical (metformin versus the OCP in adult women for the outcome of severe adverse events; the OCP versus metformin combined with the OCP for the outcome of BMI), indicating that our findings might not be influenced by publication bias. The other two funnel plots (metformin versus the OCP in adult women for the outcomes of hirsutism - Clinical F-G score and BMI) were asymmetrical indicating that our findings might be influenced by publication bias although the asymmetrical funnel plots for these two outcomes may also be due to the substantial heterogeneity (explained by mean study BMI subgroups) observed for these two outcomes.

We followed Cochrane guidelines to search for and identify all the studies eligible for this review, extract data and assess the quality and potential risks of different types of biases in all our included studies, in order to minimise the chance of error and bias by the review authors. Subjective judgements are involved in the assessment of risk of bias. This potential limitation is minimised by following the procedures in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2017), with review authors independently assessing studies and resolving any disagreement through discussion, and if required involving a third review author in the decision.

There are four studies awaiting classification (Fruzzeti 2009; NCT01573377; Spremovic-Radjenovic 2014; Vieira 2010), and three ongoing clinical trials (NCT02744131; NCT03229057; NCT03905941) that are likely to provide further useful information.

Agreements and disagreements with other studies or reviews

A recent systematic review and meta-analysis of RCTs assessed the efficacy and safety of therapeutic approaches for adult patients with PCOS not seeking fertility including OCPs, antiandrogens (AA) and/or insulin sensitisers including metformin (Luque-Ramirez 2018). This review did not compare metformin versus the OCP as such, but compared metformin versus the OCP and/or AAs, and agreed with our main outcome findings of a reduction in hirsutism and an improved menstrual pattern for the OCP and/or AAs along with a reduction in BMI for metformin. The outcomes of acne and adverse events were not reported. The global quality of evidence was low for the review, which was consistent with our findings. This review also compared the OCP and/or AAs versus metformin combined with the OCP and/or AAs, and reported on only one of our main outcomes for our review which was BMI. The review found a reduction in BMI with the addition of metformin to the OCP and/or AAs (low-quality evidence), which disagreed with our findings, although our review did not compare exactly the same interventions as we did not include studies combining AA's with the OCP.

A more recent systematic review and meta-analysis of RCTs aimed to investigate the effect of the OCP and/or metformin in the management of hormonal and clinical features of PCOS in order to inform the recent international guidelines on PCOS (Teede 2018; Teede 2019). This review restricted studies to those published in English language, used stricter study selection criteria in terms of including a RCT in terms of random sequence generation and allocation concealment and different methods for data synthesis (random-effects model), and grading the quality of evidence (according to risk of bias with each study being allocated a risk of bias rating of either low, moderate of high) compared to our review, and did not report on the outcome of acne. For the comparison metformin versus the OCP, this review's findings agreed with our main outcomes findings with respect to an improvement in menstrual pattern with the OCP, but differed by finding no statistically significant difference between the two interventions for hirsutism, BMI, gastro-intestinal side effects (all i.e. minor and severe) and other side effects (all i.e. minor and severe). The majority of included studies for this comparison were judged to be of moderate quality, and therefore, the authors recommended that the findings should be interpreted with some degree of caution.

The review by Teede and colleagues also compared metformin versus metformin + the OCP and reported on only one of our review main outcomes, BMI, and found no statistically significant difference between the two interventions and thus differed from our review. There were no RCTs reporting on adverse events (Teede 2019). The authors advised caution when interpreting the results due to moderate risk of bias across all studies. The same review also compared the OCP versus metformin + the OCP and found, in terms of the main outcomes for our review, no statistically significant difference between the two interventions

for BMI (agreement with our review), gastro-intestinal severe side effects (disagreement with our review) and gastro-intestinal minor side effects (disagreement with our review). The RCTs were of moderate quality with small sample sizes (only 1 RCT reporting on gastro-intestinal severe or minor side-effects) with the authors advising that the results should be interpreted with caution.

The recently published international guideline on PCOS made recommendations based on Grading of Recommendations, Assessment, Development and Evaluation (GRADE) framework covering evidence quality, balance between desirable versus undesirable effects (direction and magnitude), feasibility, acceptability, cost, implementation and ultimately recommendation strength (Teede 2018). In terms of pharmacological treatment for non-fertility indications in adult women with PCOS, the guideline, based on low-quality evidence, made a strong recommendation for the use of the OCP alone for management of hyperandrogenism and/or irregular menstrual cycles, conditional recommendation for the use of metformin (in addition to lifestyle) for the treatment of weight, hormonal and metabolic outcomes, and a strong recommendation for the use of the OCP combined with metformin for management of metabolic features where the OCP and lifestyle changes do not achieve desired goals. Our review findings are likely to be consistent with these recommendations given that the recommendations are based on more factors than just the evidence (both direction and magnitude in addition to quality), but also take into consideration the balance between desirable versus undesirable consequences, feasibility, acceptability, cost, and implementation.

In terms of specifically adolescents with PCOS, two recent systematic review and meta-analysis of RCTs have evaluated the use of metformin versus the OCP (but not combinations of metformin and/or the OCP) for the treatment of PCOS in adolescents (Al Khalifah 2016, Teede 2019). Both of these reviews included the same RCTs that were included for our review and both assessed the quality of evidence as very low or low, and as a result identified the need for future high-quality RCTs to address several questions for the treatment of adolescents with PCOS, which is in accordance with our conclusions. The latter review also advised caution in the interpretation of the results due to the low quality of the evidence, whilst the former review, in light of the very low/low-quality evidence, concluded that treatment choice should be guided by patient values and preferences while balancing potential adverse events. This conclusion was performed in order to formulate the recommendations in the recently published international guideline on PCOS which resulted in conditional recommendations (low-quality evidence) to consider the use of both metformin alone or the OCP alone in adolescents with PCOS as pharmacological treatment for non-fertility indications (Teede 2018). These conditional recommendations of the international guideline in PCOS reflect our review's findings of there being uncertainty as to whether there is a difference between metformin, when compared to the OCP, in the main outcomes in adolescents with PCOS.

AUTHORS' CONCLUSIONS

Implications for practice

In adult women with PCOS, metformin may be less effective in improving hirsutism compared to the oral contraceptive pill (OCP) in the subgroup body mass index (BMI) > 25 /kg2 < BMI 30 kg/m^2,but we are uncertain if there was a difference between metformin and the OCP in subgroup BMI < 25 kg/m^2 and subgroup BMI > 30 kg/m^2. Compared to the OCP, metformin may result in a higher incidence of severe gastro-intestinal adverse events and lower incidence of severe other adverse events with no trials reporting on minor adverse events. Either metformin alone or the OCP alone may be less effective in improving hirsutism compared to metformin combined with the OCP. We are uncertain as to whether there is a difference between the OCP alone and metformin alone compared to metformin combined with the OCP for adverse events (severe or minor; gastro-intestinal or other), except for the OCP versus metformin combined with the OCP where the OCP may result in a lower incidence of severe and minor gastro-intestinal adverse events .

In adolescent women with PCOS, we are uncertain as to whether there is a difference between any of the three comparisons for this review for the primary outcomes of hirsutism and adverse events (both severe requiring stopping medication and minor) due to either no evidence or very low-quality evidence, primarily due to very serious imprecision as the evidence for all reported primary outcomes derived from a single randomised controlled trial (RCT). Further large well-designed and conducted RCTs that stratify for BMI are needed to evaluate metformin versus the OCP and combinations of these interventions in women with PCOS, in particular adolescent women.

Implications for research

The evidence quality for this review ranged from very low to low based on GRADE assessment with the vast majority of the evidence quality being low or very low limiting the certainty of all of the results in this review. Therefore, further high-quality adequately-powered RCTs are required comparing metformin versus the OCP, metformin versus metformin combined with the OCP, and the OCP versus metformin combined with the OCP in adult and adolescent women with PCOS stratified by BMI to assess the important clinical outcomes of hirsutism, improvement of menstrual pattern, acne, BMI, diagnosis of type 2 diabetes and adverse events in order to increase the quality of the body of evidence on these comparisons and outcomes, and provide more certainty in the benefit and harms of these long-term medical treatment interventions.

ACKNOWLEDGEMENTS

The authors wish to thank Cochrane Gynaecology and Fertility for their assistance with updating this systematic review, in particular Helen Nagels for her advice and guidance in writing this review. The authors also wish to thank Dr Jinzhu Liu for her help in translating the RCTs in Chinese language and Mohammed Saleem Sultani for his contribution to selection and extraction of the data in addition to evaluating the quality of the trials.

We would like to thank peer reviewers Jack Wilkinson, Andrew Watson, Harry Siristatidis, Euphrasia Ebai-Atuh Ndi, and co-ordinating editor Madelon van Waly for their valuable comments.

REFERENCES

References to studies included in this review

Aghamohammadzadeh 2010 *{published data only}*

Aghamohammadzadeh N, Aliasgarzadeh A, Baglar L, Abdollahifard S, Bahrami A, Najafipour F, et al. Comparison of metformin and cyproterone-estrodiol compound effect on HS C-reactive protein and serum androgen levels in patients with polycystic ovary syndrome. *Pakistan Journal of Medical Science* 2010;**26**(2):347-51.

Allen 2005 *{published data only}*

Allen HF, Mazzoni C, Heptulla RA, Murray MA, Miller N, Koenigs L, et al. Randomized controlled trial evaluating response to Metformin versus standard therapy in the treatment of adolescents with polycystic ovary syndrome. *Journal of Pediatric Endocrinology and Metabolism* 2005;**18**:761-8.

Al-Zubeidi 2015 *{published data only}*

Al-Zubeidi H, Klein KO. Randomized clinical trial evaluating Metformin versus oral contraceptive pills in the treatment of adolescents with polycystic ovarian syndrome. *Journal of Pediatric Endocrinology and Metabolism* 2015;**28**:853-8.

Bhattacharya 2016 *{published data only}*

Bhattacharya S M, Biswas B, Basu A. Comparative study of the therapeutic effects of drospirenone (DRSP, 3mg) with ethinyl estradiol (EE, 20 mcg) pill with or without metformin 500 mg in polycystic ovary syndrome (PCOS). In: Fertility and Sterility. 2016, poster P-401:e257.

Bodur 2018 *{published data only}*

Bodur S, Dundar O, Kanat-Pektas M, Kinci MF, Tutuncu L. The effects of different therapeutic modalities on cardiovascular risk factors in women with polycystic ovary syndrome: a randomized controlled study. *Taiwanese Journal of Obstetrics and Gynecology* 2018;**57**:411-6.

Cetinkalp 2009 *{published data only}*

Cetinkalp S, Karadeniz M, Erdogan M, Ozgen G, Saygili F, Yilmaz C. The effects of rosiglitazone, metformin, and estradiol-cyproterone acetate on lean patients with polycystic ovary syndrome. *Endocrinologist* 2009;**19**:94-7.

Christakou 2014 *{published data only}*

Christakou C, Kollias A, Piperi C, Katsikis I, Panidis D, Diamanti-Kandarakis E. The benefit-to-risk ratio of common treatments in PCOS: effect of oral contraceptives versus metformin on atherogenic markers. *Hormones* 2014;**13**:488-97.

Cibula 2005 *{published data only}*

Cibula D, Fanta M, Vrbikova J, Stanicka S, Dvorakova K, Hill M, et al. The effect of combination therapy with metformin and combined oral contraceptives (COC) versus COC alone on insulin sensitivity, hyperandrogenaemia, SHBG and lipids in PCOS patients. *Human Reproduction* 2005;**20**(1):180-4.

Dardzinska 2014 *{published data only}*

Dardzinska JA, Rachon D, Kuligowska-Jakuboswka M, Aleksandrowicz-Wrona E, Ploszynski A, Wyrzykowski B, et al. Effects of metformin or an oral contraceptive containing cyproterone acetate on serum C-reactive protein, Interleukine-6 and Soluble Vascular Cell Adhesion Molecule-1 concentrations in women with polycystic ovary syndrome. *Experimental and Clinical Endocrinology & Diabetes* 2014;**122**:118-25.

El Maghraby 2015 *{published data only}*

El Maghraby HA, Nafee T, Guiziry D, Elnashar A. Randomized controlled trial of the effects of metformin versus combined oral contraceptives in adolescent PCOS women through a 24 months follow up period. *Middle East Fertility Society Journal* 2015;**20**:131-7.

Elter 2002 *{published data only}*

Elter K, Imir G, Durmusoglu F. Clinical, endocrine and metabolic effects of metformin added to ethinyl-estradiol-cyproterone acetate in non-obese women with polycystic ovarian syndrome: a randomized controlled study. *Human Reproduction* 2002;**17**(7):1729-37.

Essah 2011 *{published data only}*

Essah PA, Arrowood A, Cheang KI, Adawadkar SS, Stovall DW, Nestler JE. Effect of combined metformin and oral contraceptive therapy on metabolic factors and endothelial function in overweight and obese women with polycystic ovary syndrome. *Fertility and Sterility* 2011;**96**(2):501-4.

Feng 2016 *{published data only}*

Feng W, Jia Y-Y, Zhang D-Y, Shi H-R. Management of polycystic ovarian syndrome with Diane-35 or Diane-35 plus metformin. *Gynecological Endocrinology* 2016;**32**(2):147-50.

Glintborg 2014a *{published data only}*

Glintborg D, Altinok ML, Mumm H, Hermann AP, Ravn P, Andersen M. Body composition is improved during 12 month's treatment with Metformin alone or combined with oral contraceptives compared with treatment with oral contraceptives in polycystic ovary syndrome. *Endocrine Research* 2014;**99**:2584-91.

Harborne 2003 *{published and unpublished data}*

Harborne L, Fleming R, Lyall H, Sattar N, Norman J. Metformin or antiandrogen in the treatment of hirsutism in polycystic ovary syndrome. *Journal of Clinical Endocrinology & Metabolism* 2003;**88**(9):4116-23.

Hoeger 2008a *{published data only}*

Hoeger K, Davidson K, Kochman L, Cherry T, Kopin L, Guzick S. The impact of metformin, oral contraceptives, and lifestyle modification on polycystic ovary syndrome in obese adolescent women in two randomized, placebo-controlled clinical trials. *Journal of Clinical Endocrinology and Metabolism* 2008;**93**:4299-306.

Hoeger 2008b *{published data only}*

Hoeger K, Davidson K Kochman L, Cherry T, Kopin L, Guzick S. The impact of metformin, oral contraceptives, and lifestyle modification on polycystic ovary syndrome in obese adolescent women in two randomized, placebo-controlled clinical

trials. *Journal of Clinical Endocrinology and Metabolism* 2008;**93**:4299-306.

Jin 2006 *{published data only}*

Jin LN, Yu Q, Ma L K. The effect of combination therapy with Metformin and combined Cyproterone Acetate/ Ethinyl Estradiol in PCOS patients. *Chinese Journal of Postgraduate Medicine* 2006;**29**:10-5.

Kaya 2015 *{published data only}*

Kaya MG, Yildirim S, Calapkorur B, Akpek M, Unluhizarci K, Kelestimur F. Metformin improves endothelial function and carotid intima media thickness in patients with PCOS. *Gynecological Endocrinology* 2015;**31**:401-5.

Kebapcilar 2009a *{published data only}*

Kebapcilar L, Taner CE, Kebapcilar AG, Alacacioglu A, Sari I. Comparison of four different treatment regimens on coagulation parameters, hormonal and metabolic changes in women with polycystic ovary syndrome. *Archives of Gynecology and Obstetrics* 2009;**281**:35-42.

Kebapcilar 2009b *{published data only}*

Bilgir O, Kebapcilar L, Taner C, Bilgir F, Kebapcilar A, Bozkaya G, et al. The effect of ethinylestradiol (EE)/cyproterone acetate (CA) and EE/CA plus metformin treatment on adhesion molecules in cases with polycystic ovary syndrome (PCOS). *International Medicine* 2009;**48**:1193-9.

Kebapcilar L, Yuksel A, Bozkaya G, Taner CE, Kebapcilar AG, Bilgir O, et al. Effects of an EE/CA compared with EE/CA-metformin on serum ADMA levels in women with polycystic ovary syndrome. *Central European Journal of Medicine* 2009;**4**:423-7.

Kilic 2011 *{published data only}*

Kilic S, Yilmaz N, Zulfikaroglu E, Erdogan G, Aydin M, Batioglu S. Inflammatory-metabolic parameters in obese and nonobese normoandrogenemic polycystic ovary syndrome during metformin and oral contraceptive treatment. *Gynecological Endocrinology* 2011;**27**(9):622-9.

Kuek 2011 *{published data only}*

Kuek S, Wang W-J, Gui S-Q. Efficacy of chinese patent medicine Tian Gui Capsule in patients with polycystic ovary syndrome: a randomized controlled trial. *Journal of Chinese Integrative Medicine* 2011;**9**:965-72.

Kumar 2018 *{published data only}*

Kumar Y, Kotwal N, Singh Y, Upreti V, Somani S, Hari Kumar KV. A randomized, controlled trial comparing the metformin, oral contraceptive pills and their combination in patients with polycystic ovarian syndrome. *Journal of Family Medicine and Primary Care* 2018;**7**:551-6.

Liu 2006 *{published data only}*

Liu JX, Dang JM, Wu YZ, Zheng N, Wu XK. The effect of combination therapy with Metformin and combined Cyproterone Acetate/Ethynil Estradiol on PCOS patients. *Chinese Journal of Postgraduate Medicine* 2006;**19**:530-3.

Luque-Ramirez 2007a *{published data only}*

Luque-Ramirez M, Alvarez-Blasco F, Botella-Carretero JI, Martinez-Bermejo E, Lasuncion MA, Escobar-Morreale HF. Comparison of ethinyl-estradiol plus cyproterone acetate versus metformin effects on classic metabolic cardiovascular risk factors in women with polycystic ovary syndrome. *Journal of Clinical Endocrinology and Metabolism* 04/2007;**92**:2453-61.

Luque-Ramirez 2007b *{published data only}*

Luque-Ramirez M, Alvarez-Blasco F, Botella-Carretero J I, Sanchon R, San Millan JL, Escobar-Morreale HF. Increased body iron stores of obese women with polycystic ovary syndrome are a consequence of insulin resistance and hyperinsulinism and are not a result of reduced menstrual losses. *Diabetes Care* 09/2007;**30**:2309-13.

Luque-Ramirez 2008a *{published data only}*

Luque-Ramirez M, Alvarez-Blasco F, Uriol Rivera MG, Escobar-Morreale HF. Serum uric acid concentration as non-classic cardiovascular risk factor in women with polycystic ovary syndrome: effect of treatment with ethinyl-estradiol plus cyproterone acetate versus metformin. *Human Reproduction* 2008;**23**:1594-601.

Luque-Ramirez 2009b *{published data only}*

Luque-Ramirez M, Mendieta-Azcona C, Alvarez-Blasco F, Escobar-Morreale HF. Effects of Metformin versus Ethinyl-estradiol plus cyproterone acetate on ambulatory blood pressure monitoring and carotid intima media thickness in women with the polycystic ovary syndrome. *Fertility and Sterility* 2008;**91**:2527-36.

Lv 2005 *{published data only}*

Lv L, Liu Y, Sun Y, Tan K. Effects of Metformin combined with cyproterone acetate on clinical features, endocrine and metabolism of non-obese women with polycystic ovarian syndrome. *Journal of Huazhong University of Science and Technology* 2005;**25**:194-7.

Meyer 2007 *{published data only}*

Meyer C, McGrath BP, Teede HJ. Effect of medical therapy on insulin resistance and the cardiovascular system in polycystic ovary syndrome. *Diabetes Care* 2007;**30**(3):471-8.

Mhao 2015 *{published data only}*

Mhao NS, Al-Hilli SA, Hadi NR, Jamil DA, Al-Aubaidy HA. A comparative study to illustrate the benefits of using ethinyl estradiol-cyproterone acetate over metformin in patients with polycystic ovarian syndrome. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews* 2015;**10S**:S95-S98.

Moran 2010 *{published data only}*

Moran LJ, Meyer C, Hutchison SK, Zoungas S, Teede HJ. Novel inflammatory markers in overweight women with or without polycystic ovary syndrome and following pharmacological intervention. *Journal of Endocrinological Investigation* 2010;**33**:258-65.

Morin-Papunen 2000 *{published data only}*

Morin-Papunen LC, Vauhkonen I, Koivunen RM, Ruokonen A, Martikainen HK, Tapanainen JS. Endocrine and metabolic

effects of metformin versus ethinyl-estradiol-cyproterone acetate in obese women with polycystic ovary syndrome: a randomized study. *Journal of Clinical Endocrinology & Metabolism* 2000;**85**(9):3161-8.

Morin-Papunen 2003 *{published data only}*

Morin-Papunen L, Vauhkonen I, Koivunen R, Ruokonen A, Martikainen H, Tapanainen JS. Metformin versus ethinyl estradiol-cyproterone acetate in the treatment of nonobese women with polycystic ovary syndrome: a randomized study. *The Journal of Clinical Endocrinology & Metabolism* 2003;**88**(1):148-56.

Moro 2013 *{published data only}*

Moro F, Morciano A, Tropea A, Sagnella F, Palla C, Scarinci E, et al. Effects of drospirenone-ethinylestradiol and/or metformin on $CD4^{+}$ $CD28^{null}$ T lymphocytes frequency in women with hyperinsulinemia having polycystic ovary syndrome: a randomized clinical trial. *Reproductive Science* 2013;**20**(12):1508-17.

Ozgurtas 2008 *{published data only}*

Ozgurtas T, Oktenli C, Dede M, Tapan S, Kenar L, Sanisoglu SY, et al. Metformin and oral contraceptive treatment reduced circulating asymmetric dimethylarginine (ADMA) levels in patients with polycystic ovary syndrome (PCOS). *Atherosclerosis* 2008;**200**:336-44.

Rautio 2005 *{published data only}*

Rautio K, Tapanainen JS, Ruokonen A, Morin-Papunen LC. Effects of metformin and ethinyl-estradiol-cyproterone acetate on lipid levels in obese and non-obese women with polycystic ovary syndrome. *European Journal of Endocrinology* 2005;**152**:269-75.

Ruan 2018 *{published data only}*

Ruan X, Song J, Gu M, Wang L, Wang H, Mueck AO. Effect of Diane-35, alone or in combination with orlistat or metformin in Chinese polycystic ovary syndrome patients. *Archives of Gynecology and Obstetrics* 2018;**297**:1557-63.

Sahu 2018 *{published data only}*

Sahu A, Tripathy P, Mohanty J, Nagy A. Doppler analysis of ovarian stromal blood flow changes after treatment with metformin versus ethinyl estradiol-cyproterone acetate in women with polycystic ovarian syndrome: a randomized controlled trial. *Journal of Gynecology Obstetrics and Human Reproduction* 2018;**48**(5):335-9.

Song 2017 *{published data only}*

Song J, Ruan X, Gu M, Wang L, Wang H, Mueck AO. Effect of orlistat or metformin in overweight and obese polycystic ovary syndrome patients with insulin resistance. *Gynecological Endocrinology* 2018;**34**:413-7.

Teng 2007 *{published data only}*

Teng YQ, Zheng JH, Han XY. Evaluation of clinical efficacy of Metformin in treating polycystic ovary syndrome. *Journal of Harbin Medical University* 2007;**41**:405-6.

Wei 2012 *{published data only}*

Wei W, Zhao H, Wang A, Sui M, Liang K, Deng H, et al. A clinical study on the short-term effect of berberine in comparison to metformin on the metabolic characteristics of women with polycystic ovary syndrome. *European Journal of Endocrinology* 2012;**166**:99-105.

Wu 2008 *{published data only}*

Wu J, Zhu Y, Jiang Y, Cao Y. Effects of metformin and ethinyl estradiol-cyproterone acetate on clinical, endocrine and metabolic factors in women with polycystic ovary syndrome. *Gynecological Endocrinology* 2008;**24**:392-8.

References to studies excluded from this review

Alpanes 2017 *{published data only}*

Alpanes M, Alvarez-Blasco F, Fernandez-Duran E, Luque-Ramirez M, Escobar-Morreale HF. Combined oral contraceptives plus spironolactone compared with Metformin in women with polycystic ovary syndrome: a one-year randomized clinical trial. *European Journal of Endocrinology* 2017;**177**:399-408.

Altinok 2018 *{published data only}*

Altinok ML, Ravn P, Andersen M, Glintoborg D. Effect of 12-months treatment with metformin and/or oral contraceptives on health-related quality of life in polycystic ovary syndrome. *Gynecological Endocrinology* 2018;**34**(10):859-63.

Bachani 2016 *{published data only}*

Bachani S, Ganju Malla V. Identifying risk factors and screening tests for extent of metabolic risk in urban adolescent girls in a developing nation. *International journal of child and adolescent health* 2016;**9**(3):347-53.

Bhattacharya 2012 *{published data only}*

Battacharya SM, Jha A. Comparative study of the therapeutic effects of oral contraceptive pills containing desogestrel, cyproterone acetate, and drospirenone in patients with polycystic ovary syndrome. *Fertility and sterility* 2012;**98**(4):1053-9.

Bredella 2013 *{published data only}*

Bredella MA, McManus S, Misra M. Impact of metformin monotherapy versus metformin with oestrogen-progesterone on lipids in adolescent girls with polycystic ovarian syndrome. *Clinical Endocrinology* 2013;**79**(2):199-203.

Burchall 2015 *{published data only}*

Burchall GF, Piva TJ, Linden MD, Gibson-Helm ME, Ranasinha S, Teede HJ. Comprehensive assessment of the hemostatic system in polycystic ovarian syndrome. *Seminars in Thrombosis and Hemostasis* 2016;**42**:55-62.

Cakiroglu 2013 *{published data only}*

Cakiroglu Y, Vural B, Isgoren S. The effects of drospirenone-ethinyl estradiol and drospirenone-ethinyl estradiol + metformin on ovarian ultrasonographic markers, body fat mass index, leptin, and ghrelin. *Archive of Gynecology and Obstetrics* 2013;**288**:213-20.

Diaz 2016 *{published data only}*

Diaz M, Gallego-Escuredo JM, de Zegher F. Effect of ethinylestradiol-cyproterone acetate vs. pioglitazone-flutamide-metformin on plasma FGF21 levels in adolescent girls with androgen excess. *Diabetes and Metabolism* 2016;**42**:196-9.

Glintborg 2014b *{published data only}*

Glintborg D, Mumm H, Altinok ML, Richelsen B, Bruun JM, Andersen M. Adiponectin, interleukin-6, monocyte chemoattractant protein-1, and regional fat mass during 12-months randomized treatment with metformin and/or contraceptives in polycystic ovary syndrome. *Journal of Endocrinal Investigation* 2014;**37**:757-64.

Glintborg 2015 *{published data only}*

Glintborg D, Sidelmann JJ, Altinok ML, Mumm H, Andersen M. Increased thrombin generation in women with polycystic ovary syndrome. A pilot study on the effect of metformin and oral contraceptives. *Metabolism* 2015;**64**:1272-8.

Glintborg 2017 *{published data only}*

Glintborg D, Mumm H, Holst JJ, Andersen M. Effect of oral contraceptives and/or metformin on GLP-1 secretion and reactive hypoglycemia in polycystic ovary syndrome. *Endocrine connections* 2017;**6**:267-277.

Hadziomerovic-Pekic 2010 *{published data only}*

Hadziomerovic-Pekic D, Wildt L, Weiss JM, Moeller K, Mattle V, Seeber BE. Metformin, naltrexone, or the combination of prednisolone and antiandrogenic oral contraceptives as first-line therapy in hyperinsulinemic women with polycystic ovary syndrome. *Fertility and Sterility* 2010;**94**(6):2385-8.

Harris-Glocker 2009 *{published data only}*

Harris-Glocker M, Davidson K, Kochman L, Guzick D, Hoeger K. Improvement in quality of life questionnaire measures (PCOSQ) in obese adolescent females with PCOS treated with Lifestyle changes and oral contraceptives, with or without Metformin. *Fertility and Sterility* 2009;**93**(3):1016-9.

Hu 2010 *{published data only}*

Hu Z, Wang Y, Qiao J, Li M, Chi H, Chen X. The role of family history in clinical symptoms and therapeutic outcomes of women with polycystic ovary syndrome. *International Journal of Gynecology and Obstetrics* 2010;**108**:35-9.

Hutchison 2008 *{published data only}*

Hutchison SK, Harrison C, Stepto N, Meyer C, Teede HJ. Retinol-binding protein 4 and insulin resistance in polycystic ovary syndrome. *Cardiovascular and Metabolic Risk* 2008;**31**:1427-32.

Ibanez 2010 *{published data only}*

Ibanez L, Lopez-Bermejo A, Diaz M, Enriquez G, del Rio L, de Zegher F. Low-dose pioglitazone, flutamide, metformin plus an oestro-progestagen for non-obese young women with polycystic ovary syndrome: increasing efficacy and persistent safety over 30 months. *Gynecological Endocrinology* 2010;**26**(12):869-73.

Ibanez 2017 *{published data only}*

Ibanez L, del Rio L, Diaz M, Sebastiani G, Pozo OJ, Lopez-Bermejo A, et al. Normalizing ovulation rate by preferential reduction of hepato-visceral fat in adolescent girls with polycystic ovary syndrome. *Journal of Adolescent Health* 2017;**61**:446-53.

Kebapcilar 2010 *{published data only}*

Kebapcilar L, Kebapcilar AG, Bilgir O, Taner CE, Bozkaya G, Yilidiz Y, et al. Metformin plus oral contraceptive may decrease plasma sCD40 ligand in women with PCOS patients. *Gynecological Endocrinology* 2010;**27**(2):91-5.

Kim 2010 *{published data only}*

Kim C-H, Jeon G-H, Kim S-R, Kim S-H, Chae H-D, Kang B-M. Effects of pioglitazone on ovarian stromal blood flow, ovarian stimulation, and in vitro fertilization outcome in patients with polycystic ovary syndrome. *Fertility and Sterility* 2010;**94**:236-41.

Ladson 2011 *{published data only}*

Ladson G, Dodson WC, Sweet SD, Archibong AE, Kunselman AR, Demers LM, et al. Effects of metformin in adolescents with polycystic ovary syndrome undertaking lifestyle therapy: a pilot randomized double-blind study. *Fertility and Sterility* 2011;**95**(8):2595-8.

Lazaro 2011 *{published data only}*

Lazaro I, Diaz M, Cabre A, Masana L, Ibanez L. Fatty acid-blinding protein-4 plasma levels are associated to metabolic abnormalities and response to therapy in girls and young women with androgen excess. *Gynecological Endocrinology* 2011;**27**(11):935-9.

Lemay 2006 *{published data only}*

Lemay A, Dodin S, Turcot L, Déchêne F, Forest J-C. Rosiglitazone and ethinyl estradiol/cyproterone acetate as single and combined treatment of overweight women with polycystic ovary syndrome and insulin resistance. *Human Reproduction* 2006;**21**(1):121-8.

Luque-Ramirez 2008c *{published data only}*

Luque-Ramirez M, Alvarez-Blasco F, Escobar-Morreale HF. Antiandrogenic contraceptives increase serum adiponectin in obese polycystic ovary syndrome patients. *Obesity* 2008;**17**(1):3-9.

Luque-Ramirez 2009 *{published data only}*

Luque-Ramirez M, Mendieta-Azcona C, del Rey Sanchez JM, Maties M, Escobar-Morreale HF. Effects of an antiandrogenic oral contraceptive pill compared with metformin on blood coagulation tests and endothelial function in women with the polycystic ovary syndrome: influence of obesity and smoking. *European Journal of Endocrinology* 2009;**160**:469-80.

Luque-Ramirez 2010a *{published data only}*

Luque-Ramirez M, Escobar-Morreale HF. Treatment of polycystic ovary syndrome (PCOS) with metformin ameliorates insulin resistance in parallel with decrease of serum interleukine-6 concentrations. *Hormone and Metabolic Research* 2010;**42**:815-20.

Luque-Ramirez 2011 *{published data only}*

Luque-Ramirez M, Alvarez-Blasco F, Alpanes M, Escobar-Morreale HF. Role of decreased circulating hepcidin concentrations in the iron excess of women with polycystic ovary syndrome. *Journal of Clinical Endocrinology and Metabolism* 2011;**96**(3):846-52.

Mehrabian 2016 *{published data only}*

Mehrabian F, Ghasemi-Tehrani H, Mohamadkhani M, Moeinoddini M, Karimzadeh P. Comparison of the effects of metformin, flutamide plus oral contraceptives, and simvastatin on the metabolic consequences of polycystic ovary syndrome. *Journal of Research in Medical Sciences* 2016;**21**:7.

Mitkov 2005 *{published data only}*

Mitkov M, Pehlivanov B, Terzieva D. Combined use of metformin and ethinyl estradiol-cyproterone acetate in polycystic ovary syndrome. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 2005;**118**:209-13.

Moghtadaei 2009 *{published data only}*

Moghtadaei P, Sardari F. Metabolic and endocrine effect of metformin and metformin plus cyclic dydrogestrel in women with polycystic ovarian syndrome. *Cell Journal* 2009;**11**:67-8.

Moretti 2016 *{published data only}*

Moretti CG, Guccione L, di Giacinto P, Cannuccia A, Meleca C, Lanzolla G, et al. Efficacy and safety of myo-inositol supplementation in the treatment of obese hirsute PCOS women: comparative evaluation with OCP+bicalutamide therapy. In: Endocrine Society Journal, poster SUN-153. 2016.

NCT02866786 *{published data only}*

NCT02866786. The effect of OCP and Metformin on Clinical, hormonal, metabolic and ultrasonographic characteristics in PCOS. clinicaltrials.gov/show/NCT02866786 (first received 15 August 2016).

Orbetzova 2011 *{published data only}*

Orbetzova MM, Pehlivanov BK, Mitko MM, Atanassova IB, Kamenov ZA, Kolarov GB, et al. Effect of short-term standard therapeutic regimens on neuropeptide Y and adipose tissue hormones in overweight insulin-resistant women with polycystic ovary syndrome. *Folia Medica* 2011;**53**(3):15-24.

Panidis 2011 *{published data only}*

* Panidis D, Georgopoulos NA, Piouka A, Katsikis I, Saltamavros AD, Decavalas G, et al. The impact of oral contraceptives and metformin on anti-Müllerian hormone serum levels in women with polycystic ovary syndrome and biochemical hyperandrogenemia. *Gynecological Endocrinology* 2011;**27**(8):587-92.

Pedersen 2018 *{published data only}*

Pedersen AJ, Stage TB, Glintborg D, Andersen M, Christensen MM. The pharmacogenetics of metformin in women with polycystic ovary syndrome: a randomized trial. *Basic and Clinical Pharmacology and Toxicology* 2018;**122**:239-44.

Romualdi 2010 *{published data only}*

Romualdi D, Giuliani M, Cristello F, Fulghesu AM, Selvaggi L, Lanzone A, et sl. Metformin effects on ovarian ultrasound appearance and steroidogenic function in normal-weight normoinsulinemic women with polycystic ovary syndrome; a randomized double-blind placebo-controlled clinical trial. *Fertility and Sterility* 2010;**93**(7):2303-10.

Suvarna 2016 *{published data only}*

Suvarna Y, Maity N, Kalra P, Shivamurthy MC. Comparison of efficacy of metformin and oral contraceptive combination of ethinyl estradiol and drospirenone in polycystic ovary syndrome. *Journal of the Turkish-German Gynecological Association* 2016;**17**:6-9.

Teede 2010b *{published data only}*

Teede HJ, Meyer C, Hutchison SK, Zoungas S, McGrath BP, Moran LJ. Endothelial function and insulin resistance in polycystic ovary syndrome: the effects of medical therapy. *Fertility and Sterility* 2010;**93**(1):184-91.

Wang 2016 *{published data only}*

Wang Q-Y, Song Y, Huang W, Xiao L, Wang Q-S, Feng G-M. Comparison of drospirenone with cyproterone acetate containing oral contraceptives, combined with metformin and lifestyle modifications in women with polycystic ovary syndrome and metabolic disorders: a prospective randomized control trial. *Chinese Medical Journal* 2016;**129**(8):883-90.

References to studies awaiting assessment

Fruzzeti 2009 *{published data only}*

Fruzzetti F, Perini D, Lazzarini V, Parrini D, Gambacciani M, Genazzani RA. Comparison of effects of 3 mg drospirenone plus 20 µg ethinyl estradiol alone or combined with metformin or cyproterone acetate on classic metabolic cardiovascular risk factors in nonobese women with polycystic ovary syndrome. *Fertility and Sterility* 2009;**94**:1793-8.

NCT01573377 *{published data only}*

Spremovic-Radjenovic 2014 *{published data only}*

Spremovic-Radjenovic S, Sreckovic S, Gudovic A, Djakonovic M, Bila J, Radonjic-Lazovic G. Comparison of the effects of dianogest/ethinyl estradiol and metformin on metabolic parameters in non-obese women with polycystic ovary syndrome. In: Poster, European Society of Contraception and Reproductive Health, S207. 2014.

Vieira 2010 *{published data only}*

Vieira CS, Fernandes JB, Soares GM, Martins WP, Silva-de-Sà MF, Ferriani RA. Addition of low dose of metformin to the oral contraceptive has no great advantage in terms of cardiovascular risk markers in women with polycystic ovary syndrome regardless insulin resistance. *Fertility and Sterility* 2010;**94**(4 Suppl 1):S63-4. [Oral communication O-217]

References to ongoing studies

NCT02744131 *{published data only}*

OCP vs metformin for Improvement in clinical symptoms and metabolic markers in Indian PCOS women (OCP). Ongoing study. 04/2016. Contact author for more information.

NCT03229057 *{published data only}*

Comparing the effects of oral contraceptive pills versus metformin (COMET-PCOS). Ongoing study. Juillet 2017. Contact author for more information.

NCT03905941 *{published data only}*

Relative desirability of metformin vs. birth control pill in treating PCOS in women of later reproductive age. Ongoing study. March 2020. Contact author for more information.

Additional references

Al Khalifah 2016

Al Khalifah RA, Florez ID, Dennis B, Thsbsnr L, Bassilious E. Metformin or oral contraceptives for adolescents with polycystic ovarian syndrome: a meta-analysis. *Peditrics* 2016;**137**(5):e20154089.

Buzney 2014

Buzney E, D, Sheu J, Buzney C, Reynolds RV. Polycystic ovary syndrome: a review for dermatologists Part II. Treatment. *Journal of the American Academy of Dermatology.* 2014;**71**(5):859.e1-859.e15; quiz 873-4.

Conway 1990

Conway GS, Jacobs HS, Holly JM, Wass JA. Effects of luteinizing hormone, insulin, insulin-like growth factor-1 and insulin-like growth factor small binding protein-1 in the polycystic ovary syndrome. *Clinical Endocrinology* 1990;**33**:593-603.

Costello 2003

Costello MF, Eden JA. A systematic review of the reproductive system effects of Metformin in patients with polycystic ovary syndrome. *Fertility and Sterility* 2003;**79**(1):1-13.

Deeks 2017

Deeks JJ, Higgins JP, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). The Cochrane Collaboration, 2017. Available from www.training.cochrane.org/handbook.

Diamanti-Kandarakis 2003

Diamanti-Kandarakis E, Baillargeon, JP, Iuorno MJ, Jakubowicz DJ, Nestler JE. A modern medical quandary: polycystic ovary syndrome, insulin resistance, and oral contraceptive pills. *Journal of Clinical Endocrinology & Metabolism* 2003;**88**:1927-32.

Diamanti-Kandarakis 2012

Diamanti-Kandarakis E, Dunaif A. Insulin resistance and the polycystic ovary syndrome revisited: an update on mechanisms and implications. *Endocrine Reviews* 2012;**33**(6):981–1030.

Dunaif 1989

Dunaif A, Segal KR, Futterweit W, Dobrjansky A. Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. *Diabetes* 1989;**38**:1165-74.

Dunaif 1994

Dunaif A. Molecular mechanisms of insulin resistance in the polycystic ovary syndrome. *Seminars in Reproductive Endocrinology* 1994;**12**:15-20.

ESHRE/ASRM 2004

The Rotterdam ESHRE/ASRM-sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Human Reproduction* 2004;**19**:41-7.

Fauser 2012

Fauser BC, Tarlatzis BC, Rebar RW, Legro RS, Balen AH, Lobo R, et al. Consensus on women's health aspects of polycystic ovary syndrome (PCOS): the Amsterdam ESHRE/ ASRM-Sponsored 3rd PCOS Consensus Workshop Group. *Fertil Steril* 2012;**97**:28-38.e25.

Freitas de Medeiros 2017

Freitas de Medeiros S. Risks, benefits size and clinical implications of combined oral contraceptive use in women with polycystic ovary syndrome. *Reproductive Biology and Endocrinology* 2017;**15**:93-109.

Gallo 2013

Gallo MF, Nanda K, Grimes DA, Lopez LM, Schulz KF. 20 µg versus >20 µg estrogen combined oral contraceptives for contraception.. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No: CD003989. [DOI: 10.1002/14651858.CD003989.pub5]

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Hamilton (ON): McMaster University (developed by Evidence Prime), accessed 6 August 2019. Available at gradepro.org.

Hasegawa 1999

Hasegawa I, Murakawa H, Suzuki M, Yamamoto Y, Kurabayashi T, Tanaka K. Effect of troglitazone on endocrine and ovulatory performance in women with insulin resistance-related polycystic ovary syndrome. *Fertility and Sterility* 1999;**71**:323-7.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557–60.

Higgins 2017

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 5.2 (updated 2017). The Cochrane Collaboration, 2011. Available from www.training.cochrane.org/handbook.

Homburg 1996

Homburg R. Polycystic ovary syndrome - from gynaecological curiosity to multisystem endocrinopathy. *Human Reproduction* 1996;**11**:29-39.

Hundal 2003

Hundal RS, Inzucchi SE. Metformin: new understandings, new uses. *Drugs* 2003;**63**(18):1879-94.

Katsiki 2010

Katsiki N, Hatzitolios AI. Insulin-sensitizing agents in the treatment of polycystic ovary syndrome: an update. *Current Opinion in Obstetrics & Gynecology* 2010;**22**(6):466-76.

Luque-Ramirez 2018

Luque-Ramirez M, Nattero-Chavez L, Ortiz Flores AE, Escobar-Morreale HF. Combined oral contraceptive and/or antiandrogens versus insulin sensitizers for polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction Update* 2018;**24**:225-41.

Martin 2018

Martin KA, Anderson RR, Chang RJ, Ehrmann DA, Lobo RA, Murad MH, et al. Evaluation and treatment of hirsutism in premenopausal women: an Endocrine Society* Clinical Practice Guideline. *Journal of Clinical and Endocrinology Metabolism* 2018;**103**(4):1–25.

Ovalle 2002

Ovalle F, Azziz R. Insulin resistance, polycystic ovary syndrome, and type 2 diabetes mellitus. *Fertility and Sterility* 2002;**77**:1095-105.

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rotterdam ESHRE 2004

The Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Human Reproduction* 2004;**19**:41-7.

Ryan 2016

Ryan R, Cochrane Consumers and Communication Review Group. Cochrane Consumers and Communication Group: meta-analysis. http://cccrg.cochrane.org accessed December 2016.

Teede 2010a

Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Medicine* 2010;**8**:41.

Teede 2018

Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al, International PCOS Network. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction* 2018;**33**(9):1602-18.

Teede 2019

Teede H, Tassone EC, Piltonen T, Malhotra J, Mol BW, Pena A, et al. Effect of the combined oral contraceptive pill and/or Metformin in the management of polycystique ovary syndrome: a systematic review with meta-analyses. *Clinical Endocrinology* 2019;**91**:479-89.

Utiger 1996

Utiger R. Insulin and the polycystic ovary syndrome. *New England Journal of Medicine* 1996;**335**:657-8.

Wild 2002a

Wild RA. Long-term health consequences of PCOS. *Human Reproduction Update* 2002;**8**:231-41.

Wild 2002b

Wild RA. Polycystic ovary syndrome: a risk for coronary artery disease? *American Journal of Obstetrics & Gynaecology* 2002;**186**:35-43.

References to other published versions of this review

Costello 2005

Costello MF, Shrestha B, Eden J, Sjoblom P, Johnson N. Insulin-sensitising drugs versus the combined oral contraceptive pill for hirsutism, acne, and risk of diabetes, cardiovascular disease, and endometrial cancer in polycystic ovary syndrome. *Cochrane Database of Systematic Reviews* 2005, Issue 4. Art. No: CD005552. [DOI: 10.1002/14651858.CD005552]

Costello 2007

Costello MF, Shrestha B, Eden J, Johnson N, Moran LJ. Insulin-sensitising drugs versus the combined oral contraceptive pill for hirsutism, acne and risk of diabetes, cardiovascular disease, and endometrial cancer in polycystic ovary syndrome. *Cochrane Database of Systematic Reviews* 2007, Issue 1. Art. No: CD005552. [DOI: 10.1002/14651858.CD005552.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Aghamohammadzadeh 2010

Study characteristics

Aghamohammadzadeh 2010 *(Continued)*

Methods	Randomised controlled trial **Location of the trial:** quote: Tabriz, Iran **Method of randomisation:** unclear: quote: “Patients randomly divided in two equal groups.” **Method of allocation concealment:** not stated. **Source of funding:** Quote: "This study was financed by Research Vice Chancellor of Tabriz University of Medical Sciences."
Participants	**Inclusion criteria:** PCOS as per National Institutes of Health criteria. **Exclusion criteria:** rheumatic disease, infective disease, therapy for hirsutism and scalp hair loss such as spironolactone and finasteride and therapy for acne such as antibiotics. **Number of women randomised:** 70: MET: 35 OCP: 35 **Number of women analysed:** 60: MET: 30 OCP: 30 **Number of withdrawal and reasons:** MET: Pregnancy: 3/ 8.6% Epigastric pain: 2/ 5.7% OCP: nausea, increased body weight drug intolerance: 5/ 14.3% **Summary characteristics:** PCOS Young patients: mean (±SD) age (years): 23.4 +/-8.1 Mean weight (±SD) kg: 63.4 +/- 14.7 Mostly hirsute (76.6%) **Age (years):** MET mean (±SD): 24.9 (11) OCP mean (±SD): 22 (5.2) **BMI (kg/m^2):** MET mean (±SD): 26.5 (5.7) OCP mean (±SD): 24.6 (4.9)
Interventions	**Treatment:** MET 1000 mg/day **Control:** OCP (Ethinyl Estradiol 35 Cyproterone acetate 2 mg) /days /21 days per months **Duration:** 6 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** none **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Serum total testosterone (nmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Body weight (kg) 2. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L)
Notes	**Power calculation:** unclear

Aghamohammadzadeh 2010 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients randomly divided in two equal groups." Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	High risk	Reason for missing data imbalanced and related to intervention. MET: epigastric pain 2/ 5.7%. OCP: nausea, increased body weight drug intolerance 5/ 14.3%: more than twice.
Selective reporting (reporting bias)	Low risk	The study protocol is not available, but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Al-Zubeidi 2015

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: San Diego, USA **Method of randomisation:** quote: "Each patient was randomly assigned to group 1 or group 2 using concealed assignments from random numbers that were computer generated." **Method of allocation concealment:** not stated. **Source of funding:** not stated.
Participants	**Inclusion criteria:** inclusion of patients in the study was based on the NIH criteria for the diagnoses of PCOS in this age group, which includes irregular menses or amenorrhoea and elevated free or total testosterone. Irregular menses or amenorrhoea was defined as less than or equal to eight menses per year. In addition, other causes of hyperandrogenism were excluded including: adrenal tumours, late congenital adrenal hyperplasia, and prolactinomas. Patients were only enrolled 2 years post-menarche given the potential for menstrual irregularity in normal maturation post-menarche. **Exclusion criteria:** exclusion criteria included: current treatment or treatment within the last 3 months with metformin or OCP and personal or family history of blood clotting disorders, breast cancer, stroke, severe migraines with aura, elevated BP, defined as either systolic and/or diastolic BP ≥ 95th percentile measured upon three or more occasions, and smoking defined by more than one pack per day for the past 6 months. **Number of women randomised:** 34

Al-Zubeidi 2015 *(Continued)*

	MET: 16 OCP: 18 **Number of women analysed:** 22 MET: 10 OCP: 12 **Number of withdrawal and reasons:** MET: no shows and unreachable: 5 /31.2% No insurance: 1/ 6.2% OCP: no shows: 4/ 22.2% No insurance: 1/ 5.6% Decline treatment: 1/ 5.6% **Summary characteristics:** PCOS, adolescent 12 to 18 years, 68% Hispanic **Age (years):** MET: range: 14 to 18 OCP: range: 15 to 17 **BMI (kg/m^2):** MET mean (± SD): 33.7 (6) OCP mean (± SD): 33.4 (9)
Interventions	**Treatment:** MET 1g twice a day gradual increase maximum dose was reached over 3-week period **Control:** OCP (Ethinyl Estradiol 30mcg Noresthisterone 1mg) **Duration:** 6 months **Co-intervention(s):** routine counselling about diet and exercise was done in clinic, but no specific exercise or diet prescription was offered.
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting HDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. BMI ((kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting HDL cholesterol (mmol/L)

Al-Zubeidi 2015 *(Continued)*

	3. Fasting triglycerides (mmol/L)
Notes	**Power calculation:** yes, a priori, on FT and BMI (secondary outcomes)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Each patient was randomly assigned to group 1 or group 2 using concealed assignments from random numbers that were computer generated.”
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of ‘low risk’ and ‘high risk’.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "No show, unreachable" unclear reasons for withdrawal and could be related to the intervention. Insufficient report of attrition/exclusions to permit a judgement of ‘low risk’ or ‘high risk’.
Selective reporting (reporting bias)	High risk	Measurement of hirsutism and improvement in menstrual pattern described in methods but not reported in the result or reported without giving the number of patient with improvement in menstrual pattern. HDL cholesterol, SD not reported for OCP group therefore could not be included in the meta-analysis. Hirsutism is one of our primary outcomes, not all of the study's prespecified primary outcomes have been reported.
Other bias	Low risk	The study appears to be free of other sources of bias.

Allen 2005

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: Massachussets, USA **Method of randomisation:** quote: "Random numbers table" **Method of allocation concealment:** not stated **Source of funding:** not stated
Participants	**Inclusion criteria:** hyperandrogenaemia (total T > 60 ng/dL and free T > 1.1 pg/mL), no evidence of androgen secreting tumour, oligomenorrhoea (< 6 menses in the previous 6 months) and obesity (> 95th percentile BMI for age), hyperinsulinaemic (fasting insulin > 20 µU/mL) but not diabetic (fasting glucose < 126 mg/dL). PCOS diagnosis consistent with NIH.

Allen 2005 *(Continued)*

	Exclusion criteria: reported current or past sexual activity, OCP use within previous 6 months, positive urine pregnancy test (performed on all participants at baseline), abnormal BUN, Cr, AST or positive personal or family history of thrombosis, primary adrenal source of androgeneses. **Number of women randomised:** 35 MET: 18 OCP: 17 **Number of women analysed:** 31 MET: 16 OCP: 15 **Number of withdrawal and reasons:** MET: 2/ 11% OCP: 2/ 12% All unreachable or refused to come to further visits. No breakdown by MET versus OCP. **Summary characteristics:** PCOS, 12 to 21 years, obese, hyperinsulinaemic, oligoamenorrhoeic and hyperandrogenaemia **Age (years):** MET Mean (range): 15.4 years (13.1 to 18.4) OCP Mean (range): 15.3 years (12.5 to 21) **BMI (kg/m^2):** MET (mean ± SEM): 37.3 (1.3) OCP (mean ± SEM): 40.1 (2.1)
Interventions	**Treatment:** MET 500 mg twice a day 2 weeks and dose increased to 1 g twice a day if tolerated **Control:** OCP (Ethinyl Estradiol 35mcg norgestimate 0.25 mg) **Duration:** 6 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Acne score Body weight (kg) BMI (kg/m^2) Serum total testosterone (ng/dL) Fasting insulin (µU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mg/dL) Fasting HDL cholesterol (mg/dL) Fasting LDL cholesterol (mg/dL) Fasting triglycerides (mg/dL)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score 2. Acne score
Objective outcomes	(a) Clinical parameters 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles 2. Body weight (kg) and/or BMI (kg/m^2)

Allen 2005 *(Continued)*

	(b) Hormonal parameters 1. Serum total testosterone (ng/dL) (c) Metabolic parameters 1. Fasting insulin (µU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mg/dL) 4. Fasting HDL cholesterol (mg/dL) 5. Fasting LDL cholesterol (mg/dL) 6. Fasting triglycerides (mg/dL)
Notes	**Authors contacted about:** Ferriman-Gallwey score, menstrual rate and acne score because expressed with graph. Authors were also contacted to ask about free testosterone because calculated in the methods as total testosterone/SHBG. **Power calculation:** yes, a priori on free testosterone

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Random numbers table".
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "All unreachable or refused to come to further visits" unclear reason for withdrawal could be related to intervention. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk.
Selective reporting (reporting bias)	High risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of blood pressure, Ferriman-Gallwey score, acne score, menstrual improvement. These 3 last outcomes are presented with graphs. One or more outcomes of interest in the review have been reported incompletely so that they cannot be entered in a meta-analysis.
Other bias	Low risk	The study appears to be free of other sources of bias.

Bhattacharya 2016

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Kolkata, India **Method of randomisation:** quote: "Computer-generated randomization table, interventions were sealed in sequentially numbered identical opaque containers." **Method of allocation concealment:** quote: "Interventions were sealed in sequentially numbered identical opaque containers" **Source of fundings:** not stated
Participants	**Inclusion criteria:** PCOS 2003 Rotterdam criteria **Exclusion criteria:** not stated **Number of women randomised:** 129: OCP + MET arm: 63 OCP arm: 66 **Number of women analysed:** 85: OCP + MET : 46 OCP: 39 **Number of withdrawal and reasons:** OCP + MET: Lost to F/U: 17/ 27% OCP: Lost to F/U: 27/ 41% **Summary characteristics:** PCOS **Age (years):** OCP + MET mean (± SD): 20.68 (3.36) OCP mean (± SD): 21.24 (4.01) **BMI (kg/m^2):** OCP + MET mean (± SD) 26.23 (4.06) OCP mean (± SD): 25.11 (4.33)
Interventions	**Treatment:** OCP (Ethinyl Estradiol 20 mcg Drospirenone 3 mcg) daily cyclically 24/4 + MET 500 mg/day **Control:** OCP (Ethinyl Estradiol 20 mcg Drospirenone 3 mcg) daily cyclically 24/4 **Duration:** 12 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:** Acne BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L)

Bhattacharya 2016 *(Continued)*

	Sex hormone-binding globulin (SHBG) Free androgen index (FAI) (%)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score 2. Acne
Objective outcomes	(a) Clinical parameters 1. BMI (kg/m^2) 2. Blood pressure (systolic) (mm Hg) 3. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%)
Notes	Abstract only **Authors contacted about:** their results **Power calculation:** done, a priori, unclear on FAI (secondary outcome)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Computer-generated randomization table”.
Allocation concealment (selection bias)	Low risk	Quote: “Interventions were sealed in sequentially numbered identical opaque containers”.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Patients were masked but investigator and data analysis not”.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Patients were masked but investigator and data analysis not”.
Incomplete outcome data (attrition bias) All outcomes	High risk	Reason for missing data unclear quote: "lost to F/U". Missing values were imputed by carrying last observation forward. Quote: “use of LOCF might be appropriate if most people for whom outcomes are carried forward had a genuine measurement relatively recently.” 6 months to 12 is not judged as “recently” by the authors: inappropriate use of imputation
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified.
Other bias	Unclear risk	Insufficient information. Abstract only.

Bodur 2018

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Istanbul, Turkey**Method of randomisation:** quote: "The remaining subjects were randomized into four subgroups by a computerized method"**Method of allocation concealment:** Not stated**Source of funding:** Quote: "This study was financially supported by Gulhane Military Medical Academy as it was originally conducted as a graduation thesis in obstetrics and gynaecology."
Participants	**Inclusion criteria:** the diagnosis of PCOS was made according to the Rotterdam criteria. Clinical or biochemical hyperandrogenism was detected if the modified Ferriman-Gallwey score was higher than 8, acne was present and serum concentrations of total and/or free testosterone were increased. Oligomenorrhoea was defined when menstruation occurred six times a year at most. Polycystic ovaries were visualised by ultrasonography when increased stromal echogenicity was peripherally surrounded by more than 10 follicles with a diameter of 2 mm to 8 mm. **Exclusion criteria:** those patients who had systemic disorders and were receiving therapies that could affect carbohydrate and lipid metabolism were excluded from the study. Patients reporting contraindications for oral contraceptives and metformin were also excluded from the study. **Number of women randomised:** 87/70 MET + OCP: 20 MET: 29 OCP: 21 **Number of women analysed:** 63/46 MET+ OCP: 12 MET: 17 OCP: 17 **Number of withdrawal and reasons:** MET + OCP: severe nausea: 1/ 5% Headache: 1 / 5% Dizziness: 2 / 10% Unknown: 2 / 10% Hirsutism: 1/ 5% Unwilling weight loss: 1/ 5% MET: pregnancy: 4/ 13.8% Severe nausea: 2/ 6.9% Feeling hunger: 1 / 3.4% Unknown: 1/ 3.4% Hirsutism: 1 / 3.4% Unwilling weight loss: 1/ 3.4% Dizziness: 1 / 3.4% Hypothyroidism: 1 / 3.4% OCP: pregnancy: 1 / 4.8% Severe nausea: 1/ 4.8% Sexual reluctant: 1/ 4.8% Hirsutism: 1/ 4.8% **Summary characteristics:** PCOS, 18 to 39, non-obese **Age (years):** MET + OCP mean (± SD): 27.35 (5.65) MET mean (± SD): 26.2 (3.96) OCP mean (± SD): 26.6 (4.92) **BMI (kg/m^2):** MET + OCP mean (± SD): 25.11 (3.75) MET mean (± SD): 25.06 (3.08) OCP mean (± SD): 23.45 (3.4)

Bodur 2018 *(Continued)*

Interventions	MET 1700 mg/day (850 mg twice a day) + OCP (Ethinyl Estradiol 30 mcg Drospirenone 3 mg) MET 1700 mg/day (850mg twice a day) OCP (Ethinyl Estradiol 30 mcg / Drospirenone 3mg) **Duration:** 6 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** BMI (kg/m^2) Fasting glucose (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) (c) Metabolic parameters 2. Fasting glucose (mmol/L)
Notes	**Power calculation:** unclear 3-arm study: OCP V MET V MET + OCP

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The remaining subjects were randomized into four subgroups by a computerized method"
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data well described, and percentage of missing data related to intervention low.
Selective reporting (reporting bias)	Unclear risk	BMI described in the method not reported in the results. But not a primary or key outcomes.

Bodur 2018 *(Continued)*

		Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Cetinkalp 2009

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Izmir, Turkey**Method of randomisation:** quote: "Patients were randomized".**Method of allocation concealment:** nNot stated**Source of funding:** not stated
Participants	**Inclusion criteria:** Rotterdam PCOS Consensus criteria**Exclusion criteria:** DM, hyperprolactinaemia, congenital adrenal hyperplasia (through ACTH test), thyroid disorders, Cushing syndrome, hypertension, hepatic or renal dysfunction, confounding medications (OCP), antihypertensive medications, insulin sensitising drugs)**Number of women randomised:** unclear: 100/ 99/ 94 MET: 47 OCP: 33**Number of women analysed:** 94 MET: 47 OCP: 33**Number of withdrawal and reasons:** unclear. Could be 0/94 (0%), 5/99 (5%) or 6/100 (6%). **Summary characteristics:** PCOS, lean young women **BMI (kg/m^2):** MET Mean (± SEM): 25.82 (6.12) OCP Mean (±S EM): 24.72 (4.1)
Interventions	**Treatment:** MET 2g/day **Control:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2 mg) **Duration:** 4 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Acne – subjective Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (ng/mL) Fasting insulin (mIU/mL) Fasting glucose (mg/dL) Fasting total cholesterol (mg/dL) Fasting HDL cholesterol (mg/dL) Fasting LDL cholesterol (mg/dL) Fasting triglycerides (mg/dL)

Cetinkalp 2009 *(Continued)*

Subjective outcomes	(a) Clinical parameters 1. Acne – subjective 2. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles 2. Body weight (kg) 3. BMI (kg/m^2) 4. Blood pressure (systolic) (mm Hg) 5. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (ng/mL) (c) Metabolic parameters 1. Fasting insulin (mIU/mL) 2. Fasting glucose (mg/dL) 3. Fasting total cholesterol (mg/dL) 4. Fasting HDL cholesterol (mg/dL) 5. Fasting LDL cholesterol (mg/dL) 6. Fasting triglycerides (mg/dL)
Notes	**Authors contacted about:** fasting insulin because expressed in wrong unit. **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients were randomized". Insufficient information about the sequence generation process available to permit judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	States both n = 100, n = 99 and n = 94 randomised (n = 47 metformin, n = 14 rosiglitazone, n = 33 OCP). Withdrawal could be (0% or n = 5 (5%) or n = 6 (6%), no reasons provided for any withdrawals. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'

Cetinkalp 2009 *(Continued)*

Selective reporting (reporting bias)	High risk	Acne, hirsutism and improvement of menstrual patterns are described in the results but not in the methods. One or more primary outcomes have been reported using measurements, analysis methods or subset data THAT WERE NOT PRE-SPECIFIED. Moreover, from results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of BP, but not a primary or key outcomes.
Other bias	High risk	Baseline imbalance, MET n = 47, rosiglitazone n = 14, OCP n = 33 which cannot be explained by method of randomisation or allocation concealment or incomplete outcome data.

Christakou 2014

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Thessaloniki, Greece **Method of randomisation** quote: "Randomization was performed by random number tables" **Method of allocation concealment:** quote: "The patient number treatment codes were held by a third party" **Source of funding:** None
Participants	**Inclusion criteria:** the diagnosis of PCOS was based on the presence of irregular menstrual cycles (eight or fewer menses per year) as well as elevated serum levels of testosterone and/or clinical symptoms of hyperandrogenism, according to the National Institute of Child Health and Human Development conference. **Exclusion criteria:** non-classical congenital adrenal hyperplasia, androgen-secreting neoplasms, hyperprolactinaemia and thyroid disease were excluded by appropriate tests in all women. Patients were excluded from participation if they were pregnant or planning to become pregnant, were breastfeeding, had a history of current or recent (within 6 months) use of oral contraceptives, antidiabetics, or antiandrogens, or had any contraindications to metformin therapy including renal or hepatic impairment. **Number of women randomised:** 120 MET: 40 OCP 1: 40 OCP 2: 40 **Number of women analysed:** 109 MET: 35 OCP 1: 38 OCP 2: 36 **Number of withdrawal and reasons:** MET: GI side effect 5/ 12.5% OCP 1: loss to F/U 2 /10% OCP 2: loss to F/U 4/10% **Summary characteristics:** PCOS All the participants were lean (BMI < 25 kg/m^2), in good health and nonsmokers or had quit smoking for more than a year at baseline. **Age (years):** MET mean (± SE): 21.5 (0.5)

Christakou 2014 *(Continued)*

	OCP 1 mean (± SE): 22 (0.6) OCP 2 mean (± SE): 23.2 (0.6) **BMI (kg/m^2):** MET mean (± SE): 23.03 (0.67) OCP 1 mean (± SE): 21.8 (0.35) OCP 2 mean (±SE): 22.37 (0.48)
Interventions	**Treatment:** MET 425 mg twice a day one week then 850 mg twice a day**Control:** OCP 1 (Ethinyl Estradiol 35 mcg / Cyproterone acetate 2 mg) 21 days stop 7 days OCP 2 (Ethinyl estradiol 30 mcg / drospirenone 3 mg) 21 days stop 7 days**Duration:** 6 months**Co-intervention(s):** none
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2.BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%)
Notes	**Power calculation:** unclear 3 arms study MET V OCP (drospirenone) V OCP (cyproterone acetate), authors have decided to combine both OCP versus MET

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomization was performed by random number tables."
Allocation concealment (selection bias)	Low risk	Quote: "The patient number treatment codes were held by a third party."
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'.

Christakou 2014 *(Continued)*

		Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	High risk	MET: GI side effect 5/ 12.5% versus OCP 1: loss to F/U 2 /10% and OCP 2 loss to F/U 4/10%. Reason for missing outcome data is likely to be related to true outcome, with imbalance reasons for missing data across intervention group. Loss to F/U versus GI adverse events.
Selective reporting (reporting bias)	Unclear risk	Fasting glucose and fasting insulin pre-specified in the methods not stated in the results. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Cibula 2005

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Prague, Czech Republic **Method of randomisation:** quote: "All patients were randomly assigned to two groups using a computer generator of random values with a uniform distribution within the interval 0 to 1. The values obtained were transformed into rank values, ranks 1-15 were assigned to the OCP group and remaining 15 were assigned to identical OCP in combination with metformin" **Method of allocation concealment:** not stated **Source of funding:** not stated
Participants	**Inclusion criteria:** PCOS: (i) oligomenorrhoea from menarche (menstrual cycle > 35 days); (ii) an increased concentration of at least one androgen above the upper reference limit (testosterone 0.5 to 2.63 n mol/L, androstenedione 1.57 to 5.4 n mol/L, dehydroepiandrosterone (DHEA) 0.8 to 10.5 n mol/L); and (iii) clinical manifestation of hyperandrogenism (acne, hirsutism or both). **Exclusion criteria:** secondary endocrine disorder, such as hyperprolactinaemia, thyroid dysfunction or a non-classical form of congenital adrenal hyperplasia, those wishing to conceive within the next 6 months, or contraindications to oral contraceptive use. **Number of women randomised:** 30: OCP + MET arm: 15 OCP arm: 15 **Number of women analysed:** 28: OCP + MET: 13 OCP 15 **Number of withdrawal and reasons:** OCP + MET: 2: GIT problem: 1/ 6.7% Non compliance: 1/ 6.7% OCP: 0 **Summary characteristics:** PCOS **Age (years)**:

Cibula 2005 *(Continued)*

	MET + OCP Mean (± SD): 23.8 (5.4) OCP Mean (± SD) 23.2 (4.6) **BMI (kg/m^2):** MET + OCP Mean (± SD): 24.7(4.9) OCP Mean (± SD): 22.1(3.1)
Interventions	**Treatment(s):** MET 500 mg three times a day + OCP (Ethinyl estradiol 35 mcg norgestimate 250 mcg) once daily (21 days per month followed by 7 days pill-free period) **Control:** OCP (Ethinyl estradiol 35 mcg norgestimate 250 mcg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 6 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2)) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (pmol/L) Fasting glucose (mmol/L) Total Cholesterol (mmol/L) HDL Cholesterol (mmol/L) LDL Cholesterol (mmol/L) Triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Body weight (kg) 3. (BMI (kg/m^2) (b) Hormonal parameters 1.Serum total testosterone (nmol/L) 3.Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting insulin (pmol/L) 2. Fasting glucose (mmol/L) 3. Total Cholesterol (mmol/L) 4. HDL Cholesterol (mmol/L) 5. LDL Cholesterol (mmol/L)

Cibula 2005 *(Continued)*

	6. Triglycerides (mmol/L)
Notes	**Authors contacted** about power calculation, method of randomisation and co-intervention: Calculated statistical power of study not sufficient as due to wide physiological range for insulin sensitivity the sufficient number is extremely high (personal communication with author). Method of randomisation and co intervention kindly provided by the authors that was not in the original paper **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "All patients were randomly assigned to two groups using a generator of random values with a uniform distribution within the interval 0 to 1 (statistical software NCSS 2002). The values obtained were transformed into rank values. The subjects with ranks 1–15 were assigned to the COC group and received a monophasic COC (EE 35 mg/NGM 250mg) in a cyclic regimen (21 days of active pills followed by 7 days of pill-free interval) for 6 months. The remaining 15 subjects received an identical COC in combination with metformin (1500 mg/day) for 6 months (METOC group)."
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'Yes' or 'No'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	High risk	Withdrawal OCP 0% versus metformin + OCP 13% due to n = 1 GIT problem and n = 1 study protocol non-compliance. Reason for missing outcome data is likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention.
Selective reporting (reporting bias)	Low risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
Other bias	Low risk	The study appears to be free of other sources of bias.

Dardzinska 2014

Study characteristics

Methods	Randomised controlled trial, cross-over study **Location of the trial:** quote: Gdynia, Poland **Method of randomisation:** quote: "Women were treated in a computer-randomized order"

Dardzinska 2014 *(Continued)*

	Method of allocation concealment: not stated **Sources of fundings:** Quote: "This study has been founded by the Polish National Centre of Science (grant number 2 P05E 077 30) and the Medical University of Gdańsk Grant (ST-101)"
Participants	**Inclusion criteria:** the diagnosis of PCOS was made according to the Rotterdam criteria where the study participants had to present at least 2 of the 3 following features: oligo-/amenorrhoea, clinical/biochemical indices of hyperandrogenism or polycystic ovaries on transvaginal ultrasonography. **Exclusion criteria:** hyperprolactinaemia, non-classical congenital adrenal hyperplasia and androgen producing neoplasms. Participants who had received any medication such as oral contraceptives, antiandrogens, neuroleptics, antidepressants or corticosteroids in the preceding 3 months were not included into the study. Furthermore, women with diagnosed diabetes, overt thyroid disease, chronic inflammatory disorders or a history of infection preceding 1 month before the study were also excluded. Pregnancy, age more than 40 years and contraindications to oral contraception or metformin were additional exclusion criteria. **Number of women randomised:** 42 OCP arm: 24 MET arm: 18 **Number of women analysed:** 34 OCP: 21 MET: 13 **Number of withdrawal and reasons:** OCP: 3: Loss to F/U: 1/ 4.2% Intolerance: 2/ 8.3% MET: 5 Loss to F/U: 4/ 22% Depression: 1/ 5.5% **Summary characteristics:** PCOS, age range 18 to 36, 20% smokers **Age (years):** OCP mean (95%CI): 24.9 (23.5; 26.4) MET mean (95%CI): 24.6 (23.0; 26.3) **BMI (kg/m^2):** OCP: mean (±SD): 24.9 (4.4) MET: mean (±SD): 25.1 (9.8)
Interventions	**Treatment:** MET 850 mg twice a day **Control:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2mg) **Duration:** 4 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg)

Dardzinska 2014 *(Continued)*

	BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Body weight (kg) 3. BMI (kg/m^2) 4. Blood pressure (systolic) (mm Hg) 5. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting glucose (mmol/L) 2. Fasting total cholesterol (mmol/L) 3. Fasting HDL cholesterol (mmol/L) 4. Fasting LDL cholesterol (mmol/L) 5. Fasting triglycerides (mmol/L)
Notes	**Authors contacted** about the results, they send us the results before cross-over **Power calculation:** unclear

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: “Women were treated in a computer-randomized order”.
Allocation concealment (selection bias)	Unclear risk	Method of concealment not stated. Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.

Dardzinska 2014 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"loss of F/U" unclear reason for withdrawal could be related to intervention. 2 "intolerance" in OCP group so related to the intervention, 1 "depression" in MET group unlikely related to the intervention. Insufficient reporting of attrition/ exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	Fasting glucose pre-specified in the methods not stated in the results. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	High risk	Number of patients randomised in each group is different 24 versus 18 and more withdrawal in the smallest group. Baseline imbalance.

El Maghraby 2015

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Alexandria, Egypt**Method of randomisation:** quote: "computer-generated random-number tables"**Method of allocation concealment:** not stated**Source of funding:** Not stated
Participants	**Inclusion criteria:** PCOS was defined in these girls by the presence of oligomenorrhoea (< 6 cycles/year) and serum testosterone >1 mcg/mL (The Rotterdam consensus workshop group, 2004). **Exclusion criteria:** suprarenal dysfunction, hyperprolactinaemia, thyroid dysfunction, or recent intake of medications likely to affect hormonal profile was a reason for exclusion from the study. **Number of women randomised:** 80 MET: 40 OCP: 40 **Number of women analysed:** 65 MET: 32 OCP: 33 **Number of withdrawal and reasons:** MET: adverse events 4/ 10% Non compliant 2 /5% Not satisfied 1 / 2.5% Loss F/U 1 / 2.5% OCP: side effect 2 / 5% Non compliant 1/ 2.5% Weight gain 4 / 10% **Summary characteristics:** PCOS, adolescent 15 to 20 years, all presented with hirsutism, acne and menstrual disorder **Age (years):** MET mean (±SD):

El Maghraby 2015 *(Continued)*

	17.2 (2) OCP mean (±SD): 16.9 (1.6)
Interventions	**Treatment:** MET 1700 mg/day**Control:** OCP (Ethinyl Estradiol 30 mcg progestin 15 mg)**Duration:** 24 months**Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Body weight (kg) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. Body weight (kg) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L)
Notes	**Authors contacted about:** results, because total testosterone and fasting insulin were in wrong unit. **Power calculation:** done, a priori, on glucose/insulin ratio outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “computer-generated random-number tables”
Allocation concealment (selection bias)	Unclear risk	Method of concealment not stated. Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "The study was not blinded”.

El Maghraby 2015 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The study was not blinded".
Incomplete outcome data (attrition bias) All outcomes	High risk	10% drop out for side effect in MET group/ 10% drop out for weight gain in OCP group, both different yet related to intervention. Quote "statistical analysis of data was done on an intention-to-treat basis" but method used not described.
Selective reporting (reporting bias)	High risk	Fasting glucose and Ferriman-Gallwey score not reported in the results. Instead of Ferriman-Gallwey score: subjective assessment of hirsutism. Total testosterone and fasting insulin cannot be used because are expressed with wrong units. Not all of the study's prespecified primary outcomes have been reported.
Other bias	Low risk	The study appears to be free of other sources of bias.

Elter 2002

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Istanbul, Turkey **Method of randomisation:** quote: "Randomization was produced from a computer-generated random list, where even and odd numbers were allocated OCP and OCP+metformin treatments respectively." **Method of allocation concealment:** not stated **Source of fundings:** not stated
Participants	**Inclusion criteria:** PCOS defined as the presence of (i) bilateral polycystic ovaries on ultrasound examination (multiple >10, small 2 mm to 8 mm diameter in 1 plane in the periphery and increased stromal echogenicity); (ii) chronic oligomenorrhoea (< 6 menstrual periods in the previous year) or amenorrhoea; and (iii) manifestations of hyperandrogenism and/or hyperandrogaenemia, such as a hirsutism score of > 8 Ferriman-Gallwey); acne; elevated serum testosterone and/or androstenedione and/or free testosterone levels. All women were either of normal weight or thin (BMI = 26 kg/m^2). **Exclusion criteria:** not euthyroid, hyperprolactinaemia, serum testosterone = 7 n mol/L, serum DHEAS = 19 µmol/L, late-onset congenital adrenal hyperplasia, Cushing syndrome, adnexal mass seen during pelvic sonography, diabetes, any other known endocrinological disease, those taking drugs known to affect carbohydrate or lipid metabolism and OGTT results during the 6 months preceding the study **Number of women randomised:** 40: MET + OCP arm: 20 OCP arm: 20 **Number of women analysed:** 40: MET+ OCP arm: 20 OCP arm: 20 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, normal or thin women. **Age (years):** MET+OCP Mean (± SD):

Elter 2002 *(Continued)*

	24.9 (6.62) OCP Mean (± SD) 23.45 (6.07) **BMI (kg/m^2):** MET+OCP Mean (± SD): 22.74(2.66) OCP Mean (± SD): 21.83(1.4)
Interventions	**Treatment(s):** MET 500 mg three times a day (for the first 15 days, MET 500 mg twice a day for adequate compliance) + OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) once daily (21 days per month followed by 7 days pill-free period) **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 4 months **Co-interventions:** None
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** BM (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (pmol/L) Fasting glucose (mmol/L) Total Cholesterol (mmol/L) HDL Cholesterol (mmol/L) LDL Cholesterol (mmol/L) Triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score (Ferriman-Gallwey)
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. (BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (n mol/L) (c) Metabolic parameters 1. Fasting insulin (pmol/L) 2. Fasting glucose (mmol/L) 3. Total Cholesterol (mmol/L) 4. HDL Cholesterol (mmol/L) 5. LDL Cholesterol (mmol/L) 6. Triglycerides (mmol/L)

Elter 2002 *(Continued)*

Notes	**Power calculation:** done, a priori, on BMI (secondary outcome)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomization was produced from a computer-generated random list where even and odd numbers were allocated OCP and OCP + metformin treatments respectively"
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "No attempt was made to mask the treatments from the subjects, and placebo was not used."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Clinical parameters of the subjects were evaluated by the same person, who was blind to the type of treatment"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Low risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
Other bias	Low risk	The study appears to be free of other sources of bias.

Essah 2011

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: Virginia, USA **Method of randomisation:** quote: "The VCU Investigational Pharmacy prepared metformin and matched placebo capsules, randomized the women with a computerized random number generator using a 1:1 allocation ratio with blocks of random size (three and four women in a block) and concealed the allocation." **Method of allocation concealment:** not stated **Source of fundings:** quote:"NIH grant and clinical research centre grant"
Participants	**Inclusion criteria:** PCOS described by the modified Rotterdam criteria, after excluding other endocrine disorders, as 1) the presence of clinical and/or biochemical signs of hyperandrogenism, and 2) at least one of the following: oligo- ovulation or anovulation or polycystic ovaries. All participants in the study had hyperandrogenaemia. **Exclusion criteria:** women were excluded from the study if they were pregnant; had diabetes or history of thromboembolism; had used tobacco within the previous 6 months; or had used OCs, diabetes/hyperlipidaemia medications, or antiandrogens within the previous 3 months.

Essah 2011 *(Continued)*

	Number of women randomised: 23: OCP + MET arm: 11 OCP + PBO arm: 12 **Number of women analysed:** 19: OCP + MET: 9 OCP + PBO: 10 **Number of withdrawal and reasons:** OCP + MET: 2: Loss of F/U 1/ 9% Moving away: 1/ 9% OCP + PBO: 2: Loss of F/U 1/ 8.3% Menorrhagia: 1/ 8.3% **Summary characteristics:** PCOS, all participants of the study had hyperandrogenaemia **BMI (kg/m^2):** OCP + PBO mean (±SD): 32.6 (2.3) OCP + MET mean (±SD): 36.2 (2.5)
Interventions	**Treatment:** OCP (Ethinyl estradiol 35 norgestimate 0.18/0.215/0.25) + MET 500 mg/day/1 week then twice a day/1 week then three times a day **Control:** OCP (Ethinyl estradiol 35 norgestimate 0.18/0.215/0.25) + PBO **Duration:** 3 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters: 1. Adverse events: severe (requiring stopping of medication) and minor 2. Body weight (kg) and BMI (kg/m^2)

Essah 2011 *(Continued)*

	3. Blood pressure (systolic, diastolic) (mm Hg) (b) Hormonal parameters: 1. Serum total testosterone (nmol/L) (c) Metabolic parameters: 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Power calculation:** done, a priori, on insulin sensitivity

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The VCU Investigational Pharmacy prepared metformin and matched placebo capsules, randomized the women with a computerized random number generator using a 1:1 allocation ratio with blocks of random size (three and four women in a block), and concealed the allocation."
Allocation concealment (selection bias)	Unclear risk	Quote: "The VCU Investigational Pharmacy prepared metformin and matched placebo capsules, randomized the women with a computerized random number generator using a 1:1 allocation ratio with blocks of random size (three and four women in a block), and concealed the allocation". Method of allocation concealment not stated. Insufficient information available to permit a judgement of "low risk" or "high risk"
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: “single-center, randomized, double-blind, placebo- controlled trial” but did not state who was double blinded "matched placebo capsule" Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: “single-center, randomized, double-blind, placebo- controlled trial” but did not state who was double blinded "matched placebo capsule" Insufficient information available to permit a judgement of “low risk” or “high risk”.
Incomplete outcome data (attrition bias) All outcomes	High risk	One withdrawal/ 12: 8% in OCP+ PBO is menorrhagia which is a serious adverse event which can have clinical impact.
Selective reporting (reporting bias)	High risk	Except for fasting insulin and fasting glucose all the other outcomes are not described in the methods and given in an “annexe”. One or more primary outcomes have been reported using measurements, analysis, methods or subset of the data THAT WERE NOT PRE-SPECIFIED
Other bias	Low risk	The study appears to be free of other sources of bias.

Feng 2016

Study characteristics	
Methods	Randomised controlled trial **Location of the trial**: quote: Zhengzhou, China **Method of randomisation:** quote: "The study participators were randomly divided into two groups using a table of random numbers." **Method of allocation concealment:** quote: **"**Preparation of the medication was performed in a centralized manner, and labeling, with the exception of the relevant randomization code, was identical in all presentations." **Source of fundings:** not stated
Participants	**Inclusion criteria:** PCOS: Rotterdam criteria: (1) oligo/anovulation, (2) signs of hyperandrogenism (i.e. hirsutism and acne) or (3) enhanced ovaries (at least 12 discrete follicles of 2 mm to 9 mm in diameter in one ovary or the ovarian volume > 10 cm3 observed by transvaginal ultrasonography)**Exclusion criteria:** women with androgen-excess disorders or patients with specific aetiologies such as congenital adrenal hyperplasia, Cushing's syndrome, thyroid hormone abnormalities, hyperprolactinaemia, or ovarian/adrenal tumours. Furthermore, all patients had no history of previous first-trimester miscarriage or pregnancy. The patients with more than one outlier baseline parameter data using the Grubbs test were also excluded in the following clinical trial study.**Number of women randomised:** 82: OCP + MET arm: 41 OCP arm: 41**Number of women analysed:** not stated**Number of withdrawal and reasons:** not stated**Summary characteristics:** PCOS, mean age for the whole cohort: 29.0 years**Age (years):** OCP + MET mean (± SD): 27.86 (3.79) OCP mean(± SD): 28.57 (3.04)**BMI (kg/m^2):** OCP + MET mean (± SD) 29.46 (4.43) OCP mean(± SD) 27.77 (4.23)
Interventions	**Treatment:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2 mg) cyclically 21 days stop 7 days + MET 425 mg twice a day then 850 mg twice a day **Control:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2 mg) cyclically 21 days stop 7 days **Duration:** 3 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:** Acne score BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg)

Feng 2016 *(Continued)*

	Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1.Hirsutism score 2.Acne score
Objective outcomes	(a) Clinical parameters 1. BMI (kg/m^2) 2. Blood pressure (systolic) (mm Hg) 3. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1.Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about** the number of patient randomised in each group and number of withdrawal **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “The study participators were randomly divided into two groups using a table of random numbers”
Allocation concealment (selection bias)	Low risk	Quote: “Preparation of the medication was performed in a centralized manner, and labelling, with the exception of the relevant randomization code, was identical in all presentations”
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "This study was carried out in double-blind conditions, so neither the patient nor the doctor was aware of the composition of the treatment administered.” The “doctor” could be the clinician or outcome assessor. Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: “This study was carried out in double-blind conditions, so neither the patient nor the doctor was aware of the composition of the treatment administered.” The “doctor” could be the clinician or outcome assessor. Insufficient information available to permit a judgement of “low risk” or “high risk”.

Feng 2016 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Paper does not provide any information on participant after randomisation.
Selective reporting (reporting bias)	Unclear risk	Body weight prespecified in the methods not reported in the results. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Glintborg 2014a

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Odense, Denmark**Method of randomisation:** not stated**Method of allocation concealment:** not stated**Source of funding:** quote: "This work was supported by the Jacob Madsen's and Olga Madsen's Foundation, the Institute of Clinical Research, the Odense University Hospital, Kolding Hospital, the A. P. Møller's Foundation, the Bernhard and Marie Kleins Foundation, The Novo Nordisk Foundation, and The Danish Medical Association. Oral contraceptive pills and metformin tablets were sponsored by Sandoz."
Participants	**Inclusion criteria:** aged 18 to 39 years fulfilling the Rotterdam criteria for PCOS were included in the study. The included patients fulfilled two of the criteria: 1) irregular periods during more than a year in combination with a cycle length longer than 35 days; 2) total- or free-T levels above the reference interval (upper limits: total T >1.8 nmol/L, free T >0.035 nmol/L) and/or hirsutism; and 3) transvaginal ultrasound with polycystic ovaries. **Exclusion criteria:** Serious endocrine diseases were excluded. Patients with diabetes (fasting plasma glucose ≥ 7.0 mmol/L and/or HbA1c ≥ 44 mmol/mol), elevated liver enzymes, renal dysfunction, congestive heart disease, depression, and eating disorders were not included in the study. Obese patients (BMI) ≥ 35 kg/m^2) and patients with other contraindications for OCP(previous or family history of thrombosis or breast cancer, coagulatory defects, and heavy smokers) were not included in the study. Patients were not included if they were pregnant or expressed a wish for conception during the study period. Patients paused OCP for at least 3 months and metformin for at least 1 month before evaluation, and no patients were treated with medicine known to affect hormonal or metabolic parameters. **Number of women randomised:** 90 MET + OCP: 30 MET: 30 OCP: 30 **Number of women analysed:** 65 MET + OCP: 23 MET: 19 OCP: 23 **Number of withdrawal and reasons:** MET + OCP: 7 Nausea: 3/ 10% Regrets: 2/ 6.7%

Glintborg 2014a *(Continued)*

	Lost F/U: 2/ 6.7% MET:11 pregnant: 1/ 3.3% wants OCP: 2/ 6.7% depression: 1/ 3.3% nausea: 1/3.3% regrets: 4/ 13.3% lost F/U: 2/ 6.7% OCP: 7 wants pregnancy: 2/ 6.7% adverse events: 3/ 10% lost F/U: 2/ 6.7% **Summary characteristics:** PCOS, aged 18 to 39 years **Age (years):** MET + OCP mean (±SD): 28 (4.9) MET mean (±SD): 28 (4.8) OCP mean (±)SD: 28 (4.7) **BMI (kg/m^2):** MET + OCP mean (±SD): 27.6 (3.9) MET mean (±SD): 25.9 (4.4) OCP mean (±SD): 27.2 (4.8)
Interventions	MET 1000 mg twice a day + OCP (Ethinyl estradiol 30 mcg desogestrel 150 mg) Metformin 1000 mg twice a day OCP (Ethinyl estradiol 30 mcg desogestrel 150 mg) **Duration:** 12 months**Co-intervention(s):** general advice on lifestyle intervention, laser treatment was offered to patient with moderate or severe facial hirsutism and patients were allowed to shave/wax
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (systolic) (mm Hg) Serum total testosterone (nmol/L) Fasting insulin (mIU/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Body weight (kg) 3. BMI (kg/m^2)

Glintborg 2014a *(Continued)*

	4. Blood pressure (systolic) (mm Hg) 5. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. FAI (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** their results because they were expressed in median (25th and 75th quartiles) therefore not applicable, they kindly gave us theirs results presented in the study and a few more outcomes: blood pressure, FAI, fasting glucose, fasting total cholesterol (mmol/L), fasting HDL cholesterol (mmol/L), fasting LDL cholesterol (mmol/L), fasting triglycerides (mmol/L). We could finally calculate the mean ± SD. **Power calculation:** no 3 arms study MET versus OCP versus OCP + MET

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Included patients were randomized to 12 months of treatment with metformin (10001000 mg/d) or OCP (150 mg desogestrel/ 30 mcg ehinylestradiol) or combined treatment (metformin OCP)". Method of randomisation not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reason for missing outcome data "regrets, lost F/U, wants OCP" are unclear and could be related to intervention. insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	Total cholesterol, HDL and triglycerides described in the method not stated in the results part. Improvment of menstrual cycles described for MET group not for OCP group.

Glintborg 2014a *(Continued)*

		But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	High risk	The use of laser treatment, shaving or waxing to treat hirsutism is an important confounding factor to evaluate Ferriman-Gallwey score.

Harborne 2003

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Glasgow, Scotland, UK **Method of randomisation:** quote "Block-randomized (n = 10/block) in a 1:1 ratio to receive either OCP or metformin. Randomization was by random number tables, the patient number treatment codes were held by a third party." **Method of allocation concealment:**quote "The patient number treatment codes were held by a third party and were allocated individually after obtaining written consent" **Source of fundings:** not stated
Participants	**Inclusion criteria:** women with PCOS, whose primary complaint was hirsutism (Ferriman-Gallwey score > 8). PCOS: at least two of the three following features: oligomenorrhoea/ amenorrhoea, polycystic ovaries on ultrasound, or an elevated FAI. **Exclusion criteria:** Contraindications to either metformin or Dianette (including BMI > 38), use of oral contraception or metformin within the previous 3 months, and those taking medication known to affect gonadal or adrenal function, or carbohydrate or lipid metabolism. Thyroid dysfunction, hyperprolactinaemia, diabetes mellitus, or late-on-set congenital adrenal hyperplasia **Number of women randomised:** 52: MET arm: 26 OCP arm: 26 **Number of women analysed:** 34: MET: 18 OCP: 16 **Number of withdrawal and reasons:** 18 MET: 8: Pregnant 3/ 11.5% GI 3/ 11.5% Lost of F/U 2/ 7.7% OCP: 10: Weight gain 5/ 19% BP 1/ 3.8% Depression 1/ 3.8% Chest pain 1/ 3.8% Loss of F/U 2/ 7.7% **Summary characteristics:** PCOS, hirsute women **Age (years):** MET mean (95% CL):

Harborne 2003 *(Continued)*

	31.3 (27.9-34.7) OCP mean (95% CL): 31.7 (26.8-36.5) **BMI (kg/m^2):** MET mean (95% CL): 31.7 (29.5-35.5) OCP mean (95% CL) 31.8 (28.4-34.4)
Interventions	**Treatment(s):** MET 500 mg three times a day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2mg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 12 months **Co-intervention:** none
Outcomes	**Primary outcomes:** Hirsutism score (Ferriman-Gallwey) Hirsutism - subjective (VAS) Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Acne - subjective (VAS) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (pmol/L) Fasting glucose (mmol/L) Total Cholesterol (mmol/L) HDL Cholesterol (mmol/L) LDL Cholesterol (mmol/L) Triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism - subjective (VAS) 2. Hirsutism score (Ferriman-Gallwey) 3. Acne - subjective (VAS)
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) 3. Blood pressure (systolic) (mm Hg) 4. Blood pressure (diastolic) (mm Hg)

Harborne 2003 *(Continued)*

	(b) Hormonal parameters 1. Serum total testosterone (n mol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting insulin (pmol/L) 2. Fasting glucose (mmol/L) 3. Total Cholesterol (mmol/L) 4. HDL Cholesterol (mmol/L) 5. LDL Cholesterol (mmol/L) 6. Triglycerides (mmol/L)
Notes	**Authors contacted** about: BMI result and co intervention kindly provided by the authors that was not in the original paper. **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Patients were block-randomized (n=10/block) in a 1:1 ratio. Randomization was by random number tables"
Allocation concealment (selection bias)	Low risk	Quote: "The patient number treatment codes were held by a third party and were allocated individually after obtaining written consent."
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"Loss to F/U" reason is unclear and could be related to intervention. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Low risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
Other bias	Low risk	The study appears to be free of other sources of bias.

Hoeger 2008a

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: New York, USA **Method of randomisation:** quote:"'Subjects were randomized to one of four arms by random number assignment: placebo, metformin, OC, or lifestyle management" **Method of allocation concealment:** not stated

Hoeger 2008a *(Continued)*

	Source of funding: this work was supported by Grants K23 HD043881-01A1 and 5R03 HD41989-02 from the National Institutes of Health (NIH), and Grant UL1RR024160 from the National Center for Research Resources (NCRR), a component of the NIH and NIH Roadmap for Medical Research. Desogen was supplied by Organon Pharmaceuticals, Inc.
Participants	**Inclusion criteria:** from methods section of paper and Clinical Trial Registry Postmenarchal adolescent women between the ages of 12 and 18 years withBMI above the 95th percentile (or BMI > 25 kg/m^2) and evidence of menstrual irregularity (fewer than eight menses in the preceding year or cycle length > 45 days) and clinical or biochemical evidence of hyperandrogenism (testosterone > 50 ng/dL), currently healthy, studied off any hormonal therapy or insulin sensitisers for at least 2 months, non-smokers. PCOS: Consistent with NIH. **Exclusion criteria:** from methods section of paper and Clinical Trial Registry Exclusion of other causes of hyperandrogenism and menstrual irregularity Exclusion of diabetes mellitus, Cushing's disease, hyperprolactinaemia, untreated hypo or hyperthyroidism, history of adrenal hyperplasia, significant renal impairment, received oral contraceptives, oestrogen or progestin or other drugs known to effect lipoprotein metabolism within 2 months of starting study, exercise > 10 hours/week. **Number of women randomised:** 21 MET: 10 OCP: 11 **Number of women analysed:** 16 MET: 6 OCP: 10 **Number of withdrawal and reasons:** MET: 4 lost to follow-up: 2/ 20% personal reasons: 2/ 20% OCP: 1 lost to follow-up: 1/ 9% **Summary characteristics:** PCOS, young (12 to 18years), obese **Age (years):** MET mean (±SD): 16 (1.7) OCP mean (±SD): 15.4 (1.4) **BMI (kg/m^2):** MET mean (± SD): 34.3 (6.5) OCP mean (± SD): 37.8 (6.1)
Interventions	**Treatment:** MET 850 mg twice a day, starting as single doses of 425 mg and building gradually to two capsules twice a day. **Control:** OCP (Ethinyl estradiol 30 mcg desogestrel 0.15 mg) taken in a cyclic fashion. **Duration:** 6 months **Co-intervention(s):** "Subjects who were not assigned to lifestyle treatment received standard office advice on nutrition and exercise for healthy living and were seen monthly."
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Blood pressure (systolic, diastolic) (mm Hg)

Hoeger 2008a *(Continued)*

	Serum total testosterone (ng/dL) Free androgen index (FAI) (%) Fasting insulin (IU/mL) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles 2. BMI (kg/m^2) 3. Blood pressure (systolic, diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (ng/dL) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting insulin (IU/mL) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** fasting insulin because expressed in wrong unit. **Power calculation:** unclear 4 arms treatment: PBO V MET versus OCP versus LS

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Subjects were randomized to one of four arms by random number assignment: placebo, metformin, OC, or lifestyle management"
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of ‘low risk’ or ‘high risk’
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Quote: "Assignment to metformin or placebo was blinded to subject and investigator." "outcome assessors and data analysts were unaware of study assignment"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Assignment to metformin or placebo was blinded to subject and investigator." "outcome assessors and data analysts were unaware of study assignment"

Hoeger 2008a *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Lost F/U, personal reason" are both unclear reason of withdrawal and could be related to intervention. Insufficient reporting attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	From results section of paper and clinical trial registry, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of weight which was only reported as part of BMI and menstrual rate which was not reported for OCP. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Hoeger 2008b

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: New York, USA **Method of randomisation:** quote: "Subjects were assigned by a random number table." **Method of allocation concealment:** not stated **Source of funding:** this work was supported by Grants K23 HD043881-01A1 and 5R03 HD41989-02 from the National Institutes of Health (NIH), and Grant UL1RR024160 from the National Center for Research Resources (NCRR), a component of the NIH and NIH Roadmap for Medical Research. Desogen was supplied by Organon Pharmaceuticals, Inc..
Participants	**Inclusion criteria:** Postmenarchal adolescent women between the ages of 12 and 18 yr with BMI above the 95th percentile and evidence of menstrual irregularity (fewer than eight menses in the preceding year) and clinical or biochemical evidence of hyperandrogenism, currently healthy, studied off any hormonal therapy or insulin sensitisers for at least 2 months, non-smokers, must be able to swallow capsules, at least 6 months since onset of first menstrual cycle. PCOS diagnosis consistent with NIH. **Exclusion criteria:** exclusion of other causes of hyperandrogenism and menstrual irregularity; diabetes; kidney or lLiver disease; tobacco use; depression or bipolar disease; contraindication to exercise; weight > 300 lbs **Number of women randomised:** 36 MET + OCP: 18 OCP: 18 **Number of women analysed:** 32 MET + OCP: 16 OCP: 16 **Number of withdrawal and reasons:** MET + OCP: 2, GI symptoms: 1/ 5.6% not stated: 1/ 5.6% OCP: 2, GI symptoms: 1/ 5.6% not stated: 1/ 5.6% **Summary characteristics:** PCOS, young (12 to 18years), obese

Hoeger 2008b *(Continued)*

	Age (years): MET + OCP mean(± SD): 14.7 (1.6) OCP mean(± SD): 15.8 (1.6) **BMI (kg/m^2):** MET + OCP OCP mean (± SD): 34.1 (4.3) OCP mean (± SD): 35.7 (4.9)
Interventions	**Treatment:** MET 2000 mg/day (divided into four doses with a gradual build up at study start) + OCP (Ethinyl estradiol 30 mcg drosporenone 3 mg) **Control:** OCP (Ethinyl estradiol 30 mcg drospirenone 3 mg) + PBO **Duration:** 6 months **Co-intervention(s):** co-intervention with lifestyle (24-week program with adult family member in classes for training for diet, exercise and behaviour modification skills in open group format with rolling admission). Frequent contact with participants and a combination of structured lectures and flexible personal strategies emphasised self-esteem and social support. The format included a core curriculum taught in separate adolescent and parent groups repeated over the 24 weeks, interspersed with individual appointments. Exercise was monitored in a group setting on a weekly basis. Contact between participants was encouraged outside the weekly setting with electronic communications such as Internet chat groups to support continued lifestyle changes. Structured group exercise was included weekly. Goals of therapy were based on initial caloric intake by diet record aiming for a 500 kcal/day deficit. Exercise was recommended to include 30 minutes/day of moderate to intense activity.
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (IU/mL) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. Body weight (kg) and/or BMI (kg/m^2) 4. Blood pressure (systolic, diastolic) (mm Hg)

Hoeger 2008b *(Continued)*

	(b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting insulin (IU/mL) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** fasting insulin because expressed in wrong unit. **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Subjects were assigned by a random number table to metformin or placebo."
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of ‘low risk’ or ‘high risk’
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "Double Blind (Subject, Investigator)". Unclear if investigator blinding refers to blinding for clinician reported outcomes. Potential for lack of clinician blinding to impact on results reporting and result in bias for clinician reported outcomes.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "Double Blind (Subject, Investigator)". Unclear if investigator blinding refers to blinding for clinician reported outcomes. Potential for lack of clinician blinding to impact on results reporting and result in bias for clinician reported outcomes.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reason for missing outcome data "not stated" in 2 cases could be related to intervention. Insufficient reporting of attrition/exclusion to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	From results section of paper and clinical trial registry, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of weight which was only reported as part of BMI, menstrual rate. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Jin 2006

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Beijing, China **Method of randomisation:** quote: "Randomly divided into 4 groups" **Method of allocation concealment:** not stated **Source of funding:** Not stated
Participants	**Inclusion criteria:** women with PCOS according to Rotterdam criteria, normal liver and kidney function, no treatment provided 3 months prior to the study. **Exclusion criteria:** congenital adrenal hyperplasia (CAH); Cushing's syndrome; hyperprolactinaemia; thyroid disease; hormone-secreting tumours. **Number of women randomised:** 24 MET: 15 OCP: 10 **Number of women analysed:** 24 MET: 15 OCP: 10 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS **Age (years):** MET mean (±SD): 23.3 (6.53) OCP mean (±SD): 23 (6.11)
Interventions	**Treatment:** MET oral, 500 mg/time, three times a day **Control:** OCP (Ethynil Estradiol 35 mcg/ Cyproterone acetate 25 mg) start on day 5 of menstrual cycle, continued for 21 days; one pill a day **Duration:** 3 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism (assessed clinically and subjectively) **Secondary outcomes:** Improved menstrual pattern (i.e. an initiation of menses or significant shortening of cycles or intermenstrual days) Acne (assessed clinically and subjectively) BMI (kg/m^2) Serum total testosterone (n mol/L) Fasting insulin (mIU/L)
Subjective outcomes	(a) Clinical parameters Hirsutism (assessed clinically and subjectively) Acne (assessed clinically and subjectively)
Objective outcomes	(a) Clinical parameters Improved menstrual pattern (i.e. an initiation of menses or significant shortening of cycles or intermenstrual days)

Jin 2006 *(Continued)*

	BMI (kg/m^2)
	(b) Hormonal parameters
	Serum total testosterone (n mol/L)
	(c) Metabolic parameters
	Fasting insulin (mIU/L)
Notes	**Power calculation:** no

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomly divided into 4 groups" Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	High risk	Hirsutism and acne described with score in the methods (Ferriman-Gallwey score; acne score) but reported in the results as number of patients who had an improvement. BMI not reported in the results. One or more primary outcomes have been reported using measurements, analysis methods or subset of the data that were not prespecified.
Other bias	High risk	Baselin imbalance.

Kaya 2015

Study characteristics	
Methods	Randomised controlled trial

Kaya 2015 *(Continued)*

	Location of the trial: quote: Kayseri, Turkey**Method of randomisation:** quote: "The study population randomized into two different groups."**Method of allocation concealment:** not stated**Source of funding:** Not stated
Participants	**Inclusion criteria:** PCOS was defined as the presence of two of the following criteria after the exclusion of other aetiologies: (i) polycystic ovaries on ultrasound examination, (ii) chronic oligomenorrhoea or amenorrhoea, (iii) clinical or biochemical evidence of hyperandrogenism **Exclusion criteria:** exclusion criteria included diabetes mellitus, corticosteroid use, use of drugs affecting insulin resistance, hyperlipidaemia, hypertension, oral contraceptive use, evidence of ongoing infection, presence of severe valve disease, pregnancy, systemic disease (hepatic, renal, cardiac), smoking, aortic disease (coarctation, aneurism, Marfan syndrome or history of aortic surgery). **Number of women randomised:** 50 MET + OCP: 25 OCP: 25 **Number of women analysed:** 50 MET + OCP: 25 OCP: 25 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, aged 17-37 years **Age (years):** MET + OCP mean (±SD): 24 (4) OCP mean (±SD): 23 (5) **BMI (kg/m^2):** MET + OCP mean (±SD): 29.8 (6.9) OCP mean (±SD): 26.7 (5.7)
Interventions	**Treatment:** MET 850 mg twice a day + OCP (Ethinyl estradiol 3mcg drospirenone 3 mg) 21days then 7 days placebo **Control:** OCP (Ethinyl estradiol 3 mcg drospirenone 3 mg) 21days then 7 days placebo **Duration:** 6 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** None **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None

Kaya 2015 *(Continued)*

Objective outcomes	(a) Clinical parameters 1. Body weight (kg) 2. (BMI (kg/m^2) 3. Blood pressure (systolic) (mm Hg) 4. Blood pressure (diastolic) (mm Hg) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting total cholesterol (mmol/L) 2. Fasting HDL cholesterol (mmol/L) 3. Fasting LDL cholesterol (mmol/L) 4. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** testosterone because expressed in the wrong unit. **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Then, the study population randomized into two different groups." Method of randomisation not described. Insufficient information about the sequence generation process available to permit a judgment of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Unclear risk	Insufficient information available to permit a judgement of 'low risk' and 'high risk'. Protocole not available and measurements are not described in the methods.
Other bias	Low risk	The study appears to be free of other sources of bias.

Kebapcilar 2009a

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Izmir, Turkey**Method of randomisation:** unclear, quote: "Patients were randomly assigned to one of four treatment groups and each group contained 12 participants."**Method of allocation concealment:** not stated**Source of funding:** none
Participants	**Inclusion criteria:** PCOS was diagnosed according to the criteria of the Rotterdam European Society of Human Reproduction and Embryology–American Society for Reproductive Medicine PCOS consensus workshop group. All patients fulfilled the PCOS criteria. All women had normal thyroid, renal and hepatic functions. **Exclusion criteria:** None of them had late-onset congenital adrenal hyperplasia. Women who had any medication (e.g., antihypertensives, oral anti-diabetics, oral contraceptives, antiandrogens, statins, warfarin, antidepressant medication, and GnRH agonists and antagonists) in the preceding 3 months were not included. Women who are current smokers and with hypertension, diabetes, history of coronary heart disease, known coagulation abnormalities were also excluded. Fasting glucose levels were normal in all patients (< 100 mg/dL). None of the participants had chronic alcohol consumption. **Number of women randomised:** 36 OCP + MET: 12 OCP: 12 MET: 12**Number of women analysed:** 36 OCP + MET: 12 OCP: 12 MET: 12 **Number of withdrawal and reasons:** 0**Summary characteristics:** PCOS, 24.0 + 5.4 years; BMI 27.9 + 5.28 kg/m^2 **Age (years):** MET+ OCP mean (+ SD): 24.9 (4.8) MET mean (+ SD): 24.4 (6.2) OCP mean (+ SD): 23.2 (5.1)**BMI (kg/m^2):** MET+ OCP mean (+ SD): 27.6 (3) MET mean (+ SD): 27.8 (4) OCP mean (+ SD): 28.7 (6)
Interventions	MET 850 mg twice a day + OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21days stop 7 days OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21days stop 7 days MET 850 mg twice a day **Duration:** 3 months**Co-intervention(s):** the participants were instructed to limit fat intake and improve dietary behaviour without giving any calory restricted diet program.
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** BMI (kg/m^2) Fasting insulin (mIU/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)

Kebapcilar 2009a *(Continued)*

Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting total cholesterol (mmol/L) 3. Fasting HDL cholesterol (mmol/L) 4. Fasting LDL cholesterol (mmol/L) 5. Fasting triglycerides (mmol/L)
Notes	**Power calculation:** yes, a priori, on fasting insulin 3 arms study MET versus OCPversusV OCP + MET

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients were randomly assigned to one of four treatment groups and each group contained 12 participants." Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Unclear risk	Method of concealment not stated. Insufficient information available to permit a judgment of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Unclear risk	The study protocol is not available and clinical measurements are not described in the methods.
Other bias	Low risk	The study appears to be free of other sources of bias.

Kebapcilar 2009b

Study characteristics

Methods	Randomised controlled trial

Kebapcilar 2009b *(Continued)*

	Location of the trial: quote: Izmir, Turkey**Method of randomisation:** quote: "Patients were randomly assigned"**Method of allocation concealment:** not stated**Source of funding:** not stated
Participants	**Inclusion criteria:** PCOS consistent with Rotterdam criteria PCOS according to the criteria of the Rotterdam European Society of Human Reproduction and Embryology– American Society for Reproductive Medicine 1) oligomenorrhoea or amenorrhoea, 2) clinical (hirsutism, acne) and/ or biochemical signs of hyperandrogenism, and 3) polycystic ovaries) All women had normal thyroid, renal and hepatic functions. None of them had late-onset congenital adrenal hyperplasia. All patients were normal (< 100 mg/dL) fasting glucose levels.**Exclusion criteria:** any medication (e.g., antihypertensives, oral antidiabetics, oral contraceptives, antiandrogens, statins, vitamins, warfarin, antidepressant medication, and GnRH agonists and antagonists) in the preceding 3 months. Hypertension, diabetes, history of coronary heart disease, known coagulation abnormalities and current smokers and chronic alcohol users.**Number of women randomised:** 43 MET + OCP: 21 OCP: 22**Number of women analysed:** 43 MET + OCP: 21 OCP: 22**Number of withdrawal and reasons:** 0**Summary characteristics:** PCOS, mean age 24.6 (5) / mean BMI 27.9 (5.4) **Age (years):** MET + OCP mean (±SD): 26.1 (4.4) OCP mean (±SD): 24.1(5.6) **BMI (kg/m^2):** MET + OCP mean (± SD): 28.7(4.4) OCP mean (± SD): 27.2 (6.2)
Interventions	**Treatment:** MET 850 mg three times a day + OCP (Ethynil Estradiol 35 mcg Cyproterone acetate 2 mg) 21 days per month followed by a 7-day pill-free period **Control:** OCP (Ethynil estradiol 35 mcg Cyproterone acetate 2 mg) for 3 months 21 days per month followed by a 7-day pill-free period**Duration:** 3 months.**Co-intervention(s):** none
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** BMI (kg/m^2) Fasting insulin (µIU/L) Fasting glucose (mg/dL) Fasting total cholesterol (mg/dL) Fasting HDL cholesterol (mg/dL) Fasting LDL cholesterol (mg/dL) Fasting triglycerides (mg/dL)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) (c) Metabolic parameters 1. Fasting insulin (µIU/L)

Kebapcilar 2009b *(Continued)*

	2. Fasting glucose (mg/dL) 3. Fasting total cholesterol (mg/dL) 4. Fasting HDL cholesterol (mg/dL) 5. Fasting LDL cholesterol (mg/dL) 6. Fasting triglycerides (mg/dL)
Notes	**Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients were randomly assigned". Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "no subject were lost during the F/U and none of the 43 patients had stopped the therapies because of adverse effects" No missing outcome data.
Selective reporting (reporting bias)	Unclear risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of glucose which was only reported as HOMA. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Kilic 2011

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: Ankara, Turkey **Method of randomisation:** quote: "computer-based randomized prospective analyses" **Method of allocation concealment:** not stated **Source of funding:** not stated

Kilic 2011 *(Continued)*

Participants	**Inclusion criteria:** PCOS according to the menstrual, laboratory, and ultrasonographic criteria of Rotterdam 2003. (Menstrual criteria included oligo- or amenorrhoea, irregular cycle length, > 45 days or less than six periods a year, ultrasound criteria used for diagnosis). Patients with normal androgen levels (the normal range values for serum-free testosterone in our laboratory <3.17 ng/mL) with non obese and obese women with PCOS were selected. **Exclusion criteria:** for all participants included age less than 18 years or greater than 35 years, smoking, folic acid and vitamin B12 deficiency, pregnancy, hypothyroidism, hyperprolactinaemia, Cushing's syndrome, nonclassical congenital adrenal hyperplasia (17 OHP <5 ng/dl) and current or previous (within the last 6 months) use of oral contraceptives, glucocorticoids, antiandrogens, ovulation induction agents, antidiabetic and anti-obesity drugs, or other hormonal drugs. Women with clinical and/or biochemical hyperandrogenism alone were excluded from the study. **Number of women randomised:** 105 MET: 54: 28 BMI ≥ 25 / 26 BMI < 25 OCP: 51: 26 BMI ≥ 25 / 25 BMI < 25 **Number of women analysed:** 96 MET: 47 OCP: 49 **Number of withdrawal and reasons:** MET: Loss of F/U: 5 / 9.3% Discontinu: 2 / 3.7% OCP: Discontinu: 2 / 3.9% **Summary characteristics:** PCOS, normoandrogenaemic and oligoamenorrhoeic with impaired glucose tolerance **Age (years):** MET mean (± SD): BMI ≥ 25: Age (years): 28.7 (3.7) BM I< 25: Age (years): 26.3 (3) OCP mean (± SD): BMI ≥ 25: Age (years): 29 (3.5) BMI < 25: Age (years): 26.7 (3.8) **BMI (kg/m^2):** MET mean (± SD): BMI ≥ 25 BMI (kg/m^2): 31.5 (2.1) BMI < 25 BMI (kg/m^2): 23,37 (1.64) OCP mean (± SD): BMI ≥ 25 BMI (kg/m^2): 27.7 (0.9) BM I< 25 BMI (kg/m^2): 21.63 (1.47)
Interventions	**Treatment:** MET 850 mg twice a day **Control:** OCP (Ethinyl estradiol 30 mcg desogestrel 0.15 mg) **Duration:** 6 months **Co-intervention(s):** B-groups vitamins vitamin B1.250 mg; vitamin B6.250 mg; vitamin B12.1000 mg, twice daily
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** BMI (kg/m^2)
Subjective outcomes	None

Kilic 2011 *(Continued)*

Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2)
Notes	**Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “computer-based randomized prospective analyses”.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"Loss of F/U" or "discontinue treatment" unclear reason of withdrawal and could be related to intervention. Insufficient report of attrition/ exclusions to permit a judgement of ‘low risk’ or ‘high risk’
Selective reporting (reporting bias)	Unclear risk	Testosterone is not reported in the results. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	High risk	Baseline characteristic were not similar: BMI in metformin group higher than in OCP group in the subgroup BM I≥ 25kg/m^2 MET mean (± SD): 31.5 (2.1) / OCP mean (±SD): 27.7 (0.9)

Kuek 2011

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: Shanghai, China **Method of randomisation:** unclear quote: “randomly assigned” **Method of allocation concealment:** not stated **Source of funding:** not stated
Participants	**Inclusion criteria:** Rotterdam criteria of PCOS, the diagnostic and inclusion criteria are as follows. (1) irregular ovulation: oligomenorrhoea, amenorrhoea or irregular menstrual cycle. (2) obvious signs of

Kuek 2011 *(Continued)*

	hyperandrogen or increased androgen level or anomalies of circulation androgens. (3) polycystic ovary: ultrasound showed more than 12 follicles of 2 mm to 9 mm in diameter per side or 10 cm3 of ovarian volume in 3 to 5 days of menstruation. Patient with any 2 of the above-mentioned signs or symptoms are then diagnosed with PCOS. **Exclusion criteria:** (1) not meeting the above diagnostic criteria for PCOS. (2) other reasons causing hyperandrogenism: late-onset congenital hyperplasia, Cushing's syndrome, androgen secretion of ovarian or adrenal tumours. (3) hyperprolactinaemia or hypothalamic amenorrhoea. (4) thyroid dysfunction. **Number of women randomised:** 28 MET: 17 OCP: 11 **Number of women analysed:** 28 MET: 17 OCP: 11 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, aged from 15 to 40 years old **BMI (kg/m^2):** MET mean (± SD): 25.7 (2.96) OCP mean (± SD): 19.27 (10.46)
Interventions	**Treatment:** MET 500 mg twice a day **Control:** OCP (Ethinyl Estradiol 35 Cyproterone acetate 2 mg)/day, 21days on the 5th day of menstruations **Duration:** 3 months **Co-intervention(s):** quote: "patient in group A and B who faced amenorrhoea within 60d during treatment were given medroxyprogesterone acetate for a total of 5 days to induce menstruation"
Outcomes	**Primary outcomes:** none **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (mIU/L) Fasting glucose (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Body weight (kg) 2. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%)

Kuek 2011 *(Continued)*

	(c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L)
Notes	**Authors contacted about:** the outcome fasting insulin because expressed in a wrong unit. **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “randomly assigned”. insufficient information available to permit a judgement of “low risk” or “high risk”.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified.
Other bias	High risk	Number of patients randomised in each group is different 17 versus 11. Baseline imbalance.

Kumar 2018

Study characteristics

Methods	Randomised controlled trial **Locationof the trial:** quote: New Delhi, India **Method of randomisation:** quote: "Computer-generated random sequence numbering based on the intervention received" **Method of concealment:** not stated **Source of fundings:** none
Participants	**Inclusion criteria:** PCOS Rotterdam criteria for the diagnosis, aged between 18 and 40 years, symptom duration > 6 months, premenopausal, normal thyroid function.

Kumar 2018 *(Continued)*

	Exclusion criteria: patients with a history of using any drug therapy (insulin sensitisers, hormone therapy, calcium, and Vitamin D), use of other drugs that affect the body composition and androgen levels (insulin, glucocorticoids) pregnancy or lactation **Number of women randomised:** 90 OCP + MET arm: 30 MET arm: 30 OCP arm: 30 **Number of women analysed:** 87 OCP + MET: 29 MET: 30 OCP: 28 **Number of withdrawal and reasons:** 3 OCP + MET: 1 reason NA OCP: 2 reason NA **Summary characteristics:** The study population had a mean age 23.2 ± 4.4 year and BMI of 28.4 ± 6.1 kg/m^2. The entire study population falls into the obese category and had significant hirsutism. **Age (years):** OCP + MET mean (± SD): 24.1 (5.9) MET mean (± SD): 22 (5.2) OCP mean (± SD): 22.9 (5) **BMI (kg/m^2):** OCP + MET mean (± SD): 30.1 (5.5) MET mean (± SD): 27.14 (6) OCP mean (± SD): 26.15 (4.9)
Interventions	OCP (Ethinyl estradiol 35 mcg, Cyproterone acetate 2 mg) standard combination pill MET 500 mg/day gradually increased to 2000 mg MET 500 mg/day gradually increased to 2000 mg + OCP (Ethinyl estradiol 35 mcg, Cyproterone acetate 2 mg) standard combination pill **Duration:** 6 months **Co-intervention(s):** all the patients were explained about healthy lifestyle measures and dietary advice to reduce weight
Outcomes	**Primary outcomes:** Hirsutism score **Secondary outcomes:**

Kumar 2018 *(Continued)*

	Menstrual cyclicity, initiation of menses or significant shortening of cycles Acne score Body weight (kg) and/or BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) clinical parameters 1.Hirsutism score 2.Acne score
Objective outcomes	(a) clinical parameters 1.Menstrual cyclicity, initiation of menses or significant shortening of cycles 2.Body weight (kg) and/or BMI (kg/m^2) (b) hormonal parameters 1.Serum total testosterone (nmol/L) (c) metabolic parameters 1.Fasting insulin (mIU/L) 2.Fasting glucose (mmol/L) 3.Fasting total cholesterol (mmol/L) 4.Fasting HDL cholesterol (mmol/L) 5.Fasting LDL cholesterol (mmol/L) 6.Fasting triglycerides (mmol/L)
Notes	**Power calculation:** unclear 3 arms study MET versus OCP versus OCP + MET

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “computer-generated random sequence numbering based on the intervention received”
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated Insufficient information available to permit a judgement of “low risk” or “high risk”
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.

Kumar 2018 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reason of the dropout of the patient in the MET + OCP and OCP group not stated. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	Regarding acne score the results are given only for severe acne score not for all the patients. Regarding the menstrual cyclicity described in the methods but nothing is described in the results But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Liu 2006

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Nanjin, China **Method of randomisation:** quote: "randomly assigned to three groups" **Method of allocation concealment:** not stated **Source of funding:** not stated
Participants	**Inclusion criteria:** Women with PCOS without any precision **Exclusion criteria:** no hormonal or insulin secretion treatment three months prior entering study. **Number of women randomised:** 40 MET: 20 OCP: 20 **Number of women analysed:** 40 MET: 20 OCP: 20 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, age: 28.4 +/- 3.5 years
Interventions	**Treatment:** MET 0.25 g three times a day + OCP (Ethinyl Estradiol 35 mcg/ Cyproterone acetate 25 mg),1pill/day, 21 days, starts day 5 of the menstrual cycle. **Control:** MET only for 6 month: 0.25 g, three times a day **Duration:** 6 months **Co-intervention(s):**nNone
Outcomes	**Primary outcomes:** none **Secondary outcomes:** BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Total cholesterol (mmol/L) HDL cholesterol (mmol/L)

Liu 2006 *(Continued)*

	Triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters BMI (kg/m^2) (b) Hormonal parameters Serum total testosterone (nmol/L) Free androgen index (FAI) (%) (c) Metabolic parameters Fasting insulin (mIU/L) Fasting glucose (mmol/L) Total cholesterol (mmol/L) HDL cholesterol (mmol/L) Triglycerides (mmol/L)
Notes	**Power calculation:** no Third arm treatment not included: MET (6 months) + OCP only 3 months

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomly divided into 4 groups" Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Low risk	The study protocol is not available, but it is clear that the published reports include all expected outcomes, including those that were pre-specified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Luque-Ramirez 2007a

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: Madrid, Spain**Method of randomisation:** quote: “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane35 Diario and five patients to receive metformin”**Method of allocation concealment:** “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “sealed opaque envelopes” **Source of funding:** "This study was supported by the Spanish Ministry of Health and Consumer Affairs, Instituto de Investigacion Carlos III Grants Fondo de Investigacion Sanitaria and Red de Diabetes Enfermedades Metabolicas Asociadas; Ministry of Education and Science Grant; and by economic aid from Hospital Ramon y Cajal."
Participants	**Inclusion criteria:** the diagnosis of PCOS was based on the presence of clinical and/or biochemical hyperandrogenism, oligo-ovulation, and exclusion of secondary aetiologies. Hirsutism was defined by a modified Ferriman-Gallwey score above 7 and oligomenorrhoea (more than six cycles longer than 36 days in the previous year) or amenorrhoea (absence of menstruation for three consecutive months), or luteal phase progesterone measurements less than 4 ng/ml (12.72 nmol/L) in women with regular menstrual cycles were considered indicative of oligo-ovulation. **Exclusion criteria:** secondary aetiologies, including hyperprolactinaemia, thyroid dysfunction, Cushing’s syndrome, congenital adrenal hyperplasia, and virilising tumours, were actively ruled out in all the patients. None of the patients had a personal history of hypertension, diabetes mellitus, or cardiovascular events, or received treatment with oral contraceptives, antiandrogens, insulin sensitisers, or drugs that might interfere with blood pressure regulation, lipid profile, or carbohydrate metabolism for the previous 6 months. **Number of women randomised:** 34 MET: 19 OCP: 15 **Number of women analysed:** 27 MET: 12 OCP: 15 **Number of withdrawal and reasons:** MET: protocol violation: 3/ 15.8% GI adverse events: 2/ 10.5% Pregnancy: 1/ 5.3% Lost F/U: 1/ 5.3% OCP: 0 **Summary characteristics:** PCOS, hyperandrogenic (precised in an other study with same population) **Age (years):** MET mean (± SD): 25.1 (6.6) OCP mean (± SD): 23.4 (5.6) **BMI (kg/m^2):** MET mean (± SD): 30.5 (6.9) OCP mean (± SD): 29.2 (5.7)
Interventions	**Treatment:** MET 425 mg twice a day 1 week then 850 mg twice a day**Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days followed by 7 days placebo pills**Duration:** 6 months**Co-intervention(s):** all the patients were instructed to maintain a diet containing 25 to 30 kcal per kg of body weight per day and moderate physical activity throughout the trial, although these measures were not stressed thereafter.

Luque-Ramirez 2007a *(Continued)*

Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Free androgen index (FAI) (%) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** the results because it was presented as a graph.**Power calculation:** yes, a priori, on secondary outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “patients were randomized to receive an antiandrogenic low-dose oral contraceptive pill or 850 mg of metformin twice daily for 24 wk”, “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane 35 Diario and five patients to receive metformin” Insufficient information about the sequence generation process available to permit a judgement of ‘low risk’ or ‘high risk’.
Allocation concealment (selection bias)	Low risk	Quote: “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment” and “sealed opaque envelopes”.

Luque-Ramirez 2007a *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “No masking method was used after randomization”.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “No masking method was used after randomization”.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	7 patients in MET group did not complete the entire study/ 0 in OCP group. Quote: "protocol violation and lost to F/U" unclear if related to intervention. Insufficient reporting of attrition /exclusions to permit a judgement of 'low risk' or 'high risk."
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Luque-Ramirez 2007b

Study characteristics	
Methods	Randomised controlled trial Same population as Luque-Ramirez 2007a **Location of the trial:** quote: Madrid, Spain**Method of randomisation:** quote: “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane35 Diario and five patients to receive metformin”**Method of allocation concealment:** “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “sealed opaque envelopes” **Source of funding:** "This study was supported by the Spanish Ministry of Health and Consumer Affairs, Instituto de Investigacion Carlos III Grants Fondo de Investigacion Sanitaria and Red de Diabetes Enfermedades Metabolicas Asociadas; Ministry of Education and Science Grant; and by economic aid from Hospital Ramon y Cajal."
Participants	**Inclusion criteria:** the diagnosis of PCOS was based on the presence of clinical and/or biochemical hyperandrogenism, oligo-ovulation, and exclusion of secondary aetiologies. Hirsutism was defined by a modified Ferriman-Gallwey score above 7 and oligomenorrhoea (more than six cycles longer than 36 days in the previous year) or amenorrhoea (absence of menstruation for three consecutive months), or luteal phase progesterone measurements less than 4 ng/ml (12.72 nmo/L) in women with regular menstrual cycles were considered indicative of oligo-ovulation. **Exclusion criteria: s**econdary aetiologies, including hyperprolactinaemia, thyroid dysfunction, Cushing’s syndrome, congenital adrenal hyperplasia, and virilising tumours, were actively ruled out in all the patients. None of the patients had a personal history of hypertension, diabetes mellitus, or cardiovascular events, or received treatment with oral contraceptives, antiandrogens, insulin sensitisers, or drugs that might interfere with blood pressure regulation, lipid profile, or carbohydrate metabolism for the previous 6 months. **Number of women randomised:** 34 MET: 19 OCP: 15 **Number of women analysed:** 27 MET: 12 OCP: 15 **Number of withdrawal and reasons:** MET: protocol violation: 3/ 15.8% GI adverse events: 2/ 10.5%

Luque-Ramirez 2007b *(Continued)*

	Pregnancy: 1/ 5.3% Lost F/U: 1/ 5.3% OCP: 0 **Summary characteristics:** PCOS, hyperandrogenic (precised in an other study with same population) **Age (years):** MET mean (±SD): 25.1 (6.6) OCP mean (± SD): 23.4 (5.6) **BMI (kg/m^2):** MET mean (± SD): 30.5 (6.9) OCP mean (± SD): 29.2 (5.7)
Interventions	**Treatment:** MET 425 mg twice a day 1 week then 850 mg twice a day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days followed by 7 days placebo pills **Duration:** 6 months **Co-intervention(s):** All the patients were instructed to maintain a diet containing 25 to 30 kcal per kg of body weight per day and moderate physical activity throughout the trial, although these measures were not stressed thereafter.
Outcomes	**Primary outcomes:** None **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters: 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles
Notes	**Authors contacted about:** the results because it was presented as a graph. **Power calculation:** yes, a priori, on secondary outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “patients were randomized to receive an antiandrogenic low-dose oral contraceptive pill or 850 mg of metformin twice daily for 24 wk”, “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane 35 Diario and five patients to receive metformin” Insufficient information about the sequence generation process available to permit a judgement of ‘low risk’ or ‘high risk’
Allocation concealment (selection bias)	Low risk	Quote: “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment” and “sealed opaque envelopes”

Luque-Ramirez 2007b *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “open label RCT”.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	7 patients in MET group did not complete the entire study/ 0 in OCP group Quote: "protocol violation and lost to F/U" unclear if related to intervention. Insufficient reporting of attrition /exclusions to permit a judgement of 'low risk' or 'high risk.".
Selective reporting (reporting bias)	High risk	Menstrual cyclicity is the only outcome of interest in this study and was not described in the methods. One or more primary outcomes have been reported using measurements, analysis, methods or subset of the data THAT WERE NOT PRE-SPECIFIED
Other bias	Low risk	The study appears to be free of other sources of bias.

Luque-Ramirez 2008a

Study characteristics	
Methods	Randomised controlled trial Same population as Luque-Ramirez 2007a **Location of the trial:** quote: Madrid, Spain**Method of randomisation:** quote: “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane35 Diario and five patients to receive metformin”**Method of allocation concealment:** “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “sealed opaque envelopes” **Source of funding:** "This study was supported by the Spanish Ministry of Health and Consumer Affairs, Instituto de Investigacion Carlos III Grants Fondo de Investigacion Sanitaria and Red de Diabetes Enfermedades Metabolicas Asociadas; Ministry of Education and Science Grant; and by economic aid from Hospital Ramon y Cajal."
Participants	**Inclusion criteria:** the diagnosis of PCOS was based on the presence of clinical and/or biochemical hyperandrogenism, oligo-ovulation, and exclusion of secondary aetiologies. Hirsutism was defined by a modified Ferriman-Gallwey score above 7 and oligomenorrhoea (more than six cycles longer than 36 days in the previous year) or amenorrhoea (absence of menstruation for 3 consecutive months), or luteal phase progesterone measurements less than 4 ng/ml (12.72 nmol/L) in women with regular menstrual cycles were considered indicative of oligo-ovulation. **Exclusion criteria:** secondary aetiologies, including hyperprolactinaemia, thyroid dysfunction, Cushing’s syndrome, congenital adrenal hyperplasia, and virilizing tumours, were actively ruled out in all the patients. None of the patients had a personal history of hypertension, diabetes mellitus, or cardiovascular events, or received treatment with oral contraceptives, antiandrogens, insulin sensitisers, or drugs that might interfere with blood pressure regulation, lipid profile, or carbohydrate metabolism for the previous 6 months. **Number of women randomised:** 34 MET: 19 OCP: 15 **Number of women analysed:** 27 MET: 12 OCP: 15

Luque-Ramirez 2008a *(Continued)*

	Number of withdrawal and reasons: MET: protocol violation: 3/ 15.8% GI adverse events: 2/ 10.5% Pregnancy: 1/ 5.3% Lost F/U: 1/ 5.3% OCP: 0 **Summary characteristics:** PCOS, hyperandrogenic (precised in an other study with same population) **Age (years):** MET mean (± SD): 25.1 (6.6) OCP mean (± SD): 23.4 (5.6) **BMI (kg/m^2):** MET mean (± SD): 30.5 (6.9) OCP mean (± SD): 29.2 (5.7)
Interventions	**Treatment:** MET 425 mg twice a day 1 week then 850 mg twice a day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days followed by 7 days placebo pills **Duration:** 6 months **Co-intervention(s):** all the patients were instructed to maintain a diet containing 25 to 30 kcal per kg of body weight per day and moderate physical activity throughout the trial, although these measures were not stressed thereafter.
Outcomes	**Primary outcomes:** none **Secondary outcomes:** Free androgen index (FAI) (%)
Subjective outcomes	None
Objective outcomes	(b) Hormonal parameters 1. Free androgen index (FAI) (%)
Notes	**Authors contacted about:** the results because it was presented as a graph. **Power calculation:** yes, a priori, on secondary outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “patients were randomized to receive an antiandrogenic low-dose oral contraceptive pill or 850 mg of metformin twice daily for 24 wk”, “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane 35 Diario and five patients to receive metformin” Insufficient information about the sequence generation process available to permit a judgement of ‘low risk’ or ‘high risk’

Luque-Ramirez 2008a *(Continued)*

Allocation concealment (selection bias)	Low risk	Quote: “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment” and “sealed opaque envelopes”
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “open label RCT”.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	7 patients in MET group did not complete the entire study/ 0 in OCP group Quote: "protocol violation and lost to F/U" unclear if related to intervention. Insufficient reporting of attrition /exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	High risk	FAI is the only outcome of interest in this study and was not described in the methods. One or more primary outcomes have been reported using measurements, analysis, methods or subset of the data THAT WERE NOT PRESPECIFIED
Other bias	Low risk	The study appears to be free of other sources of bias.

Luque-Ramirez 2009b

Study characteristics	
Methods	Randomised controlled trial Same population as Luque-Ramirez 2007a **Location of the trial:** quote: Madrid, Spain**Method of randomisation:** quote “Simple randomization was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane35 Diario and five patients to receive metformin”**Method of allocation concealment:** “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “sealed opaque envelopes” **Source of funding:** "This study was supported by the Spanish Ministry of Health and Consumer Affairs, Instituto de Investigacion Carlos III Grants Fondo de Investigacion Sanitaria and Red de Diabetes Enfermedades Metabolicas Asociadas; Ministry of Education and Science Grant; and by economic aid from Hospital Ramon y Cajal."
Participants	**Inclusion criteria:** the diagnosis of PCOS was based on the presence of clinical and/or biochemical hyperandrogenism, oligo-ovulation, and exclusion of secondary aetiologies. Hirsutism was defined by a modified Ferriman-Gallwey score above 7 and oligomenorrhoea (more than six cycles longer than 36 d in the previous year) or amenorrhoea (absence of menstruation for 3 consecutive months), or luteal phase progesterone measurements less than 4 ng/ml (12.72 nmol/L) in women with regular menstrual cycles were considered indicative of oligo-ovulation. **Exclusion criteria:** secondary aetiologies, including hyperprolactinaemia, thyroid dysfunction, Cushing’s syndrome, congenital adrenal hyperplasia, and virilizing tumours, were actively ruled out in all the patients. None of the patients had a personal history of hypertension, diabetes mellitus, or cardiovascular events, or received treatment with oral contraceptives, antiandrogens, insulin sensitisers, or drugs that might interfere with blood pressure regulation, lipid profile, or carbohydrate metabolism for the previous 6 months. **Number of women randomised:** 34 MET: 19

Luque-Ramirez 2009b *(Continued)*

OCP: 15
Number of women analysed: 27
MET: 12
OCP: 15
Number of withdrawal and reasons:
MET: protocol violation: 3/ 15.8%
GI adverse events: 2/ 10.5%
Pregnancy: 1/ 5.3%
Lost F/U: 1/ 5.3%
OCP: 0
Summary characteristics: PCOS, hyperandrogenic (precised in an other study with same population)
Age (years):
MET mean (± SD):

25.1 (6.6)

OCP mean (± SD):

23.4 (5.6)
BMI (kg/m^2):

MET mean (± SD):

30.5 (6.9)

OCP mean (± SD):

29.2 (5.7)

Interventions	**Treatment:** MET 425 mg twice a day 1 week then 850 mg twice a day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days followed by 7 days placebo pills **Duration:** 6 months **Co-intervention(s):** All the patients were instructed to maintain a diet containing 25 to 30 kcal per kg of body weight per day and moderate physical activity throughout the trial, although these measures were not stressed thereafter.
Outcomes	**Primary outcomes:** none **Secondary outcomes:** Blood pressure (systolic, diastolic) (mm Hg)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters Blood pressure (systolic, diastolic) (mm Hg)
Notes	**Authors contacted about:** the results because it was presented as a graph. **Power calculation:** yes, a priori, on secondary outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “patients were randomized to receive an antiandrogenic low-dose oral contraceptive pill or 850 mg of metformin twice daily for 24 wk”, “One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment”, “Simple randomization

Luque-Ramirez 2009b *(Continued)*

		was conducted using blocks of 10 sealed opaque envelopes assigning five patients to receive Diane 35 Diario and five patients to receive metformin" Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Low risk	Quote: "One investigator generated the randomization envelopes, whereas another enrolled and assigned the participants to their arm of treatment" and "sealed opaque envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "No masking method was used after randomization".
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "No masking method was used after randomization".
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	7 patients in MET group did not complete the entire study/ 0 in OCP group Quote: "protocol violation and lost to F/U" unclear if related to intervention. Insufficient reporting of attrition /exclusions to permit a judgement of 'low risk' or 'high risk.".
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Lv 2005

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: Huazhong, China **Method of randomisation:** unclear, quote: "The subjects were randomized to either the CPA group (n=25) or to the CPA + metformin group (n=25)." **Method of allocation concealment:** not stated **Source of funding:** Not stated
Participants	**Inclusion criteria:** PCOS was defined as the presence of: (1) chronic anovulatory disorders such as oligomenorrhoea, anovulatory cycles, or secondary amenorrhoea; (2) the ratio of LH/FSH was > 2 and (or) the plasma testosterone (T) level was > 2.6 nmol/L; (3) 10 or more follicles (2 mm to 8 mm in diameter) in one or both ovaries by transvaginal ultrasound examination. All the women were euthyroid and had normal prolactin levels. **Exclusion criteria:** women who had any other known endocrinological disease, and those taking drugs known to affect carbohydrate or lipid metabolism and OGTT results during the 6 months preceding the study were excluded. **Number of women randomised:** 50 MET + OCP: 25 OCP: 25 **Number of women analysed:** not stated **Number of withdrawal and reasons:** not stated **Summary characteristics:** PCOS, aged 16 to 36 years, all women were either of normal weight or thin BMI (≤ 25 kg/m^2). **Age (years):**

Lv 2005 *(Continued)*

	OCP + MET mean (+ SD): 24.5 (5.6) OCP mean (+ SD): 24.35 (5.11) **BMI (kg/m^2):** OCP + MET mean (+ SD): 22.1 (2.46) OCP mean (+ SD): 21.81 (1.37)
Interventions	**Treatment:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2mg) 21 days stop 7 days + MET 500 mg/day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days stop 7 days **Duration:** 6 months **Co-intervention(s):** Medroxyprogesterone for amenorrhoeic patients
Outcomes	**Primary outcomes:** none**Secondary outcomes:** BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** the number of patient randomised in each group and number of withdrawal.**Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “The subjects were randomized to either the CPA group (n=25) or to the CPA + metformin group (n=25).”

Lv 2005 *(Continued)*

		Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Paper does not provide any information on participant after randomisation.
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Meyer 2007

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Melbourne, Australia"**Method of randomisation:** quote: "computer-generated random numbers"**Method of allocation concealment:** not stated**Source of funding:** pharmaceutical, quote: "This work was an investigator-initiated trial funded by a competitive CVL grant sponsored by Pfizer Australia and through internal department funds. Pfizer Australia provided the aldactone, and Douglas Pharmaceuticals Australia provided the metformin. H.J.T is an National Health and MedicalResearch Council and National Heart Foundation Fellow. C.M. is an National Health and Medical Research Council PhD Scholar."
Participants	**Inclusion criteria:** PCOS diagnosis was based on perimenarchal onset of irregular cycles (< 21 days or > 35 days) and clinical manifestations of hyperandrogenism (hirsutism or acne) or biochemical hyperandrogenism with elevation of at least one circulating ovarian androgen (according to 1990 National Institutes of Health criteria). **Exclusion criteria:** secondary causes of amenorrhoea and hyperandrogenism were excluded with clinical screening and early follicular 17-hydroxyprogesterone levels. Diabetes was excluded on OGTTs. Pregnancy tests were negative before enrolment. **Number of women randomised:** 72: MET : 37 OCP: 35 **Number of women analysed:** 67 MET: 36 OCP: 31 **Number of withdrawal and reasons:** MET: Personal reason: 1/ 2.7%

Meyer 2007 *(Continued)*

	OCP: Personal reason: 3/ 9.7% Swinging mood: 1/ 3.2% **Summary characteristics:** PCOS, Overweight BMI > 27 kg/m^2, Mean age 31 years. **BMI (kg/m^2):** MET mean: BMI (kg/m^2): 36.3 OCP mean: BMI (kg/m^2): 36.5
Interventions	**Treatment:** MET (1000 mg twice a day 500 mg twice a day 4 weeks)**Control:** OCP (Ethinyl estradiol 35mcg Cyproterone acetate 2 mg)**Duration:** 6 months**Co-intervention(s):** exercise and diet
Outcomes	**Primary outcomes:** Hirsutisme score Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (mIU/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles 2. BMI (kg/m^2) 3. Blood pressure (systolic) (mm Hg) 4. Adverse events: severe (requiring stopping of medication) and minor (b) Hormonal parameters 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting total cholesterol (mmol/L) 3. Fasting HDL cholesterol (mmol/L) 4. Fasting LDL cholesterol (mmol/L) 5. Fasting triglycerides (mmol/L)

Meyer 2007 *(Continued)*

Notes	**Power calculation:** yes, a priori, on insulin sensitivity (secondary outcome)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “At randomization, 110 subjects were allocated to one of three groups based on computer-generated random numbers”.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “This was an open-label study”.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "End point data collection was completed by the research nurse, who was blinded to treatment allocation”.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Personnal reason" unclear reason of withdrawal. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified.
Other bias	High risk	Population not comparable regarding androgenism status. Quote: "Groups were well matched at baseline except for Ferriman-Gallwey score (8.8 + 0.8 vs. 6.7 + 0.7, P< 0.05) and testosterone levels (2.5 +0.1vs. 2.1 +0.1) in the metformin and high -dose OCP" Baseline imbalance.

Mhao 2015

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote:" Najaf, Iraq" **Method of randomisation:** not stated **Method of allocation concealment:** not stated **Source of funding:** none
Participants	**Inclusion criteria:** the diagnostic criteria of PCOS include menstrual cycle disturbance; infertility; clinical and/or biochemical signs of hyperandrogenism; and the presence of polycystic ovaries in ultrasound. The presence of two of those key features must be present to allow the diagnosis of PCOS **Exclusion criteria:** not stated **Number of women randomised:** 26 MET: 16 OCP: 10 **Number of women analysed:** 26 MET: 16

Mhao 2015 *(Continued)*

	OCP: 10 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, patients quote: "seeking treatment for their infertility and/or cycle abnormalities", "aging 14-40" **BMI (kg/m^2):** MET mean (±SD): 27.23 (5.44) OCP mean (± SD): 30.5 (5.3)
Interventions	**Treatment:** MET 500 mg twice a day **Control:** OCP (Ethniyl estradiol 35 mcg Cyproterone acetate 2 mg) 21 days stop 7days **Duration:** 3 months **Co-intervention(s):** dydrogesterone 10 mg/day for 10 days in amenorrhoeic women
Outcomes	**Primary outcomes:** Hirsutism score**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Acne – subjective BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Acne – subjective 2. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Menstrual cyclicity, initiation of menses or significant shortening of cycles 2. BMI (kg/m^2) (b) Hormonal parameters Serum total testosterone (nmol/L) (c) Metabolic parameters Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** total testosterone because expressed in the wrong unit and Ferriman-Gallwey score because SD not available **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Random sequence generation not stated.

Mhao 2015 *(Continued)*

		Insufficient information about the sequence generation process available to permit a judgement of “low risk” or “high risk”
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of “low risk” or “high risk”
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	High risk	Ferriman-Gallwey score cannot be used because SD is not precised. One or more primary outcomes of interest in the review have been reported incompletely so that they cannot be entered in meta-analysis.
Other bias	High risk	Important difference between the group regarding the number of patients (randomisation). Baseline imbalance.

Moran 2010

Study characteristics	
Methods	Randomised controlled trial Same population as Meyer 2007 **Location of the trial:** quote: Melbourne, Australia**Method of randomisation:** quote: "Allocated to one of two open-label groups based on computer-generated random numbers"**Method of allocation concealment:** not described**Source of funding:** pharmaceutical
Participants	**Inclusion criteria:** quote "PCOS was diagnosed based on National Institute of Health criteria. All subjects with PCOS also met the European Society for Human Reproduction and Embryology/American Society for Reproductive Medicine criteria." **Exclusion criteria:** quote: "Secondary causes of amenorrhea and hyperandrogenism (based on clinical screening, hyperprolactinemia, thyroid disease, and early follicular 17-hydroxyprogesterone levels), DM2 (based on World Health Organisation criteria), smoking, pregnancy, and use of anti-obesity, anti-hypertensive and anti-inflammatory drugs. Participants were required to cease oral contraceptives, endocrine hormonal treatment or insulin-sensitising agents and all participants received standard diet and lifestyle advice." **Number of women randomised:** 66 MET: 36 OCP: 30 **Number of women analysed:** 56

Moran 2010 *(Continued)*

	MET: 30 OCP: 26 **Number of withdrawal and reasons:** MET: 6/ 17% OCP: 4 / 13% Quote: “Data has been previously reported on study withdrawals” in Meyer et al 2007, however the number of patients in each group is different, therefore the number of withdrawal is different and the reasons are not clearly stated. **Summary characteristics:** PCOS Overweight women BMI > 25 kg/m^2/ BMI (kg/m^2) 36.1 (7.2) Age (years) 33.5 (6.7) Weight (kg) 97.1 (21.1)
Interventions	**Treatment:** MET 1 g twice a day with doses titrated up over 4 weeks starting at 500 mg twice a day **Control:** OCP (Ethinyl estradiol 35 mcg/ Cyproterone acetate 2 mg) **Duration:** 6 months **Co-intervention(s):** standard diet and lifestyle advice
Outcomes	**Primary outcomes:** none **Secondary outcomes:** Body weight (kg) Fasting glucose (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Body weight (kg) (c) Metabolic parameters 1. Fasting glucose (mmol/L)
Notes	**Authors contacted:** the study population, they confirm it was a subset of a previous article of Meyer et al. **Power calculation:** yes, a priori, on letpin outcome

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "patient were allocated to one of two open-label groups based on computer-generated random numbers"
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not sated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “This was an open-label study”.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Blinding of outcome assessor not described in this article but described in the previous study which authors have confirmed to be the same study using the same methods:

Moran 2010 *(Continued)*

		Quote: “end point data collection was completed by the research nurse, who was blinded to treatment allocation” blinding for the outcome assessor
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: “Data has been previously reported on study withdrawals” in Meyer et al 2007, however the number of patients in each group is different, therefore the number of withdrawal is different and the reasons are not clearly stated. Moreover, some patients are missing for a few outcomes and it is not precised in which group the patient are missing Quote: “the 56 subjects who completed the intervention except for lipids, insulin, glucose and HOMA-IR (no.=55); testosterone and SHBG (no.=53); FAI, leptin, adiponectin, and L/A (no.=52); OGTT insulin (no.=51); OGTT glucose (no.=50); and hsCRP (no.=36)." Insufficient reporting of attrition/exclusions to permit a judgement of ‘low risk’ or ‘high risk’ (no reason for missing data provided).
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	High risk	It is not described in this article but described in the previous study which authors have confirmed to be the same study using the same methods: Population not comparable regarding androgenism status quote: “Groups were well matched at baseline except for Ferriman-Gallwey score (8.8 + 0.8 versus 6.7 + 0.7, P< 0.05) and testosterone levels (2.5 +0.1 vs. 2.1 +0.1) in the metformin and high -dose OCP". Baseline imbalance.

Morin-Papunen 2000

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote "Oulu, Finland" **Method of randomisation:** computer generated **Method of allocation concealment:** not stated **Source of funding:** this research was supported by grants provided by the university of Oulu, the Finish Gynecological Association, the Sigrid Juselius Foundation and the Academy of Finland
Participants	**Inclusion criteria:** Obese (BMI > 27 kg/m^2) women. PCOS (Homburg 1996) i.e. PCO shown by vaginal ultrasonography (=8 subcapsular follicles of 3 mm to 8 mm diameter in 1 plane in 1 ovary and increased stroma) and at least one of the following symptoms: (a) oligomenorrhoea or amenorrhoea; (b) clinical manifestations of hyperandrogenism, such as a Ferriman-Gallway hirsutism score of more than 7; acne; and/or (c) an elevated serum T level (> 2.7 n mol/L) **Exclusion criteria:** women with diabetes, smokers, alcohol users, and those taking sex hormones or drugs known to affect lipid metabolism during the two months preceding the study **Number of women randomised:** 32: MET arm: 16 OCP arm: 16 **Number of women analysed:** 18: MET: 8 OCP: 10 **Number of withdrawal and reasons:**

Morin-Papunen 2000 *(Continued)*

	3 did not receive allocated treatment (2 in MET group, 1 in OCP group) because of type 2 diabetes discover before starting the intervention. MET: 6: nausea and diarrhoea: 1/ 6.3% personal reasons: 1/ 6.3 pregnancy: 2/13% Loss F/U: 2/ 13% OCP: 5: headache and high blood pressure: 1/ 6.3% try for pregnancy: 3/ 19% Loss F/U 1/ 6.3% **Summary characteristics:** PCOS, obese women **Age (years):** MET mean (±SE): 29.9 (1.5) OCP mean (±SE): 29.8(1.0) **BMI (kg/m^2):** MET Mean (±SE): 32.5(1.1) OCP Mean (±SE): 37.2(1.8)
Interventions	**Treatment(s):** MET 500 mg twice a day for 3 months, then 1000 mg twice a day for next 3 months **Control:** OCP (Ethinyl estradiol 35mcg Cyproterone acetate 2 mg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 6 months **Co-interventions:** progestin to induce menses if necessary
Outcomes	**Primary outcomes:** Hirsutism score (Ferriman-Gallwey) Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Serum total testosterone (n mol/L) Free androgen index (FAI) (%) Fasting insulin (pmol/L) Fasting glucose (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score (F-G score)
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor

Morin-Papunen 2000 *(Continued)*

	2. Diagnosis of Type II diabetes mellitus 3. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (n mol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1.Fasting insulin (pmol/L) 2.Fasting glucose (mmol/L)
Notes	**Authors contacted** about: method of randomisation, reason for withdrawal and co-intervention kindly provided by the authors that was not in the original paper **Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The subjects were randomized to either the metformin group or to the OCP group". Insufficient information to permit judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'Low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Total number of drop off quite important (50% for MET group/ 37.5% for OCP group) unclear if this could have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of weight which was only measured as BMI. Although period (days) reported for both groups, 'menstrual cyclicity' only reported for metformin. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Morin-Papunen 2003

Study characteristics

Morin-Papunen 2003 *(Continued)*

Methods	Randomised controlled trial **Location of the trial:** quote: "Oulu, Finland" **Method of randomisation:** computer generated **Method of allocation concealment:** not stated **Source of fundings:** quote: "This research was supported by grants provided by the university of Oulu, the Finish Gynecological Association, the Sigrid Juselius Foundation and the Academy of Finland"
Participants	**Inclusion criteria:** non-obese (BMI < 25 kg/m^2) women. PCOS (Homburg 1996) i.e. PCO shown by vaginal ultrasonography (= 8 subcapsular follicles of 3 mm to 8 mm diameter in 1 plane in 1 ovary and increased stroma) and at least one of the following symptoms: (a) oligomenorrhoea or amenorrhoea; (b) clinical manifestations of hyperandrogenism, such as a F-G hirsutism score of more than 7; acne; and/or (c) an elevated serum T level (> 2.7 n mol/L) **Exclusion criteria:** women with diabetes, smokers, alcohol users, and those taking sex hormones or drugs known to affect lipid metabolism during the two months preceding the study **Number of women randomised:** 20: MET arm: 10 OCP arm: 10 **Number of women analysed:** 17: MET arm: 8 OCP arm: 9 **Number of withdrawal and reasons:** 3 discontinued medication (2 in metformin arm, 1 in OCP arm). MET: 2: nausea and diarrhoea: 1/ 10% personal reasons 1/ 10% OCP: 1 headache and high blood pressure 1/ 10% **Summary characteristics:** PCOS, non-obese women. **Age (years):** MET mean (± SE): 28.2(1.4) OCP mean (± SE): 28.5(1.7) **BMI (kg/m^2):** MET mean (± SE): 22.5(0.8) OCP mean (± SE): 21.8(0.7)
Interventions	**Treatment(s):** MET 500 mg twice a day for 3 months, then 1000 mg twice a day for next 3 months **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 6 months **Co-interventions:** progestin to induce menses if necessary
Outcomes	**Primary outcomes:**

Morin-Papunen 2003 *(Continued)*

	Hirsutism score (F-G) Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting insulin (pmol/L) Fasting glucose (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score (F-G)
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2.Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (n mol/L) 2. Free androgen index (FAI) (%) (c) Metabolic parameters 1.Fasting insulin (pmol/L) 2.Fasting glucose (mmol/L)
Notes	**Authors contacted** about: method of randomisation, reason for withdrawal and co-intervention kindly provided by the authors that was not in the original paper **Power calculation:** NA

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The subjects were randomized to either the metformin group or to the EE=CA pill group". Insufficient information to permit judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"discontinue treatment, personal reason" unclear reason for withdrawal. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.

Morin-Papunen 2003 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the with the exception of weight which was only measured as BMI, BP and lipids. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Moro 2013

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Roma, Italy" **Method of randomisation:** quote: **"**computer-generated randomization list" **Method of allocation concealment:** quote: "staff member independent of the study controlled the randomization" **Source of funding:** None
Participants	**Inclusion criteria:** the PCOS was diagnosed on the basis of the presence of 2 of the 3 following criteria: oligomenorrhoea, biochemical/clinical hyperandrogenism, and polycystic ovary (PCO) on ultrasound, according to the Rotterdam criteria. **Exclusion criteria:** the exclusion criteria were inflammatory/autoimmune disease, cancer, treatment with clomiphene citrate, OC, antiandrogens, anorexic, or insulin-sensitizing drugs during the last 6 months prior to our evaluation, DM2, hypertension, major surgery in the last 3 months, or other hormonal dysfunctions (hypothalamic–pituitary, thyroidal, or adrenal causes). Patients with normoinsulinaemic PCOS, with insulinaemic area under the curve (AUCi) < 7000, were also excluded. **Number of women randomised:** 93 MET + OCP: 31 MET: 31 OCP: 31 **Number of women analysed:** 76 MET + OCP: 25 MET: 25 OCP: 26 **Number of withdrawal and reasons:** MET + OCP: Incomplete data: 2 /6.5% GI adverse events: 2/ 6.5% Voluntary drop out: 1/ 3.2% MET: Incomplete data: 3/9.7% GI adverse events: 2/ 6.5% Voluntary drop out: 1/ 3.2% OCP: Incomplete data: 5/16% Voluntary drop out: 1/ 3.2%

Moro 2013 *(Continued)*

	Summary characteristics: PCOS Non normoinsulinaemic **Age (years):** MET + OCP mean (±SD): 25 (4) MET mean (±SD): 25 (5) OCP mean (±SD): 26 (3) **BMI (kg/m^2):** MET + OCP median (range): 26.5 (21.3 to 30) MET median (range): 25.1 (21.9 to 28.3) OCP median (range): 23.7 (20.8 to 28.6)
Interventions	OCP (Ethinyl Estradiol 30 mcg drospirenone 3 mg) + MET 500 mg three times a day OCP (Ethinyl Estradiol 30 mcg drospirenone 3 mg) MET 500 mg three times/day **Duration:** 6 months **Co-intervention(s): i**nduced menstrual cycle with medroxyprogesterone acetate 10m g/day for 7 days
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Free androgen index (FAI) (%) 2. Fasting total cholesterol (mmol/L) 3. Fasting HDL cholesterol (mmol/L) 4. Fasting LDL cholesterol (mmol/L) 5. Fasting triglycerides (mmol/L)
Notes	**Power calculation:** unclear
Risk of bias	
Bias	**Authors' judgement** **Support for judgement**

Moro 2013 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote: "A total of 93 consecutive eligible patients, who accepted to participate to the study, were randomly assigned to 1 of the 3 treatment groups, after age and BMI matching by giving them a code number from a computer-generated randomization list, in order of enrolment."
Allocation concealment (selection bias)	Low risk	Quote: "Staff member independent of the study controlled the randomization."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Both the patients and the gynaecologists were informed of the assigned treatment."
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Incomplete data, voluntary drop out" unclear reason for withdrawal. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Ozgurtas 2008

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote:**"**Ankara, Turkey"**Method of randomisation:** quote: "the patients were randomized"**Method of allocation concealment:** not stated**Source of funding:** not stated
Participants	**Inclusion criteria:** non-obese (BMI < 25 kg/m^2), aged >18 years women with PCOS. PCOS consistent with ESHRE/ASRM: All eligible patients presented with at least two of the three following criteria: (1) oligomenorrhoea or amenorrhoea, (2) clinical (hirsutism, acne) and/or biochemical signs of hyperandrogenism, and (3) polycystic ovaries. Also states all had PCO on ultrasound. Non-smokers with regular daily activity, normotensive (<120/80 mmHg in two measurements) and not regular consumers of alcoholic beverages. For at least 3 months before the study, all participants refrained from using steroid hormones or any other medications likely to interfere with ovarian function, insulin sensitivity, or lipid metabolism. All participants in the study were off any regular medication, like aspirin, statins, etc. that could affect the outcome of the study. None of the participants contemplated pregnancy during the study period.**Exclusion criteria:** possible ovarian tumours, congenital adrenal hyperplasia, BMI greater than 25 kg/m^2, any chronic disease that could interfere with the absorption, distribution, metabolism or excretion of metformin or EE/CPA, renal or liver disease.**Number of women randomised:** 44 MET: 22 OCP: 22**Number of women analysed:** 41 MET: 20 OCP: 21**Number of withdrawal and reasons:** MET: lost to follow-up: 2/ 9% OCP: lost to follow-up: 1/ 4.5%**Summary characteristics:** PCOS, non-obese**BMI (kg/m^2):**

Ozgurtas 2008 *(Continued)*

	MET mean (± SD): 21.81 ± 1.27 OCP mean (± SD): 21.72 ± 1.24
Interventions	**Treatment:** MET 850 mg twice a day**Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) taken daily 21 out of 28 days for a period of 3 months.**Duration:** 3 months**Co-intervention(s):** none
Outcomes	**Primary outcomes:** none**Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters: 1. Body weight (kg) 2. (BMI (kg/m^2) (b) Hormonal parameters: 1. Serum total testosterone (nmol/L) (c) Metabolic parameters: 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Power calculation:** unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The patients were randomized." Insufficient information about the sequence generation process available to permit judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.

Ozgurtas 2008 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Lost to F/U" unclear reason for withdrawal. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'
Selective reporting (reporting bias)	Unclear risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of insulin and glucose which were only reported as HOMA, weight which was only reported as BMI. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears free of other sources of bias.

Rautio 2005

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Oulu, Finland" **Method of randomisation:** quote: "Computer generated" **Method of allocation concealment:** not stated **Source of fundings:** quote|: "This research was supported by grants provided by the university of Oulu, the Finish Gynecological Association, the Sigrid Juselius Foundation and the Academy of Finland"
Participants	**Inclusion criteria:** non-obese (BMI < 27 kg/m^2) and obese (BMI = 27 kg/m^2) women. PCOS (Homburg 1996) i.e. PCO shown by vaginal ultrasonography (= 8 subcapsular follicles of 3 mm to 8 mm diameter in 1 plane in 1 ovary and increased stroma) and at least one of the following symptoms: (a) oligomenorrhoea or amenorrhoea; (b) clinical manifestations of hyperandrogenism, such as a F-G hirsutism score of more than 7; acne; and/or (c) an elevated serum T level (> 2.7 nmol/L)**Exclusion criteria:** women with diabetes, smokers, alcohol users, and those taking sex hormones or drugs known to affect lipid metabolism during the two months preceding the study **Obese:** **Number of women randomised:** 32: MET arm: 16 OCP arm: 16 **Number of women analysed:** 18: MET: 8 OCP: 10 **Number of withdrawal and reasons:** 3 did not receive allocated treatment (2 in MET group, 1 in OCP group) because of type 2 diabetes discover before starting the intervention.

Rautio 2005 *(Continued)*

MET: 6: nausea and diarrhoea: 1/ 6.3%

personal reasons: 1/ 6.3

pregnancy: 2/13%

Loss F/U: 2/ 13%

OCP: 5: headache and high blood pressure: 1/ 6.3%

try for pregnancy: 3/ 19%

Loss F/U 1/ 6.3%

Age (years):
MET mean (± SE):

29.9 (1.5)

OCP mean (± SE):

29.8(1.0)
BMI (kg/m^2):

MET Mean (± SE):

32.5(1.1)

OCP Mean (± SE):

37.2(1.8)**Non obese:**

Number of women randomised: 20:

MET arm: 10

OCP arm: 10
Number of women analysed: 17:

MET arm: 8

OCP arm: 9
Number of withdrawal and reasons:

3 discontinued medication (2 in metformin arm, 1 in OCP arm).

MET: 2: nausea and diarrhoea 1/ 10%

personal reasons 1/10%

OCP: headache and high blood pressure 1/10%

Age (years):

MET mean (± SE):

28.2(1.4)

OCP mean (± SE):

28.5(1.7)
BMI (kg/m^2):

MET mean (± SE):

22.5(0.8)

Rautio 2005 *(Continued)*

	OCP mean (± SE): 21.8(0.7)**Summary characteristics:** PCOS, obese and non-obese women, the patients included in this study (obese and non-obese PCOS women) are the same patients from two previous studies (obese PCOS women in Morin-Papunen 2000 and non-obese PCOS women in Morin-Papunen 2003) on the effects of metformin and the EE-CA pill on insulin sensitivity, glucose tolerance, hormonal parameters and adverse events.
Interventions	**Treatment(s):** MET 500 mg twice a day for 3 months, then 1000 mg twice a day for next 3 months **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) once daily (21 days per month followed by 7 days pill-free period) **Duration:** 6 months **Co-interventions:** progestin to induce menses if necessary
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Total Cholesterol (mmol/L) HDL Cholesterol (mmol/L) LDL Cholesterol (mmol/L) Triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor (b) Metabolic parameters 1. Total Cholesterol (mmol/L) 2. HDL Cholesterol (mmol/L) 3. LDL Cholesterol (mmol/L) 4. Triglycerides (mmol/L)
Notes	**Authors contacted** about: method of randomisation, reason for withdrawal and co-intervention kindly provided by the authors that was not in the original paper **Power calculation:** this study was powered according to serum triglyceride level changes seen with the OCP asked to the authors not sated in the paper

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The subjects were randomized to either the metformin group or the OC group". Insufficient information to permit judgement of 'low risk' or 'high risk'.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.

Rautio 2005 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Total number of drop off quite important (50% for MET group/ 37.5% for OCP group) unclear if this could have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	From results section of paper, all of the studies pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the with the exception of weight, waist circumference and WHR which were reported in Morin-Papunen 2000 and Morin-Papunen 2003 (although in these studies waist circumference was only measured as WHR and weight which was only measured as BMI). Although key expected primary and secondary outcomes (e.g. hyperandrogenism, menstrual function, insulin resistance) were not reported in this paper, they were reported in the original papers (Morin-Papunen 2000 and Morin-Papunen 2003). But not primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Ruan 2018

Study characteristics	
Methods	Randomised controlled trial Same population as Song 2017 **Location of the trial:** quote: "Beijing, China **Method of randomisation:** quote: "Randomization was performed using a random number table." **Method of allocation concealment:** Not stated **Source of funding:** quote: "This study was supported by Capital Characteristic Clinic Project of China; Beijing Capital Foundation for Medical Science Development and Research; Clinical Technique Innovation Project of Beijing Municipal Administration of Hospitals; Beijing Municipality Health Technology High-level Talent; Foreign technical and administrative talent introduction project in 2017, State Administration of Foreign Experts A airs, the P. R. of China."
Participants	**Inclusion criteria:** being Chinese, diagnosis of PCOS, aged between 18 and 40 years, BMI ≥ 24 kg/ m2, fasting insulin (FINS) > 10 mIU/L and no history of taking medication or dietary modification currently or for the preceding 3 months. The diagnosis of PCOS was made according to the Rotterdam criteria with the presence of at least two of the following three features: oligo- or anovulation, clinical and/or biochemical hyperandrogenism and ultrasound finding of polycystic ovaries (presence of 12 or more follicles in each ovary measuring 2 mm to 9 mm in diameter, and/or increased ovarian volume > 10 mL) **Exclusion criteria:** exclusion of other aetiologies (congenital adrenal hyperplasia, Cushing's syndrome, androgen-secreting neoplasms, hyperprolactinaemia, and thyroid disease). Exclusion criteria were ischaemic heart disease, clinically evident vascular disease, type-2 diabetes with ketoacidosis, renal or hepatic impairment, severe infection, and malignant tumour. **Number of women randomised:** 240/120

Ruan 2018 *(Continued)*

	OCP + MET: 60 OCP: 60 **Number of women analysed:** 120 OCP + MET: 60 OCP: 60 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, aged between 18 and 40 years, BMI ≥ 24 kg/ m2, fasting insulin (FINS) > 10 mIU/L **Age (years):** OCP + MET mean (±SD): 28.63 (5.12) OCP mean (±SD): 27.68 (4.99) **BMI (kg/m^2):** OCP + MET mean (±SD): 27.00 (3.47) OCP mean (±SD): 28.64 (4.89)
Interventions	**Treatment:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) from the 5th day of menstruation for a period of 21 days + MET (500mg/d/1 week then 500mg twice a day/1 week then 500mg three times a day) **Control:** OCP (Ethinyl estradiol 35 mcg, Cyproterone acetate 2 mg) from the 5th day of menstruation for a period of 21 days **Duration:** 3 months **Co-intervention(s):** low fat diet and moderate daily physical activity
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) clinical parameters: 1. Body weight (kg) and/or BMI (kg/m^2) 2. Adverse events: severe (requiring stopping of medication) and minor (c) Metabolic parameters: 1. Fasting insulin (mIU/L)

Ruan 2018 *(Continued)*

	2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about** the number of patient randomised in each group and number of withdrawal **Power calculation:** unclear RCT with 4 arms: OCP versus OCP + MET versus OCP + orlistat versus OCP + MET + orlistat

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Randomization was performed using a random number table.”
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described in this article but described as a “a randomized, open-label, and parallel study” in the previous study. Which authors have confirmed to be the same population using same methods.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data outcome.
Selective reporting (reporting bias)	Unclear risk	Body weight and BMI pre-specified in the methods not stated in the results. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Sahu 2018

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** quote: "Cuttack, India"**Method of randomisation:** quote: “random number table”**Method of allocation concealment:** not stated**Source of funding:** none
Participants	**Inclusion criteria:** Women diagnosed with PCOS according to the Rotterdam criteria **Exclusion criteria:**

Sahu 2018 *(Continued)*

	Current or previous use of oral contraceptives, glucocorticoids, antiandrogens, ovulation induction agents, antidiabetic and anti-obesity drugs or other hormonal drugs. Medical or surgical treatment of PCOS during the previous 3 months Presence of other endocrinopathies; except treated hypothyroidism on stable replacement doses of thyroid hormone Pregnancy, breastfeeding or desire for pregnancy during study interval (6 months) Inability to understand the proposal of the study precluding effective informed consent **Number of women randomised:** 101 MET: 50 OCP: 51 **Number of women analysed:** 86 MET: 42 OCP: 44 **Number of withdrawal and reasons:** MET: 8/ 16% Discontinued drug: 2 / 4% Lost F/U: 5/ 10% Pregnancy: 1/ 2% OCP: 7/ 13.7% Discontinued drug: 2/ 3.9% Lost F/U: 5/ 9.8% **Summary characteristics:** PCOS, with menstrual irregularities, 18 and 35 years. **Age (years):** MET mean (±SD): 27 (5.2) OCP mean (±SD): 26.8 (4.2) **BMI (kg/m^2):** MET mean (±SD): 25.7 (2.6) OCP mean (±SD): 25.6 (2.7)
Interventions	**Treatment:** MET 500 mg twice a day/day **Control:** OCP (Ethinyl Estradiol 35 mcg / Cyproterone acetate 2 mg) 21 days/ stop 7 days **Duration:** 6 months **Co-intervention(s):** medroxyprogesterone acetate (5 mg/day) for 5 days was administered in order to induce vaginal bleeding. All the patients were instructed to perform moderate physical activity and maintain a low carbohydrate diet throughout the trial.
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	(a) Clinical parameters

Sahu 2018 *(Continued)*

	1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Authors contacted about:** to have the publication references. **Power calculation:** unclear, a priori, on ovarian stromal PI. Fasting insulin and total testosterone expressed in wrong unit in the article but with appropriate unit in the protocol.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “random number table”.
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "open-label"
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Lost of follow up, discontinued drug" unclear reason of withdrawal. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Low risk	The study protocol is available and it is clear that the published reports include all expected outcomes, including those that were prespecified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Song 2017

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Beijing, China" **Method of randomisation:** quote: "Randomization was performed using a random number table." **Method of allocation concealment:** not stated **Source of funding:** quote: "This study was supported by Capital Characteristic Clinic Project of China; Beijing Capital Foundation for Medical Science Development and Research; Clinical Technique Innovation Project of Beijing Municipal Administration of Hospitals; Beijing Municipality Health Technology High-level Talent; Foreign technical and administrative talent introduction project in 2017, State Administration of Foreign Experts A airs, the P. R. of China."
Participants	**Inclusion criteria:** being Chinese, diagnosis of PCOS, aged between 18 and 40 years, BMI ≥ 24 kg/ m2, fasting insulin (FINS) > 10 mIU/L and no history of taking medication or dietary modification currently or for the preceding 3 months. The diagnosis of PCOS was made according to the Rotterdam criteria with the presence of at least two of the following three features: oligo- or anovulation, clinical and/or biochemical hyperandrogenism and ultrasound finding of polycystic ovaries (presence of 12 or more follicles in each ovary measuring 2 mm to 9 mm in diameter, and/or increased ovarian volume > 10 mL) **Exclusion criteria:** exclusion of other aetiologies (congenital adrenal hyperplasia, Cushing's syndrome, androgen-secreting neoplasms, hyperprolactinaemia, and thyroid disease). Exclusion criteria were ischaemic heart disease, clinically evident vascular disease, type-2 diabetes with ketoacidosis, renal or hepatic impairment, severe infection, and malignant tumour. **Number of women randomised:** 240/120 OCP + MET: 60 OCP: 60 **Number of women analysed:** 120 OCP + MET: 60 OCP: 60 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS, aged between 18 and 40 years, BMI ≥ 24 kg/ m2, fasting insulin (FINS) > 10 mIU/L **Age (years):** OCP + MET mean (± SD): 28.63 (5.12) OCP mean (± SD): 27.68 (4.99) **BMI (kg/m^2):** OCP + MET mean (± SD): 27.00 (3.47) OCP mean (± SD): 28.64 (4.89)
Interventions	**Treatment:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2 mg) from the 5th day of menstruation for a period of 21 days + MET (500 mg/day/1 week then 500 mg twice a day/1 week then 500 mg three times a day) **Control:** OCP (Ethinyl Estradiol 35 mcg Cyproterone acetate 2 mg) from the 5th day of menstruation for a period of 21 days

Song 2017 *(Continued)*

	Duration: 3 months **Co-intervention(s):** low-fat diet and moderate daily physical activity
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Serum total testosterone (nmol/L) Free androgen index (FAI) (%)
Subjective outcomes	None
Objective outcomes	(a) clinical parameters: 1. Adverse events: severe (requiring stopping of medication) and minor 2. Body weight (kg) and/or BMI (kg/m^2) (b) hormonal parameters: 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%)
Notes	**Authors contacted about** the number of patient randomised in each group and number of withdrawal **Power calculation:** unclear RCT with 4 arms: OCP versus OCP + MET versus OCP + orlistat versus OCP + MET + orlistat

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Randomization was performed using a random number table.”
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated Insufficient information available to permit a judgement of “low risk” or “high risk”.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “The trial was a randomized, open-label, and parallel study”
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Unclear risk	Body weight, BMI, BP pre-specified in the methods not stated in the results But not a primary or key outcomes.

Song 2017 *(Continued)*

		Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Teng 2007

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Harbin" **Method of randomisation:** quote: "Randomly assigned to two groups" **Method of allocation concealment:** not stated **Source of funding:** not stated
Participants	**Inclusion criteria:** Rotterdam criteria **Exclusion criteria: ?** **Number of women randomised:** 32 MET: 16 OCP: 16 **Number of women analysed:** 32 MET: 16 OCP: 16 **Number of withdrawal and reasons:** 0 **Summary characteristics:** PCOS
Interventions	**Treatment:** MET from Day 5, 500 mg/time, three times a day, ½ hour after meal **Control:** OCP (Ethinyl Estradiol 35 mcg/ Cyproterone acetate 2 mg) from Day 5, one pill per day, continues for 21 days **Duration:** 3 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** none **Secondary outcomes:** BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) Clinical parameters BMI (kg/m^2) (b) Hormonal parameters Serum total testosterone (nmol/L) (c) Metabolic parameters Fasting insulin (mIU/L)

Teng 2007 *(Continued)*

	Fasting glucose (mmol/L)
Notes	**Power calculation:** no

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomly assigned to two groups" Insufficient information about the sequence generation process available to permit a judgement of 'low risk' or 'high risk'
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit a judgement of 'low risk' or 'high risk'
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data.
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified.
Other bias	Low risk	The study appears to be free of other sources of bias.

Wei 2012

Study characteristics

Methods	Randomised controlled trial **Location of the trial:** Harbin, China **Method of randomisation:** quote: "Randomization was based on a computer-generated code in blocks of six." **Method of allocation concealment:** quote: "A copy of the code was stored in a sealed envelope by personnel not involved in the trial." **Source of fundings:** none
Participants	**Inclusion criteria:** 2003 Rotterdam ESHRE/ASRM criteria. Oligomenorrhea (interval between menstrual periods ≥ 35 days) or amenorrhoea (absence of vaginal bleeding for at least 6 months) and clinical (a F-G score ≥ 6) and/or biochemical hyperandrogenism (total testosterone (TT) ≥ 58 ng/dL (2 nmol/L)) were used to assess PCOS. The phenotype of polycystic ovaries was detected by vaginal ultrasound examination presenting 12 follicles or more in one or both ovaries and/or increased ovarian volume (> 10mL) .

Wei 2012 *(Continued)*

	All participants fulfilled at least two of the three diagnostic criteria. IR was assessed by homeostasis model assessment for IR (HOMA-IR) ≥ 3.8 or fasting glucose insulin ratio (FGIR) ≤ 4.5 **Exclusion criteria:** congenital adrenal hyperplasia, Cushing's syndrome, thyroid dysfunction, hyperprolactinaemia, diabetes mellitus, coronary artery disease, and spontaneous abortion. Furthermore, participants accepting treatment with medications known to alter insulin haemodynamics, ovulation induction, anti-obesity, or oral contraceptives (OCs) within 3 months were excluded from the study **Number of women randomised:** 67: MET + OCP arm: 36 OCP arm: 31 **Number of women analysed:** 58: OCP + MET: 30 OCP: 28 **Number of withdrawal and reasons:** 9 MET + OCP: 6: 2 loss of F/U: 5.6% 1 personal reason: 2.8% 2 travel difficulties: 5.6% 1 not described in the text (31 in the text/30 in the result table): 2.8% OCP 3: 2 loss of F/U: 6.5% 1 travel difficulties: 3.2% **Summary characteristics:** PCOS with Insulin resistance **Age (years):** MET + OCP mean (±SD): 26.03 (2.82) OCP mean (±SD): 26.75 (2.62) **BMI (kg/m^2):** MET + OCP mean (±SD): 24.74 (1.85) OCP mean (±SD): 24.91 (1.66)
Interventions	**Treatment:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) cyclically + MET 500 mg twice a day then three times a day **Control:** OCP (Ethinyl estradiol 35 mcg Cyproterone acetate 2 mg) cyclically + PBO * 2/day **Duration:** 3 months **Co-intervention(s):** diet and exercise
Outcomes	**Primary outcomes:** Adverse events: severe (requiring stopping of medication) and minor **Secondary outcomes:** Body weight (kg) BMI (kg/m^2) Blood pressure (systolic) (mm Hg) Blood pressure (diastolic) (mm Hg) Serum total testosterone (nmol/L)

Wei 2012 *(Continued)*

	Free androgen index (FAI) (%) Fasting insulin (mIU/L) Fasting glucose (mmol/L) Fasting total cholesterol (mmol/L) Fasting HDL cholesterol (mmol/L) Fasting LDL cholesterol (mmol/L) Fasting triglycerides (mmol/L)
Subjective outcomes	None
Objective outcomes	(a) clinical parameters: 1. Body weight (kg) and/or BMI (kg/m^2) 2. Blood pressure (systolic, diastolic) (mm Hg) 3. Adverse events: severe (requiring stopping of medication) and minor (b) hormonal parameters: 1. Serum total testosterone (nmol/L) 2. Free androgen index (FAI) (%) (c) metabolic parameters: 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L) 3. Fasting total cholesterol (mmol/L) 4. Fasting HDL cholesterol (mmol/L) 5. Fasting LDL cholesterol (mmol/L) 6. Fasting triglycerides (mmol/L)
Notes	**Power calculation**: unclear

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Randomization was based on a computer-generated code in blocks of six”
Allocation concealment (selection bias)	Low risk	Quote: “A copy of the code was stored in a sealed envelope by personnel not involved in the trial”
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: “Placebo (provided by pharmaceutical preparation section) was administered as one tablet twice a day.” Whereas MET is given 500 two times/week and then three times/week. Blinding of key study participants and personnel attempted, but likely that blinding could have been broken and the outcome was likely to be influenced by lack of blinding.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of ‘low risk’ or ‘high risk’. Blinding not stated.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Personnal reason, lost of F/U" unclear reason, could be related to intervention. Insufficient reporting of attrition/exclusion to permit a judgement of 'low risk' or 'high risk'.

Wei 2012 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	BP stated in the method not in the result. But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Wu 2008

Study characteristics	
Methods	Randomised controlled trial **Location of the trial:** quote: "Nanjing, China"**Method of randomisation:** Quote: "Randomization was produced from a computer generated random list"**Method of allocation concealment:** not stated**Source of funding:** Not stated
Participants	**Inclusion criteria:** Age 19 to 35 years PCOS consistent with Rotterdam. (1) The presence of polycystic ovaries (10 or more subcapsular follicles of 2 mm to 8 mm diameter in one plane in one ovary and increased stroma) by pelvic ultrasonography; (2) oligomenorrhoea (more than six cycles longer than 36 days in the previous year) or amenorrhoea (absence of menstruation for three consecutive months) and/or hirsutism, acne (evaluated by the modified Ferriman-Gallwey method); (3) elevated serum testosterone level (>2.5 nmol/L) and/or the ratio of serum luteinising hormone (LH) to follicle-stimulating hormone (FSH) >2. **Exclusion criteria:** exclusion criteria consisted of patients with secondary aetiologies, including hyperprolactinaemia, thyroid dysfunction, Cushing's syndrome, congenital adrenal hyperplasia, diabetes mellitus, smokers, alcohol users and those taking sex hormones for 3 months preceding the study. **Number of women randomised:** 60 MET: 20 OCP: 20 OCP + MET: 20 **Number of women analysed:** 53 MET: 18 OCP: 19 OCP + MET: 16 **Number of withdrawal and reasons:** MET: 2 GI adverse events 2/ 10% OCP:1 body weight increase 1/ 5% OCP + MET: 4 GI adverse events 4/ 20% **Summary characteristics:** PCOS **Age (years):** OCP mean (±SD): obese 25.0 (4.3) non-obese 26.1 (4.6) MET mean (±SD): obese 25.6 (3.6) non-obese 25.6 (4.2) OCP + MET: obese 24.5 (2.4) non-obese 25.8 (4.0) **BMI (kg/m^2):**

Wu 2008 *(Continued)*

	OCP mean (±SD): obese 25.3 (0.8) non-obese 21.4 (1.6) MET: obese 25.6 (0.6) non-obese 21.5 (1.8) OCP + MET: obese 25.2 (1.0) non-obese 21.6 (1.4)
Interventions	OCP (Ethynil estradiol 35 mcg Cyproterone acetate 2 mg) for 21 days per month at the first day of bleeding for 3 months MET 500 mg metformin three times a day at the first day of bleeding for 3 months MET 500 mg MET three times a day + OCP (Ethynil estradiol 35 mcg Cyproterone acetate 2 mg) for 21 days per month at the first day of bleeding for 3 months **Duration:** 3 months **Co-intervention(s):** none
Outcomes	**Primary outcomes:** Hirsutism score Adverse events: severe (requiring stopping of medication) and minor**Secondary outcomes:** Menstrual cyclicity, initiation of menses or significant shortening of cycles Body weight (kg) BMI (kg/m^2) Serum total testosterone (nmol/L) Fasting insulin (mIU/L) Fasting glucose (mmol/L)
Subjective outcomes	(a) Clinical parameters 1. Hirsutism score
Objective outcomes	(a) Clinical parameters 1. Adverse events: severe (requiring stopping of medication) and minor 2. Menstrual cyclicity, initiation of menses or significant shortening of cycles 3. Body weight (kg) and/or BMI (kg/m^2) (b) Hormonal parameters 1. Serum total testosterone (nmol/L) (c) Metabolic parameters 1. Fasting insulin (mIU/L) 2. Fasting glucose (mmol/L)
Notes	**Power calculation:** unclear
Risk of bias	
Bias	**Authors' judgement** **Support for judgement**

Wu 2008 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote: "Randomization was produced from a computer generated random list"
Allocation concealment (selection bias)	Unclear risk	Method of allocation concealment not stated. Insufficient information available to permit judgement of 'low risk' or 'high risk'.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information available to permit a judgement of 'low risk' or 'high risk'. Blinding not stated.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "All clinical parameters were recorded by the same investigator, who was blinded to the type of treatment in the study."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Discrepancy between the explanations for the number of withdrawals for the metformin group. The description for Figure 1 states that two people withdrew from the metformin arm due to gastrointestinal side effects, whereas the text states that although two people experienced GIT adverse events in the metformin arm, this did not necessitate their withdrawal. Unclear. Insufficient reporting of attrition/exclusions to permit a judgement of 'low risk' or 'high risk'.
Selective reporting (reporting bias)	Unclear risk	From the results section of the paper, all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way with the exception of weight and glucose, which were only reported as BMI and glucose/insulin ratio, respectively. Fasting glucose is only reported as quote "fasting glucose levels showed no changes in the three groups at the end of the study (data not shown)" Menstrual bleeding was inadequately reported for the OCP and OCP/metformin arms quote: "Menstrual regularity was restored as expected in all PCOS patients from the Diane 35-treated group and Diane 35/metformin group, but only approximately 28% of subjects in the metformin-treated group resumed regular menses." But not a primary or key outcomes. Insufficient information available to permit a judgment of 'low risk' or 'high risk'.
Other bias	Low risk	The study appears to be free of other sources of bias.

Abbreviations used:
BMI: Body mass index; **BP:** Blood pressure;**BUN:** blood urea nitrogen; **CI:** Confidence interval;**CPA:** Cyproterone acetate;**EE:** Ethinyl estradiol;**FAI:** Free androgen index; **F/U:** follow-up; **GI:** gastro-intestinal; **HDL:** High-density lipoprotein; **lDLC:** Low-density lipoprotein; **MET:** metformin; **NA:** Not available; **NIH:** National Institutes of Health; **OCP:** Oral contraceptive pill;**OGTT:** oral glucose tolerance test; **PCOS:** Polycystic ovary syndrome;**RCT:** Randomised controlled trial; **D:** Standard deviation;**SE:** Standard error of the mean;**SHBG:** Sex hormone-binding globulin; **T:** Testosterone; **T2DM:** Type II diabetes mellitus; **VAS:** visual analogue scale; **WHR:** Waist:hip ratio; difference

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Alpanes 2017	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).

Study	Reason for exclusion
Altinok 2018	No primary or secondary outcomes of interest for us.
Bachani 2016	Not a RCT
Bhattacharya 2012	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Bredella 2013	Not a RCT.
Burchall 2015	Not a RCT.
Cakiroglu 2013	Objective of the study was to look at outcomes according to the BMI and not per intervention. Email sent to the author asking for their results according to the treatment: no answer.
Diaz 2016	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Glintborg 2014b	Study excluded because population and results described in the first study of the same author.
Glintborg 2015	No primary or secondary outcomes of interest for us.
Glintborg 2017	Study excluded because population and results described in the first study of the same author.
Hadziomerovic-Pekic 2010	No intervention of interest for us (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Harris-Glocker 2009	No primary or secondary outcomes of interest for us.
Hu 2010	Not a RCT.
Hutchison 2008	Study excluded because population and results described in an other study. (Meyer 2007).
Ibanez 2010	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Ibanez 2017	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Kebapcilar 2010	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Kim 2010	Duration less than 3 months.
Ladson 2011	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Lazaro 2011	No intervention of interest (OCP versus Metformin or OCP versus Metformin + OCP or Metformin versus Metformin + OCP).
Lemay 2006	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Luque-Ramirez 2008c	No primary or secondary outcomes of interest for us.
Luque-Ramirez 2009	No primary or secondary outcomes of interest for us.

Study	Reason for exclusion
Luque-Ramirez 2010a	No primary or secondary outcomes of interest for us.
Luque-Ramirez 2011	No primary or secondary outcomes of interest for us.
Mehrabian 2016	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Mitkov 2005	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Moghtadaei 2009	Abstract only. No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Moretti 2016	Abstract only. No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
NCT02866786	Information given by the authors, no further publication.
Orbetzova 2011	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Panidis 2011	Not an RCT: quasi-randomised study.
Pedersen 2018	Objective of the study was to look at outcomes per phenotype and not per intervention. Email sent to the author asking for their results according to the treatment: results presented in an another article by Glinborg.
Romualdi 2010	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).
Suvarna 2016	Not a RCT.
Teede 2010b	Study excluded because population and results described in another study.
Wang 2016	No intervention of interest (OCP versus metformin or OCP versus metformin + OCP or metformin versus metformin + OCP).

BMI: body mass index; **OCP:** oral contraceptive pill; **RCT:** randomised controlled trial.

Characteristics of studies awaiting classification *[ordered by study ID]*

Fruzzeti 2009

Methods	Randomised controlled trial
Participants	**Inclusion criteria:** women with PCOS who attended to the Outpatient Clinic of Reproductive Endocrinology of the University of Pisa were enrolled prospectively for the study during 2008 (from March to November).The mean age was 24 years (range, 15 to 34 years). Women were diagnosed with PCOS on the basis of the presence of chronic oligomenorrhoea and hirsutism, according to the Rotterdam and National Institutes of Health criteria. Hirsutism was defined as a Ferriman-Gallwey score >8. The hirsutism score in these women ranged from 9 to 22.

Fruzzeti 2009 *(Continued)*

	Exclusion criteria: none of the patients were affected by hypertension, glucose intolerance, or diabetes mellitus, and none of the participants had a personal history of cardiovascular events or received treatment with OCs, antiandrogens, insulin sensitisers, or drugs that might interfere with blood pressure regulation, lipid profile, or carbohydrate metabolism for the previous 6 months. Participants with hyperprolactinaemia, hypo- or hyperthyroidism, congenital adrenal hyperplasia, Cushing's syndrome, or androgen-secreting tumours were excluded from this study.
Interventions	Treatment: OCP (Ethinyl Estradiol 20mcg, drospirenone 3mg) 21 days/ 7 days off + MET 500 mg three times daily Control: OCP (Ethinyl Estradiol 20mcg, drospirenone 3mg) 21 days/7 days off
Outcomes	(a) Clinical parameters BMI (kg/m^2) (b) Metabolic parameters: 1.Fasting glucose (mmol/L) 2.Fasting insulin (IU/ml) 3.Fasting total cholesterol (mmol/L) 4.Fasting HDL cholesterol (mmol/L) 5.Fasting LDL cholesterol (mmol/L) 6.Fasting triglycerides (mmol/L
Notes	Authors contacted because results given as a graph without real values.

NCT01573377

Methods	Randomised controlled trial
Participants	**Inclusion criteria:** at least two of the following features: (i) oligo-amenorrhoea or chronic anovulation; (ii) clinical and/or biochemical hyperandrogenism; (iii) ultrasound appearance of polycystic ovaries **Exclusion criteria:** other known causes of hyperandrogenaemia and ovulatory dysfunction, including 21-hydroxylase deficiency, congenital adrenal hyperplasia, Cushing's syndrome, androgen-secreting tumours, thyroid disease, and hyperprolactinaemia. use of hormone medications (including oral contraceptives) within the past month and the use of medicines that affect insulin sensitivity (e.g. metformin or thiazolidinediones) within the past three months
Interventions	Treatment: MET 425 mg twice daily/ one week then 850 mg twice daily Control: OCP (Ethinyl Estradiol 35 mcg/cyproterone acetate 2 mg)
Outcomes	(a) Clinical parameters Menstrual cycle (b) Metabolic parameters Fasting insulin (IU/mL)
Notes	Completed, email sent to the authors to have the publication references because not found in databases.

Spremovic-Radjenovic 2014

Methods	Randomised controlled trial
Participants	PCOS Rotterdam criteria, non-obese
Interventions	Treatment: MET no information about the dose Control: OCP (Ethinyl Estradiol/dianogest) no information about the dose
Outcomes	(a) metabolic parameters: 1.Fasting glucose (mmol/L) 2.Fasting insulin (IU/ml) 3.Fasting total cholesterol (mmol/L) 4.Fasting HDL cholesterol (mmol/L) 5.Fasting LDL cholesterol (mmol/L) 6.Fasting triglycerides (mmol/L
Notes	Abstract only Authors contacted, said they will send their results, still waiting.

Vieira 2010

Methods	Randomised controlled trial
Participants	PCOS No other information
Interventions	Treatment: OCP (Ethinyl Estradiol 30 mcg/chlormadinone acetate 2 mg)/day + MET 875 mg/day Control: OCP (Ethinyl Estradiol 30 mcg/chlormadinone acetate 2 mg)/day
Outcomes	(a) Clinical parameters 1.Body weight (kg) and/or BMI (kg/m^2) 2.Waist circumference (cm) and/or waist-hip ratio (WHR) 3.Blood pressure (systolic, diastolic) (mm Hg) (b) Metabolic parameters 1.Fasting LDL cholesterol (mmol/L) 2.Fasting triglycerides (mmol/L)
Notes	Abstract only Authors contacted, said they will send their results, still waiting.

BMI: body mass index; **HDL:** high-density lipoprotein; **IU:** international units; **LDL:** low-density lipoprotein;**MET:** metformin; **OCP:** oral contraceptive pill; **PCOS:** polycystic ovary syndrome;

Characteristics of ongoing studies *[ordered by study ID]*

NCT02744131

Study name	OCP vs metformin for Improvement in clinical symptoms and metabolic markers in Indian PCOS women (OCP)
Methods	Randomised controlled trial
Participants	**Inclusion criteria** - Women with PCOS diagnosed by Rotterdam criteria - Age > 16 or menarche of at least 2 years up to age 40. - Diagnosed by Rotterdam Anovulation Ferriman-Gallwey score ≥ 8 or Total testosterone ≥ 40 ngm/dL ovarian volume ≥ 10 cc BMI > 23 - Women not attempting pregnancy at present. **Exclusion criteria** - Women with PCOS with BMI ≤ 23 - Undergoing any treatment for acne hirsutism. Including Homeopathic, Ayurvedic - Known diabetic or hypertension .
Interventions	Group a: OCP (Ethinyl Estradiol 20 mcg/ 2 mg Cyproterone acetate). Group b: MET 1500 g/day
Outcomes	Improvements in menstrual cycle Improvement in hirsutism Improvement in acne score Weight loss following treatment Reduction in total testosterones Reduction in serum fasting insulin Improvements in lipid profile
Starting date	04/2016
Contact information	Responsible party: Sujata Kar,Odisha, India
Notes	Recruiting

NCT03229057

Study name	Comparing the effects of oral contraceptive pills versus metformin (COMET-PCOS)
Methods	Three-arm, double-blind, double-dummy, multicentre, prospective, randomised clinical trial
Participants	**Inclusion criteria** - Women ≥ 18 to ≤ 40 years of age (at the time of screening), with hyperandrogenic PCOS. - Participants will be diagnosed with PCOS defined by the Rotterdam criteria 20 based on: A history of chronic anovulation (8 or fewer periods per year) androgen excess (defined as an elevated serum T level or hirsutism, based on a Ferriman-Gallwey score > 8 (note: > 2 for women of Asian descent) +/- polycystic ovaries. - BMI ≥ 25 kg/m^2 to ≤ 45 kg/m^2 obtained at screening visit. - In good general health. - Willing to avoid pregnancy for the duration of the study. **Exclusion criteria** - Current pregnancy or desire of pregnancy during course of study

NCT03229057 *(Continued)*

	- Currently breastfeeding - Known 21 hydroxylase deficiency - Untreated thyroid disease (TSH < 0.45 mIU/mL and > 4.5 mIU/mL) - Untreated hyperprolactinaemia (2 Levels>30 ng/mL at least one week apart) - Type 1 or type 2 diabetes mellitus (elevated fasting serum glucose >126 mg/dL on two occasions, poorly controlled diabetes (HbA1C > 6.5%), currently receiving anti-diabetic agents, or currently receiving metformin for treatment of diabetes - Liver disease (AST/ALT>2 times normal or a total bilirubin >2.5 mg/dL) - Renal disease (BUN>30 mg/dL or serum creatinine >1.4 mg/dL) - Anemia (Hb < 10 mg/dL) - History of deep venous thrombosis, pulmonary embolus, or cerebrovascular accident - Current history of alcohol abuse (>14 drinks/week) - Poorly controlled hypertension defined as average SBP >= 150 mm Hg or average DBP >= 100 mm Hg obtained on three measurements obtained 5 minutes apart. If treated, average SBP >=140 mm Hg or average DBP >= 90 mm Hg - Patients with a history of, or suspected cervical carcinoma, endometrial carcinoma, or breast carcinoma - TG>200 mg/dL - Use of lipid-lowering or weight loss agents (participants may wash out from weight loss agents) - Current use of oral contraceptives, depo progestin, or hormonal implants - Participation in any study of an investigational drug or device or biological agent within 30 days - Suspected adrenal or ovarian tumour secreting androgens - Suspected Cushing's syndrome - Bariatric surgery procedure in the recent past (<12 months) - Absolute contraindications to the use of hormonal contraceptives or metformin - Participants who are unable to comply with the study procedures, for instance due to mental illness, substance abuse, or participation in other studies.
Interventions	Group a: MET 2000 mg/day + placebo Group b: OCP (Ethinyl Estradiol 20 mcg/desogestrel 0.15 mg) + placebo Group c: MET + OCP
Outcomes	HDL-C function
Starting date	Juillet 2017
Contact information	Responsible party: Anuja Dokras, University of Pennsylvania, USA
Notes	Recruiting

NCT03905941

Study name	Relative desirability of metformin vs. birth control pill in treating PCOS in women of later reproductive age
Methods	Randomised controlled trial
Participants	**Inclusion criteria** - Women with PCOS aged 40-49 years. Participant is considered to have PCOS if she has current or verifiable history of: a) clinical and/or biochemical evidence of hyperandrogenism plus b) oligomenorrhoea or irregular menstruation (substantially inconsistent menstrual cycle length). Participants with fewer than 10 menses/year or average menstrual cycle length > 35 days are allowed to participate if they have a compelling past history of oligomenorrhoea (average menstrual cycle length > 45 days or fewer than 9 menses/year) or irregular menstruation. - Screening safety labs within normal reference ranges although mild abnormalities that are common in obesity and/or hyperandrogenism will not be grounds for exclusion (see exclusion criteria).

NCT03905941 *(Continued)*

- Participants must be willing and able to provide written informed consent.
- Willingness to strictly avoid pregnancy (using non-hormonal methods) during the time of the study
- Willingness and ability to comply with scheduled visits and study procedures

Exclusion criteria

- Postmenopausal status (i.e. absence of periods for previous year plus elevated follicle stimulating hormone [FSH] level)
- Biochemical evidence for perimenopause as defined by an anti-Mullerian hormone <0.5 ng/mL. As an alternative, cycle day 3 FSH > 9 IU/L (with concomitant estradiol level >80 pg/mL), if this testing is available, will serve as evidence of perimenopause status. NOTE: If FSH >9 IU/L on screening (but it is not cycle day 3), FSH and estradiol will be repeated on cycle day 3
- History of hysterectomy and/or bilateral oophorectomy
- BMI ≥ 40 kg/m^2
- Inability to comprehend what will be done during the study or why it will be done.
- Being a study of older women with PCOS, children and men will be excluded.
- Pregnancy or lactation within the past 6 months. Participants with a positive pregnancy test will be informed of the result by the screening physician.
- Prisoners.
- History of (or clinical evidence for) Cushing's syndrome or adrenal insufficiency.
- History of congenital adrenal hyperplasia or 17-hydroxyprogesterone (17-OHP) >200 ng/dL, which suggest the possibility of congenital adrenal hyperplasia. 17-OHP will be collected during follicular phase. NOTE: if a 17-OHP >200 ng/dL and is confirmed on repeat testing, an ACTH-stimulated 17-OHP <1000 ng/dL will be required for study participation.
- Total testosterone >150 ng/dL, which suggests the possibility of virilising neoplasm.
- DHEA-S greater than 1.5 times the upper limit of normal range (mild elevations may be seen in PCOS, so elevations < 1.5 times the upper limit of normal will be accepted in these groups).
- Virilisation
- Diagnosis of diabetes mellitus (DM), fasting glucose ≥ 126 mg/dL, or a HbA1c of ≥ 6.5%.
- Abnormal thyroid stimulating hormone (TSH). Participants with stable and adequately-treated hypothyroidism, reflected by normal TSH values, will not be excluded.
- Moderate to severe hyperprolactinaemia. Mild prolactin elevations may be seen in PCOS, and elevations < 1.5 times the upper limit of normal will be accepted in this group.
- Persistent liver abnormalities, with the exception that mild bilirubin elevations will be accepted in the setting of known Gilbert's syndrome. Mild transaminase elevations may be seen in women with obesity, so elevations < 1.5 times the upper limit of normal will be accepted in this group.
- Persistent HCT <36% and Hb <12 g/dL.
- Abnormal sodium, potassium, or bicarbonate concentrations or elevated creatinine concentration.
- Significant history of pulmonary dysfunction (e.g. asthma or chronic obstructive pulmonary disease (COPD) requiring intermittent systemic corticosteroid, pulmonary hypertension, etc.).
- History of known or suspected congestive heart failure.
- History of known or suspected ischaemic heart disease or cerebrovascular disease.
- History of hypertension.
- History of uncontrolled/untreated dyslipidaemia. Subjects with stable and adequately treated dyslipidaemia reflected by normal lipid panel values will not be excluded.
- History of complicated valvular heart disease (e.g. pulmonary hypertension, risk of atrial fibrillation, history of subacute bacterial endocarditis)
- History of stroke
- History of smoking
- History of severe cirrhosis or liver tumour (e.g. hepatocellular adenoma or malignant hepatoma).
- Use of anticonvulsants, rifampicin or rifabutin therapy. The interaction of these drugs with OCs will not be harmful to the subjects, but it will reduce the effectiveness of OCs.
- History of venous thromboembolism (e.g. deep venous thrombosis (DVT), pulmonary embolism (PE)).
- Personal history of blood clotting disorders (e.g. protein C, protein S, positive antiphospholipid antibodies).
- First-degree relative history with blood clotting disorder, unless the same disorder has been formally excluded for the study participant.
- History of migraine headaches.
- History of breast, ovarian, or endometrial cancer

NCT03905941 *(Continued)*	
	- No medications known to affect the reproductive system can be taken in the 2 months prior to screening and in the 3 months prior to the study. Such medications include oral contraceptive pills, metformin, progestins, glucocorticoids, anti-psychotics, and/or mood stabilisers that are known to cause hormone abnormalities
Interventions	Group a: OCP (Ethinyl Estradiol 20 mcg/ norethindrone acetate 1 mg). Group b: MET 2000 g/day
Outcomes	Weight kg Blood pressure mmHg BMI kg/m^2 Total testosterone ng/dL LDL cholesterol level mg/dL HDL cholesterol level mg/dL TG level mg/dL Fasting insulin µIU/mL Fasting glucose mg/dL
Starting date	March 2020
Contact information	Responsible party: Su Hee Kim, University of Virginia
Notes	Recruiting

ALT: alanine aminotransferase; **AST:** Aspartate transaminase; **BMI:** body mass index; **BUN: blood** urea nitrogen; **COPD:** chronic obstructive pulmonary disease, **DBP:** diastolic blood pressure; **FSH:** follicle-stimulating hormone; **Hb:** haemoglobin; **HbA1C:** haemoglobin A1C; **HCT:** haematocrit; **HDL:** high-density lipoprotein; **IU:** international units; **LDL:** low-density lipoprotein;**MET:** metformin; **OCP:** oral contraceptive pill; **PCOS:** polycystic ovary syndrome; **SBP:** systolic blood pressure; **TG**: triglyceride**; TSH:** thyroid stimulating hormone;

DATA AND ANALYSES

Comparison 1. Adult - Metformin versus OCP (Clinical parameters)

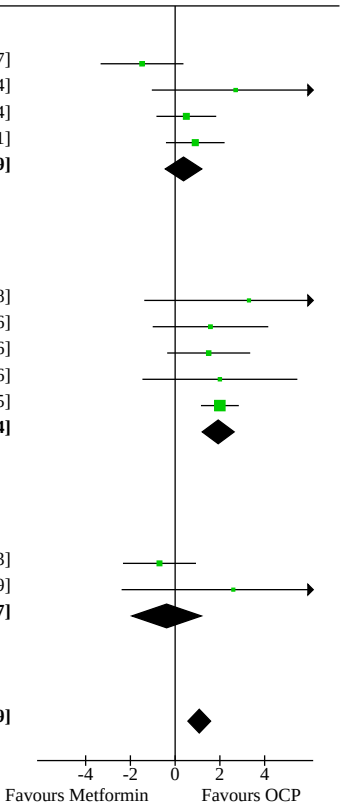
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Hirsutism - Clinical F-G score	10	473	Mean Difference (IV, Fixed, 95% CI)	1.08 [0.57, 1.59]
1.1.1 BMI ≤ 25 kg/m^2	3	134	Mean Difference (IV, Fixed, 95% CI)	0.38 [-0.44, 1.19]
1.1.2 BMI > 25 kg/m^2 < 30 kg/m^2	5	254	Mean Difference (IV, Fixed, 95% CI)	1.92 [1.21, 2.64]
1.1.3 BMI ≥ 30 kg/m^2	2	85	Mean Difference (IV, Fixed, 95% CI)	-0.38 [-1.93, 1.17]
1.2 Hirsutism - Subjective visual analogue scale	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.2.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.3 Hirsutism - Subjective improvement	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.3.1 Mean BMI not stated	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.4 Adverse events - severe	12		Peto Odds Ratio (Peto, Fixed, 95% CI)	Subtotals only
1.4.1 Gastro-intestinal	11	602	Peto Odds Ratio (Peto, Fixed, 95% CI)	6.42 [2.98, 13.84]
1.4.2 Others	8	363	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.20 [0.09, 0.44]
1.5 Improved menstrual pattern (ie. shortening of intermenstrual days)	2	153	Mean Difference (IV, Fixed, 95% CI)	6.05 [2.37, 9.74]
1.5.1 BMI > 25 kg/m^2< 30 kg/m^2	1	86	Mean Difference (IV, Fixed, 95% CI)	6.10 [2.40, 9.80]
1.5.2 BMI ≥ 30 kg/m^2	1	67	Mean Difference (IV, Fixed, 95% CI)	-3.10 [-55.47, 49.27]
1.6 Improved menstrual pattern (ie. an initiation of menses or cycle regularity)	6	189	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.21 [0.11, 0.40]
1.6.1 BMI ≤ 25 kg/m^2	1	17	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.07 [0.01, 0.65]
1.6.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	129	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.15 [0.07, 0.33]
1.6.3 BMI ≥ 30 kg/m^2	1	18	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.09 [0.01, 1.62]
1.6.4 Mean BMI not stated	1	25	Peto Odds Ratio (Peto, Fixed, 95% CI)	1.95 [0.39, 9.65]
1.7 Acne - Visual analogue scale	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.7.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.8 Acne - Subjective improvement	3	131	Odds Ratio (M-H, Fixed, 95% CI)	0.30 [0.11, 0.79]
1.8.1 BMI ≤ 25 kg/m^2	1	80	Odds Ratio (M-H, Fixed, 95% CI)	0.67 [0.16, 2.91]
1.8.2 BMI > 25 kg/m^2 < 30 kg/m^2	1	26	Odds Ratio (M-H, Fixed, 95% CI)	0.21 [0.03, 1.50]
1.8.3 Mean BMI not stated	1	25	Odds Ratio (M-H, Fixed, 95% CI)	0.10 [0.01, 0.72]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.9 Diagnosis of Type II diabetes mellitus	1		Peto Odds Ratio (Peto, Fixed, 95% CI)	Totals not selected
1.9.1 BMI ≥ 30 kg/m^2	1		Peto Odds Ratio (Peto, Fixed, 95% CI)	Totals not selected
1.10 Body weight (kg)	7	358	Mean Difference (IV, Fixed, 95% CI)	-0.93 [-2.93, 1.08]
1.10.1 BMI ≤ 25 kg/m^2	3	168	Mean Difference (IV, Fixed, 95% CI)	4.02 [-0.22, 8.25]
1.10.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	134	Mean Difference (IV, Fixed, 95% CI)	-1.96 [-6.32, 2.41]
1.10.3 BMI ≥ 30 kg/m^2	1	56	Mean Difference (IV, Fixed, 95% CI)	-2.50 [-5.16, 0.16]
1.11 Body Mass Index (kg/m^2)	19	923	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-0.74, -0.06]
1.11.1 BMI ≤ 25kg/m^2	9	451	Mean Difference (IV, Fixed, 95% CI)	-0.59 [-1.02, -0.17]
1.11.2 BMI > 25kg/m^2< 30kg/m^2	8	353	Mean Difference (IV, Fixed, 95% CI)	0.11 [-0.48, 0.70]
1.11.3 BMI ≥ 30 kg/m^2	3	119	Mean Difference (IV, Fixed, 95% CI)	-2.31 [-4.40, -0.21]
1.12 Blood pressure - systolic (mm Hg)	5	209	Mean Difference (IV, Fixed, 95% CI)	-4.81 [-8.55, -1.06]
1.12.1 BMI > 25 kg/m^2< 30 kg/m^2	3	108	Mean Difference (IV, Fixed, 95% CI)	-5.48 [-9.56, -1.41]
1.12.2 BMI ≥ 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	-1.11 [-10.62, 8.41]
1.13 Blood pressure - diastolic (mm Hg)	4	142	Mean Difference (IV, Fixed, 95% CI)	-1.98 [-4.72, 0.76]
1.13.1 BMI > 25 kg/m^2< 30 kg/m^2	3	108	Mean Difference (IV, Fixed, 95% CI)	-4.25 [-7.30, -1.20]
1.13.2 BMI ≥ 30 kg/m^2	1	34	Mean Difference (IV, Fixed, 95% CI)	7.50 [1.27, 13.73]

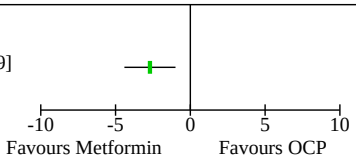
Analysis 1.1. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 1: Hirsutism - Clinical F-G score

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
1.1.1 BMI ≤ 25 kg/m2								
Cetinkalp 2009	10.55	4.1	47	12.03	4.2	33	7.5%	-1.48 [-3.33 , 0.37]
Morin-Papunen 2003	7	5.37	8	4.3	0.6	9	1.8%	2.70 [-1.04 , 6.44]
Wu 2008	7.4	2	11	6.9	1.1	12	14.4%	0.50 [-0.84 , 1.84]
Wu 2008	7.7	1.2	7	6.8	1.3	7	15.0%	0.90 [-0.41 , 2.21]
Subtotal (95% CI)			**73**			**61**	**38.7%**	**0.38 [-0.44 , 1.19]**
Heterogeneity: Chi² = 5.99, df = 3 (P = 0.11); I² = 50%								
Test for overall effect: Z = 0.90 (P = 0.37)								
1.1.2 BMI > 25 kg/m2 < 30 kg/m2								
Dardzinska 2014	13.4	7.3	13	10.1	5.8	21	1.2%	3.30 [-1.38 , 7.98]
Glintborg 2014a	5.58	4.3	19	4	4.19	23	3.9%	1.58 [-1.00 , 4.16]
Kumar 2018	7.3	3.9	30	5.8	3.3	28	7.5%	1.50 [-0.36 , 3.36]
Luque-Ramirez 2007a	9.3	5.8	19	7.3	4.5	15	2.1%	2.00 [-1.46 , 5.46]
Sahu 2018	7	2	42	5	2	44	36.0%	2.00 [1.15 , 2.85]
Subtotal (95% CI)			**123**			**131**	**50.6%**	**1.92 [1.21 , 2.64]**
Heterogeneity: Chi² = 0.63, df = 4 (P = 0.96); I² = 0%								
Test for overall effect: Z = 5.29 (P < 0.00001)								
1.1.3 BMI ≥ 30 kg/m2								
Meyer 2007	-2.7	3.54	36	-2	3.27	31	9.7%	-0.70 [-2.33 , 0.93]
Morin-Papunen 2000	10	5.37	8	7.4	5.37	10	1.0%	2.60 [-2.39 , 7.59]
Subtotal (95% CI)			**44**			**41**	**10.7%**	**-0.38 [-1.93 , 1.17]**
Heterogeneity: Chi² = 1.52, df = 1 (P = 0.22); I² = 34%								
Test for overall effect: Z = 0.48 (P = 0.63)								
Total (95% CI)			**240**			**233**	**100.0%**	**1.08 [0.57 , 1.59]**
Heterogeneity: Chi² = 19.81, df = 10 (P = 0.03); I² = 50%								
Test for overall effect: Z = 4.17 (P < 0.0001)								
Test for subgroup differences: Chi² = 11.67, df = 2 (P = 0.003), I² = 82.9%								



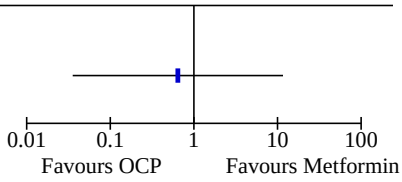
Analysis 1.2. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 2: Hirsutism - Subjective visual analogue scale

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Mean Difference IV, Fixed, 95% CI
1.2.1 BMI ≥ 30 kg/m2							
Harborne 2003	3.9	3.1	18	6.6	1.9	16	-2.70 [-4.41 , -0.99]



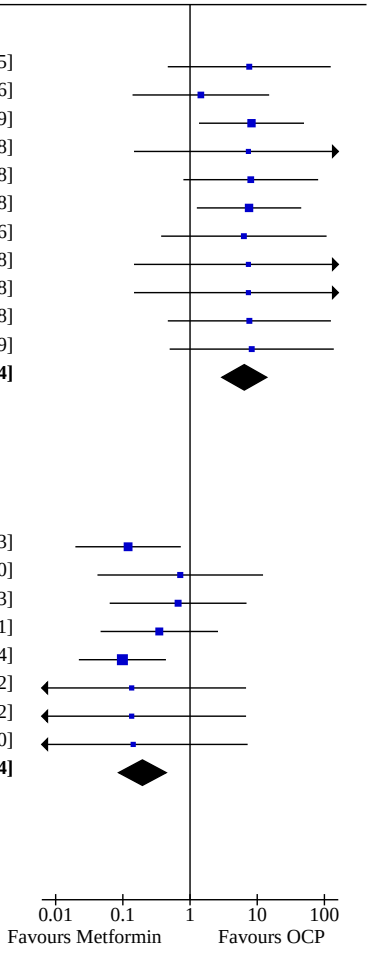
Analysis 1.3. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 3: Hirsutism - Subjective improvement

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Odds Ratio M-H, Fixed, 95% CI
1.3.1 Mean BMI not stated					
Jin 2006	1	15	1	10	0.64 [0.04 , 11.63]



Analysis 1.4. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 4: Adverse events - severe

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
1.4.1 Gastro-intestinal						
Aghamohammadzadeh 2010	2	35	0	35	7.6%	7.61 [0.47 , 124.15]
Bodur 2018	2	29	1	21	10.8%	1.45 [0.14 , 15.06]
Christakou 2014	5	40	0	40	18.2%	8.22 [1.36 , 49.69]
Glintborg 2014a	1	30	0	30	3.8%	7.39 [0.15 , 372.38]
Harborne 2003	3	26	0	26	11.1%	8.02 [0.80 , 80.68]
Kilic 2011	5	54	0	51	18.5%	7.56 [1.26 , 45.18]
Luque-Ramirez 2007a	2	19	0	15	7.3%	6.33 [0.37 , 107.76]
Morin-Papunen 2000	1	16	0	16	3.8%	7.39 [0.15 , 372.38]
Morin-Papunen 2003	1	10	0	10	3.8%	7.39 [0.15 , 372.38]
Moro 2013	2	31	0	31	7.6%	7.64 [0.47 , 124.98]
Wu 2008	2	18	0	19	7.5%	8.28 [0.50 , 137.89]
Subtotal (95% CI)		**308**		**294**	**100.0%**	**6.42 [2.98 , 13.84]**
Total events:	26		1			
Heterogeneity: Chi² = 1.77, df = 10 (P = 1.00); I² = 0%						
Test for overall effect: Z = 4.74 (P < 0.00001)						
1.4.2 Others						
Aghamohammadzadeh 2010 (1)	0	35	5	35	20.4%	0.12 [0.02 , 0.73]
Bodur 2018 (2)	1	29	1	21	8.3%	0.72 [0.04 , 12.20]
Dardzinska 2014 (3)	1	18	2	24	12.1%	0.66 [0.06 , 6.93]
Glintborg 2014a (4)	1	30	3	30	16.4%	0.35 [0.05 , 2.61]
Harborne 2003 (5)	0	26	8	26	29.9%	0.10 [0.02 , 0.44]
Morin-Papunen 2000 (6)	0	16	1	16	4.3%	0.14 [0.00 , 6.82]
Morin-Papunen 2003 (6)	0	10	1	10	4.3%	0.14 [0.00 , 6.82]
Wu 2008 (7)	0	18	1	19	4.3%	0.14 [0.00 , 7.20]
Subtotal (95% CI)		**182**		**181**	**100.0%**	**0.20 [0.09 , 0.44]**
Total events:	3		22			
Heterogeneity: Chi² = 3.35, df = 7 (P = 0.85); I² = 0%						
Test for overall effect: Z = 3.92 (P < 0.0001)						
Test for subgroup differences: Chi² = 37.31, df = 1 (P < 0.00001), I² = 97.3%						

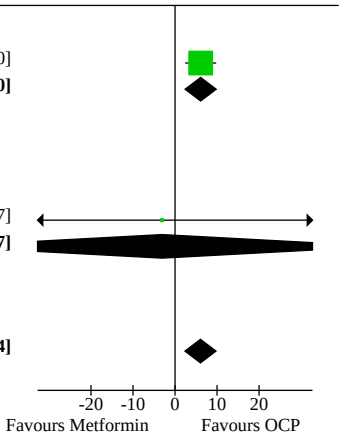


Footnotes

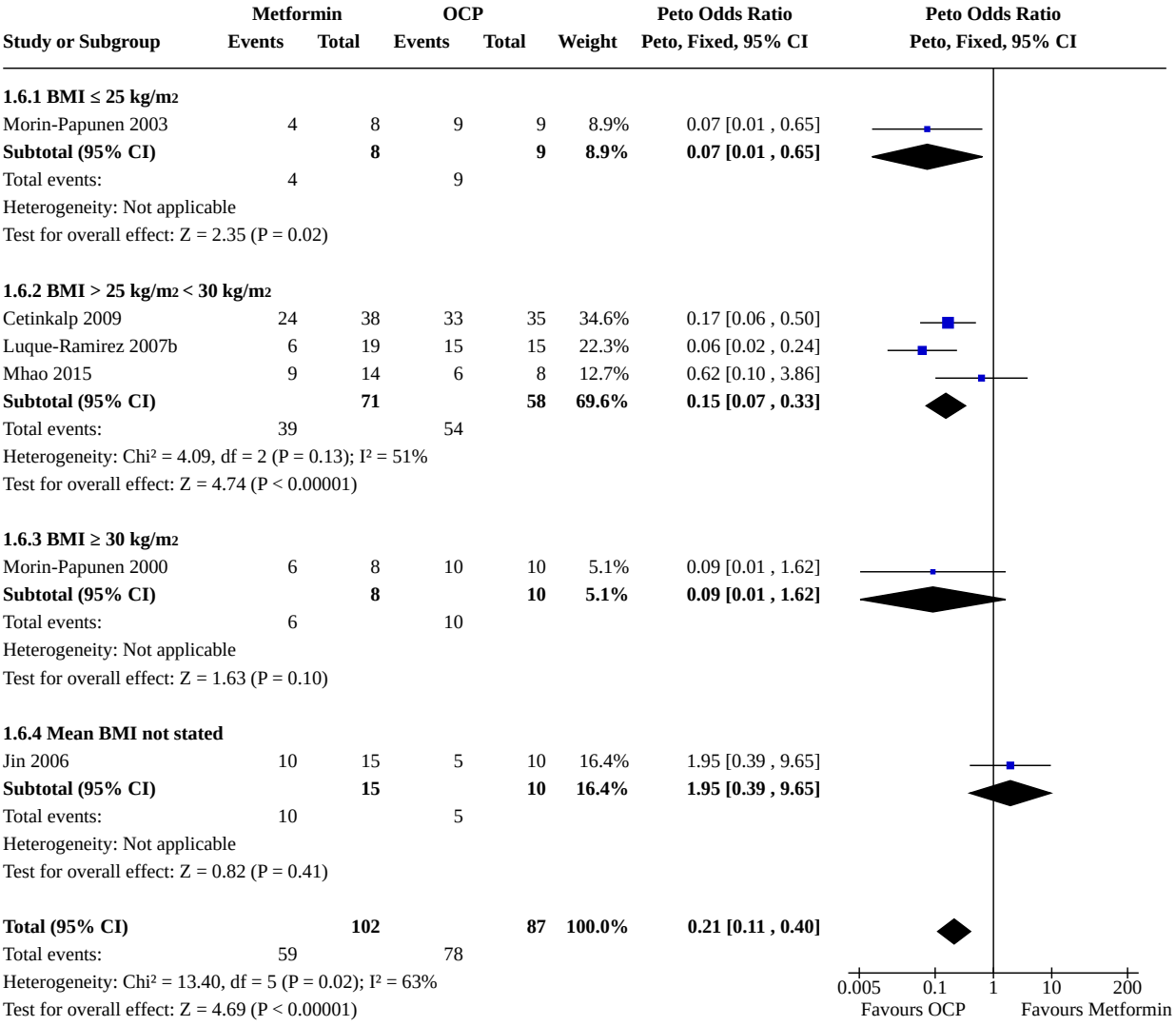
(1) OCP: combined: nausea, increasing BW, drug intolerance
(2) MET: dizziness/ OCP: sexual reluctance
(3) MET: depression/ OCP: intolerance
(4) Met: depression/ OCP: side effects
(5) Weight gain, depression, HBP, chest pain
(6) Headache, HBP
(7) Weight gain

Analysis 1.5. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 5: Improved menstrual pattern (ie. shortening of intermenstrual days)

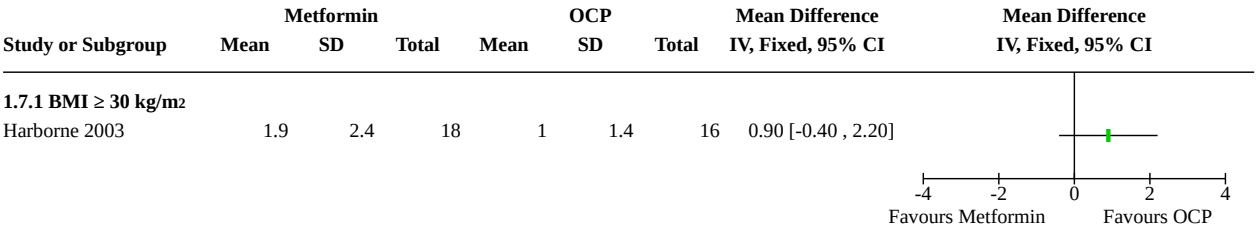
Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
1.5.1 BMI > 25 kg/m2< 30 kg/m2								
Sahu 2018	39.9	10.2	42	33.8	6.9	44	99.5%	6.10 [2.40 , 9.80]
Subtotal (95% CI)			**42**			**44**	**99.5%**	**6.10 [2.40 , 9.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.23 (P = 0.001)								
1.5.2 BMI ≥ 30 kg/m2								
Meyer 2007	-47.1	140.04	36	-44	72.4366	31	0.5%	-3.10 [-55.47 , 49.27]
Subtotal (95% CI)			**36**			**31**	**0.5%**	**-3.10 [-55.47 , 49.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.12 (P = 0.91)								
Total (95% CI)			**78**			**75**	**100.0%**	**6.05 [2.37 , 9.74]**
Heterogeneity: Chi² = 0.12, df = 1 (P = 0.73); I² = 0%								
Test for overall effect: Z = 3.22 (P = 0.001)								
Test for subgroup differences: Chi² = 0.12, df = 1 (P = 0.73), I² = 0%								



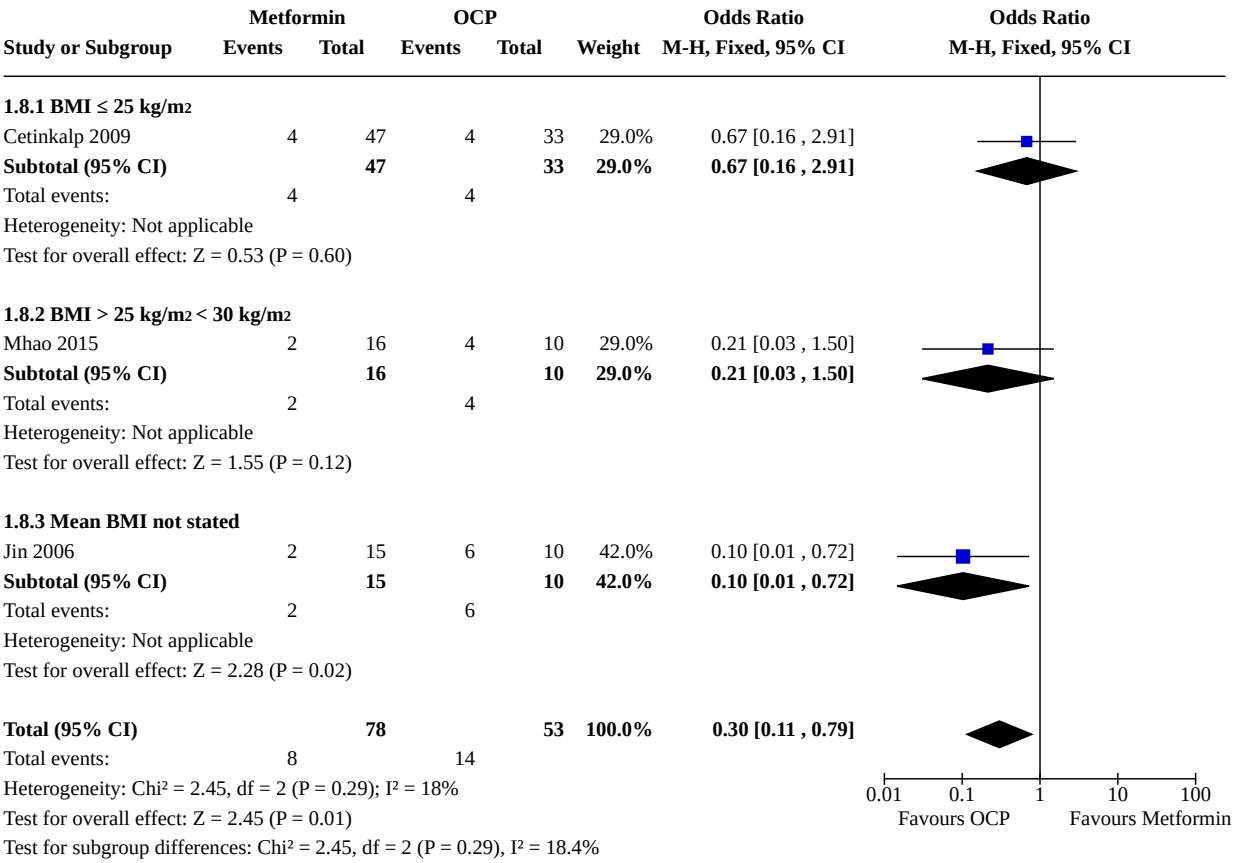
Analysis 1.6. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 6: Improved menstrual pattern (ie. an initiation of menses or cycle regularity)



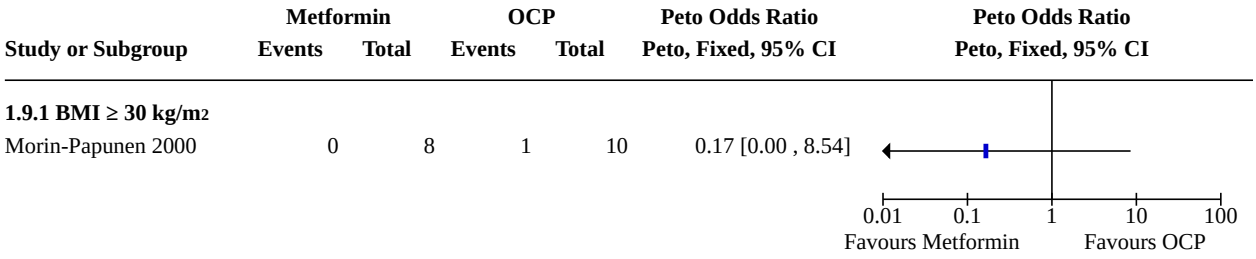
Analysis 1.7. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 7: Acne - Visual analogue scale



Analysis 1.8. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 8: Acne - Subjective improvement

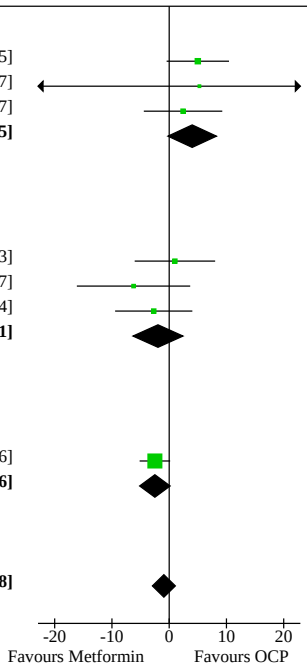


Analysis 1.9. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 9: Diagnosis of Type II diabetes mellitus

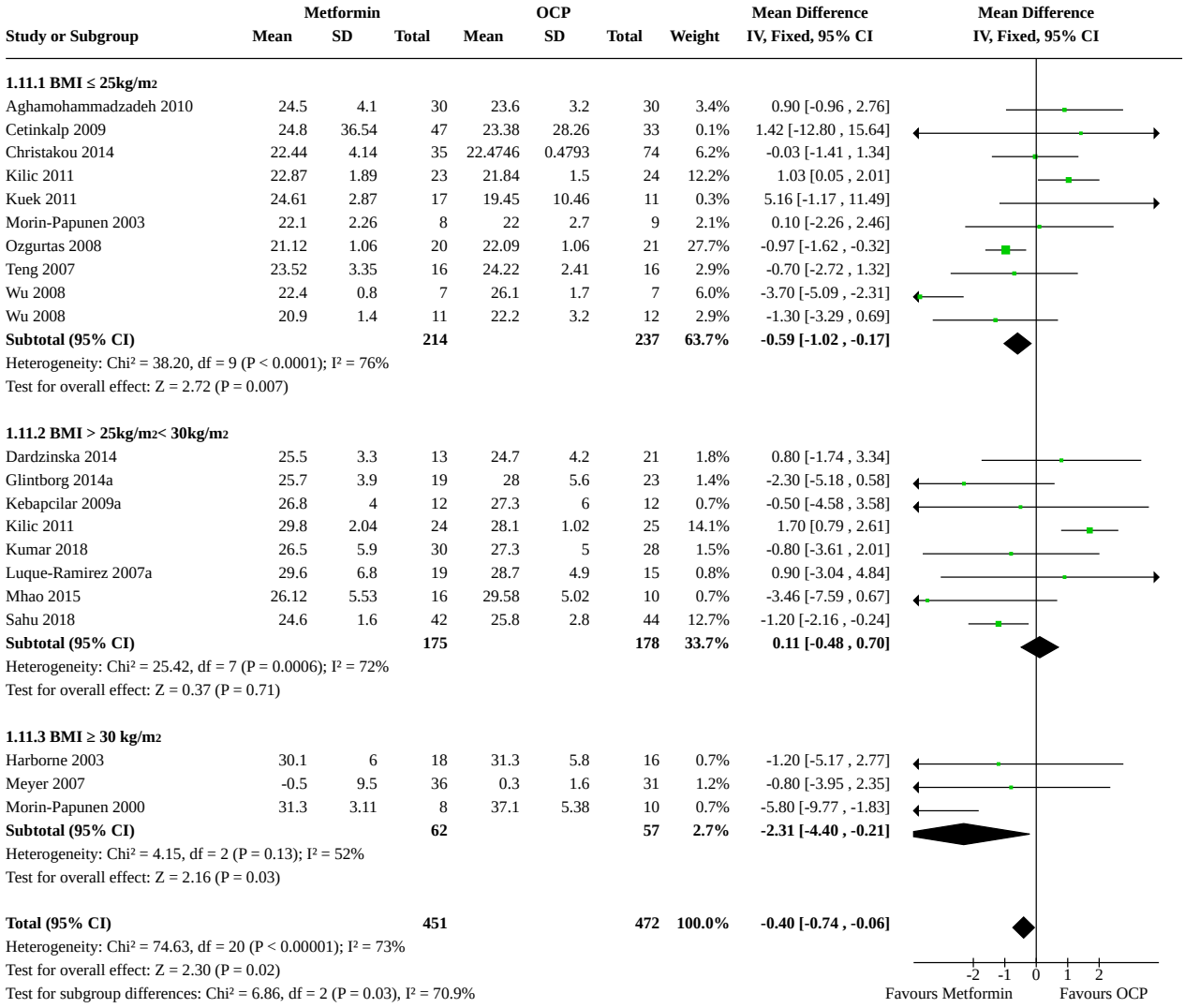


Analysis 1.10. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 10: Body weight (kg)

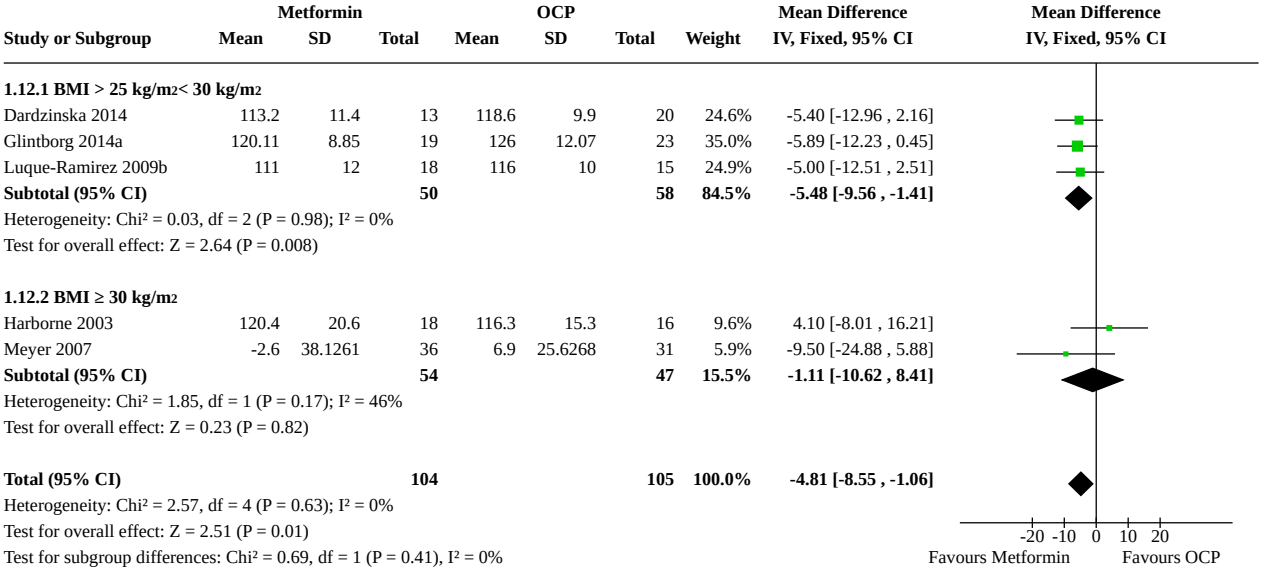
Study or Subgroup	Favours Metformin Mean	SD	Total	OCP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
1.10.1 BMI ≤ 25 kg/m2								
Aghamohammadzadeh 2010	62	12.2	30	57	9.1	30	13.5%	5.00 [-0.45 , 10.45]
Cetinkalp 2009	67.6	103.5	47	62.32	73.53	33	0.3%	5.28 [-33.51 , 44.07]
Kuek 2011	61.46	6.9	17	59.04	10.18	11	8.6%	2.42 [-4.43 , 9.27]
Subtotal (95% CI)			**94**			**74**	**22.4%**	**4.02 [-0.22 , 8.25]**
Heterogeneity: Chi² = 0.34, df = 2 (P = 0.84); I² = 0%								
Test for overall effect: Z = 1.86 (P = 0.06)								
1.10.2 BMI > 25 kg/m2 < 30 kg/m2								
Dardzinska 2014	71	9.6	13	70	11	21	8.1%	1.00 [-6.03 , 8.03]
Glintborg 2014a	73.87	13.49	19	80.1	19.14	23	4.1%	-6.23 [-16.13 , 3.67]
Kumar 2018	66.6	13.8	30	69.3	12.4	28	8.8%	-2.70 [-9.44 , 4.04]
Subtotal (95% CI)			**62**			**72**	**21.1%**	**-1.96 [-6.32 , 2.41]**
Heterogeneity: Chi² = 1.44, df = 2 (P = 0.49); I² = 0%								
Test for overall effect: Z = 0.88 (P = 0.38)								
1.10.3 BMI ≥ 30 kg/m2								
Moran 2010	-1.7	6.02	30	0.8	4.08	26	56.6%	-2.50 [-5.16 , 0.16]
Subtotal (95% CI)			**30**			**26**	**56.6%**	**-2.50 [-5.16 , 0.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.84 (P = 0.07)								
Total (95% CI)			**186**			**172**	**100.0%**	**-0.93 [-2.93 , 1.08]**
Heterogeneity: Chi² = 8.56, df = 6 (P = 0.20); I² = 30%								
Test for overall effect: Z = 0.91 (P = 0.36)								
Test for subgroup differences: Chi² = 6.78, df = 2 (P = 0.03), I² = 70.5%								



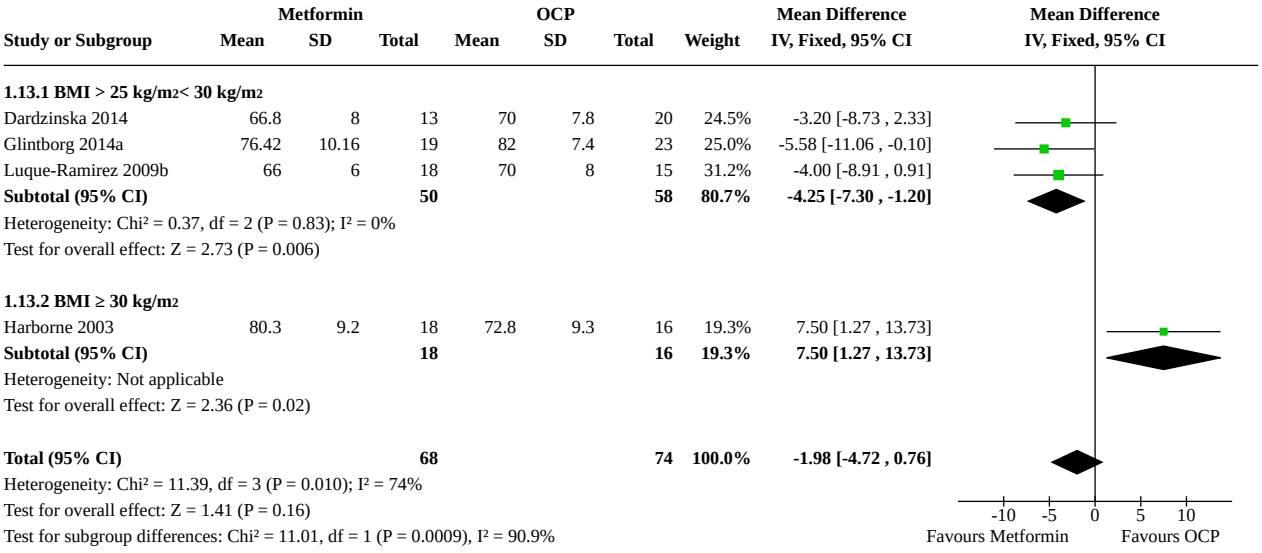
Analysis 1.11. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 11: Body Mass Index (kg/m^2)



Analysis 1.12. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 12: Blood pressure - systolic (mm Hg)



Analysis 1.13. Comparison 1: Adult - Metformin versus OCP (Clinical parameters), Outcome 13: Blood pressure - diastolic (mm Hg)

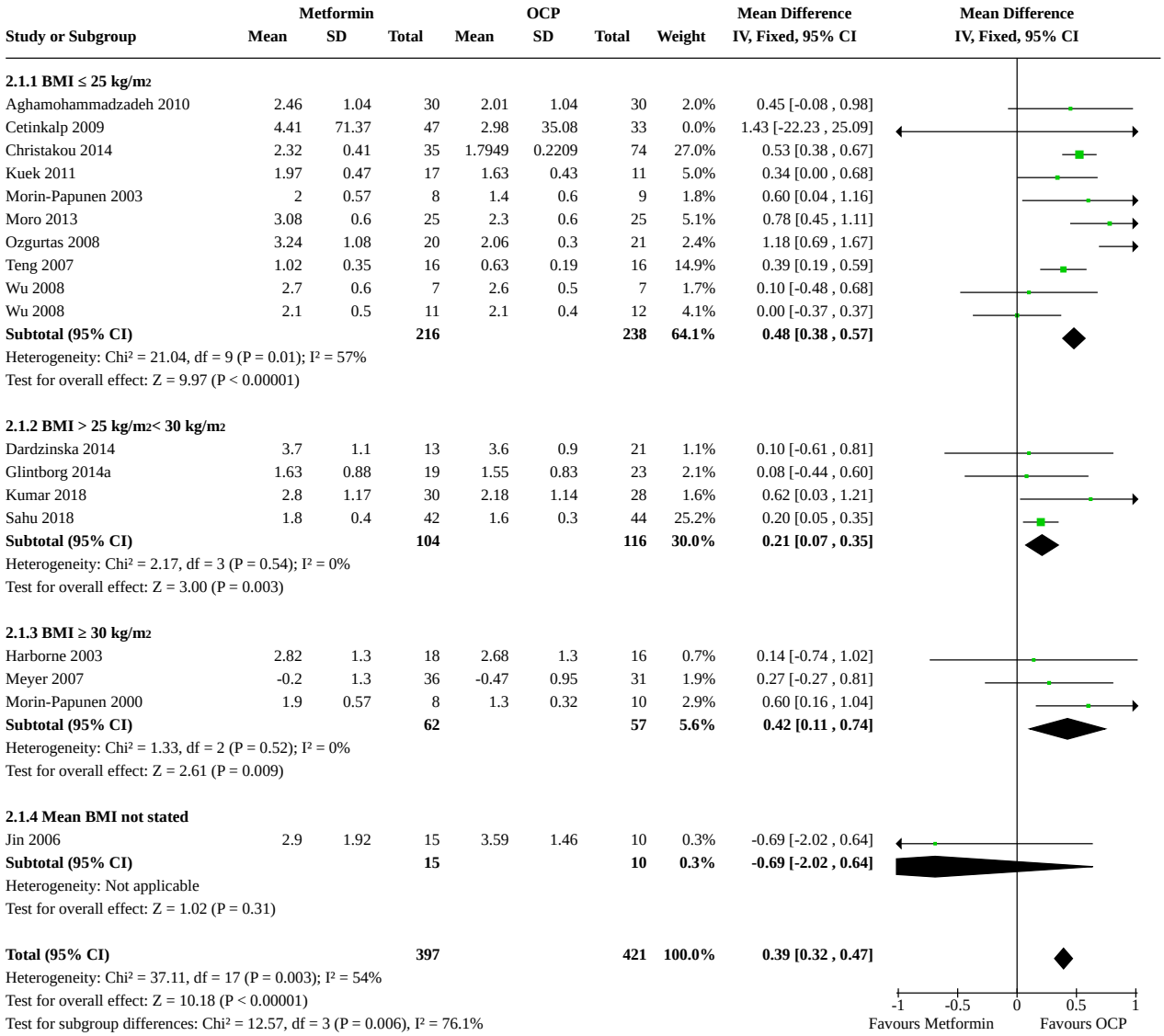


Comparison 2. Adult - Metformin versus OCP (Hormonal parameters)

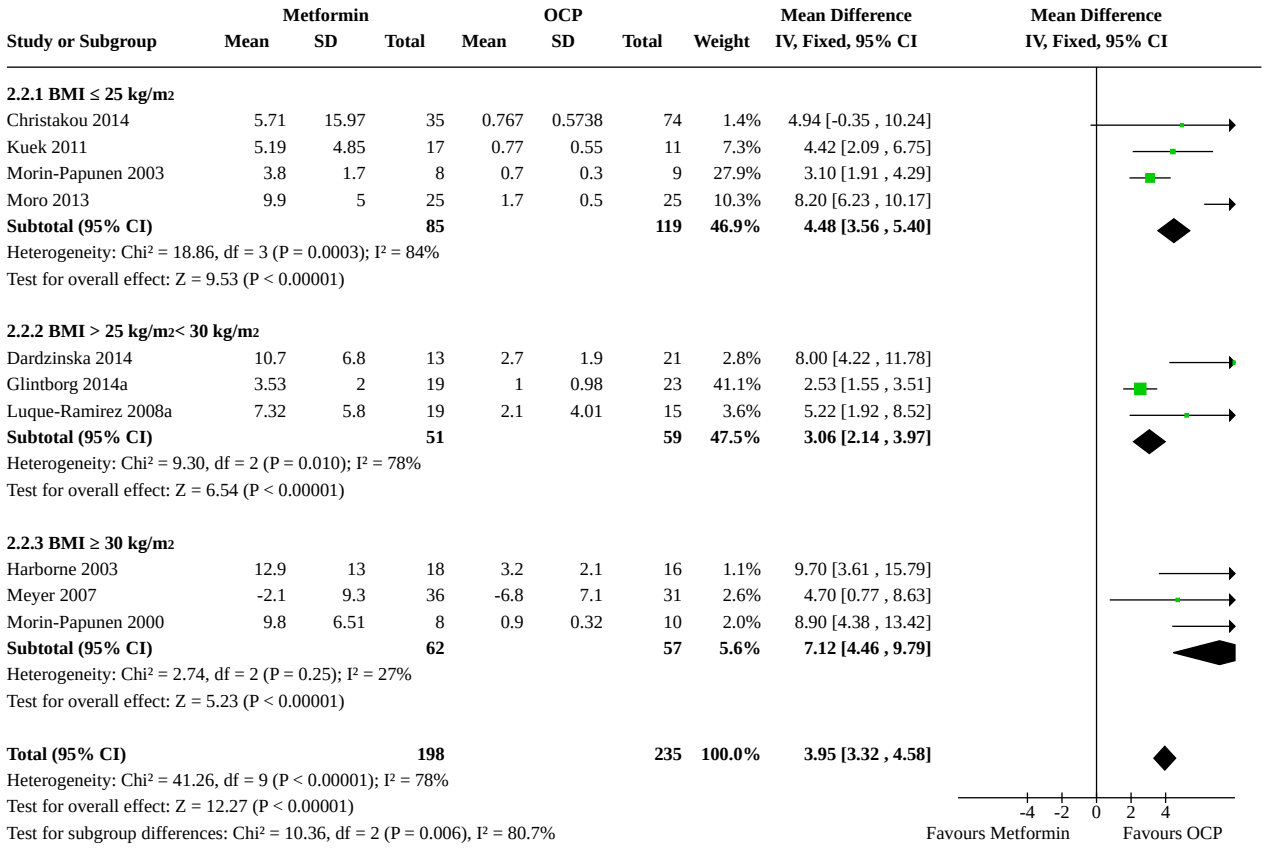
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Serum total testosterone (nmol/L)	17	818	Mean Difference (IV, Fixed, 95% CI)	0.39 [0.32, 0.47]
2.1.1 BMI ≤ 25 kg/m^2	9	454	Mean Difference (IV, Fixed, 95% CI)	0.48 [0.38, 0.57]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1.2 BMI > 25 kg/m^2< 30 kg/m^2	4	220	Mean Difference (IV, Fixed, 95% CI)	0.21 [0.07, 0.35]
2.1.3 BMI ≥ 30 kg/m^2	3	119	Mean Difference (IV, Fixed, 95% CI)	0.42 [0.11, 0.74]
2.1.4 Mean BMI not stated	1	25	Mean Difference (IV, Fixed, 95% CI)	-0.69 [-2.02, 0.64]
2.2 Free androgen index (FAI) (%)	10	433	Mean Difference (IV, Fixed, 95% CI)	3.95 [3.32, 4.58]
2.2.1 BMI ≤ 25 kg/m^2	4	204	Mean Difference (IV, Fixed, 95% CI)	4.48 [3.56, 5.40]
2.2.2 BMI > 25 kg/m^2< 30 kg/m^2	3	110	Mean Difference (IV, Fixed, 95% CI)	3.06 [2.14, 3.97]
2.2.3 BMI ≥ 30 kg/m^2	3	119	Mean Difference (IV, Fixed, 95% CI)	7.12 [4.46, 9.79]

Analysis 2.1. Comparison 2: Adult - Metformin versus OCP (Hormonal parameters), Outcome 1: Serum total testosterone (nmol/L)



Analysis 2.2. Comparison 2: Adult - Metformin versus OCP (Hormonal parameters), Outcome 2: Free androgen index (FAI) (%)

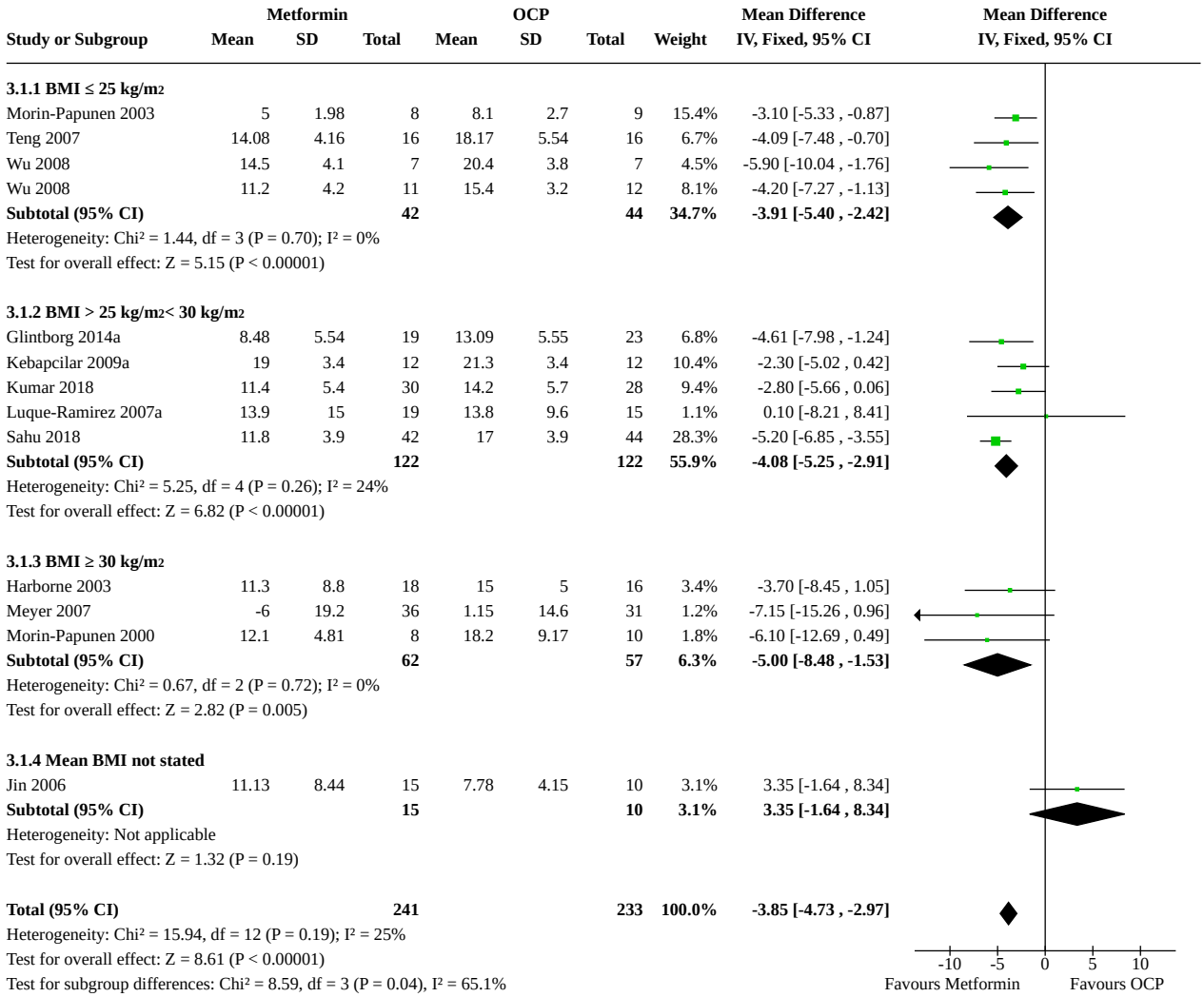


Comparison 3. Adult - Metformin versus OCP (Metabolic parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Fasting insulin (mIU/L)	12	474	Mean Difference (IV, Fixed, 95% CI)	-3.85 [-4.73, -2.97]
3.1.1 BMI ≤ 25 kg/m^2	3	86	Mean Difference (IV, Fixed, 95% CI)	-3.91 [-5.40, -2.42]
3.1.2 BMI > 25 kg/m^2< 30 kg/m^2	5	244	Mean Difference (IV, Fixed, 95% CI)	-4.08 [-5.25, -2.91]
3.1.3 BMI ≥ 30 kg/m^2	3	119	Mean Difference (IV, Fixed, 95% CI)	-5.00 [-8.48, -1.53]
3.1.4 Mean BMI not stated	1	25	Mean Difference (IV, Fixed, 95% CI)	3.35 [-1.64, 8.34]
3.2 Fasting glucose (mmol/L)	12	519	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.22, -0.07]
3.2.1 BMI ≤ 25 kg/m^2	5	191	Mean Difference (IV, Fixed, 95% CI)	-0.09 [-0.23, 0.06]
3.2.2 BMI > 25 kg/m^2< 30 kg/m^2	4	220	Mean Difference (IV, Fixed, 95% CI)	-0.21 [-0.31, -0.10]

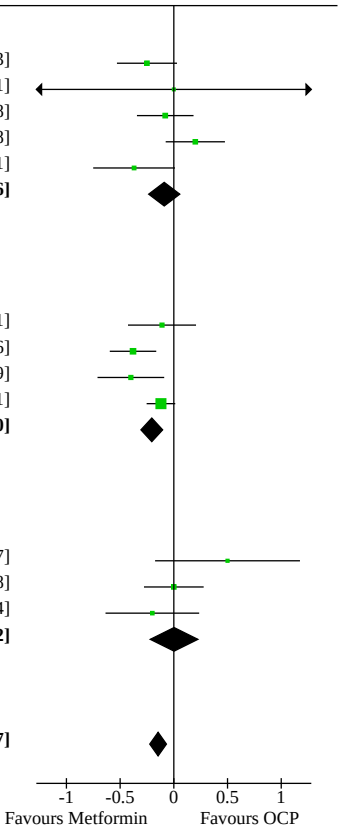
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.2.3 BMI ≥ 30 kg/m^2	3	108	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.22, 0.22]
3.3 Total Cholesterol (mmol/L)	13	610	Mean Difference (IV, Fixed, 95% CI)	-0.28 [-0.40, -0.16]
3.3.1 BMI ≤ 25 kg/m^2	4	206	Mean Difference (IV, Fixed, 95% CI)	-0.77 [-1.00, -0.53]
3.3.2 BMI > 25 kg/m^2< 30 kg/m^2	7	303	Mean Difference (IV, Fixed, 95% CI)	-0.14 [-0.30, 0.01]
3.3.3 BMI ≥ 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.32, 0.28]
3.4 HDL Cholesterol (mmol/L)	13	610	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.01, 0.06]
3.4.1 BMI ≤ 25 kg/m^2	4	206	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.10, 0.07]
3.4.2 BMI > 25 kg/m^2< 30 kg/m^2	7	303	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.02, 0.06]
3.4.3 BMI ≥ 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	0.20 [0.05, 0.35]
3.5 LDL Cholesterol (mmol/L)	13	610	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.12, 0.02]
3.5.1 BMI ≤ 25 kg/m^2	4	206	Mean Difference (IV, Fixed, 95% CI)	-0.39 [-0.54, -0.23]
3.5.2 BMI > 25 kg/m^2< 30 kg/m^2	7	303	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.06, 0.10]
3.5.3 BMI ≥ 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	0.35 [0.02, 0.67]
3.6 Triglycerides (mmol/L)	13	610	Mean Difference (IV, Fixed, 95% CI)	-0.07 [-0.12, -0.02]
3.6.1 BMI ≤ 25 kg/m^2	4	206	Mean Difference (IV, Fixed, 95% CI)	-0.45 [-0.61, -0.30]
3.6.2 BMI > 25 kg/m^2< 30 kg/m^2	7	303	Mean Difference (IV, Fixed, 95% CI)	-0.01 [-0.07, 0.05]
3.6.3 BMI ≥ 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	-0.31 [-0.64, 0.01]

Analysis 3.1. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 1: Fasting insulin (mIU/L)



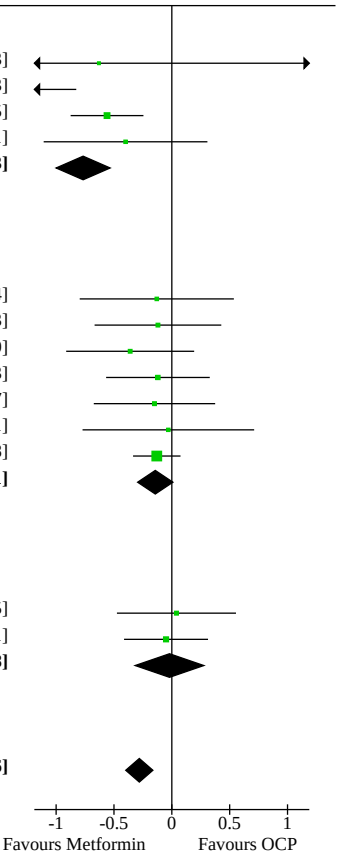
Analysis 3.2. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 2: Fasting glucose (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
3.2.1 BMI ≤ 25 kg/m2								
Bodur 2018	4.35	0.34	17	4.6	0.48	17	7.7%	-0.25 [-0.53 , 0.03]
Cetinkalp 2009	5.04	3.05	47	5.04	2.86	33	0.4%	0.00 [-1.31 , 1.31]
Kuek 2011	4.57	0.45	17	4.65	0.26	11	8.7%	-0.08 [-0.34 , 0.18]
Morin-Papunen 2003	5	0.28	8	4.8	0.3	9	7.9%	0.20 [-0.08 , 0.48]
Teng 2007	4.92	0.66	16	5.29	0.41	16	4.1%	-0.37 [-0.75 , 0.01]
Subtotal (95% CI)			**105**			**86**	**28.8%**	**-0.09 [-0.23 , 0.06]**
Heterogeneity: Chi² = 7.61, df = 4 (P = 0.11); I² = 47%								
Test for overall effect: Z = 1.21 (P = 0.23)								
3.2.2 BMI > 25 kg/m2< 30 kg/m2								
Glintborg 2014a	5.09	0.53	19	5.2	0.51	23	6.0%	-0.11 [-0.43 , 0.21]
Kumar 2018	4.86	0.45	30	5.24	0.39	28	12.8%	-0.38 [-0.60 , -0.16]
Luque-Ramirez 2007a	4.9	0.4	19	5.3	0.5	15	6.2%	-0.40 [-0.71 , -0.09]
Sahu 2018	4.75	0.3	42	4.87	0.33	44	33.9%	-0.12 [-0.25 , 0.01]
Subtotal (95% CI)			**110**			**110**	**59.0%**	**-0.21 [-0.31 , -0.10]**
Heterogeneity: Chi² = 5.94, df = 3 (P = 0.11); I² = 49%								
Test for overall effect: Z = 3.98 (P < 0.0001)								
3.2.3 BMI ≥ 30 kg/m2								
Harborne 2003	5.3	1.4	18	4.8	0.39	16	1.3%	0.50 [-0.17 , 1.17]
Moran 2010	-0.1	0.55	30	-0.1	0.51	26	7.8%	0.00 [-0.28 , 0.28]
Morin-Papunen 2000	4.9	0.28	8	5.1	0.63	10	3.2%	-0.20 [-0.64 , 0.24]
Subtotal (95% CI)			**56**			**52**	**12.3%**	**0.00 [-0.22 , 0.22]**
Heterogeneity: Chi² = 2.92, df = 2 (P = 0.23); I² = 31%								
Test for overall effect: Z = 0.02 (P = 0.98)								
Total (95% CI)			**271**			**248**	**100.0%**	**-0.15 [-0.22 , -0.07]**
Heterogeneity: Chi² = 20.10, df = 11 (P = 0.04); I² = 45%								
Test for overall effect: Z = 3.70 (P = 0.0002)								
Test for subgroup differences: Chi² = 3.64, df = 2 (P = 0.16), I² = 45.0%								



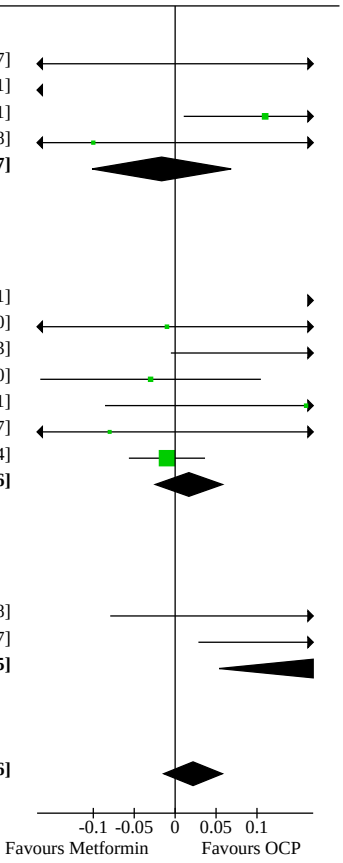
Analysis 3.3. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 3: Total Cholesterol (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
3.3.1 BMI ≤ 25 kg/m2								
Cetinkalp 2009	4.8	6.9	47	5.43	5.67	33	0.2%	-0.63 [-3.39 , 2.13]
Moro 2013	4.6	0.65	25	5.83	0.8	25	8.4%	-1.23 [-1.63 , -0.83]
Ozgurtas 2008	3.76	0.56	20	4.32	0.46	21	13.9%	-0.56 [-0.87 , -0.25]
Rautio 2005	5	0.8	16	5.4	1.31	19	2.7%	-0.40 [-1.11 , 0.31]
Subtotal (95% CI)			**108**			**98**	**25.2%**	**-0.77 [-1.00 , -0.53]**
Heterogeneity: Chi² = 7.75, df = 3 (P = 0.05); I² = 61%								
Test for overall effect: Z = 6.44 (P < 0.00001)								
3.3.2 BMI > 25 kg/m2< 30 kg/m2								
Dardzinska 2014	5.02	1.04	12	5.15	0.73	21	3.1%	-0.13 [-0.80 , 0.54]
Glintborg 2014a	4.68	0.99	19	4.8	0.78	23	4.6%	-0.12 [-0.67 , 0.43]
Kebapcilar 2009a	3.59	0.73	12	3.95	0.65	12	4.5%	-0.36 [-0.91 , 0.19]
Kumar 2018	4.13	0.9	30	4.25	0.84	28	6.8%	-0.12 [-0.57 , 0.33]
Luque-Ramirez 2007a	4.15	0.89	19	4.3	0.67	15	5.0%	-0.15 [-0.67 , 0.37]
Mhao 2015	3.74	0.85	16	3.77	0.99	10	2.5%	-0.03 [-0.77 , 0.71]
Sahu 2018	4.37	0.47	42	4.5	0.5	44	32.7%	-0.13 [-0.34 , 0.08]
Subtotal (95% CI)			**150**			**153**	**59.2%**	**-0.14 [-0.30 , 0.01]**
Heterogeneity: Chi² = 0.72, df = 6 (P = 0.99); I² = 0%								
Test for overall effect: Z = 1.84 (P = 0.07)								
3.3.3 BMI ≥ 30 kg/m2								
Harborne 2003	4.79	0.83	18	4.75	0.7	16	5.2%	0.04 [-0.47 , 0.55]
Meyer 2007	-0.17	0.7	36	-0.12	0.8	31	10.4%	-0.05 [-0.41 , 0.31]
Subtotal (95% CI)			**54**			**47**	**15.6%**	**-0.02 [-0.32 , 0.28]**
Heterogeneity: Chi² = 0.08, df = 1 (P = 0.78); I² = 0%								
Test for overall effect: Z = 0.13 (P = 0.89)								
Total (95% CI)			**312**			**298**	**100.0%**	**-0.28 [-0.40 , -0.16]**
Heterogeneity: Chi² = 31.30, df = 12 (P = 0.002); I² = 62%								
Test for overall effect: Z = 4.70 (P < 0.00001)								
Test for subgroup differences: Chi² = 22.76, df = 2 (P < 0.0001), I² = 91.2%								



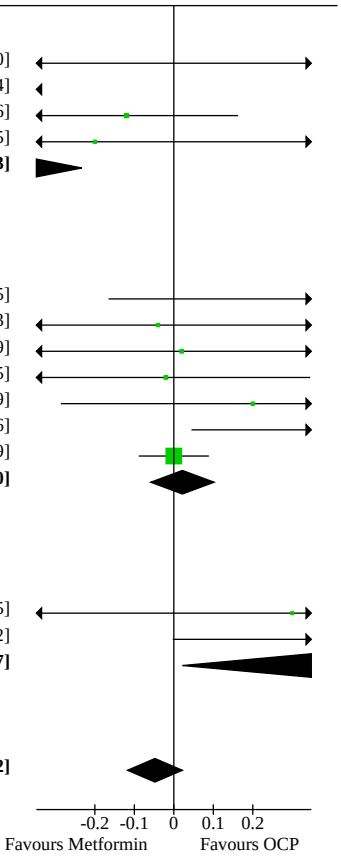
Analysis 3.4. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 4: HDL Cholesterol (mmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metfomin Mean	Metfomin SD	Metfomin Total	Weight	Mean Difference IV, Fixed, 95% CI
3.4.1 BMI ≤ 25 kg/m2								
Cetinkalp 2009	1.67	2.18	33	1.47	2.68	47	0.1%	0.20 [-0.87 , 1.27]
Moro 2013	1.39	0.45	25	1.91	0.27	25	3.0%	-0.52 [-0.73 , -0.31]
Ozgurtas 2008	1.37	0.11	21	1.26	0.2	20	13.0%	0.11 [0.01 , 0.21]
Rautio 2005	1.5	0.44	19	1.6	0.4	16	1.7%	-0.10 [-0.38 , 0.18]
Subtotal (95% CI)			**98**			**108**	**17.8%**	**-0.02 [-0.10 , 0.07]**
Heterogeneity: Chi² = 29.73, df = 3 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 0.38 (P = 0.70)								
3.4.2 BMI > 25 kg/m2< 30 kg/m2								
Dardzinska 2014	1.76	0.39	21	1.27	0.25	12	2.7%	0.49 [0.27 , 0.71]
Glintborg 2014a	1.5	0.31	23	1.51	0.38	19	2.8%	-0.01 [-0.22 , 0.20]
Kebapcilar 2009a	1.72	0.39	12	1.46	0.26	12	1.8%	0.26 [-0.01 , 0.53]
Kumar 2018	1.05	0.22	28	1.08	0.3	30	7.1%	-0.03 [-0.16 , 0.10]
Luque-Ramirez 2007a	1.27	0.34	15	1.11	0.39	19	2.1%	0.16 [-0.09 , 0.41]
Mhao 2015	0.95	1.2	10	1.03	0.19	16	0.2%	-0.08 [-0.83 , 0.67]
Sahu 2018	1.38	0.11	44	1.39	0.11	42	59.4%	-0.01 [-0.06 , 0.04]
Subtotal (95% CI)			**153**			**150**	**76.2%**	**0.02 [-0.02 , 0.06]**
Heterogeneity: Chi² = 24.38, df = 6 (P = 0.0004); I² = 75%								
Test for overall effect: Z = 0.80 (P = 0.42)								
3.4.3 BMI ≥ 30 kg/m2								
Harborne 2003	1.51	0.51	16	1.31	0.27	18	1.6%	0.20 [-0.08 , 0.48]
Meyer 2007	0.1	0.4	31	-0.1	0.3	36	4.4%	0.20 [0.03 , 0.37]
Subtotal (95% CI)			**47**			**54**	**6.0%**	**0.20 [0.05 , 0.35]**
Heterogeneity: Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 2.68 (P = 0.007)								
Total (95% CI)			**298**			**312**	**100.0%**	**0.02 [-0.01 , 0.06]**
Heterogeneity: Chi² = 60.66, df = 12 (P < 0.00001); I² = 80%								
Test for overall effect: Z = 1.20 (P = 0.23)								
Test for subgroup differences: Chi² = 6.55, df = 2 (P = 0.04), I² = 69.5%								

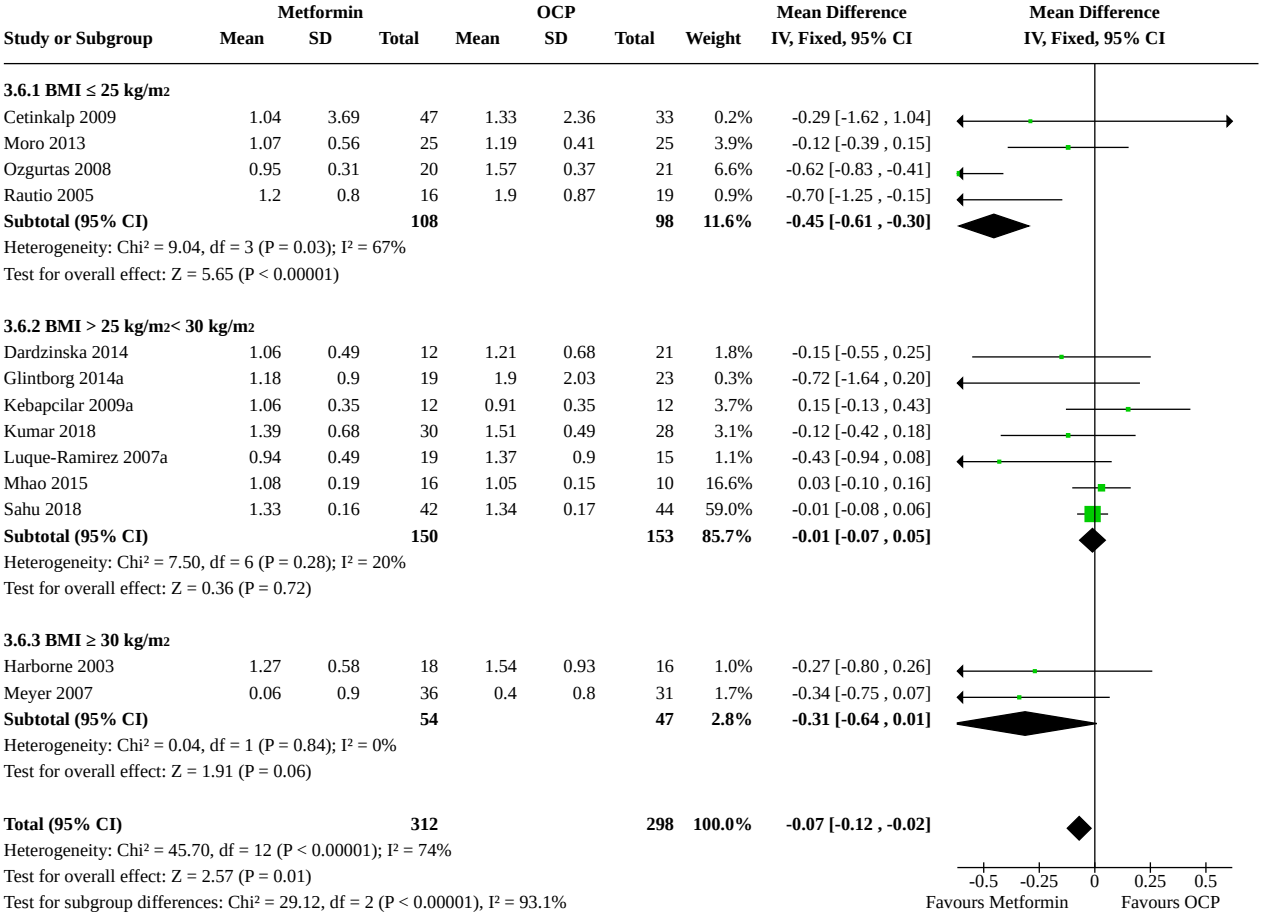


Analysis 3.5. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 5: LDL Cholesterol (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
3.5.1 BMI ≤ 25 kg/m2								
Cetinkalp 2009	2.8	6.08	47	3.22	4.51	33	0.1%	-0.42 [-2.74 , 1.90]
Moro 2013	2.46	0.31	25	2.99	0.38	25	13.0%	-0.53 [-0.72 , -0.34]
Ozgurtas 2008	2.12	0.49	20	2.24	0.43	21	6.0%	-0.12 [-0.40 , 0.16]
Rautio 2005	2.8	0.8	16	3	0.87	19	1.6%	-0.20 [-0.75 , 0.35]
Subtotal (95% CI)			**108**			**98**	**20.7%**	**-0.39 [-0.54 , -0.23]**
Heterogeneity: Chi² = 5.99, df = 3 (P = 0.11); I² = 50%								
Test for overall effect: Z = 4.95 (P < 0.00001)								
3.5.2 BMI > 25 kg/m2< 30 kg/m2								
Dardzinska 2014	3.27	0.93	12	2.83	0.7	21	1.3%	0.44 [-0.17 , 1.05]
Glintborg 2014a	2.76	0.82	19	2.8	0.72	23	2.2%	-0.04 [-0.51 , 0.43]
Kebapcilar 2009a	2	0.5	12	1.98	0.42	12	3.5%	0.02 [-0.35 , 0.39]
Kumar 2018	2.45	0.74	30	2.47	0.68	28	3.6%	-0.02 [-0.39 , 0.35]
Luque-Ramirez 2007a	2.64	0.87	19	2.44	0.57	15	2.0%	0.20 [-0.29 , 0.69]
Mhao 2015	3.28	1.06	16	2.53	0.77	10	1.0%	0.75 [0.04 , 1.46]
Sahu 2018	2.31	0.2	42	2.31	0.22	44	61.1%	0.00 [-0.09 , 0.09]
Subtotal (95% CI)			**150**			**153**	**74.7%**	**0.02 [-0.06 , 0.10]**
Heterogeneity: Chi² = 6.79, df = 6 (P = 0.34); I² = 12%								
Test for overall effect: Z = 0.53 (P = 0.60)								
3.5.3 BMI ≥ 30 kg/m2								
Harborne 2003	2.85	1	18	2.55	1.2	16	0.9%	0.30 [-0.45 , 1.05]
Meyer 2007	-0.04	0.7	36	-0.4	0.8	31	3.7%	0.36 [-0.00 , 0.72]
Subtotal (95% CI)			**54**			**47**	**4.5%**	**0.35 [0.02 , 0.67]**
Heterogeneity: Chi² = 0.02, df = 1 (P = 0.89); I² = 0%								
Test for overall effect: Z = 2.09 (P = 0.04)								
Total (95% CI)			**312**			**298**	**100.0%**	**-0.05 [-0.12 , 0.02]**
Heterogeneity: Chi² = 40.17, df = 12 (P < 0.0001); I² = 70%								
Test for overall effect: Z = 1.35 (P = 0.18)								
Test for subgroup differences: Chi² = 27.37, df = 2 (P < 0.00001), I² = 92.7%								



Analysis 3.6. Comparison 3: Adult - Metformin versus OCP (Metabolic parameters), Outcome 6: Triglycerides (mmol/L)

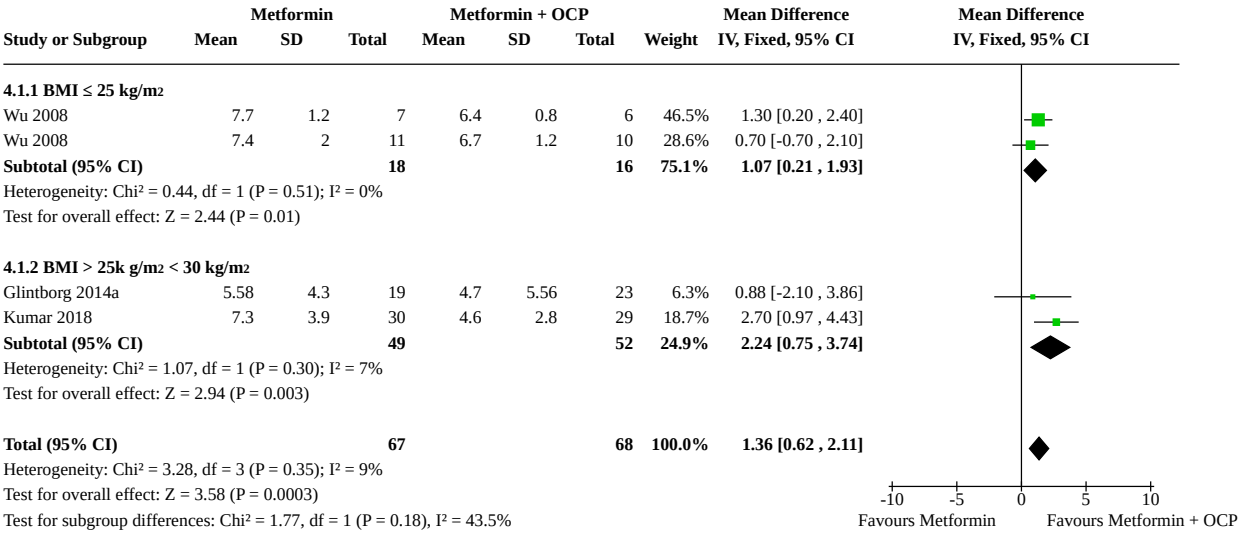


Comparison 4. Adult - Metformin versus Metformin combined with OCP (Clinical parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Hirsutism - Clinical F-G score	3	135	Mean Difference (IV, Fixed, 95% CI)	1.36 [0.62, 2.11]
4.1.1 BMI ≤ 25 kg/m^2	1	34	Mean Difference (IV, Fixed, 95% CI)	1.07 [0.21, 1.93]
4.1.2 BMI > 25k g/m^2 < 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	2.24 [0.75, 3.74]
4.2 Adverse events -severe	3		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.2.1 Gastro-intestinal	3	171	Odds Ratio (M-H, Fixed, 95% CI)	0.74 [0.21, 2.53]
4.2.2 Others	2	109	Odds Ratio (M-H, Fixed, 95% CI)	0.56 [0.11, 2.82]
4.3 Body weight (kg)	2	101	Mean Difference (IV, Fixed, 95% CI)	-5.39 [-10.70, -0.08]

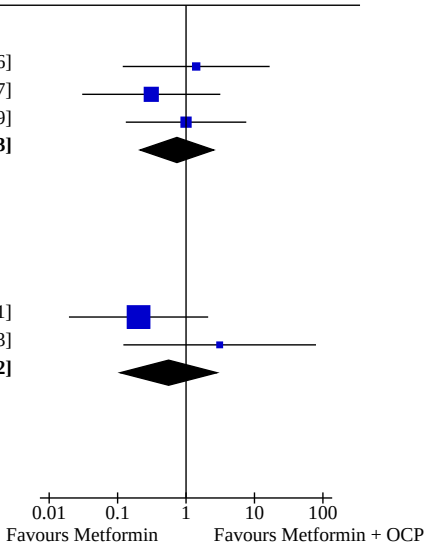
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.3.1 BMI > 25 kg/m2 < 30 kg/m^2	2	101	Mean Difference (IV, Fixed, 95% CI)	-5.39 [-10.70, -0.08]
4.4 Body Mass Index (kg/m2)	5	199	Mean Difference (IV, Fixed, 95% CI)	-1.47 [-2.27, -0.66]
4.4.1 BMI ≤ 25 kg/m^2	2	74	Mean Difference (IV, Fixed, 95% CI)	-1.50 [-2.44, -0.55]
4.4.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	125	Mean Difference (IV, Fixed, 95% CI)	-1.38 [-2.94, 0.17]
4.5 Blood pressure - systolic (mm Hg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
4.5.1 BMI > 25 kg/m2 < 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
4.6 Blood pressure - diastolic (mm Hg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
4.6.1 BMI > 25 kg/m2 < 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Analysis 4.1. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 1: Hirsutism - Clinical F-G score



Analysis 4.2. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 2: Adverse events -severe

Study or Subgroup	Metformin Events	Metformin Total	Metformin + OCP Events	Metformin + OCP Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.2.1 Gastro-intestinal						
Bodur 2018	2	29	1	20	18.8%	1.41 [0.12 , 16.66]
Glintborg 2014a	1	30	3	30	49.4%	0.31 [0.03 , 3.17]
Moro 2013	2	31	2	31	31.9%	1.00 [0.13 , 7.59]
Subtotal (95% CI)		**90**		**81**	**100.0%**	**0.74 [0.21 , 2.53]**
Total events:	5		6			
Heterogeneity: Chi² = 0.88, df = 2 (P = 0.64); I² = 0%						
Test for overall effect: Z = 0.49 (P = 0.63)						
4.2.2 Others						
Bodur 2018 (1)	1	29	3	20	87.8%	0.20 [0.02 , 2.11]
Glintborg 2014a (2)	1	30	0	30	12.2%	3.10 [0.12 , 79.23]
Subtotal (95% CI)		**59**		**50**	**100.0%**	**0.56 [0.11 , 2.82]**
Total events:	2		3			
Heterogeneity: Chi² = 1.80, df = 1 (P = 0.18); I² = 44%						
Test for overall effect: Z = 0.71 (P = 0.48)						
Test for subgroup differences: Chi² = 0.07, df = 1 (P = 0.79), I² = 0%						



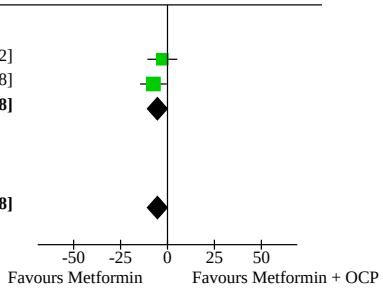
Footnotes

(1) OCP + MET dizziness/ headache/ MET dizziness

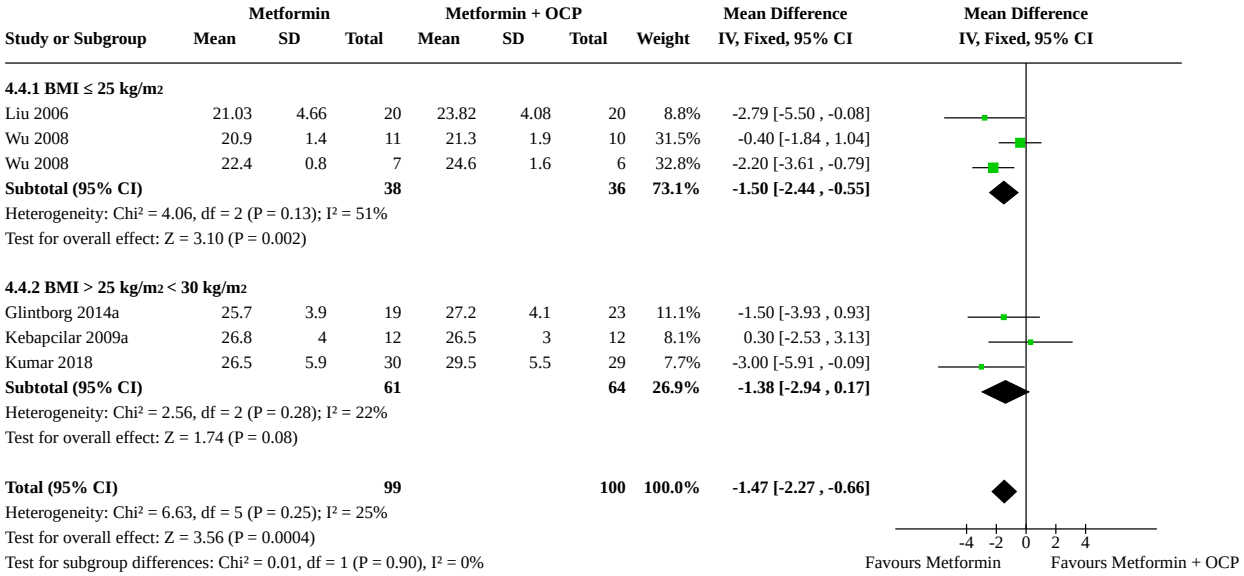
(2) MET group: depression

Analysis 4.3. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 3: Body weight (kg)

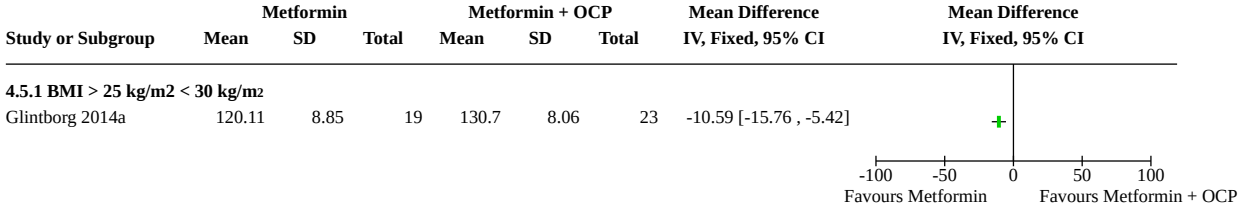
Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
4.3.1 BMI > 25 kg/m2 < 30 kg/m2								
Glintborg 2014a	73.87	13.49	19	76.62	12.64	23	44.4%	-2.75 [-10.72 , 5.22]
Kumar 2018	66.6	13.8	30	74.1	14.1	29	55.6%	-7.50 [-14.62 , -0.38]
Subtotal (95% CI)			**49**			**52**	**100.0%**	**-5.39 [-10.70 , -0.08]**
Heterogeneity: Chi² = 0.76, df = 1 (P = 0.38); I² = 0%								
Test for overall effect: Z = 1.99 (P = 0.05)								
Total (95% CI)			**49**			**52**	**100.0%**	**-5.39 [-10.70 , -0.08]**
Heterogeneity: Chi² = 0.76, df = 1 (P = 0.38); I² = 0%								
Test for overall effect: Z = 1.99 (P = 0.05)								
Test for subgroup differences: Not applicable								



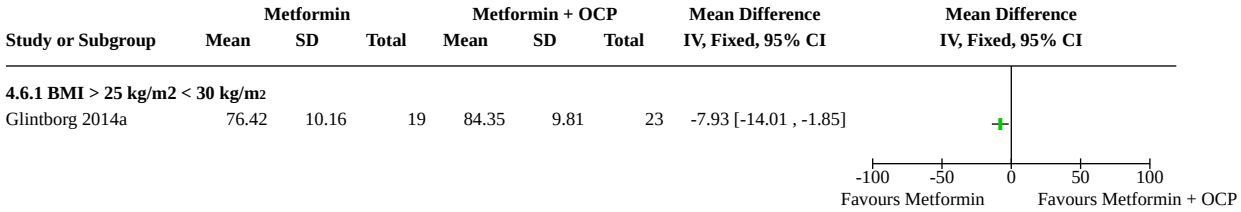
Analysis 4.4. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 4: Body Mass Index (kg/m2)



Analysis 4.5. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 5: Blood pressure - systolic (mm Hg)



Analysis 4.6. Comparison 4: Adult - Metformin versus Metformin combined with OCP (Clinical parameters), Outcome 6: Blood pressure - diastolic (mm Hg)

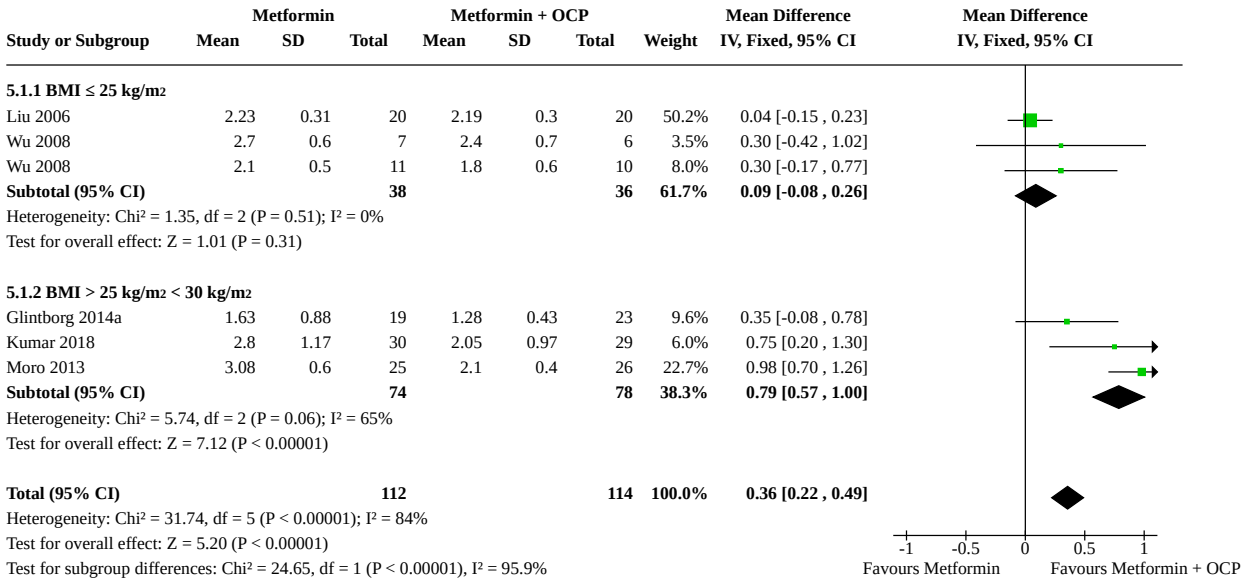


Comparison 5. Adult - Metformin versus Metformin combined with OCP (Hormonal parameters)

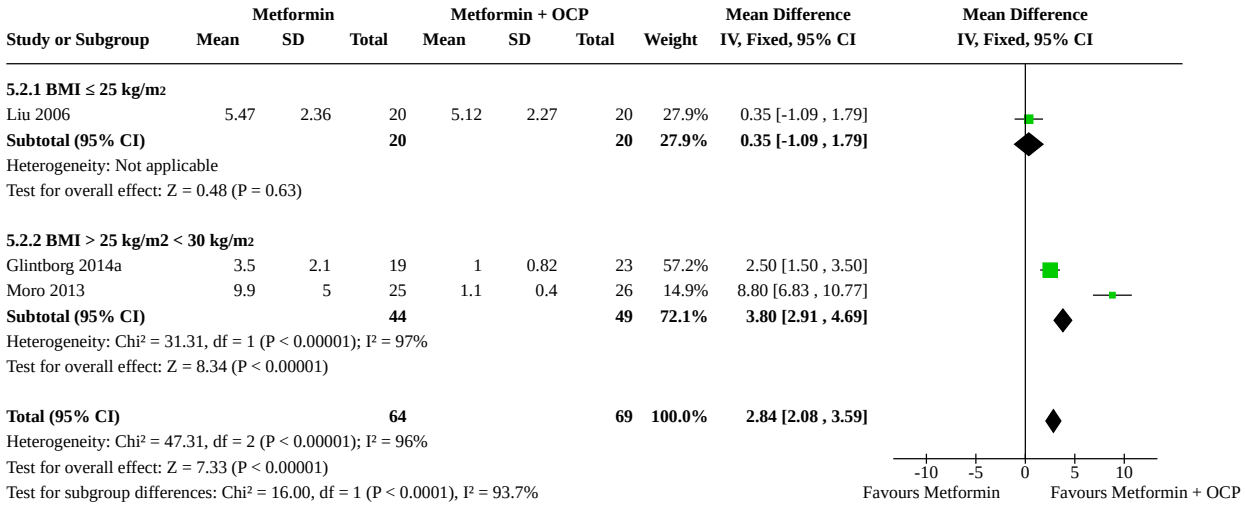
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Serum total testosterone (nmol/L)	5	226	Mean Difference (IV, Fixed, 95% CI)	0.36 [0.22, 0.49]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1.1 BMI ≤ 25 kg/m^2	2	74	Mean Difference (IV, Fixed, 95% CI)	0.09 [-0.08, 0.26]
5.1.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	152	Mean Difference (IV, Fixed, 95% CI)	0.79 [0.57, 1.00]
5.2 FAI (%)	3	133	Mean Difference (IV, Fixed, 95% CI)	2.84 [2.08, 3.59]
5.2.1 BMI ≤ 25 kg/m^2	1	40	Mean Difference (IV, Fixed, 95% CI)	0.35 [-1.09, 1.79]
5.2.2 BMI > 25 kg/m2 < 30 kg/m^2	2	93	Mean Difference (IV, Fixed, 95% CI)	3.80 [2.91, 4.69]

Analysis 5.1. Comparison 5: Adult - Metformin versus Metformin combined with OCP (Hormonal parameters), Outcome 1: Serum total testosterone (nmol/L)



Analysis 5.2. Comparison 5: Adult - Metformin versus Metformin combined with OCP (Hormonal parameters), Outcome 2: FAI (%)



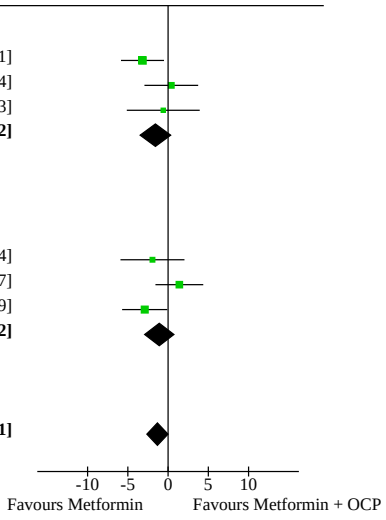
Comparison 6. Adult - Metformin versus Metformin combined with OCP (Metabolic parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Fasting insulin (mIU/L)	5	199	Mean Difference (IV, Fixed, 95% CI)	-1.32 [-2.63, -0.01]
6.1.1 BMI ≤ 25 kg/m^2	2	74	Mean Difference (IV, Fixed, 95% CI)	-1.58 [-3.47, 0.32]
6.1.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	125	Mean Difference (IV, Fixed, 95% CI)	-1.09 [-2.91, 0.72]
6.2 Fasting glucose (mmol/L)	4	170	Mean Difference (IV, Fixed, 95% CI)	-0.21 [-0.37, -0.06]
6.2.1 BMI ≤ 25 kg/m^2	1	40	Mean Difference (IV, Fixed, 95% CI)	-0.25 [-0.54, 0.04]
6.2.2 BMI > 25 kg/m^2 < 30 kg/m^2	3	130	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.38, -0.01]
6.3 Total Cholesterol (mmol/L)	5	216	Mean Difference (IV, Random, 95% CI)	-0.54 [-0.97, -0.11]
6.3.1 BMI ≤ 25 kg/m^2	1	40	Mean Difference (IV, Random, 95% CI)	-0.65 [-1.56, 0.26]
6.3.2 BMI > 25 kg/m2 < 30 kg/m^2	4	176	Mean Difference (IV, Random, 95% CI)	-0.52 [-1.02, -0.02]
6.4 HDL Cholesterol (mmol/L)	5	216	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.08, 0.09]
6.4.1 BMI ≤ 25 kg/m^2	1	40	Mean Difference (IV, Fixed, 95% CI)	-0.64 [-0.99, -0.29]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.4.2 BMI > 25 kg/m2 < 30 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.04, 0.14]
6.5 LDL Cholesterol (mmol/L)	4	176	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.32, 0.06]
6.5.1 BMI > 25 kg/m2 < 30 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.32, 0.06]
6.6 Triglycerides (mmol/L)	5	216	Mean Difference (IV, Fixed, 95% CI)	-0.22 [-0.37, -0.07]
6.6.1 BMI ≤ 25 kg/m^2	1	40	Mean Difference (IV, Fixed, 95% CI)	-0.69 [-1.14, -0.24]
6.6.2 BMI > 25 kg/m2 < 30 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	-0.17 [-0.32, -0.01]

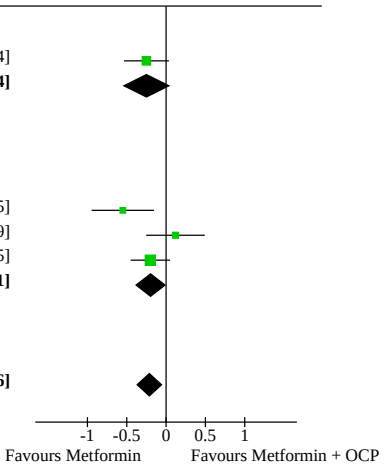
Analysis 6.1. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 1: Fasting insulin (mIU/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
6.1.1 BMI ≤ 25 kg/m2								
Liu 2006	22.79	3.83	20	25.97	4.74	20	24.1%	-3.18 [-5.85 , -0.51]
Wu 2008	11.2	4.2	11	10.8	3.6	10	15.4%	0.40 [-2.94 , 3.74]
Wu 2008	14.5	4.1	7	15.1	4.2	6	8.4%	-0.60 [-5.13 , 3.93]
Subtotal (95% CI)			**38**			**36**	**47.9%**	**-1.58 [-3.47 , 0.32]**
Heterogeneity: Chi² = 2.91, df = 2 (P = 0.23); I² = 31%								
Test for overall effect: Z = 1.63 (P = 0.10)								
6.1.2 BMI > 25 kg/m2 < 30 kg/m2								
Glintborg 2014a	8.48	5.54	19	10.42	7.58	23	10.9%	-1.94 [-5.92 , 2.04]
Kebapcilar 2009a	19	3.4	12	17.6	4	12	19.5%	1.40 [-1.57 , 4.37]
Kumar 2018	11.4	5.4	30	14.3	5.6	29	21.8%	-2.90 [-5.71 , -0.09]
Subtotal (95% CI)			**61**			**64**	**52.1%**	**-1.09 [-2.91 , 0.72]**
Heterogeneity: Chi² = 4.47, df = 2 (P = 0.11); I² = 55%								
Test for overall effect: Z = 1.18 (P = 0.24)								
Total (95% CI)			**99**			**100**	**100.0%**	**-1.32 [-2.63 , -0.01]**
Heterogeneity: Chi² = 7.51, df = 5 (P = 0.19); I² = 33%								
Test for overall effect: Z = 1.98 (P = 0.05)								
Test for subgroup differences: Chi² = 0.13, df = 1 (P = 0.72), I² = 0%								



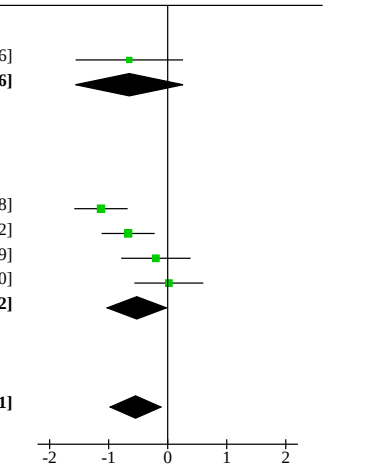
Analysis 6.2. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 2: Fasting glucose (mmol/L)

Study or Subgroup	Metformin Mean	SD	Total	Metformin + OCP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
6.2.1 BMI ≤ 25 kg/m2								
Liu 2006	4.21	0.53	20	4.46	0.38	20	29.4%	-0.25 [-0.54 , 0.04]
Subtotal (95% CI)			**20**			**20**	**29.4%**	**-0.25 [-0.54 , 0.04]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.71 (P = 0.09)								
6.2.2 BMI > 25 kg/m2 < 30 kg/m2								
Bodur 2018	4.35	0.34	17	4.9	0.64	12	15.3%	-0.55 [-0.95 , -0.15]
Glintborg 2014a	5.09	0.53	19	4.97	0.7	23	17.3%	0.12 [-0.25 , 0.49]
Kumar 2018	4.86	0.45	30	5.06	0.53	29	38.0%	-0.20 [-0.45 , 0.05]
Subtotal (95% CI)			**66**			**64**	**70.6%**	**-0.20 [-0.38 , -0.01]**
Heterogeneity: Chi² = 5.83, df = 2 (P = 0.05); I² = 66%								
Test for overall effect: Z = 2.10 (P = 0.04)								
Total (95% CI)			**86**			**84**	**100.0%**	**-0.21 [-0.37 , -0.06]**
Heterogeneity: Chi² = 5.92, df = 3 (P = 0.12); I² = 49%								
Test for overall effect: Z = 2.69 (P = 0.007)								
Test for subgroup differences: Chi² = 0.09, df = 1 (P = 0.76), I² = 0%								



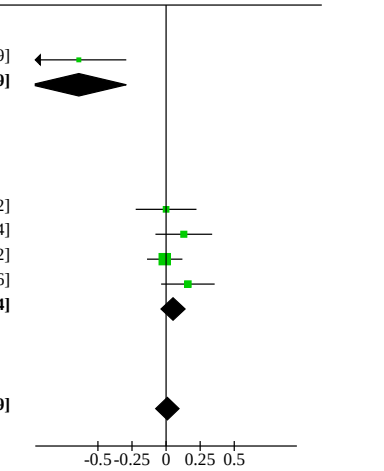
Analysis 6.3. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 3: Total Cholesterol (mmol/L)

Study or Subgroup	Favours Metformin Mean	SD	Total	Metformin + OCP Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.3.1 BMI ≤ 25 kg/m2								
Liu 2006	4.69	0.98	20	5.34	1.83	20	13.1%	-0.65 [-1.56 , 0.26]
Subtotal (95% CI)			**20**			**20**	**13.1%**	**-0.65 [-1.56 , 0.26]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.40 (P = 0.16)								
6.3.2 BMI > 25 kg/m2 < 30 kg/m2								
Moro 2013	4.6	0.65	25	5.73	0.97	26	23.5%	-1.13 [-1.58 , -0.68]
Kumar 2018	4.13	0.9	30	4.8	0.86	29	23.6%	-0.67 [-1.12 , -0.22]
Glintborg 2014a	4.68	0.99	19	4.88	0.94	23	19.9%	-0.20 [-0.79 , 0.39]
Kebapcilar 2009a	3.59	0.73	12	3.57	0.73	12	20.0%	0.02 [-0.56 , 0.60]
Subtotal (95% CI)			**86**			**90**	**86.9%**	**-0.52 [-1.02 , -0.02]**
Heterogeneity: Tau² = 0.19; Chi² = 11.49, df = 3 (P = 0.009); I² = 74%								
Test for overall effect: Z = 2.04 (P = 0.04)								
Total (95% CI)			**106**			**110**	**100.0%**	**-0.54 [-0.97 , -0.11]**
Heterogeneity: Tau² = 0.15; Chi² = 11.50, df = 4 (P = 0.02); I² = 65%								
Test for overall effect: Z = 2.48 (P = 0.01)								
Test for subgroup differences: Chi² = 0.06, df = 1 (P = 0.81), I² = 0%								



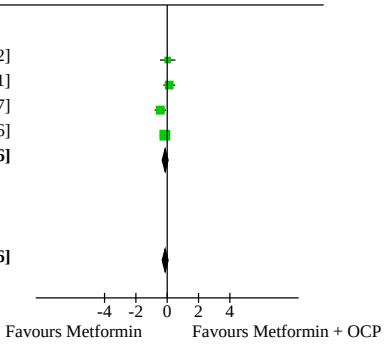
Analysis 6.4. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 4: HDL Cholesterol (mmol/L)

Study or Subgroup	Metformin + OCP Mean	SD	Total	Metformin Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
6.4.1 BMI ≤ 25 kg/m2								
Liu 2006	1.97	0.41	20	2.61	0.68	20	6.0%	-0.64 [-0.99 , -0.29]
Subtotal (95% CI)			**20**			**20**	**6.0%**	**-0.64 [-0.99 , -0.29]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.60 (P = 0.0003)								
6.4.2 BMI > 25 kg/m2 < 30 kg/m2								
Glintborg 2014a	1.51	0.35	23	1.51	0.38	19	14.7%	0.00 [-0.22 , 0.22]
Kebapcilar 2009a	1.59	0.26	12	1.46	0.26	12	16.9%	0.13 [-0.08 , 0.34]
Kumar 2018	1.07	0.2	29	1.08	0.3	30	43.4%	-0.01 [-0.14 , 0.12]
Moro 2013	2.07	0.43	26	1.91	0.27	25	19.0%	0.16 [-0.04 , 0.36]
Subtotal (95% CI)			**90**			**86**	**94.0%**	**0.05 [-0.04 , 0.14]**
Heterogeneity: Chi² = 2.79, df = 3 (P = 0.43); I² = 0%								
Test for overall effect: Z = 1.13 (P = 0.26)								
Total (95% CI)			**110**			**106**	**100.0%**	**0.01 [-0.08 , 0.09]**
Heterogeneity: Chi² = 17.02, df = 4 (P = 0.002); I² = 77%								
Test for overall effect: Z = 0.21 (P = 0.83)								
Test for subgroup differences: Chi² = 14.23, df = 1 (P = 0.0002), I² = 93.0%								

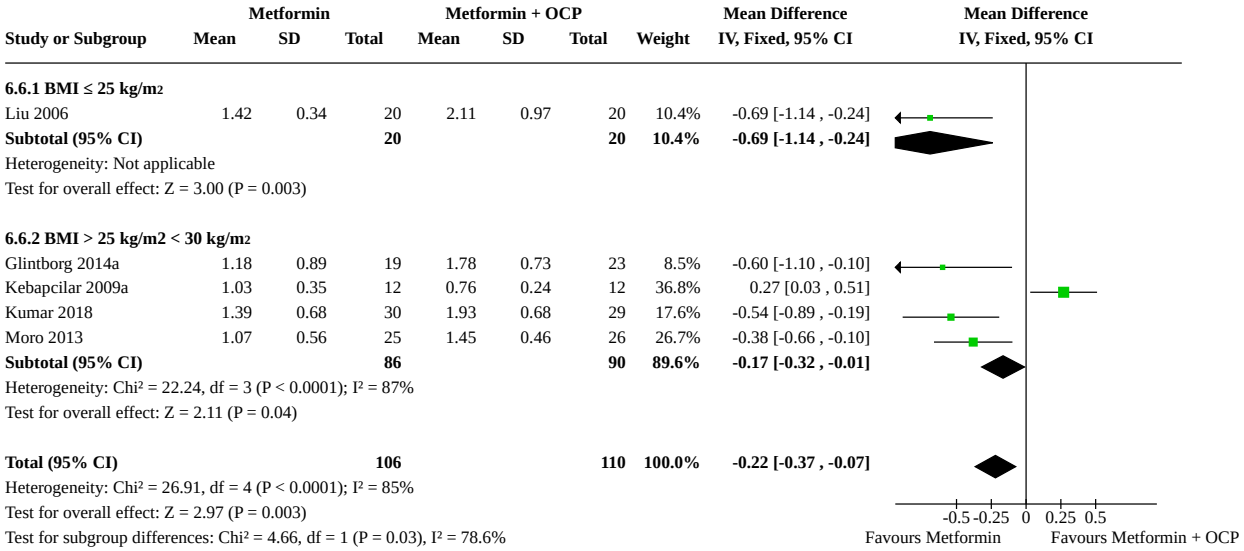


Analysis 6.5. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 5: LDL Cholesterol (mmol/L)

Study or Subgroup	Metformin Mean	SD	Total	Metformin + OCP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
6.5.1 BMI > 25 kg/m2 < 30 kg/m2								
Glintborg 2014a	2.76	0.82	19	2.73	0.8	23	14.3%	0.03 [-0.46 , 0.52]
Kebapcilar 2009a	2	0.5	12	1.87	0.44	12	24.5%	0.13 [-0.25 , 0.51]
Kumar 2018	2.45	0.74	30	2.89	0.7	29	25.7%	-0.44 [-0.81 , -0.07]
Moro 2013	2.46	0.31	25	2.61	0.75	26	35.5%	-0.15 [-0.46 , 0.16]
Subtotal (95% CI)			**86**			**90**	**100.0%**	**-0.13 [-0.32 , 0.06]**
Heterogeneity: Chi² = 4.98, df = 3 (P = 0.17); I² = 40%								
Test for overall effect: Z = 1.37 (P = 0.17)								
Total (95% CI)			**86**			**90**	**100.0%**	**-0.13 [-0.32 , 0.06]**
Heterogeneity: Chi² = 4.98, df = 3 (P = 0.17); I² = 40%								
Test for overall effect: Z = 1.37 (P = 0.17)								
Test for subgroup differences: Not applicable								



Analysis 6.6. Comparison 6: Adult - Metformin versus Metformin combined with OCP (Metabolic parameters), Outcome 6: Triglycerides (mmol/L)

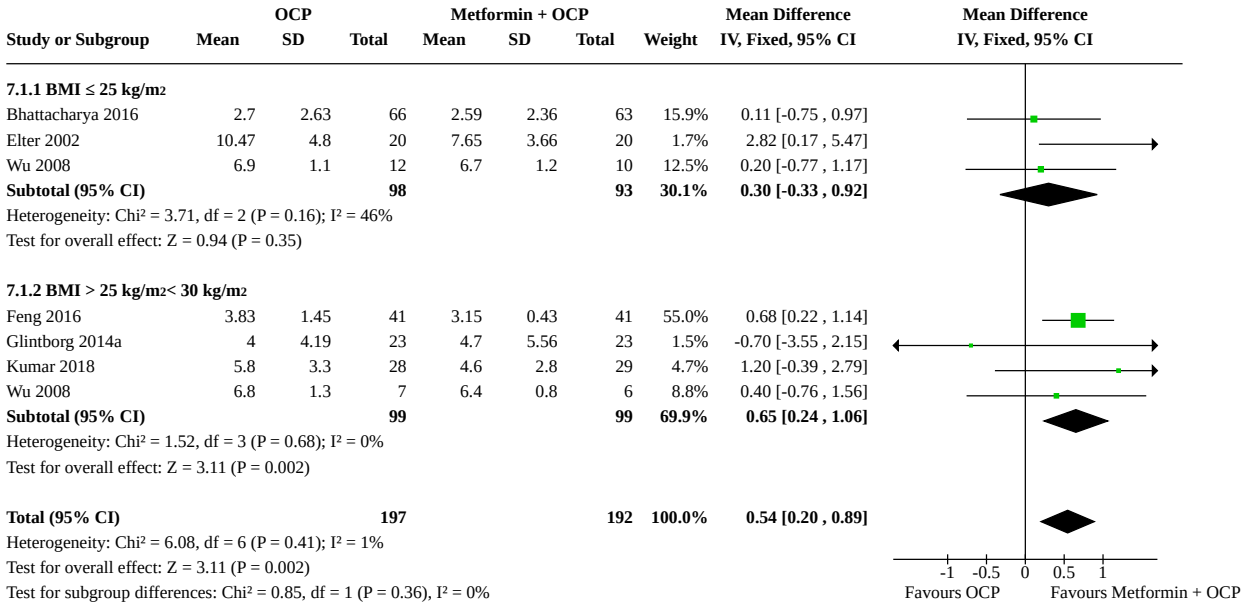


Comparison 7. Adult - OCP versus Metformin combined with OCP (Clinical parameters)

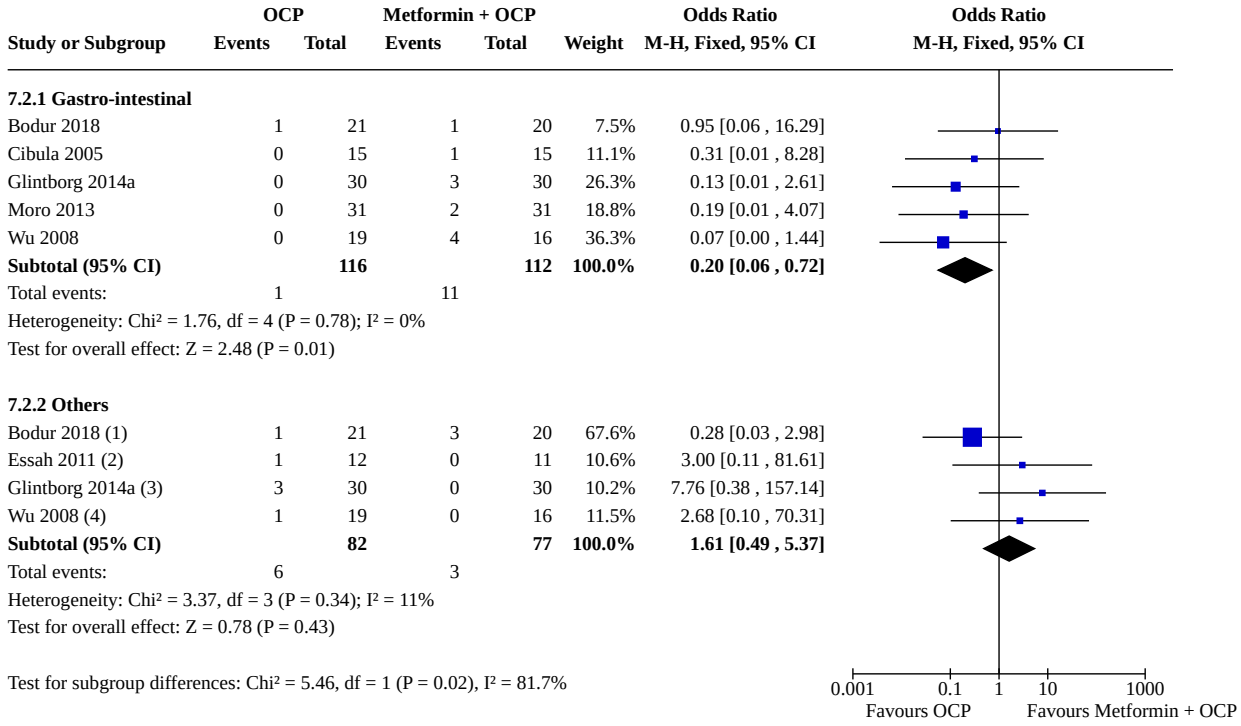
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Hirsutism - Clinical F-G score	6	389	Mean Difference (IV, Fixed, 95% CI)	0.54 [0.20, 0.89]
7.1.1 BMI ≤ 25 kg/m^2	3	191	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.33, 0.92]
7.1.2 BMI > 25 kg/m^2< 30 kg/m^2	4	198	Mean Difference (IV, Fixed, 95% CI)	0.65 [0.24, 1.06]
7.2 Adverse events - severe	6		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
7.2.1 Gastro-intestinal	5	228	Odds Ratio (M-H, Fixed, 95% CI)	0.20 [0.06, 0.72]
7.2.2 Others	4	159	Odds Ratio (M-H, Fixed, 95% CI)	1.61 [0.49, 5.37]
7.3 Adverse events - minor	2	98	Odds Ratio (M-H, Fixed, 95% CI)	0.06 [0.01, 0.44]
7.3.1 Gastro-intestinal	2	98	Odds Ratio (M-H, Fixed, 95% CI)	0.06 [0.01, 0.44]
7.4 Acne - Clinical acne score	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
7.4.1 BMI > 25 kg/m^2< 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
7.5 Acne - Subjective improvement	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
7.5.1 BMI ≤ 25 kg/m^2	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.6 Body weight (kg)	7	387	Mean Difference (IV, Fixed, 95% CI)	-0.63 [-1.58, 0.33]
7.6.1 BMI ≤ 25 kg/m^2	3	215	Mean Difference (IV, Fixed, 95% CI)	-0.57 [-1.55, 0.41]
7.6.2 BMI > 25 kg/m^2< 30 kg/m^2	3	153	Mean Difference (IV, Fixed, 95% CI)	-1.67 [-6.46, 3.12]
7.6.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-5.30 [-23.46, 12.86]
7.7 Body Mass Index (kg/m2)	13	661	Mean Difference (IV, Fixed, 95% CI)	-0.21 [-0.53, 0.12]
7.7.1 BMI ≤ 25 kg/m^2	6	327	Mean Difference (IV, Fixed, 95% CI)	-0.18 [-0.52, 0.17]
7.7.2 BMI > 25 kg/m2 < 30 kg/m^2	7	315	Mean Difference (IV, Fixed, 95% CI)	-0.37 [-1.29, 0.56]
7.7.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-3.10 [-9.61, 3.41]
7.8 Blood Pressure - Systolic (mmHg)	5	326	Mean Difference (IV, Fixed, 95% CI)	-1.75 [-4.03, 0.53]
7.8.1 BMI ≤ 25kg/m^2	1	129	Mean Difference (IV, Fixed, 95% CI)	0.52 [-3.42, 4.46]
7.8.2 BMI > 25 kg/m2 < 30 kg/m^2	3	178	Mean Difference (IV, Fixed, 95% CI)	-2.29 [-5.24, 0.66]
7.8.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-8.80 [-17.94, 0.34]
7.9 Blood Pressure - Diastolic (mmHg)	5	326	Mean Difference (IV, Fixed, 95% CI)	-1.05 [-2.79, 0.68]
7.9.1 BMI ≤ 25 kg/m^2	1	129	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-3.75, 2.75]
7.9.2 BMI > 25 kg/m2 < 30 kg/m^2	3	178	Mean Difference (IV, Fixed, 95% CI)	-0.61 [-2.84, 1.61]
7.9.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-5.20 [-10.60, 0.20]

Analysis 7.1. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 1: Hirsutism - Clinical F-G score



Analysis 7.2. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 2: Adverse events - severe

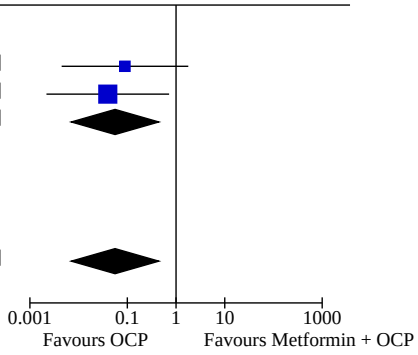


Footnotes

(1) OCP + MET dizziness/ headache/ OCP sexual reluctance
(2) Menorrhagia
(3) OCP group: side effects
(4) Weight gain

Analysis 7.3. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 3: Adverse events - minor

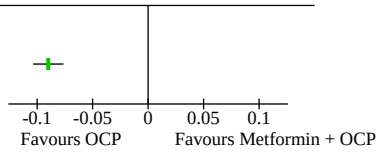
Study or Subgroup	OCP Events	OCP Total	Metformin + OCP Events	Metformin + OCP Total	Weight	Odds Ratio M-H, Fixed, 95% CI
7.3.1 Gastro-intestinal						
Elter 2002	0	20	4	20	32.7%	0.09 [0.00 , 1.78]
Wei 2012	0	28	9	30	67.3%	0.04 [0.00 , 0.72]
Subtotal (95% CI)		**48**		**50**	**100.0%**	**0.06 [0.01 , 0.44]**
Total events:	0		13			
Heterogeneity: Chi² = 0.15, df = 1 (P = 0.70); I² = 0%						
Test for overall effect: Z = 2.73 (P = 0.006)						
Total (95% CI)		**48**		**50**	**100.0%**	**0.06 [0.01 , 0.44]**
Total events:	0		13			
Heterogeneity: Chi² = 0.15, df = 1 (P = 0.70); I² = 0%						
Test for overall effect: Z = 2.73 (P = 0.006)						
Test for subgroup differences: Not applicable						



Analysis 7.4. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 4: Acne - Clinical acne score

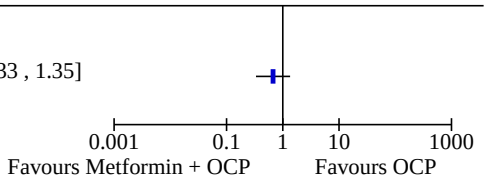
Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
7.4.1 BMI > 25 kg/m2< 30 kg/m2							
Feng 2016 (1)	0.45	0.04	41	0.54	0.02	41	-0.09 [-0.10 , -0.08]

Footnotes
(1) data subject to confirmation



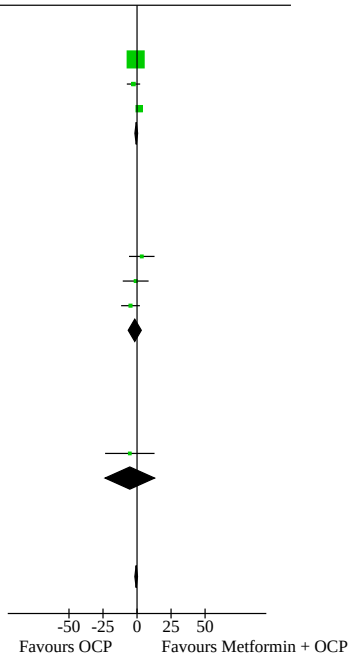
Analysis 7.5. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 5: Acne - Subjective improvement

Study or Subgroup	OCP Events	OCP Total	Metformin + OCP Events	Metformin + OCP Total	Odds Ratio M-H, Fixed, 95% CI
7.5.1 BMI ≤ 25 kg/m2					
Bhattacharya 2016	26	66	31	63	0.67 [0.33 , 1.35]



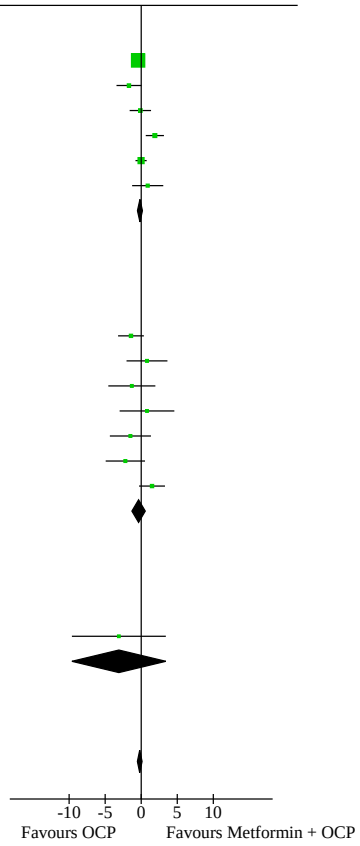
Analysis 7.6. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 6: Body weight (kg)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
7.6.1 BMI ≤ 25 kg/m2								
Bhattacharya 2016	0.76	2.4	66	1.77	3.89	63	73.2%	-1.01 [-2.13 , 0.11]
Cibula 2005	65.1	5.01	15	67.7	7.7	13	3.8%	-2.60 [-7.49 , 2.29]
Wei 2012	60.16	3.95	28	58.58	4.66	30	18.7%	1.58 [-0.64 , 3.80]
Subtotal (95% CI)			**109**			**106**	**95.7%**	**-0.57 [-1.55 , 0.41]**
Heterogeneity: Chi² = 4.86, df = 2 (P = 0.09); I² = 59%								
Test for overall effect: Z = 1.13 (P = 0.26)								
7.6.2 BMI > 25 kg/m2< 30 kg/m2								
Glintborg 2014a	80.1	19.14	23	76.62	12.64	23	1.0%	3.48 [-5.89 , 12.85]
Kaya 2015	71	15	25	72	19	25	1.0%	-1.00 [-10.49 , 8.49]
Kumar 2018	69.3	12.4	28	74.1	14.1	29	1.9%	-4.80 [-11.69 , 2.09]
Subtotal (95% CI)			**76**			**77**	**4.0%**	**-1.67 [-6.46 , 3.12]**
Heterogeneity: Chi² = 1.97, df = 2 (P = 0.37); I² = 0%								
Test for overall effect: Z = 0.68 (P = 0.49)								
7.6.3 BMI ≥ 30 kg/m2								
Essah 2011	90.4	20.24	10	95.7	20.1	9	0.3%	-5.30 [-23.46 , 12.86]
Subtotal (95% CI)			**10**			**9**	**0.3%**	**-5.30 [-23.46 , 12.86]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.57 (P = 0.57)								
Total (95% CI)			**195**			**192**	**100.0%**	**-0.63 [-1.58 , 0.33]**
Heterogeneity: Chi² = 7.28, df = 6 (P = 0.30); I² = 18%								
Test for overall effect: Z = 1.28 (P = 0.20)								
Test for subgroup differences: Chi² = 0.45, df = 2 (P = 0.80), I² = 0%								



Analysis 7.7. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 7: Body Mass Index (kg/m2)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
7.7.1 BMI ≤ 25 kg/m2								
Bhattacharya 2016	0.29	0.96	66	0.72	1.52	63	53.4%	-0.43 [-0.87 , 0.01]
Cibula 2005	22.7	1.8	15	24.4	2.7	13	3.5%	-1.70 [-3.43 , 0.03]
Elter 2002	22.08	1.9	20	22.2	2.8	20	4.7%	-0.12 [-1.60 , 1.36]
Lv 2005 (1)	22.04	1.85	25	20.15	2.63	25	6.5%	1.89 [0.63 , 3.15]
Wei 2012	23.01	1.27	28	23.03	1.72	30	17.3%	-0.02 [-0.79 , 0.75]
Wu 2008	22.2	3.2	12	21.3	1.9	10	2.2%	0.90 [-1.26 , 3.06]
Subtotal (95% CI)			**166**			**161**	**87.6%**	**-0.18 [-0.52 , 0.17]**
Heterogeneity: Chi² = 15.70, df = 5 (P = 0.008); I² = 68%								
Test for overall effect: Z = 1.00 (P = 0.32)								
7.7.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	25.81	4.24	41	27.23	4.02	41	3.2%	-1.42 [-3.21 , 0.37]
Glintborg 2014a	28	5.6	23	27.2	4.1	23	1.3%	0.80 [-2.04 , 3.64]
Kaya 2015	27.2	4.8	25	28.5	6.8	25	1.0%	-1.30 [-4.56 , 1.96]
Kebapcilar 2009a	27.3	6	12	26.5	3	12	0.7%	0.80 [-3.00 , 4.60]
Kebapcilar 2009b	25.9	5.6	22	27.4	3.8	21	1.3%	-1.50 [-4.35 , 1.35]
Kumar 2018	27.3	5	28	29.5	5.5	29	1.4%	-2.20 [-4.93 , 0.53]
Wu 2008	26.1	1.7	7	24.6	1.6	6	3.2%	1.50 [-0.30 , 3.30]
Subtotal (95% CI)			**158**			**157**	**12.1%**	**-0.37 [-1.29 , 0.56]**
Heterogeneity: Chi² = 9.15, df = 6 (P = 0.17); I² = 34%								
Test for overall effect: Z = 0.77 (P = 0.44)								
7.7.3 BMI ≥ 30 kg/m2								
Essah 2011	32.5	7.27	10	35.6	7.2	9	0.2%	-3.10 [-9.61 , 3.41]
Subtotal (95% CI)			**10**			**9**	**0.2%**	**-3.10 [-9.61 , 3.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.93 (P = 0.35)								
Total (95% CI)			**334**			**327**	**100.0%**	**-0.21 [-0.53 , 0.12]**
Heterogeneity: Chi² = 25.75, df = 13 (P = 0.02); I² = 50%								
Test for overall effect: Z = 1.25 (P = 0.21)								
Test for subgroup differences: Chi² = 0.90, df = 2 (P = 0.64), I² = 0%								

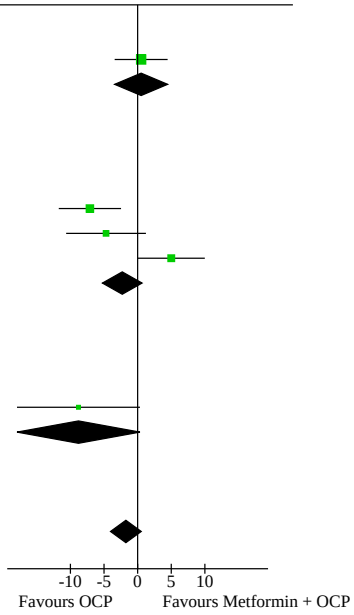


Footnotes

(1) data subject to confirmation

Analysis 7.8. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 8: Blood Pressure - Systolic (mmHg)

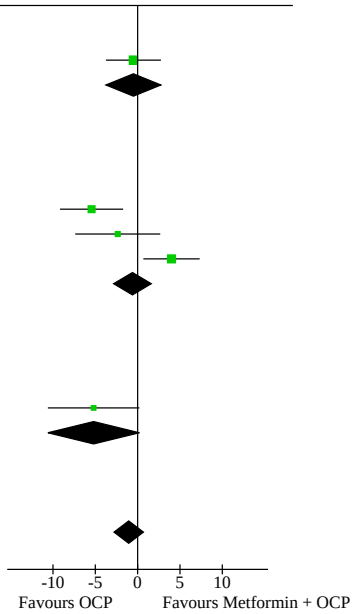
Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
7.8.1 BMI ≤ 25kg/m2								
Bhattacharya 2016	2.3	12.51	66	1.78	10.24	63	33.7%	0.52 [-3.42 , 4.46]
Subtotal (95% CI)			**66**			**63**	**33.7%**	**0.52 [-3.42 , 4.46]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.26 (P = 0.80)								
7.8.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	117.53	9.64	41	124.63	11.66	41	24.3%	-7.10 [-11.73 , -2.47]
Glintborg 2014a	126	12.07	23	130.7	8.06	23	14.8%	-4.70 [-10.63 , 1.23]
Kaya 2015	122	9	25	117	9	25	21.0%	5.00 [0.01 , 9.99]
Subtotal (95% CI)			**89**			**89**	**60.1%**	**-2.29 [-5.24 , 0.66]**
Heterogeneity: Chi² = 12.98, df = 2 (P = 0.002); I² = 85%								
Test for overall effect: Z = 1.52 (P = 0.13)								
7.8.3 BMI ≥ 30 kg/m2								
Essah 2011	117.4	10.1	10	126.2	10.2	9	6.2%	-8.80 [-17.94 , 0.34]
Subtotal (95% CI)			**10**			**9**	**6.2%**	**-8.80 [-17.94 , 0.34]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.89 (P = 0.06)								
Total (95% CI)			**165**			**161**	**100.0%**	**-1.75 [-4.03 , 0.53]**
Heterogeneity: Chi² = 16.67, df = 4 (P = 0.002); I² = 76%								
Test for overall effect: Z = 1.50 (P = 0.13)								
Test for subgroup differences: Chi² = 3.69, df = 2 (P = 0.16), I² = 45.8%								



Footnotes
(1) data subject to confirmation

Analysis 7.9. Comparison 7: Adult - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 9: Blood Pressure - Diastolic (mmHg)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
7.9.1 BMI ≤ 25 kg/m2								
Bhattacharya 2016	-1.61	9.34	66	-1.11	9.47	63	28.6%	-0.50 [-3.75 , 2.75]
Subtotal (95% CI)			**66**			**63**	**28.6%**	**-0.50 [-3.75 , 2.75]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.30 (P = 0.76)								
7.9.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	73.09	7.45	41	78.53	9.62	41	21.8%	-5.44 [-9.16 , -1.72]
Glintborg 2014a	82	7.4	23	84.35	9.81	23	12.0%	-2.35 [-7.37 , 2.67]
Kaya 2015	79	6	25	75	6	25	27.3%	4.00 [0.67 , 7.33]
Subtotal (95% CI)			**89**			**89**	**61.0%**	**-0.61 [-2.84 , 1.61]**
Heterogeneity: Chi² = 14.30, df = 2 (P = 0.0008); I² = 86%								
Test for overall effect: Z = 0.54 (P = 0.59)								
7.9.3 BMI ≥ 30 kg/m2								
Essah 2011	69.9	6	10	75.1	6	9	10.3%	-5.20 [-10.60 , 0.20]
Subtotal (95% CI)			**10**			**9**	**10.3%**	**-5.20 [-10.60 , 0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.89 (P = 0.06)								
Total (95% CI)			**165**			**161**	**100.0%**	**-1.05 [-2.79 , 0.68]**
Heterogeneity: Chi² = 16.83, df = 4 (P = 0.002); I² = 76%								
Test for overall effect: Z = 1.19 (P = 0.23)								
Test for subgroup differences: Chi² = 2.52, df = 2 (P = 0.28), I² = 20.8%								



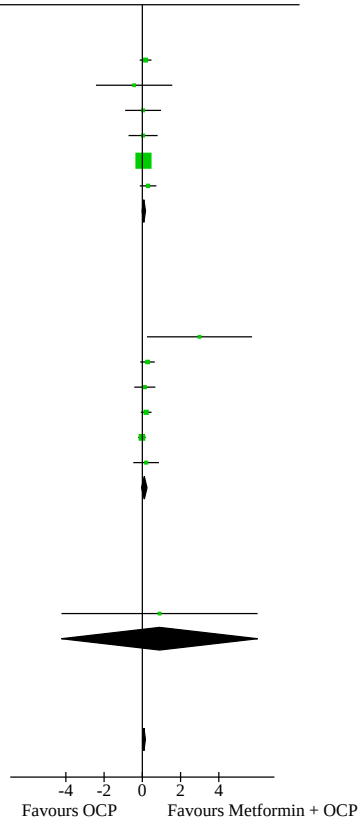
Footnotes
(1) data subject to confirmation

Comparison 8. Adult - OCP versus Metformin combined with OCP (Hormonal parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 Serum total testosterone (nmol/L)	12	715	Mean Difference (IV, Fixed, 95% CI)	0.08 [0.01, 0.16]
8.1.1 BMI ≤ 25 kg/m^2	6	327	Mean Difference (IV, Fixed, 95% CI)	0.08 [-0.01, 0.17]
8.1.2 BMI > 25 kg/m2 < 30 kg/m^2	6	369	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.04, 0.24]
8.1.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	0.90 [-4.23, 6.03]
8.2 Free androgen index (FAI) (%)	7	482	Mean Difference (IV, Fixed, 95% CI)	0.51 [0.30, 0.71]
8.2.1 BMI ≤ 25 kg/m^2	3	215	Mean Difference (IV, Fixed, 95% CI)	0.47 [-0.12, 1.06]
8.2.2 BMI > 25 kg/m2 < 30 kg/m^2	4	267	Mean Difference (IV, Fixed, 95% CI)	0.51 [0.29, 0.73]

Analysis 8.1. Comparison 8: Adult - OCP versus Metformin combined with OCP (Hormonal parameters), Outcome 1: Serum total testosterone (nmol/L)

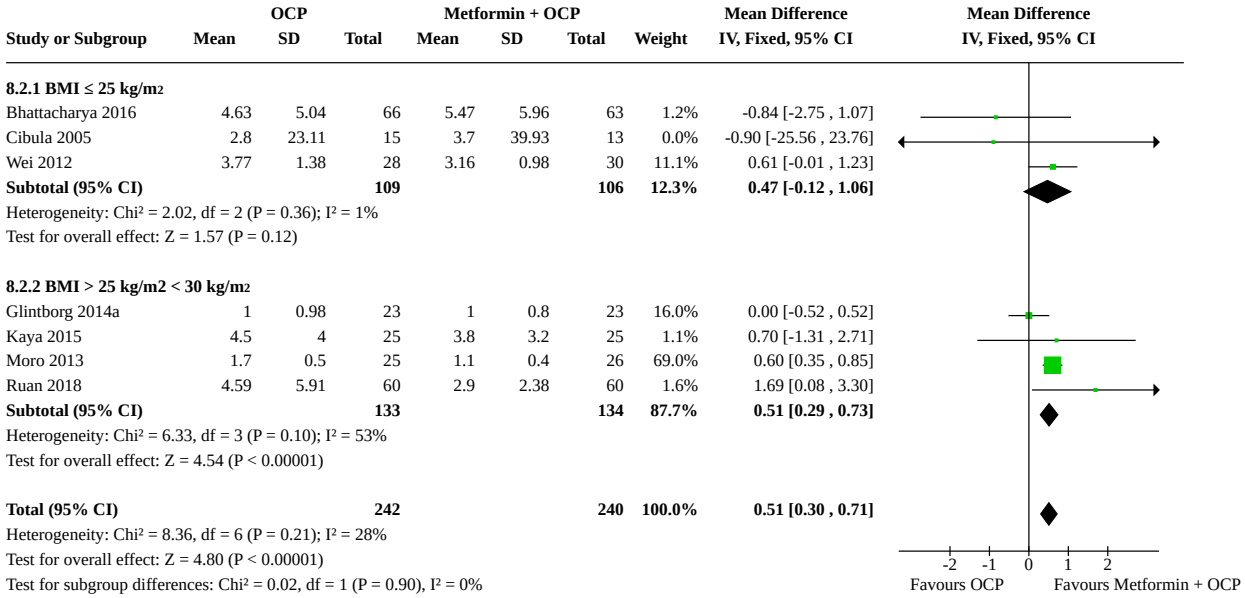
Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
8.1.1 BMI ≤ 25 kg/m2								
Bhattacharya 2016	0.45	1.01	66	0.28	0.76	63	5.9%	0.17 [-0.14 , 0.48]
Cibula 2005	3.49	1.41	15	3.92	3.43	13	0.1%	-0.43 [-2.43 , 1.57]
Elter 2002	1.63	1.4	20	1.59	1.62	20	0.6%	0.04 [-0.90 , 0.98]
Lv 2005	1.23	1.42	25	1.19	1.32	25	1.0%	0.04 [-0.72 , 0.80]
Wei 2012	1.59	0.17	28	1.53	0.2	30	61.3%	0.06 [-0.04 , 0.16]
Wu 2008	2.1	0.4	12	1.8	0.6	10	2.9%	0.30 [-0.14 , 0.74]
Subtotal (95% CI)			**166**			**161**	**71.9%**	**0.08 [-0.01 , 0.17]**
Heterogeneity: Chi² = 1.74, df = 5 (P = 0.88); I² = 0%								
Test for overall effect: Z = 1.72 (P = 0.08)								
8.1.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	18.45	6.35	41	15.46	6.35	41	0.1%	2.99 [0.24 , 5.74]
Glintborg 2014a	1.55	0.83	23	1.28	0.43	23	3.8%	0.27 [-0.11 , 0.65]
Kumar 2018	2.18	1.14	28	2.05	0.97	29	1.8%	0.13 [-0.42 , 0.68]
Moro 2013	2.3	0.6	25	2.1	0.4	26	7.1%	0.20 [-0.08 , 0.48]
Ruan 2018	1.54	0.36	60	1.56	0.7	60	14.0%	-0.02 [-0.22 , 0.18]
Wu 2008	2.6	0.5	7	2.4	0.7	6	1.2%	0.20 [-0.47 , 0.87]
Subtotal (95% CI)			**184**			**185**	**28.1%**	**0.10 [-0.04 , 0.24]**
Heterogeneity: Chi² = 6.98, df = 5 (P = 0.22); I² = 28%								
Test for overall effect: Z = 1.42 (P = 0.16)								
8.1.3 BMI ≥ 30 kg/m2								
Essah 2011	7.8	5.69	10	6.9	5.7	9	0.0%	0.90 [-4.23 , 6.03]
Subtotal (95% CI)			**10**			**9**	**0.0%**	**0.90 [-4.23 , 6.03]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.34 (P = 0.73)								
Total (95% CI)			**360**			**355**	**100.0%**	**0.08 [0.01 , 0.16]**
Heterogeneity: Chi² = 8.91, df = 12 (P = 0.71); I² = 0%								
Test for overall effect: Z = 2.22 (P = 0.03)								
Test for subgroup differences: Chi² = 0.18, df = 2 (P = 0.91), I² = 0%								



Footnotes

(1) data subject to confirmation

Analysis 8.2. Comparison 8: Adult - OCP versus Metformin combined with OCP (Hormonal parameters), Outcome 2: Free androgen index (FAI) (%)



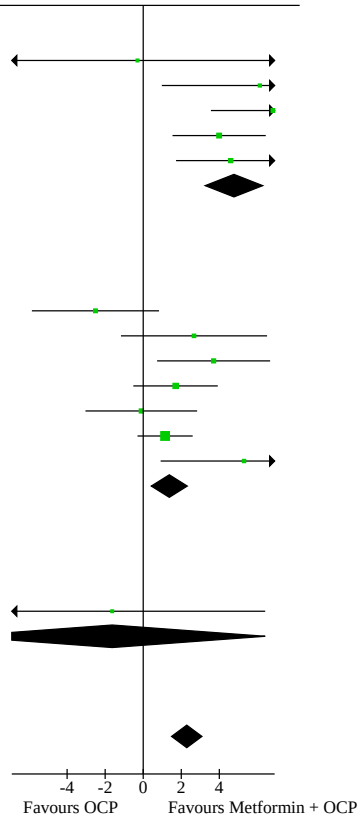
Comparison 9. Adult - OCP versus Metformin combined with OCP (Metabolic parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9.1 Fasting insulin (mIU/L)	12	602	Mean Difference (IV, Fixed, 95% CI)	2.29 [1.49, 3.09]
9.1.1 BMI ≤ 25 kg/m^2	5	198	Mean Difference (IV, Fixed, 95% CI)	4.77 [3.26, 6.28]
9.1.2 BMI > 25 kg/m2 < 30 kg/m^2	7	385	Mean Difference (IV, Fixed, 95% CI)	1.37 [0.42, 2.32]
9.1.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-1.63 [-9.67, 6.41]
9.2 Fasting glucose (mmol/L)	10	529	Mean Difference (IV, Fixed, 95% CI)	0.20 [0.11, 0.29]
9.2.1 BMI ≤ 25 kg/m^2	5	205	Mean Difference (IV, Fixed, 95% CI)	0.11 [-0.07, 0.29]
9.2.2 BMI > 25 kg/m2 < 30 kg/m^2	4	305	Mean Difference (IV, Fixed, 95% CI)	0.25 [0.14, 0.36]
9.2.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.81, 0.21]
9.3 Total Cholesterol (mmol/L)	13	668	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.05, 0.17]
9.3.1 BMI ≤ 25 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.20, 0.16]
9.3.2 BMI > 25 kg/m2 < 30 kg/m^2	8	473	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.04, 0.25]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9.3.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.31, 0.91]
9.4 HDL Cholesterol (mmol/L)	13	668	Mean Difference (IV, Fixed, 95% CI)	0.05 [0.01, 0.09]
9.4.1 BMI ≤ 25 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.00, 0.10]
9.4.2 BMI > 25 kg/m2 < 30 kg/m^2	8	473	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.02, 0.11]
9.4.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.02, 0.42]
9.5 LDL Cholesterol (mmol/L)	13	668	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.05, 0.14]
9.5.1 BMI ≤ 25 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.18, 0.21]
9.5.2 BMI > 25 kg/m2 < 30 kg/m^2	8	473	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.07, 0.16]
9.5.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.27, 0.87]
9.6 Triglycerides (mmol/L)	13	668	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.09, 0.04]
9.6.1 BMI ≤ 25 kg/m^2	4	176	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.08, 0.11]
9.6.2 BMI > 25 kg/m2 < 30 kg/m^2	8	473	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.15, 0.03]
9.6.3 BMI ≥ 30 kg/m^2	1	19	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.67, 1.27]

Analysis 9.1. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 1: Fasting insulin (mIU/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
9.1.1 BMI ≤ 25 kg/m2								
Cibula 2005	14.8	5.39	15	15.1	12.58	13	1.2%	-0.30 [-7.66 , 7.06]
Elter 2002	19.77	10.03	20	13.64	6.13	20	2.4%	6.13 [0.98 , 11.28]
Lv 2005	15.66	6.16	25	8.84	5.59	25	6.0%	6.82 [3.56 , 10.08]
Wei 2012	17.8	4.6	28	13.81	4.92	30	10.6%	3.99 [1.54 , 6.44]
Wu 2008	15.4	3.2	12	10.8	3.6	10	7.7%	4.60 [1.73 , 7.47]
Subtotal (95% CI)			**100**			**98**	**27.9%**	**4.77 [3.26 , 6.28]**
Heterogeneity: Chi² = 4.01, df = 4 (P = 0.40); I² = 0%								
Test for overall effect: Z = 6.19 (P < 0.00001)								
9.1.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	17.94	7.03	41	20.45	8.35	41	5.7%	-2.51 [-5.85 , 0.83]
Glintborg 2014a	13.09	5.55	23	10.42	7.58	23	4.3%	2.67 [-1.17 , 6.51]
Kebapcilar 2009a	21.3	3.4	12	17.6	4	12	7.2%	3.70 [0.73 , 6.67]
Kebapcilar 2009b	18.5	4.1	22	16.8	3.3	21	12.9%	1.70 [-0.52 , 3.92]
Kumar 2018	14.2	5.7	28	14.3	5.6	29	7.4%	-0.10 [-3.03 , 2.83]
Ruan 2018	22.52	4	60	21.37	4.1	60	30.3%	1.15 [-0.30 , 2.60]
Wu 2008	20.4	3.8	7	15.1	4.2	6	3.3%	5.30 [0.92 , 9.68]
Subtotal (95% CI)			**193**			**192**	**71.1%**	**1.37 [0.42 , 2.32]**
Heterogeneity: Chi² = 12.21, df = 6 (P = 0.06); I² = 51%								
Test for overall effect: Z = 2.84 (P = 0.005)								
9.1.3 BMI ≥ 30 kg/m2								
Essah 2011	9.5	8.85	10	11.13	9	9	1.0%	-1.63 [-9.67 , 6.41]
Subtotal (95% CI)			**10**			**9**	**1.0%**	**-1.63 [-9.67 , 6.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.40 (P = 0.69)								
Total (95% CI)			**303**			**299**	**100.0%**	**2.29 [1.49 , 3.09]**
Heterogeneity: Chi² = 31.11, df = 12 (P = 0.002); I² = 61%								
Test for overall effect: Z = 5.62 (P < 0.00001)								
Test for subgroup differences: Chi² = 14.89, df = 2 (P = 0.0006), I² = 86.6%								

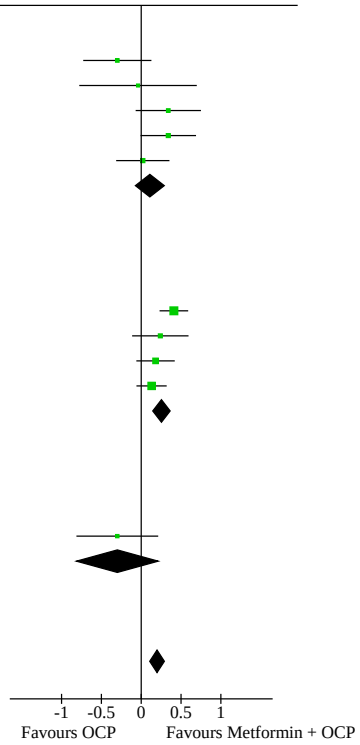


Footnotes

(1) data subject to confirmation

Analysis 9.2. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 2: Fasting glucose (mmol/L)

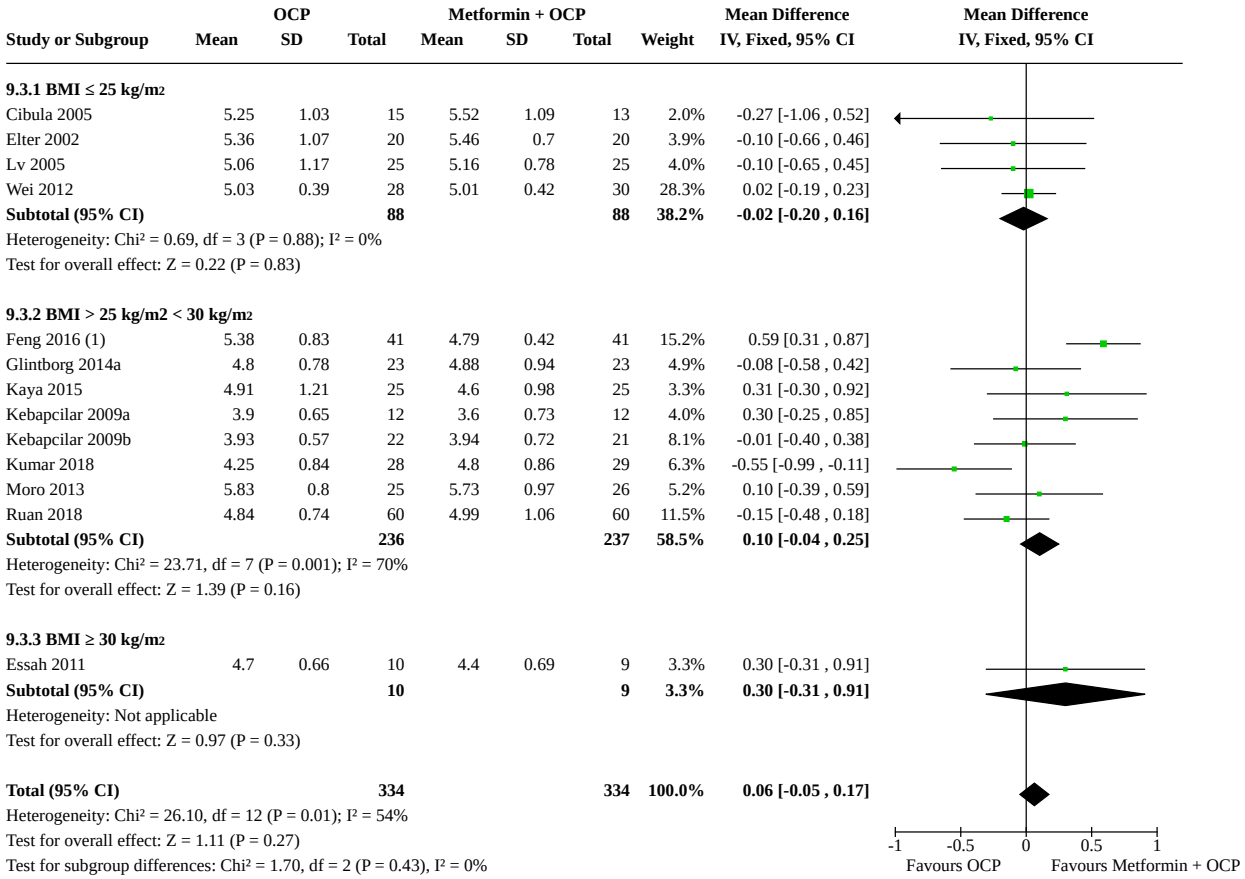
Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
9.2.1 BMI ≤ 25 kg/m2								
Bodur 2018	4.6	0.48	17	4.9	0.64	12	4.6%	-0.30 [-0.73 , 0.13]
Cibula 2005	4.38	1.03	15	4.42	0.96	13	1.6%	-0.04 [-0.78 , 0.70]
Elter 2002	4.68	0.7	20	4.34	0.62	20	5.0%	0.34 [-0.07 , 0.75]
Lv 2005	4.88	0.72	25	4.54	0.52	25	7.0%	0.34 [-0.01 , 0.69]
Wei 2012	4.63	0.73	28	4.61	0.55	30	7.6%	0.02 [-0.31 , 0.35]
Subtotal (95% CI)			**105**			**100**	**25.8%**	**0.11 [-0.07 , 0.29]**
Heterogeneity: Chi² = 6.85, df = 4 (P = 0.14); I² = 42%								
Test for overall effect: Z = 1.17 (P = 0.24)								
9.2.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	5.27	0.42	41	4.86	0.41	41	26.2%	0.41 [0.23 , 0.59]
Glintborg 2014a	5.2	0.51	23	4.96	0.7	23	6.8%	0.24 [-0.11 , 0.59]
Kumar 2018	5.24	0.39	28	5.06	0.53	29	14.6%	0.18 [-0.06 , 0.42]
Ruan 2018	5.04	0.52	60	4.91	0.54	60	23.5%	0.13 [-0.06 , 0.32]
Subtotal (95% CI)			**152**			**153**	**71.0%**	**0.25 [0.14 , 0.36]**
Heterogeneity: Chi² = 4.91, df = 3 (P = 0.18); I² = 39%								
Test for overall effect: Z = 4.56 (P < 0.00001)								
9.2.3 BMI ≥ 30 kg/m2								
Essah 2011	4.9	0.57	10	5.2	0.57	9	3.2%	-0.30 [-0.81 , 0.21]
Subtotal (95% CI)			**10**			**9**	**3.2%**	**-0.30 [-0.81 , 0.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.15 (P = 0.25)								
Total (95% CI)			**267**			**262**	**100.0%**	**0.20 [0.11 , 0.29]**
Heterogeneity: Chi² = 17.32, df = 9 (P = 0.04); I² = 48%								
Test for overall effect: Z = 4.23 (P < 0.0001)								
Test for subgroup differences: Chi² = 5.57, df = 2 (P = 0.06), I² = 64.1%								



Footnotes

(1) data subject to confirmation

Analysis 9.3. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 3: Total Cholesterol (mmol/L)

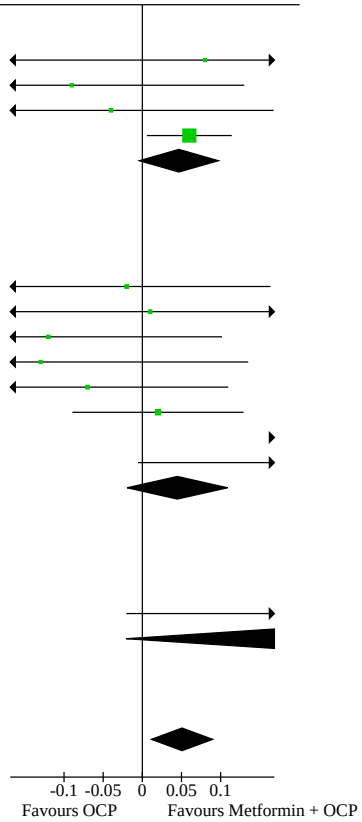


Footnotes

(1) data subject to confirmation

Analysis 9.4. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 4: HDL Cholesterol (mmol/L)

Study or Subgroup	Metformin + OCP Mean	SD	Total	OCP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
9.4.1 BMI ≤ 25 kg/m2								
Cibula 2005	1.76	0.39	13	1.68	0.51	15	1.4%	0.08 [-0.25 , 0.41]
Elter 2002	1.26	0.29	20	1.35	0.41	20	3.1%	-0.09 [-0.31 , 0.13]
Lv 2005	1.26	0.28	25	1.3	0.45	25	3.5%	-0.04 [-0.25 , 0.17]
Wei 2012	1.19	0.1	30	1.13	0.11	28	51.8%	0.06 [0.01 , 0.11]
Subtotal (95% CI)			**88**			**88**	**59.8%**	**0.05 [-0.00 , 0.10]**
Heterogeneity: Chi² = 2.42, df = 3 (P = 0.49); I² = 0%								
Test for overall effect: Z = 1.81 (P = 0.07)								
9.4.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	1.72	0.48	41	1.74	0.36	41	4.5%	-0.02 [-0.20 , 0.16]
Glintborg 2014a	1.51	0.35	23	1.5	0.31	23	4.2%	0.01 [-0.18 , 0.20]
Kaya 2015	1.37	0.41	25	1.49	0.39	25	3.1%	-0.12 [-0.34 , 0.10]
Kebapcilar 2009a	1.59	0.26	12	1.72	0.39	12	2.2%	-0.13 [-0.40 , 0.14]
Kebapcilar 2009b	1.53	0.23	21	1.6	0.36	22	4.7%	-0.07 [-0.25 , 0.11]
Kumar 2018	1.07	0.2	29	1.05	0.22	28	12.8%	0.02 [-0.09 , 0.13]
Moro 2013	2.07	0.43	26	1.39	0.45	25	2.6%	0.68 [0.44 , 0.92]
Ruan 2018	1.7	0.66	60	1.48	0.6	60	3.0%	0.22 [-0.01 , 0.45]
Subtotal (95% CI)			**237**			**236**	**37.0%**	**0.04 [-0.02 , 0.11]**
Heterogeneity: Chi² = 35.00, df = 7 (P < 0.0001); I² = 80%								
Test for overall effect: Z = 1.37 (P = 0.17)								
9.4.3 BMI ≥ 30 kg/m2								
Essah 2011	1.4	0.24	9	1.2	0.25	10	3.1%	0.20 [-0.02 , 0.42]
Subtotal (95% CI)			**9**			**10**	**3.1%**	**0.20 [-0.02 , 0.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.78 (P = 0.08)								
Total (95% CI)			**334**			**334**	**100.0%**	**0.05 [0.01 , 0.09]**
Heterogeneity: Chi² = 39.24, df = 12 (P < 0.0001); I² = 69%								
Test for overall effect: Z = 2.55 (P = 0.01)								
Test for subgroup differences: Chi² = 1.82, df = 2 (P = 0.40), I² = 0%								

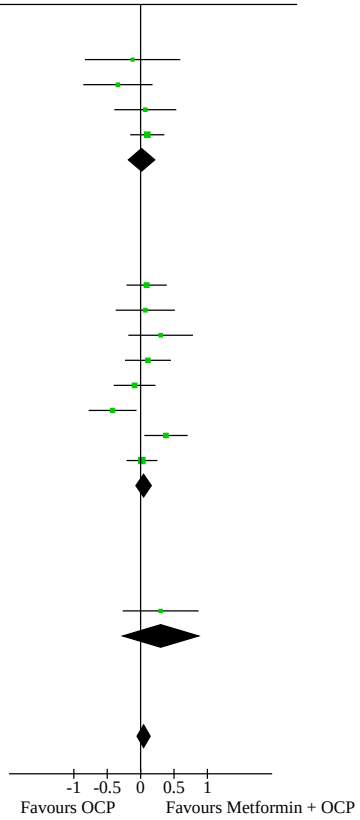


Footnotes

(1) data subject to confirmation

Analysis 9.5. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 5: LDL Cholesterol (mmol/L)

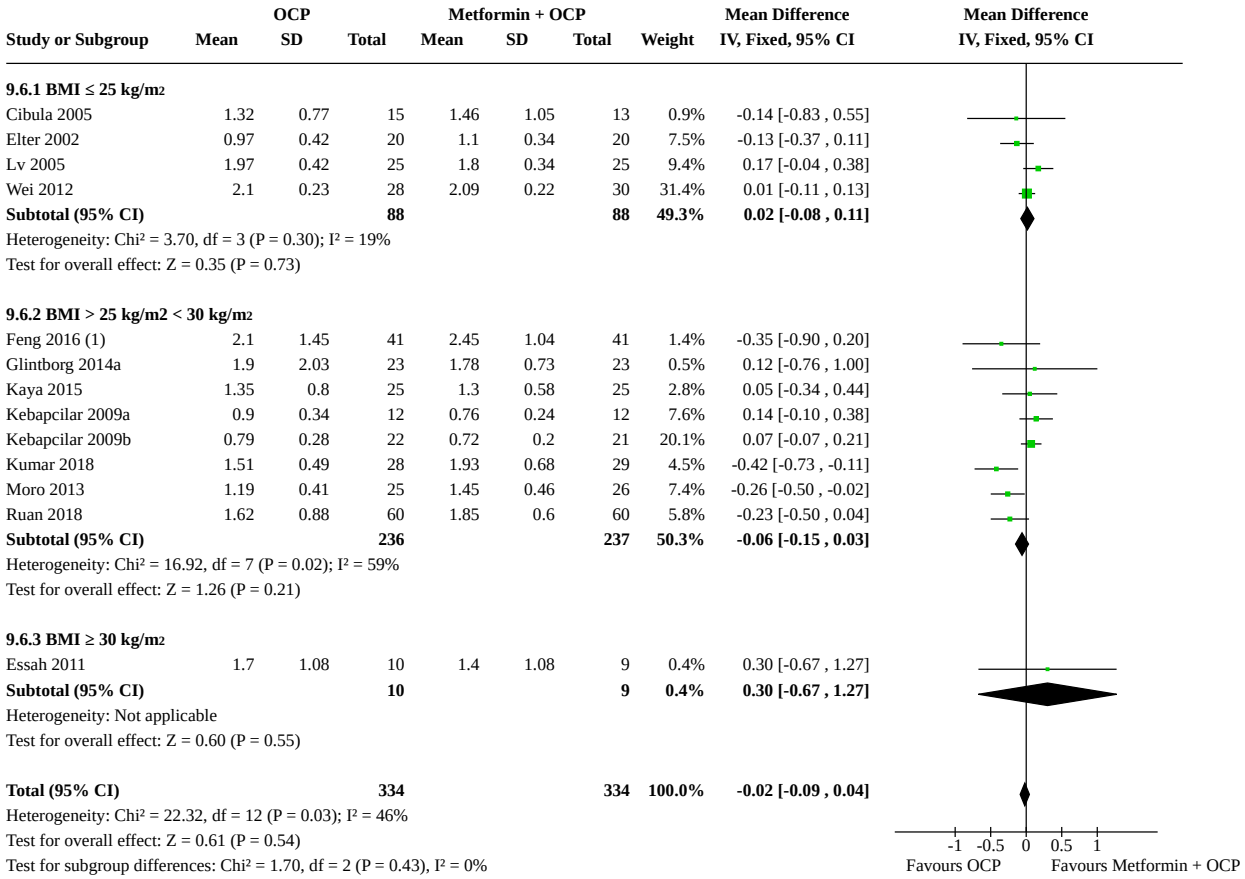
Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
9.5.1 BMI ≤ 25 kg/m2								
Cibula 2005	2.98	0.77	15	3.1	1.1	13	1.9%	-0.12 [-0.83 , 0.59]
Elter 2002	2.97	0.89	20	3.31	0.78	20	3.6%	-0.34 [-0.86 , 0.18]
Lv 2005	2.97	0.89	25	2.9	0.78	25	4.5%	0.07 [-0.39 , 0.53]
Wei 2012	4.05	0.51	28	3.95	0.48	30	14.8%	0.10 [-0.16 , 0.36]
Subtotal (95% CI)			**88**			**88**	**24.7%**	**0.01 [-0.18 , 0.21]**
Heterogeneity: Chi² = 2.42, df = 3 (P = 0.49); I² = 0%								
Test for overall effect: Z = 0.14 (P = 0.89)								
9.5.2 BMI > 25 kg/m2 < 30 kg/m2								
Feng 2016 (1)	2.23	0.75	41	2.14	0.64	41	10.6%	0.09 [-0.21 , 0.39]
Glintborg 2014a	2.8	0.72	23	2.73	0.81	23	4.9%	0.07 [-0.37 , 0.51]
Kaya 2015	2.9	1.05	25	2.6	0.65	25	4.1%	0.30 [-0.18 , 0.78]
Kebapcilar 2009a	1.98	0.42	12	1.87	0.44	12	8.1%	0.11 [-0.23 , 0.45]
Kebapcilar 2009b	2.15	0.49	22	2.24	0.55	21	9.9%	-0.09 [-0.40 , 0.22]
Kumar 2018	2.47	0.68	28	2.89	0.7	29	7.5%	-0.42 [-0.78 , -0.06]
Moro 2013	2.99	0.38	25	2.61	0.75	26	9.1%	0.38 [0.06 , 0.70]
Ruan 2018	2.97	0.65	60	2.95	0.64	60	18.1%	0.02 [-0.21 , 0.25]
Subtotal (95% CI)			**236**			**237**	**72.3%**	**0.04 [-0.07 , 0.16]**
Heterogeneity: Chi² = 12.63, df = 7 (P = 0.08); I² = 45%								
Test for overall effect: Z = 0.76 (P = 0.45)								
9.5.3 BMI ≥ 30 kg/m2								
Essah 2011	2.8	0.63	10	2.5	0.63	9	3.0%	0.30 [-0.27 , 0.87]
Subtotal (95% CI)			**10**			**9**	**3.0%**	**0.30 [-0.27 , 0.87]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.04 (P = 0.30)								
Total (95% CI)			**334**			**334**	**100.0%**	**0.04 [-0.05 , 0.14]**
Heterogeneity: Chi² = 15.92, df = 12 (P = 0.19); I² = 25%								
Test for overall effect: Z = 0.89 (P = 0.37)								
Test for subgroup differences: Chi² = 0.87, df = 2 (P = 0.65), I² = 0%								



Footnotes

(1) data subject to confirmation

Analysis 9.6. Comparison 9: Adult - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 6: Triglycerides (mmol/L)



Footnotes

(1) data subject to confirmation

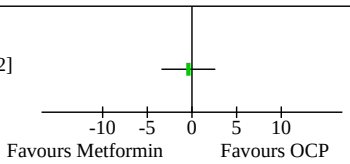
Comparison 10. Adolescent - Metformin versus OCP (Clinical parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
10.1 Hirsutism - Clinical F-G score	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
10.1.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
10.2 Hirsutism - Subjective improvement	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.2.1 Mean BMI not stated	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.3 Adverse event - severe	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.3.1 Others	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.4 Adverse event - minor	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
10.4.1 Gastro-intestinal	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.5 Improved menstrual pattern (ie. an initiation of menses or cycle regularity)	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.5.1 Mean BMI not stated	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
10.6 Body Weight (kg)	2	111	Mean Difference (IV, Fixed, 95% CI)	-18.77 [-20.57, -16.98]
10.6.1 BMI ≥ 30 kg/m^2	1	31	Mean Difference (IV, Fixed, 95% CI)	-2.60 [-17.87, 12.67]
10.6.2 Mean BMI not stated	1	80	Mean Difference (IV, Fixed, 95% CI)	-19.00 [-20.81, -17.19]
10.7 Body Mass Index (kg/m^2)	3	69	Mean Difference (IV, Fixed, 95% CI)	-1.45 [-5.08, 2.17]
10.7.1 BMI ≥ 30 kg/m^2	3	69	Mean Difference (IV, Fixed, 95% CI)	-1.45 [-5.08, 2.17]
10.8 Blood pressure - systolic (mm Hg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
10.8.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
10.9 Blood pressure - diastolic (mm Hg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
10.9.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

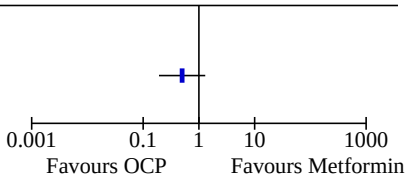
Analysis 10.1. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 1: Hirsutism - Clinical F-G score

	Metformin			OCP			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
10.1.1 BMI ≥ 30 kg/m2								
Hoeger 2008a	8.2	3.4	6	8.6	2.1	10	-0.40 [-3.42 , 2.62]	



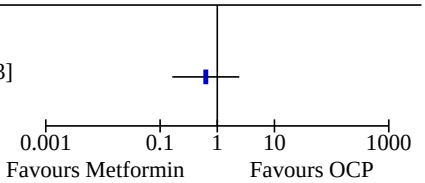
Analysis 10.2. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 2: Hirsutism - Subjective improvement

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Odds Ratio M-H, Fixed, 95% CI	Odds Ratio M-H, Fixed, 95% CI
10.2.1 Mean BMI not stated						
El Maghraby 2015	10	40	16	40	0.50 [0.19 , 1.30]	



Analysis 10.3. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 3: Adverse event - severe

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Odds Ratio M-H, Fixed, 95% CI	Odds Ratio M-H, Fixed, 95% CI
10.3.1 Others						
El Maghraby 2015 (1)	4	40	6	40	0.63 [0.16 , 2.43]	

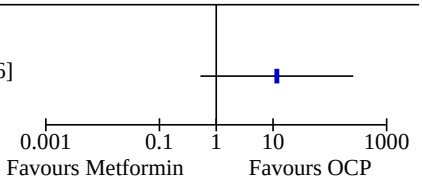


Footnotes

(1) MET group: side effects/ OCP group: 4 weight gain 2 side effects

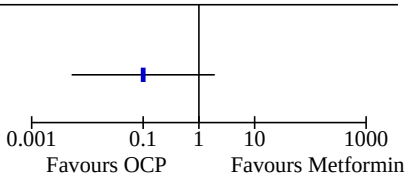
Analysis 10.4. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 4: Adverse event - minor

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Odds Ratio M-H, Fixed, 95% CI	Odds Ratio M-H, Fixed, 95% CI
10.4.1 Gastro-intestinal						
Al-Zubeidi 2015	3	10	0	12	11.67 [0.53 , 258.56]	



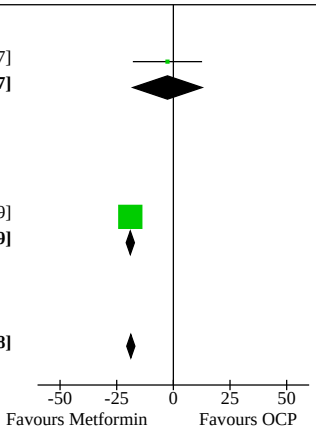
Analysis 10.5. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 5: Improved menstrual pattern (ie. an initiation of menses or cycle regularity)

Study or Subgroup	Metformin Events	Metformin Total	OCP Events	OCP Total	Odds Ratio M-H, Fixed, 95% CI	Odds Ratio M-H, Fixed, 95% CI
10.5.1 Mean BMI not stated						
El Maghraby 2015	36	40	40	40	0.10 [0.01 , 1.92]	



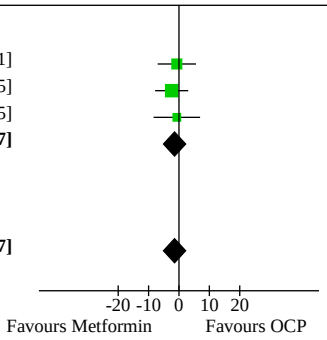
Analysis 10.6. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 6: Body Weight (kg)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
10.6.1 BMI ≥ 30 kg/m2								
Allen 2005	96.3	17.2	16	98.9	25.17	15	1.4%	-2.60 [-17.87 , 12.67]
Subtotal (95% CI)			**16**			**15**	**1.4%**	**-2.60 [-17.87 , 12.67]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.33 (P = 0.74)								
10.6.2 Mean BMI not stated								
El Maghraby 2015	72	5	40	91	3	40	98.6%	-19.00 [-20.81 , -17.19]
Subtotal (95% CI)			**40**			**40**	**98.6%**	**-19.00 [-20.81 , -17.19]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 20.61 (P < 0.00001)								
Total (95% CI)			**56**			**55**	**100.0%**	**-18.77 [-20.57 , -16.98]**
Heterogeneity: Chi² = 4.37, df = 1 (P = 0.04); I² = 77%								
Test for overall effect: Z = 20.50 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.37, df = 1 (P = 0.04), I² = 77.1%								



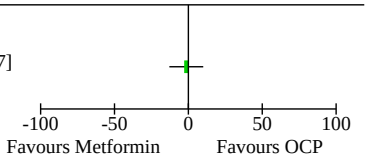
Analysis 10.7. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 7: Body Mass Index (kg/m2)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
10.7.1 BMI ≥ 30 kg/m2								
Al-Zubeidi 2015	31.9	6	10	32.6	9	12	33.1%	-0.70 [-7.01 , 5.61]
Allen 2005	36.5	5.6	16	38.9	9.3	15	44.4%	-2.40 [-7.85 , 3.05]
Hoeger 2008a	35.7	8.6	6	36.4	5.4	10	22.5%	-0.70 [-8.35 , 6.95]
Subtotal (95% CI)			**32**			**37**	**100.0%**	**-1.45 [-5.08 , 2.17]**
Heterogeneity: Chi² = 0.21, df = 2 (P = 0.90); I² = 0%								
Test for overall effect: Z = 0.79 (P = 0.43)								
Total (95% CI)			**32**			**37**	**100.0%**	**-1.45 [-5.08 , 2.17]**
Heterogeneity: Chi² = 0.21, df = 2 (P = 0.90); I² = 0%								
Test for overall effect: Z = 0.79 (P = 0.43)								
Test for subgroup differences: Not applicable								



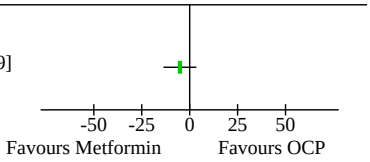
Analysis 10.8. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 8: Blood pressure - systolic (mm Hg)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Mean Difference IV, Fixed, 95% CI
10.8.1 BMI ≥ 30 kg/m2							
Hoeger 2008a	122.7	8.9	6	124.1	14.5	10	-1.40 [-12.87 , 10.07]



Analysis 10.9. Comparison 10: Adolescent - Metformin versus OCP (Clinical parameters), Outcome 9: Blood pressure - diastolic (mm Hg)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Mean Difference IV, Fixed, 95% CI
10.9.1 BMI ≥ 30 kg/m2							
Hoeger 2008a	65	8.6	6	70.1	8.3	10	-5.10 [-13.69 , 3.49]

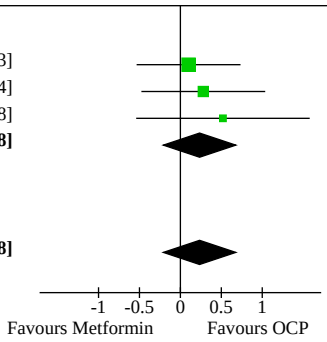


Comparison 11. Adolescent - Metformin versus OCP (Hormonal parameters)

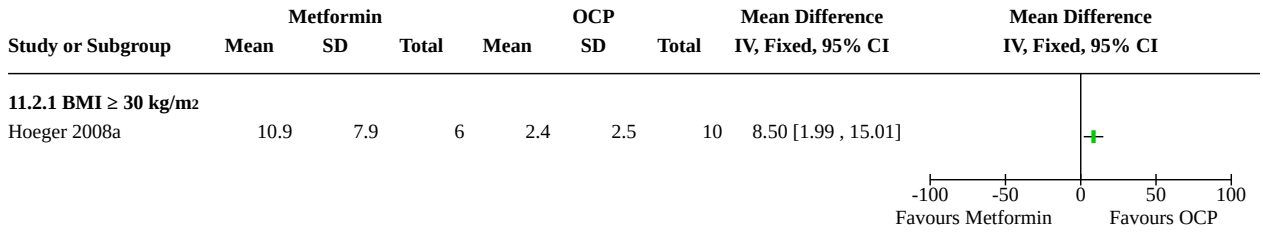
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
11.1 Serum total testosterone (nmol/L)	3	69	Mean Difference (IV, Fixed, 95% CI)	0.23 [-0.21, 0.68]
11.1.1 BMI ≥ 30 kg/m^2	3	69	Mean Difference (IV, Fixed, 95% CI)	0.23 [-0.21, 0.68]
11.2 Free androgen index (FAI) (%)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
11.2.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Analysis 11.1. Comparison 11: Adolescent - Metformin versus OCP (Hormonal parameters), Outcome 1: Serum total testosterone (nmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
11.1.1 BMI ≥ 30 kg/m2								
Al-Zubeidi 2015	1.5	0.8	10	1.4	0.7	12	48.5%	0.10 [-0.53 , 0.73]
Allen 2005	2.32	0.92	16	2.04	1.2	15	34.1%	0.28 [-0.48 , 1.04]
Hoeger 2008a	1.72	1.08	6	1.2	0.99	10	17.4%	0.52 [-0.54 , 1.58]
Subtotal (95% CI)			**32**			**37**	**100.0%**	**0.23 [-0.21 , 0.68]**
Heterogeneity: Chi² = 0.47, df = 2 (P = 0.79); I² = 0%								
Test for overall effect: Z = 1.04 (P = 0.30)								
Total (95% CI)			**32**			**37**	**100.0%**	**0.23 [-0.21 , 0.68]**
Heterogeneity: Chi² = 0.47, df = 2 (P = 0.79); I² = 0%								
Test for overall effect: Z = 1.04 (P = 0.30)								
Test for subgroup differences: Not applicable								



Analysis 11.2. Comparison 11: Adolescent - Metformin versus OCP (Hormonal parameters), Outcome 2: Free androgen index (FAI) (%)

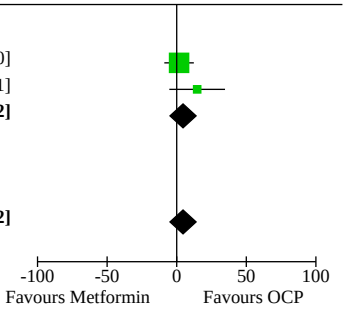


Comparison 12. Adolescent - Metformin versus OCP (Metabolic parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.1 Fasting insulin (mIU/L)	2	53	Mean Difference (IV, Fixed, 95% CI)	4.55 [-4.82, 13.92]
12.1.1 BMI ≥ 30 kg/m^2	2	53	Mean Difference (IV, Fixed, 95% CI)	4.55 [-4.82, 13.92]
12.2 Fasting glucose (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
12.2.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
12.3 Total Cholesterol (mmol/L)	2	47	Mean Difference (IV, Fixed, 95% CI)	-1.12 [-1.66, -0.58]
12.3.1 BMI ≥ 30 kg/m^2	2	47	Mean Difference (IV, Fixed, 95% CI)	-1.12 [-1.66, -0.58]
12.4 HDL Cholesterol (mmol/L)	2	47	Mean Difference (IV, Fixed, 95% CI)	0.12 [-0.10, 0.34]
12.4.1 BMI ≥ 30 kg/m^2	2	47	Mean Difference (IV, Fixed, 95% CI)	0.12 [-0.10, 0.34]
12.5 LDL Cholesterol (mmol/L)	2	47	Mean Difference (IV, Fixed, 95% CI)	-0.92 [-1.49, -0.35]
12.5.1 BMI ≥ 30 kg/m^2	2	47	Mean Difference (IV, Fixed, 95% CI)	-0.92 [-1.49, -0.35]
12.6 Triglycerides (mmol/L)	3	69	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.37, 0.10]
12.6.1 BMI ≥ 30 kg/m^2	3	69	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.37, 0.10]

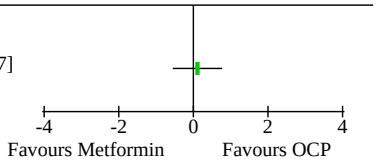
Analysis 12.1. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 1: Fasting insulin (mIU/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
12.1.1 BMI ≥ 30 kg/m2								
Al-Zubeidi 2015	18	15	10	16.3	9	12	78.1%	1.70 [-8.90 , 12.30]
Allen 2005	40.2	35.6	16	25.5	19.36	15	21.9%	14.70 [-5.31 , 34.71]
Subtotal (95% CI)			**26**			**27**	**100.0%**	**4.55 [-4.82 , 13.92]**
Heterogeneity: Chi² = 1.27, df = 1 (P = 0.26); I² = 21%								
Test for overall effect: Z = 0.95 (P = 0.34)								
Total (95% CI)			**26**			**27**	**100.0%**	**4.55 [-4.82 , 13.92]**
Heterogeneity: Chi² = 1.27, df = 1 (P = 0.26); I² = 21%								
Test for overall effect: Z = 0.95 (P = 0.34)								
Test for subgroup differences: Not applicable								



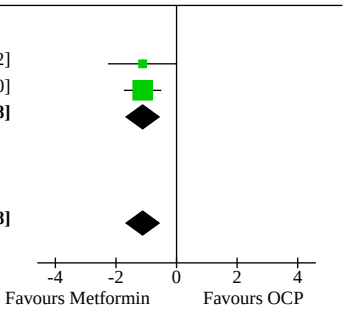
Analysis 12.2. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 2: Fasting glucose (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Mean Difference IV, Fixed, 95% CI
12.2.1 BMI ≥ 30 kg/m2							
Hoeger 2008a	4.75	0.71	6	4.64	0.55	10	0.11 [-0.55 , 0.77]



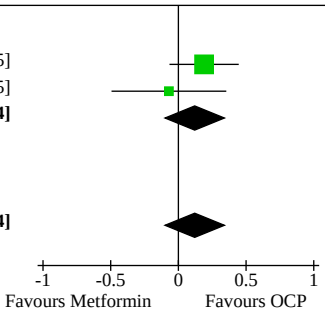
Analysis 12.3. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 3: Total Cholesterol (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
12.3.1 BMI ≥ 30 kg/m2								
Allen 2005	4.39	0.83	16	5.51	2.11	15	22.7%	-1.12 [-2.26 , 0.02]
Hoeger 2008a	3.78	0.65	6	4.9	0.54	10	77.3%	-1.12 [-1.74 , -0.50]
Subtotal (95% CI)			**22**			**25**	**100.0%**	**-1.12 [-1.66 , -0.58]**
Heterogeneity: Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 4.04 (P < 0.0001)								
Total (95% CI)			**22**			**25**	**100.0%**	**-1.12 [-1.66 , -0.58]**
Heterogeneity: Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 4.04 (P < 0.0001)								
Test for subgroup differences: Not applicable								



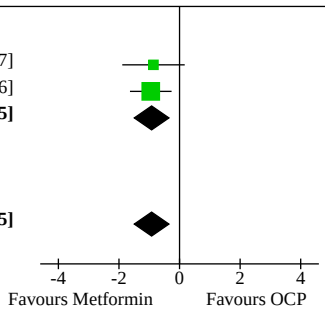
Analysis 12.4. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 4: HDL Cholesterol (mmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin Mean	Metformin SD	Metformin Total	Weight	Mean Difference IV, Fixed, 95% CI
12.4.1 BMI ≥ 30 kg/m2								
Allen 2005	1.33	0.3	15	1.14	0.42	16	73.3%	0.19 [-0.07 , 0.45]
Hoeger 2008a	1.24	0.26	10	1.31	0.49	6	26.7%	-0.07 [-0.49 , 0.35]
Subtotal (95% CI)			**25**			**22**	**100.0%**	**0.12 [-0.10 , 0.34]**
Heterogeneity: Chi² = 1.06, df = 1 (P = 0.30); I² = 6%								
Test for overall effect: Z = 1.08 (P = 0.28)								
Total (95% CI)			**25**			**22**	**100.0%**	**0.12 [-0.10 , 0.34]**
Heterogeneity: Chi² = 1.06, df = 1 (P = 0.30); I² = 6%								
Test for overall effect: Z = 1.08 (P = 0.28)								
Test for subgroup differences: Not applicable								



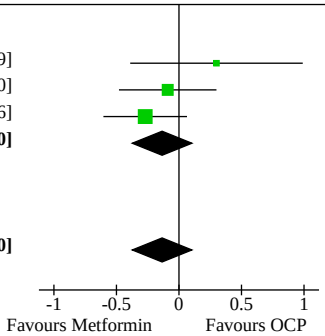
Analysis 12.5. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 5: LDL Cholesterol (mmol/L)

Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
12.5.1 BMI ≥ 30 kg/m2								
Allen 2005	2.6	0.73	16	3.46	1.91	15	30.7%	-0.86 [-1.89 , 0.17]
Hoeger 2008a	2.39	0.4	6	3.34	0.98	10	69.3%	-0.95 [-1.64 , -0.26]
Subtotal (95% CI)			**22**			**25**	**100.0%**	**-0.92 [-1.49 , -0.35]**
Heterogeneity: Chi² = 0.02, df = 1 (P = 0.89); I² = 0%								
Test for overall effect: Z = 3.16 (P = 0.002)								
Total (95% CI)			**22**			**25**	**100.0%**	**-0.92 [-1.49 , -0.35]**
Heterogeneity: Chi² = 0.02, df = 1 (P = 0.89); I² = 0%								
Test for overall effect: Z = 3.16 (P = 0.002)								
Test for subgroup differences: Not applicable								



Analysis 12.6. Comparison 12: Adolescent - Metformin versus OCP (Metabolic parameters), Outcome 6: Triglycerides (mmol/L)

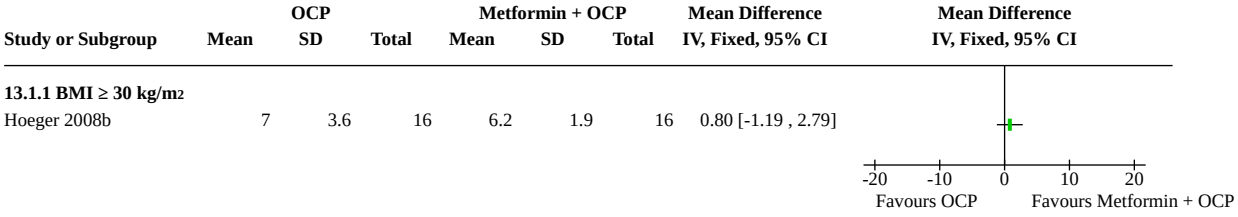
Study or Subgroup	Metformin Mean	Metformin SD	Metformin Total	OCP Mean	OCP SD	OCP Total	Weight	Mean Difference IV, Fixed, 95% CI
12.6.1 BMI ≥ 30 kg/m2								
Al-Zubeidi 2015	1.51	0.84	10	1.21	0.8	12	11.9%	0.30 [-0.39 , 0.99]
Allen 2005	1.4	0.66	16	1.49	0.43	15	37.3%	-0.09 [-0.48 , 0.30]
Hoeger 2008a	0.78	0.23	6	1.05	0.45	10	50.8%	-0.27 [-0.60 , 0.06]
Subtotal (95% CI)			**32**			**37**	**100.0%**	**-0.13 [-0.37 , 0.10]**
Heterogeneity: Chi² = 2.21, df = 2 (P = 0.33); I² = 9%								
Test for overall effect: Z = 1.11 (P = 0.27)								
Total (95% CI)			**32**			**37**	**100.0%**	**-0.13 [-0.37 , 0.10]**
Heterogeneity: Chi² = 2.21, df = 2 (P = 0.33); I² = 9%								
Test for overall effect: Z = 1.11 (P = 0.27)								
Test for subgroup differences: Not applicable								



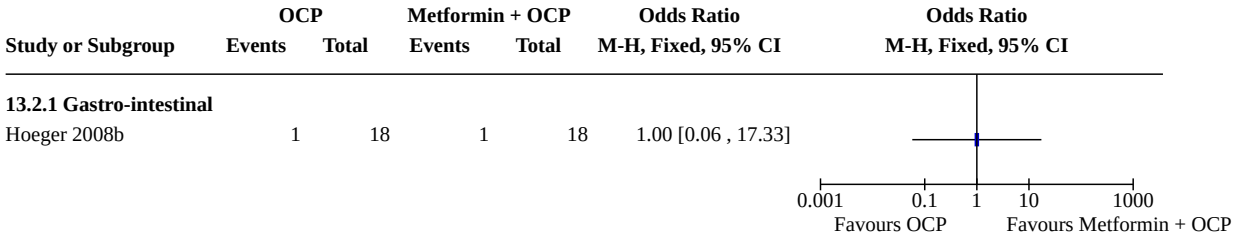
Comparison 13. Adolescent - OCP versus Metformin combined with OCP (Clinical parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.1 Hirsutism - Clinical F-G score	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.1.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.2 Adverse events - severe	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
13.2.1 Gastro-intestinal	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
13.3 Body Mass Index (kg/m2)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.3.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.4 Blood Pressure - Systolic (mmHg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.4.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.5 Blood Pressure - Diastolic (mmHg)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
13.5.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Analysis 13.1. Comparison 13: Adolescent - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 1: Hirsutism - Clinical F-G score

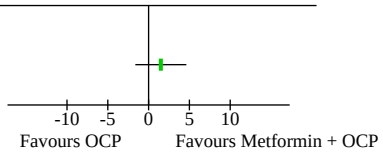


Analysis 13.2. Comparison 13: Adolescent - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 2: Adverse events - severe



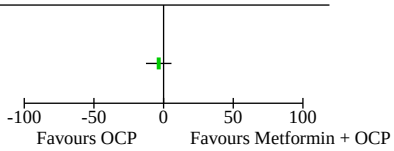
Analysis 13.3. Comparison 13: Adolescent - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 3: Body Mass Index (kg/m2)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
13.3.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	33.9	4.2	16	32.4	4.8	16	1.50 [-1.63 , 4.63]



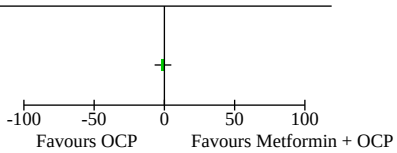
Analysis 13.4. Comparison 13: Adolescent - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 4: Blood Pressure - Systolic (mmHg)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
13.4.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	115.7	13.1	16	119.2	13.3	16	-3.50 [-12.65 , 5.65]



Analysis 13.5. Comparison 13: Adolescent - OCP versus Metformin combined with OCP (Clinical parameters), Outcome 5: Blood Pressure - Diastolic (mmHg)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
13.5.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	62.1	11	16	63.1	5.1	16	-1.00 [-6.94 , 4.94]

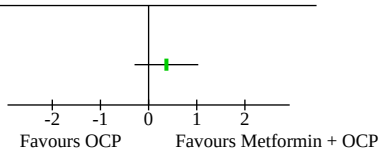


Comparison 14. Adolescent - OCP versus Metformin combined with OCP (Hormonal parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
14.1 Serum total testosterone (nmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
14.1.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
14.2 Free androgen index (FAI) (%)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
14.2.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

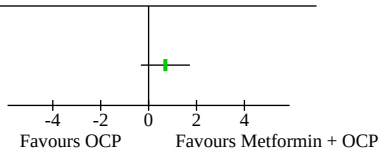
Analysis 14.1. Comparison 14: Adolescent - OCP versus Metformin combined with OCP (Hormonal parameters), Outcome 1: Serum total testosterone (nmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
14.1.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	1.93	1.13	16	1.56	0.74	16	0.37 [-0.29 , 1.03]



Analysis 14.2. Comparison 14: Adolescent - OCP versus Metformin combined with OCP (Hormonal parameters), Outcome 2: Free androgen index (FAI) (%)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
14.2.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	1.1	2	16	0.4	0.6	16	0.70 [-0.32 , 1.72]

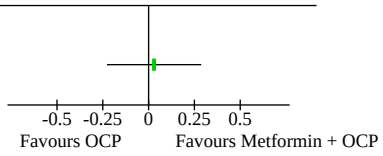


Comparison 15. Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
15.1 Fasting glucose (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.1.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.2 Total Cholesterol (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.2.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.3 HDL Cholesterol (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.3.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.4 LDL Cholesterol (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.4.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.5 Triglycerides (mmol/L)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
15.5.1 BMI ≥ 30 kg/m^2	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

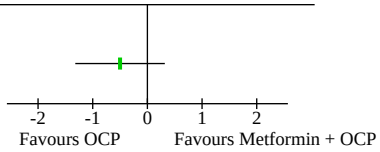
Analysis 15.1. Comparison 15: Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 1: Fasting glucose (mmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
15.1.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	4.6	0.34	16	4.57	0.4	16	0.03 [-0.23 , 0.29]



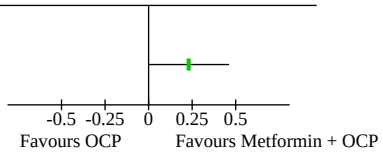
Analysis 15.2. Comparison 15: Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 2: Total Cholesterol (mmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
15.2.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	4.73	1.24	16	5.23	1.12	16	-0.50 [-1.32 , 0.32]



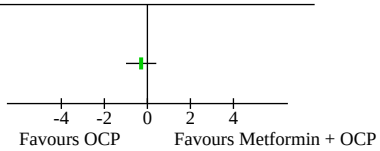
Analysis 15.3. Comparison 15: Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 3: HDL Cholesterol (mmol/L)

Study or Subgroup	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	OCP Mean	OCP SD	OCP Total	Mean Difference IV, Fixed, 95% CI
15.3.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	1.51	0.33	16	1.28	0.34	16	0.23 [-0.00 , 0.46]

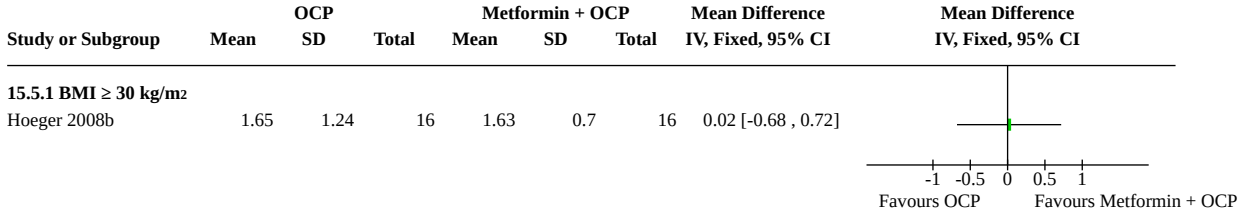


Analysis 15.4. Comparison 15: Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 4: LDL Cholesterol (mmol/L)

Study or Subgroup	OCP Mean	OCP SD	OCP Total	Metformin + OCP Mean	Metformin + OCP SD	Metformin + OCP Total	Mean Difference IV, Fixed, 95% CI
15.4.1 BMI ≥ 30 kg/m2							
Hoeger 2008b	2.71	1.03	16	3	1	16	-0.29 [-0.99 , 0.41]



Analysis 15.5. Comparison 15: Adolescent - OCP versus Metformin combined with OCP (Metabolic parameters), Outcome 5: Triglycerides (mmol/L)



ADDITIONAL TABLES

Table 1. Conversion Factors

	Convert from	Convert to	Conversion factor
Androstenedione	ng/mL	nmol/L	3.49
Cholesterol	mg/dL	mmol/L	0.026
Confidence intervals	Confidence intervals	Standard error	(upper limit - lower limit)/3.92
Glucose	mg/dL	mmol/L	0.056
Insulin	pmol/L	mIU/L (= microIU/mL)	0.1667
Sex hormone-binding globulin	mcg/dL	nmol/L	34.7
Standard deviation	Standard error	Standard deviation	Sqrt n
Testosterone	pg/mL	pmol/L	3.47
Triglycerides	mg/dL	mmol/L	0.011

APPENDICES

Appendix 1. Cochrane Gynaecology and Fertility Group specialised register search strategy

PROCITE platform

Searched 15 August 2019

Keywords CONTAINS "polycystic ovary morphology" or "polycystic ovary syndrome" or "PCOS" or Title CONTAINS "polycystic ovary morphology" or "polycystic ovary syndrome" or "PCOS"

AND

Keywords CONTAINS "metformin" or "glucophage" or Title CONTAINS "metformin" or "glucophage"

(529 records)

Appendix 2. CENTRAL search strategy

OVID platform

Issue July 2019

Searched 15 August 2019

1 exp Polycystic Ovary Syndrome/ (1343)
2 Polycystic Ovar$.tw. (3279)
3 PCO$.tw. (4732)
4 (stein-leventhal or leventhal).tw. (46)
5 (ovar$ adj2 (sclerocystic or polycystic)).tw. (3284)
6 or/1-5 (5656)
7 exp Metformin/ (3469)
8 Metformin.tw. (8817)
9 (dimethylguanyl$ or glucophage).tw. (173)
10 or/7-9 (9080)
11 6 and 10 (1127)

Appendix 3. MEDLINE search strategy

OVID platform

Searched from 1946 to 15 August 2019

1 exp Polycystic Ovary Syndrome/ (13678)
2 Polycystic Ovar$.tw. (15614)
3 PCO$.tw. (29517)
4 (stein-leventhal or leventhal).tw. (721)
5 (ovar$ adj2 (sclerocystic or polycystic)).tw. (15745)
6 or/1-5 (36955)
7 exp Metformin/ (12327)
8 Metformin.tw. (18275)
9 (dimethylguanyl$ or glucophage).tw. (117)
10 or/7-9 (20052)
11 6 and 10 (1650)
12 randomized controlled trial.pt. (487269)
13 controlled clinical trial.pt. (93212)
14 randomized.ab. (451880)
15 randomised.ab. (90104)
16 placebo.tw. (205530)
17 clinical trials as topic.sh. (187968)
18 randomly.ab. (316329)
19 trial.ti. (203246)
20 (crossover or cross-over or cross over).tw. (81309)
21 or/12-20 (1294842)
22 exp animals/ not humans.sh. (4608593)
23 21 not 22 (1191300)
24 11 and 23 (626)

Appendix 4. Embase search strategy

OVID platform

Searched from 1980 to 15 August 2019

1 exp ovary polycystic disease/ (25367)
2 Polycystic Ovar$.tw. (21864)
3 PCO$.tw. (39029)
4 (stein-leventhal or leventhal).tw. (305)
5 (ovar$ adj2 (sclerocystic or polycystic)).tw. (21950)
6 or/1-5 (51262)
7 exp Metformin/ (58473)
8 (dimethylguanyl$ or glucophage).tw. (1598)

9 Metformin.tw. (31941)
10 or/7-9 (60528)
11 6 and 10 (4078)
12 Clinical Trial/ (951648)
13 Randomized Controlled Trial/ (559952)
14 exp randomization/ (83672)
15 Single Blind Procedure/ (36128)
16 Double Blind Procedure/ (161038)
17 Crossover Procedure/ (60118)
18 Placebo/ (325743)
19 Randomi?ed controlled trial$.tw. (208563)
20 Rct.tw. (33370)
21 random allocation.tw. (1896)
22 randomly allocated.tw. (32956)
23 allocated randomly.tw. (2466)
24 (allocated adj2 random).tw. (807)
25 Single blind$.tw. (23129)
26 Double blind$.tw. (193914)
27 ((treble or triple) adj blind$).tw. (984)
28 placebo$.tw. (288469)
29 prospective study/ (540295)
30 or/12-29 (2062347)
31 case study/ (63160)
32 case report.tw. (378294)
33 abstract report/ or letter/ (1067463)
34 or/31-33 (1499125)
35 30 not 34 (2011075)
36 11 and 35 (1544)

Appendix 5. PsycINFO search strategy

OVID platform

Searched from 1806 to 15 August 2019

1 exp endocrine sexual disorders/ (1713)
2 Polycystic Ovar$.tw. (398)
3 PCO$.tw. (922)
4 (stein-leventhal or leventhal).tw. (293)
5 (ovar$ adj2 (sclerocystic or polycystic)).tw. (400)
6 or/1-5 (2912)
7 Metformin.tw. (439)
8 (dimethylguanyl$ or glucophage).tw. (2)
9 or/7-8 (439)
10 random*.ti,ab,hw,id. (190656)
11 trial*.ti,ab,hw,id. (174837)
12 controlled stud*.ti,ab,hw,id. (11886)
13 placebo*.ti,ab,hw,id. (39359)
14 ((singl* or doubl* or trebl* or tripl*) and (blind* or mask*)).ti,ab,hw,id. (28260)
15 (cross over or crossover or factorial* or latin square).ti,ab,hw,id. (29300)
16 (assign* or allocat* or volunteer*).ti,ab,hw,id. (158403)
17 treatment effectiveness evaluation/ (23280)
18 mental health program evaluation/ (2073)
19 exp experimental design/ (55841)
20 or/10-19 (499609)
21 6 and 9 and 20 (7)

Appendix 6. CINAHL search strategy

EBSCO platform

Searched from 1961 to 15 August 2019

#	Query	Results
S23	S10 AND S22	203
S22	S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21	1,342,371
S21	TX allocat* random*	10,805
S20	(MH "Quantitative Studies")	23,098
S19	(MH "Placebos")	11,405
S18	TX placebo*	58,145
S17	TX random* allocat*	10,805
S16	(MH "Random Assignment")	56,076
S15	TX randomi* control* trial*	173,085
S14	TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*))	1,027,527
S13	TX clinic* n1 trial*	248,146
S12	PT Clinical trial	86,817
S11	(MH "Clinical Trials+")	264,145
S10	S5 AND S9	500
S9	S6 OR S7 OR S8	6,907
S8	TX dimethylguanyl* or TX glucophage	40
S7	TX metformin	6,902
S6	(MM "Metformin")	2,650
S5	S1 OR S2 OR S3 OR S4	4,665
S4	TX polycystic ovar*	4,065
S3	TX stein leventhal syndrome	10
S2	TX PCOS or TX PCOD	2,507
S1	(MM "Polycystic Ovary Syndrome")	2,542

WHAT'S NEW

Date	Event	Description
9 March 2020	New search has been performed	Review updated; 38 new studies included.
9 March 2020	New citation required and conclusions have changed	The addition of 38 new studies has led to a change in conclusions.

HISTORY

Protocol first published: Issue 4, 2005
Review first published: Issue 1, 2007

Date	Event	Description
12 November 2008	Amended	Converted to new review format.
15 November 2006	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

Eloise Fraison: assessment of literature search for included relevant trials; extracted data; evaluated the quality of trials and risk of bias; analysis and interpretation of data, revised the review critically for important intellectual content; approved the final draft for publication.

Lisa Moran: assessment of literature search for included relevant trials; extracted data; evaluated the risk of bias of trials; revised the review critically for important intellectual content; approved the final draft for publication.

Sophia Bilal: extracted data; revised the review critically for important intellectual content; approved the final draft for publication.

Carolyn Ee: extracted data; revised the review critically for important intellectual content; approved the final draft for publication.

Elena Kostova: provided consultation on data analysis, evaluation of the quality of trials and risk of bias and interpretation of data, revised the review critically for important intellectual content; approved the final draft for publication.

Christos Venetis: assessment of literature search for included relevant trials; revised the review critically for important intellectual content; approved the final draft for publication.

Michael Costello: initiated and conceptualised the review; assessment of literature search for included relevant trials; evaluated the quality of trials and risk of bias; interpretation of data; revised the review critically for important intellectual content; approved the final draft for publication.

DECLARATIONS OF INTEREST

M Costello has declared shares in Virtus Health and past sponsorship from Merck Serono for scientific conference presentations.

E Fraison has no conflicts of interest to declare.

L Moran has received a National Heart Foundation Future Leader Fellowship that has funded her work on this review.

S Bilal has no conflicts of interest to declare.

C Ee declares that as a medical research institute, NICM Health Research Institute receives research grants and donations from foundations, universities, government agencies and industry. Sponsors and donors provide untied and tied funding for work to advance the vision and mission of the institute. She confirms that she personally received no industry funding relating to this review.

E Kostova has no conflicts of interest to declare.

C Venetis reports an early career Fellowship from National Health and Medical Research Council of Australia which supported him for this work, personal fees and non-financial support from Merck Sharpe & Dohme, Merck, Beisins, research grants and non-financial support from Ferring, outside the submitted work.

SOURCES OF SUPPORT

Internal sources

- University of New South Wales, Sydney, Australia

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Changes in 2020 updated review and the 2007 original review

1. The title of the review was changed from "Insulin-sensitising drugs versus the combined oral contraceptive pill for hirsutism, acne and risk of diabetes, cardiovascular disease, and endometrial cancer in polycystic ovary syndrome" to "Metformin versus the combined oral contraceptive pill for hirsutism, acne, and menstrual pattern in polycystic ovary syndrome" in order to reflect:

(i) that only the insulin-sensitising drug to be included as an intervention for the review update was metformin in view metformin being by far the most commonly used insulin-sensitising drug in PCOS and adverse cardiovascular disease risks of rosiglitazone and pioglitazone limiting their use in PCOS. Therefore, studies using the insulin-sensitising drugs rosiglitazone, pioglitazone, and D-chiro-inositol were excluded;

(ii) the removal of a number of outcomes - see point number 3 below.

2. The Methods section was revised to bring this section to current Cochrane standards.

3. The following outcomes have been removed as the old review had too many outcomes compared to current Cochrane standards:

a. cardiovascular disease (stroke, myocardial infarction) event: too rare an outcome over too long a time period (decades) and not reported in RCTs;

b. endometrial cancer event: too rare an outcome over too long a time period (decades) and not reported in RCTs;

c. waist circumference (cm): less clinically relevant than other clinical outcomes;

d. waist-Hip ratio: less clinically relevant than other clinical outcomes;

e. quality of life score: rarely reported in RCTs;

f. Serum free testosterone (pmol/L): direct measurement of free testosterone is inaccurate as per new international PCOS guidelines 2018 (Teede 2018);

g. Sex hormone-binding globulin (SHBG) (nmol/L): less hormonally relevant than other hormonal outcomes and its main purpose is to calculate free androgen index (FAI).

4. Where data were available, we conducted pre-specified subgroup analyses to determine the separate evidence within the following subgroups for all outcomes (see subgroup analysis and investigation of heterogeneity section in Methods):

(i) all outcomes (apart from the outcome of adverse events) were divided or subgrouped according to studies of women with different mean BMI (e.g. BMI ≤ 25 kg/m^2, BMI > 25 kg/m^2 but < 30 kg/m^2, BMI ≥30 kg/m^2) (see subgroup analysis and investigation of heterogeneity section in Methods);

(ii) adverse events outcomes were divided or subgrouped according to type of adverse event (gastro-intestinal or other adverse events) measured.

NOTES

Conflict of interest added

INDEX TERMS

Medical Subject Headings (MeSH)

Acne Vulgaris [drug therapy]; Body Mass Index; Cardiovascular Diseases [prevention & control]; Contraceptives, Oral, Combined [adverse effects] [*therapeutic use]; Drug Therapy, Combination; Endometrial Neoplasms [prevention & control]; Hirsutism [*drug therapy]; Hypoglycemic Agents [*therapeutic use]; Menstruation Disturbances [*drug therapy]; Metformin [adverse effects] [*therapeutic use]; Polycystic Ovary Syndrome [complications] [*drug therapy]; Randomized Controlled Trials as Topic

MeSH check words

Adolescent; Adult; Female; Humans; Young Adult