Received: 20 July 2018 | Accepted: 13 February 2019

DOI: 10.1111/aogs.13586

SYSTEMATIC REVIEW

AOGS Acta Obstetricia et Gynecologica Scandinavica

Should progesterone and estrogen receptors be assessed for predicting the response to conservative treatment of endometrial hyperplasia and cancer? A systematic review and meta-analysis

Antonio Raffone[1] | Antonio Travaglino[2] | Gabriele Saccone[1] | Antonio Mollo[1] | Giuseppe De Placido[1] | Luigi Insabato[2] | Fulvio Zullo[1]

[1]Obstetrics and Gynecology Unit, Department of Neuroscience, Reproductive Sciences and Dentistry, School of Medicine, University of Naples Federico II, Naples, Italy

[2]Anatomic Pathology Unit, Department of Advanced Biomedical Sciences, School of Medicine, University of Naples Federico II, Naples, Italy

Correspondence
Antonio Travaglino, Anatomic Pathology Unit, Department of Advanced Biomedical Sciences, School of Medicine, University of Naples Federico II, Naples, Italy.
Email: antonio.travaglino.ap@gmail.com

Abstract

Introduction: Progestins are used as conservative treatment of endometrial hyperplasia (EH) and early endometrial cancer (EEC). We aimed to assess whether immunohistochemical expression of estrogens and progesterone receptors (ER and PR) predicts the treatment response.

Material and methods: Electronic databases were searched for studies assessing ER and PR expression in EH and EEC treated with progestins. Relative risk for poor response, sensitivity, specificity, diagnostic odds ratio positive and negative likelihood ratios (LR^+ and LR^-) and area under the curve (AUC) on summary receiver operating characteristic curve were calculated. Subgroup analyses were based on administration route (oral progestin or levonorgestrel-intrauterine device) and on histological diagnosis (atypical EH/EEC or non-atypical EH). Only high accuracy (AUC > .9; LR^+ >10; LR^- <.1) was considered determining for the clinical practice.

Results: Thirteen studies with 635 patients were included in the systematic review. Studies at high risk of bias were excluded from the meta-analysis. Negative ER expression did not significantly predict poor response (P = .16), with low predictive accuracy (AUC = .637). Negative PR significantly predicted poor response (P = .01), with moderate accuracy (AUC = .806). In the oral progestin subgroup, neither ER (P = .55) nor PR (P = .18) had significant predictive value. In the levonorgestrel-intrauterine device subgroup, both ER (P < .0001) and PR (P = .02) were significantly predictive of good response, although the accuracy was suboptimal (LR^+ 6.02 and 2.48, respectively; LR^- .59 and .55, respectively). The atypical EH/EEC subgroup showed non-significant results. Data about non-atypical EH were not extractable.

Conclusions: ER and PR expressions are significantly predictive of response in EH and EEC treated with a levonorgestrel-intrauterine device but not with oral progestins. However, their accuracy is insufficient to be determining in the clinical practice.

Abbreviations: AUC, area under the curve; BMI, body mass index; CI, confidence interval; DOR, diagnostic odds ratio; EEC, early endometrial cancer; EH, endometrial hyperplasia; ER, estrogen receptor; LNG-IUD, levonorgestrel-intrauterine device; LR^+, positive likelihood ratio; LR^-, negative likelihood ratio; PR, progesterone receptor; RR, relative risk; SROC, summary receiver operating characteristic.

KEYWORDS
endometrial cancer, endometrial hyperplasia, endometrial intraepithelial neoplasia, estrogen receptor, hormonal therapy, levonorgestrel-intrauterine device, predictive markers, progesterone receptor, progestin

1 | INTRODUCTION

Endometrial hyperplasia (EH) is an irregular proliferation of endometrial glands which often precedes endometrial cancer of endometrioid type.[1-3] Several studies have shown that EH includes both hyperproliferative reactions to an unbalanced action of estrogens and true precancerous lesions.[1,2,4]

The revised 2014 WHO classification of EH differentiates between these two conditions based on the presence of cytological atypia, identifying EH without atypia (benign) and atypical EH (premalignant).[1,2] In patients with EH without atypia, the 20-year risk of progression to cancer is indeed <5%[5]; thus, these patients may be managed with observation alone and follow up biopsies, while progestins are recommended in symptomatic cases.[6]

On the other hand, atypical EH requires a total hysterectomy, although progestins may be used in women who wish to preserve their fertility or who are not suitable for surgery.[6] Such a conservative approach can still be used in the well differentiated, endometrioid type endometrial cancer at stage FIGO Ia without tumor invasion of myometrium (early endometrial cancer, EEC).[7]

Although progestins are widely used in young women with EH and EEC, they are not always effective, and patients who do not respond are at risk of progression to invasive disease.[8]

Thus, great efforts to find predictive markers of response to progestins have been made in the last years, including studies of clinical, pathological and immunohistochemical features.[9-11]

In particular, most studies focused on the estrogen receptor (ER) and progesterone receptor (PR), whose expression is easily assessable by immunohistochemistry.[12-24] Estrogens are indeed involved in the development of EH and EC, and progestins mediate their action through PR.[2,12] In spite of this, results are conflicting and the possible predictive role of ER and PR is still undefined.

The aim of this systematic review and meta-analysis was to assess whether the expression of ER and PR can predict the response to conservative treatment in EH and EEC.

Key message

In endometrial hyperplasia and cancer, estrogen or progesterone receptor expression predicts the response to a levonorgestrel-intrauterine device but not to oral progestins. Their accuracy is insufficient to be routinely used.

2 | MATERIAL AND METHODS

This study was performed following the SEDATE guidelines.[25] The study protocol was designed a priori, defining methods for collecting, extracting and analyzing data. All review stages were conducted independently by two reviewers (A.R., A.T.). The two authors independently assessed electronic search, eligibility of the studies, inclusion criteria, risk of bias, data extraction and data analysis. Disagreements were resolved by discussion with a third reviewer (G.S.).

Two reviewers (A.R., A.T.) independently conducted several researches using MEDLINE, Embase, Web of Sciences, Scopus, ClinicalTrial.gov, OVID and Cochrane Library as electronic databases. A combination of the following text words from the inception of each database to June 2018 was used: endometrial hyperplasia; endometrial cancer; endometrioid adenocarcinoma; endometrial intraepithelial neoplasia; EIN; therapy; treatment; fertility sparing; conservative; medroxyprogesterone; MPA; mirena; LNG; levonorgestrel; progestogen; progestin; response; resistance; persistence; outcome; progesterone receptor; PR; estrogen receptor; ER; marker; immunohistochemistry; immunohistochemical. Review of articles also included the abstracts of all references retrieved from the search.

This systematic review included all studies meeting the following inclusion criteria:

- study population constituted of women diagnosed with EH or EEC and conservatively treated with progestins;
- assessment of the expression of the marker (ER or PR) on pretreatment endometrial specimens by immunohistochemistry;
- assessment of the association between the expression of the marker and the response to therapy.

The revised tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews (QUADAS-2)[26] was used to assess the risk of bias in each study. Review authors' judgments were categorized as "low risk", "high risk" or "unclear risk of bias" for each of the four domains: (1) Patient selection (low risk if all eligible consecutive patients were included; unclear risk if authors did not specify the number of patients not included; high risk if the study was designed as case-control); (2) Index test (low risk if authors clearly differentiated between low expression and negative expression of hormone receptors through appropriate criteria; unclear risk if criteria used were incomplete; high risk if there was no differentiation between low expression and negative expression or if individual patient data were not reported); (3)

TABLE 1 Characteristics of the included studies

Year	Study [Ref]	Country	Study design	Period of enrollment	Sample size	Patients' features		Sampling method	Histology			Progestin administered						Follow-up duration	Response	
						Age	BMI		HWA	AH	EEC	MG	NT	MP	OR	LNG	MIX		Good	Poor
2003	Utsunomiya et al[12]	Japan	Retrospective	1994-2001	16	26-38	n.r.	Curettage	—	—	16	—	—	16	—	—	—	6-12 mo	11	5
2006	Vereide et al[13]	Norway	Prospective	n.r.	50	30-70	18-43	Curettage, suction	37	13	—	—	—	29	—	21	—	3 mo	36	14
2007	Minaguchi et al[14]	Japan	Retrospective	1989-2003	31	19-60	n.r.	Curettage	—	12	19	—	—	31	—	—	—	2-18 mo	26	5
	Yamazawa et al[15]	Japan	Prospective	1999-2005	9	28-40	n.r.	Curettage	—	—	9	—	—	9	—	—	—	6-9 mo	7	2
2009	Kashima et al[16]	Japan	Retrospective	1996-2003	15	20-49	n.r.	n.r.	2	8	5	—	—	15	—	—	—	4-6 mo	10	5
2010	Akesson et al[17]	UK	Prospective	1999-2004	34	36-77	21-49	n.r.	29	5	—	—	—	—	—	34	—	≥6 mo	28	6
	Orbo et al[18]	Norway	Retrospective	1999-2004	41	32-55	20-43	Curettage, pipelle	39	2	—	—	—	16	—	25	—	6 mo	36	5
2011	Kamoi et al[19]	Japan	Retrospective	1997-2006	7	20-36	23-49	Curettage	—	—	7	—	—	7	—	—	—	3-8 mo	5	2
2012	Upson et al[20]	USA	Retrospective	1985-2005	114	<39 to >70	<25 to >30	n.r.	73	41	—	46	9	50	—	—	9	n.r.	81	33
2013	Gallos et al[21]	UK	Prospective	1998-2007	174	<40 to >60	n.r.	Office biopsy	155	19	—	—	—	—	—	174	—	n.r.	164	10
2014	Gunderson et al[22]	USA	Retrospective	1997-2012	46	24-63	18-70	n.r.	—	17	29	41	—	12	5	14	20	1-84 mo	30	16
2015	Yang et al[23]	China	Retrospective	2001-2010	88	24-39	17-45	Curettage, biopsy	—	37	51	45	11	29	—	31[e]	—	5-14 mo	77	11
	Reyes et al[24]	USA	Retrospective	n.r.	10	28-63	39-72	Curettage	—	8	2	—	—	—	—	10	—	6-50 mo	7	3
Total				1985-2012	635	19-77	17-70	—	335	162	138	132	20	214	5	309	29	1-84 mo	518	117

BMI, body mass index; HWA, hyperplasia without atypia; AH, atypical hyperplasia; EEC, early endometrial cancer; MG, megestrol acetate; NT, norethindrone acetate; MP, medroxyprogesterone acetate; OR, not specified oral progestins; LNG, levonorgestrel-intrauterine device; MIX, mixture of two or more progestins.

Reference test (low risk if good response was defined as absence of any lesions; unclear risk if histologic criteria defining good response were incomplete; high risk if authors considered even a partial regression as good response); (4) Flow and timing (low risk if the follow up was at least 3 months; unclear risk if the follow up was 1-3 months; high risk if follow up was <1 months or if hormone receptors were assessed on biopsies withdrawn after the beginning of the treatment). Domain 2 was assessed separately for ER and PR.

Data were extracted from each study without modification and reported in 2 × 2 contingency tables. For each marker, two dichotomous qualitative variables were assessed: immunohistochemical expression on pretreatment biopsy ("positive" vs "negative") and response to conservative therapy ("good" vs "poor").

The expression of the marker was considered "negative" if the percentage of immunostained cells was <10%; otherwise, the expression was considered "positive".

"Good response" indicated a complete regression of the lesion, whereas "poor response" indicated partial or no regression.

If discrepancies between text and tables were found, values from the tables were used.

Data were also subdivided into subgroups based on:

- administration route of progestins (oral or intrauterine), due to the recognized superiority of levonorgestrel-medicated intrauterine device (LNG-IUD)[27-30];
- histological diagnosis (atypical EH/EEC or non-atypical EH), since the first two are neoplastic lesions and the latter one is a reactive condition.[31-34]

The predictive value of ER and PR was assessed as relative risk (RR) for failure of therapy, with 95% confidence interval (CI). RR was calculated for each study and as pooled estimate and reported graphically on a Forest plot. $P < .05$ was considered significant.

The statistical heterogeneity among studies was assessed using the inconsistency index (I^2): heterogeneity was considered insignificant for I^2 <25%, low for I^2 <50%, moderate for I^2 <75% and high for I^2 ≥75%. If I^2 was <50%, a fixed effect model was adopted; otherwise, a random effects model was preferred.

The risk of bias across studies (publication bias) was assessed by reporting the results on a funnel plot.

We calculated the predictive accuracy of the immunohistochemical assessment of ER and PR to define its clinical usefulness in predicting the response to conservative treatment of EH and EEC. Sensitivity, specificity, positive likelihood ratio (LR^+), negative likelihood ratio (LR^-) and diagnostic odds ratio (DOR) were calculated for each study and as pooled estimate, and reported graphically on Forest plots, with 95% confidence intervals (CI).

Area under the curve (AUC) was calculated on summary receiver operating characteristic (SROC) curves. The predictive accuracy was considered low for AUC ≤.75, moderate for .75 < AUC ≤ .9, high for .9 < AUC < .97, and very high for AUC ≥.97.[25]

The random effect model of DerSimonian and Laird was chosen a priori, since an actual heterogeneity is expected in a meta-analysis of diagnostic and predictive accuracy.[25]

We defined a priori that only a high predictive accuracy (AUC >.9) would have indicated an actual clinical usefulness, since conservatively treated patients with atypical EH and EEC are already closely followed with endometrial biopsies every 3-6 months due to the risk of progression to invasive disease. If SROC calculation was not available due to an insufficient number of studies, we adopted LR^+ and LR^- as surrogates, which would have been >10 and <.1, respectively, to indicate an actual clinical usefulness.[25]

Data analysis was performed using REVIEW MANAGER 5.3 (Copenhagen: The Nordic Cochrane Centre, Cochrane Collaboration, 2014) and META-DISC version 1.4 (Clinical Biostatistics Unit, Ramon y Cajal Hospital, Madrid, Spain).

3 | RESULTS

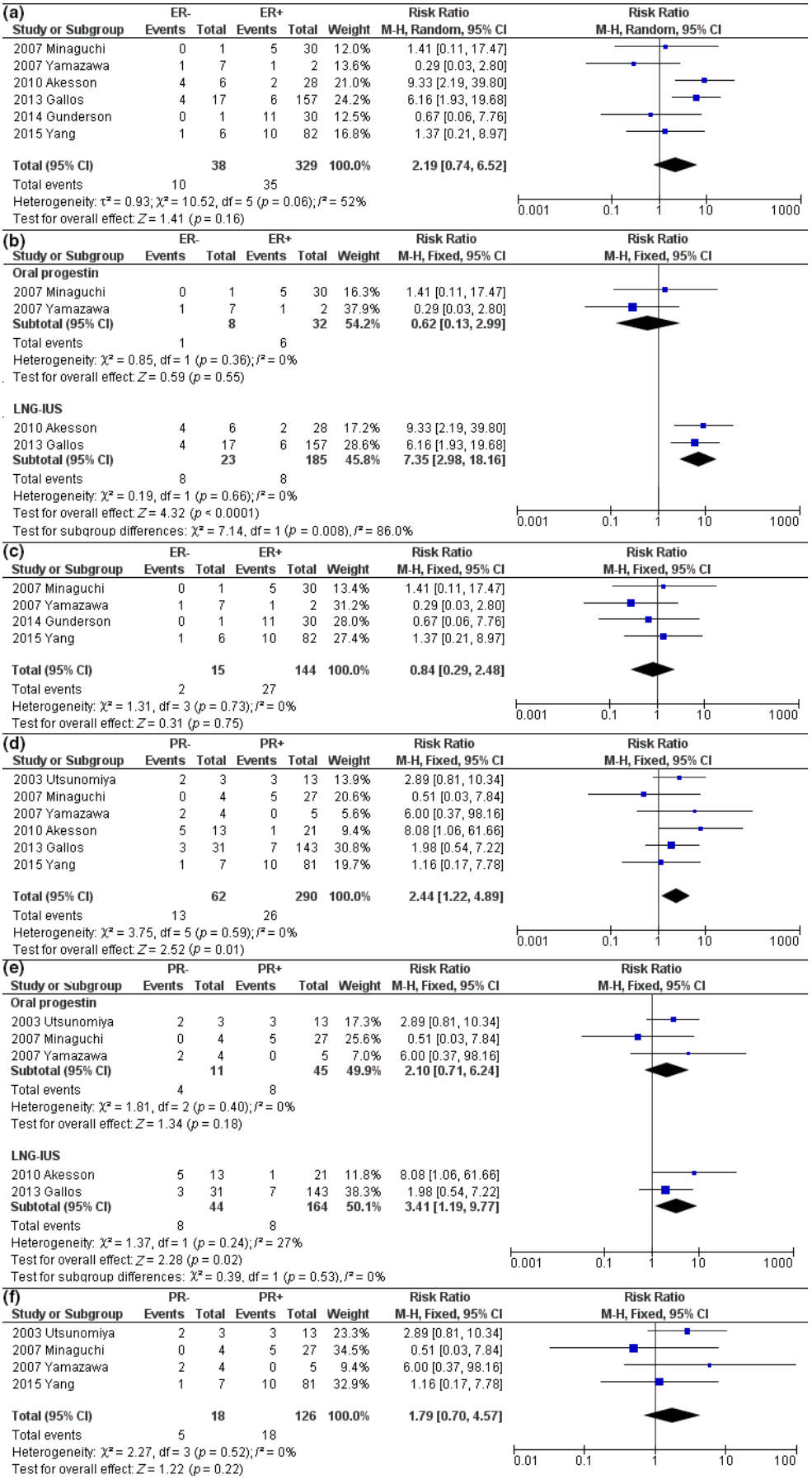
Thirteen studies with a total of 635 patients were included in the systematic review.[12-24] The whole process of study selection is reported in Supporting Information Figure S1.

Four studies were prospective and 9 were retrospective. Sample assessed included 138 EEC, 162 atypical EH and 335 non-atypical EH. All EEC were G1 endometrioid adenocarcinoma. Twelve of 13 studies evaluated ER, and all 13 studies evaluated PR. Patient age ranged between 19 and 77 years; body mass index (BMI) ranged between 17 and 70 kg/m^2. Sampling methods included curettage, Pipelle biopsy and hysteroscopic biopsy. Progestins administered included LNG-IUD (n = 309), medroxyprogesterone acetate (n = 214), megestrol acetate (n = 132), norethindrone acetate (n = 20) or a mixture of two or more progestins (n = 29).

Characteristics of each included studies are reported in detail in Table 1.

Regarding risk of bias within study assessment, for the "Patient selection" domain, 3 studies were considered to be low risk of bias and 10 were considered unclear risk (because they did clearly not state that patients were consecutive[12-14,16,18,20,22-24] and/or that only cancers with initial myometrial invasion were excluded[16,19,22,24]).

FIGURE 1 Forest plots reporting relative risk for progestin therapy failure in estrogen and progesterone receptor negative endometrial hyperplasia and early endometrial cancer. (a) Overall analysis for estrogen receptor. (b) Subgroup analysis for estrogen receptor according to the administration route of progestins. (c) Subgroup analysis for estrogen receptor in patient with atypical hyperplasia and cancer. (d) Overall analysis for progesterone receptor. (e) Subgroup analysis for progesterone receptor according to the administration route of progestins. (f) Subgroup analysis for progesterone receptor in patient with atypical hyperplasia and cancer [Colour figure can be viewed at wileyonlinelibrary.com]


(a)
ER- ER+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Random, 95% CI M-H, Random, 95% CI
2007 Minaguchi 0 1 5 30 12.0% 1.41 [0.11, 17.47]
2007 Yamazawa 1 7 1 2 13.6% 0.29 [0.03, 2.80]
2010 Akesson 4 6 2 28 21.0% 9.33 [2.19, 39.80]
2013 Gallos 4 17 6 157 24.2% 6.16 [1.93, 19.68]
2014 Gunderson 0 1 11 30 12.5% 0.67 [0.06, 7.76]
2015 Yang 1 6 10 82 16.8% 1.37 [0.21, 8.97]
Total (95% CI) 38 329 100.0% 2.19 [0.74, 6.52]
Total events 10 35
Heterogeneity: τ² = 0.93; χ² = 10.52, df = 5 (p = 0.06); I² = 52%
Test for overall effect: Z = 1.41 (p = 0.16)
0.001 0.1 1 10 1000
(b)
ER- ER+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI
Oral progestin
2007 Minaguchi 0 1 5 30 16.3% 1.41 [0.11, 17.47]
2007 Yamazawa 1 7 1 2 37.9% 0.29 [0.03, 2.80]
Subtotal (95% CI) 8 32 54.2% 0.62 [0.13, 2.99]
Total events 1 6
Heterogeneity: χ² = 0.85, df = 1 (p = 0.36); I² = 0%
Test for overall effect: Z = 0.59 (p = 0.55)
LNG-IUS
2010 Akesson 4 6 2 28 17.2% 9.33 [2.19, 39.80]
2013 Gallos 4 17 6 157 28.6% 6.16 [1.93, 19.68]
Subtotal (95% CI) 23 185 45.8% 7.35 [2.98, 18.16]
Total events 8 8
Heterogeneity: χ² = 0.19, df = 1 (p = 0.66); I² = 0%
Test for overall effect: Z = 4.32 (p < 0.0001)
Test for subgroup differences: χ² = 7.14, df = 1 (p = 0.008), I² = 86.0%
0.001 0.1 1 10 1000
(c)
ER- ER+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI
2007 Minaguchi 0 1 5 30 13.4% 1.41 [0.11, 17.47]
2007 Yamazawa 1 7 1 2 31.2% 0.29 [0.03, 2.80]
2014 Gunderson 0 1 11 30 28.0% 0.67 [0.06, 7.76]
2015 Yang 1 6 10 82 27.4% 1.37 [0.21, 8.97]
Total (95% CI) 15 144 100.0% 0.84 [0.29, 2.48]
Total events 2 27
Heterogeneity: χ² = 1.31, df = 3 (p = 0.73); I² = 0%
Test for overall effect: Z = 0.31 (p = 0.75)
0.001 0.1 1 10 1000
(d)
PR- PR+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI
2003 Utsunomiya 2 3 3 13 13.9% 2.89 [0.81, 10.34]
2007 Minaguchi 0 4 5 27 20.6% 0.51 [0.03, 7.84]
2007 Yamazawa 2 4 0 5 5.6% 6.00 [0.37, 98.16]
2010 Akesson 5 13 1 21 9.4% 8.08 [1.06, 61.66]
2013 Gallos 3 31 7 143 30.8% 1.98 [0.54, 7.22]
2015 Yang 1 7 10 81 19.7% 1.16 [0.17, 7.78]
Total (95% CI) 62 290 100.0% 2.44 [1.22, 4.89]
Total events 13 26
Heterogeneity: χ² = 3.75, df = 5 (p = 0.59); I² = 0%
Test for overall effect: Z = 2.52 (p = 0.01)
0.001 0.1 1 10 1000
(e)
PR- PR+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI
Oral progestin
2003 Utsunomiya 2 3 3 13 17.3% 2.89 [0.81, 10.34]
2007 Minaguchi 0 4 5 27 25.6% 0.51 [0.03, 7.84]
2007 Yamazawa 2 4 0 5 7.0% 6.00 [0.37, 98.16]
Subtotal (95% CI) 11 45 49.9% 2.10 [0.71, 6.24]
Total events 4 8
Heterogeneity: χ² = 1.81, df = 2 (p = 0.40); I² = 0%
Test for overall effect: Z = 1.34 (p = 0.18)
LNG-IUS
2010 Akesson 5 13 1 21 11.8% 8.08 [1.06, 61.66]
2013 Gallos 3 31 7 143 38.3% 1.98 [0.54, 7.22]
Subtotal (95% CI) 44 164 50.1% 3.41 [1.19, 9.77]
Total events 8 8
Heterogeneity: χ² = 1.37, df = 1 (p = 0.24); I² = 27%
Test for overall effect: Z = 2.28 (p = 0.02)
Test for subgroup differences: χ² = 0.39, df = 1 (p = 0.53), I² = 0%
0.001 0.1 1 10 1000
(f)
PR- PR+ Risk Ratio Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI
2003 Utsunomiya 2 3 3 13 23.3% 2.89 [0.81, 10.34]
2007 Minaguchi 0 4 5 27 34.5% 0.51 [0.03, 7.84]
2007 Yamazawa 2 4 0 5 9.4% 6.00 [0.37, 98.16]
2015 Yang 1 7 10 81 32.9% 1.16 [0.17, 7.78]
Total (95% CI) 18 126 100.0% 1.79 [0.70, 4.57]
Total events 5 18
Heterogeneity: χ² = 2.27, df = 3 (p = 0.52); I² = 0%
Test for overall effect: Z = 1.22 (p = 0.22)
0.01 0.1 1 10 100

For the "Index test" domain, with regard to ER, 6 studies were considered low risk, 2 unclear risk, and 4 high risk (absence of individual data[13,16,18]). With regard to PR, 7 studies were considered low risk, 2 unclear risk and 4 high risk (3 for absence of individual data[13,16,18] and 1 for not differentiating between low expression and negative expression[20]).

For the "Reference test" domain, 9 studies were considered low risk, 3 studies unclear risk and 1 study high risk (not differentiating between regression and persistence of disease[24]).

For the "Flow and timing" domain, 9 studies were considered low risk and 4 unclear risk.

The studies at high risk of bias were excluded from the meta-analysis.

Authors' judgments about risks of bias are summarized in Supporting Information Figure S2.

In the main analysis, a negative expression of ER showed no significant predictive value, with a RR of poor response of 2.19 (95% CI .74-6.52; P = .16), with moderate heterogeneity among studies (I^2 = 52%) (Figure 1a).

Sensitivity and specificity of ER were .22 (95% CI .11-.37) and .91 (95% CI .88-.94), respectively, with LR^+ and LR^- of 2.28 (95% CI .81-6.42) and .91 (95% CI .72-1.16), respectively, and a DOR of 2.80 (95% CI .70-11.18). SROC analysis showed low predictive accuracy and no actual clinical usefulness, with an AUC of .637. Heterogeneity among studies was low for DOR (I^2 = 47.3%), moderate for sensitivity (I^2 = 70.8%), LR^+ (I^2 = 53.6%) and LR^- (I^2 = 64.2%) and high for specificity (I^2 = 80.6%) (Figure 2).

A negative expression of PR was significantly predictive of poor response, with an RR of 2.44 (95% CI 1.22-4.89; P = .01), without heterogeneity among studies (I^2 = 0%) (Figure 1d).

Sensitivity and specificity of PR were .33 (95% CI .19-.50) and .84 (95% CI .80-.88), respectively, with LR^+ and LR^- of 2.37 (95% CI 1.47-3.80) and .88 (95% CI .65-1.29), respectively, and a DOR of 2.81 (95% CI 1.14-6.93). SROC analysis showed moderate predictive accuracy and no clinical usefulness, with an AUC of .806. Heterogeneity among studies was absent for DOR and LR^+ (I^2 = 0%), low for specificity (I^2 = 42.3%), and moderate for sensitivity (I^2 = 73.1%) and LR^- (I^2 = 58.1%) (Figure 3).

In the subgroup of patients treated with oral progestins, ER was not significantly predictive of response, with an RR of .62 (95% CI .13-1.41; P = .55) and without heterogeneity (I^2 = 0%) (Figure 1b). Sensitivity and specificity were .14 (95% CI .00-.58) and .79 (95%

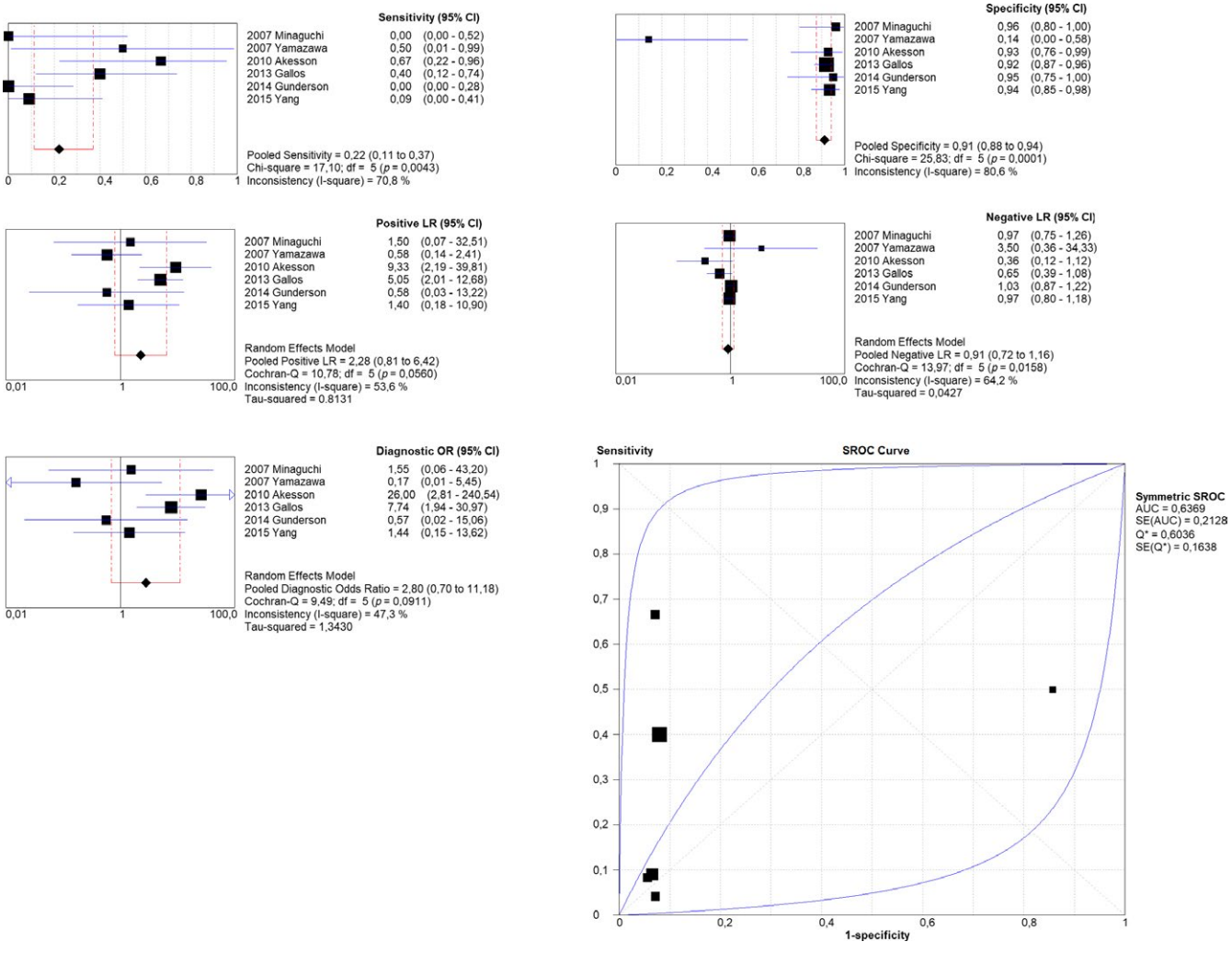


FIGURE 2 Plots reporting prognostic accuracy metrics for estrogen receptor [Colour figure can be viewed at wileyonlinelibrary.com]

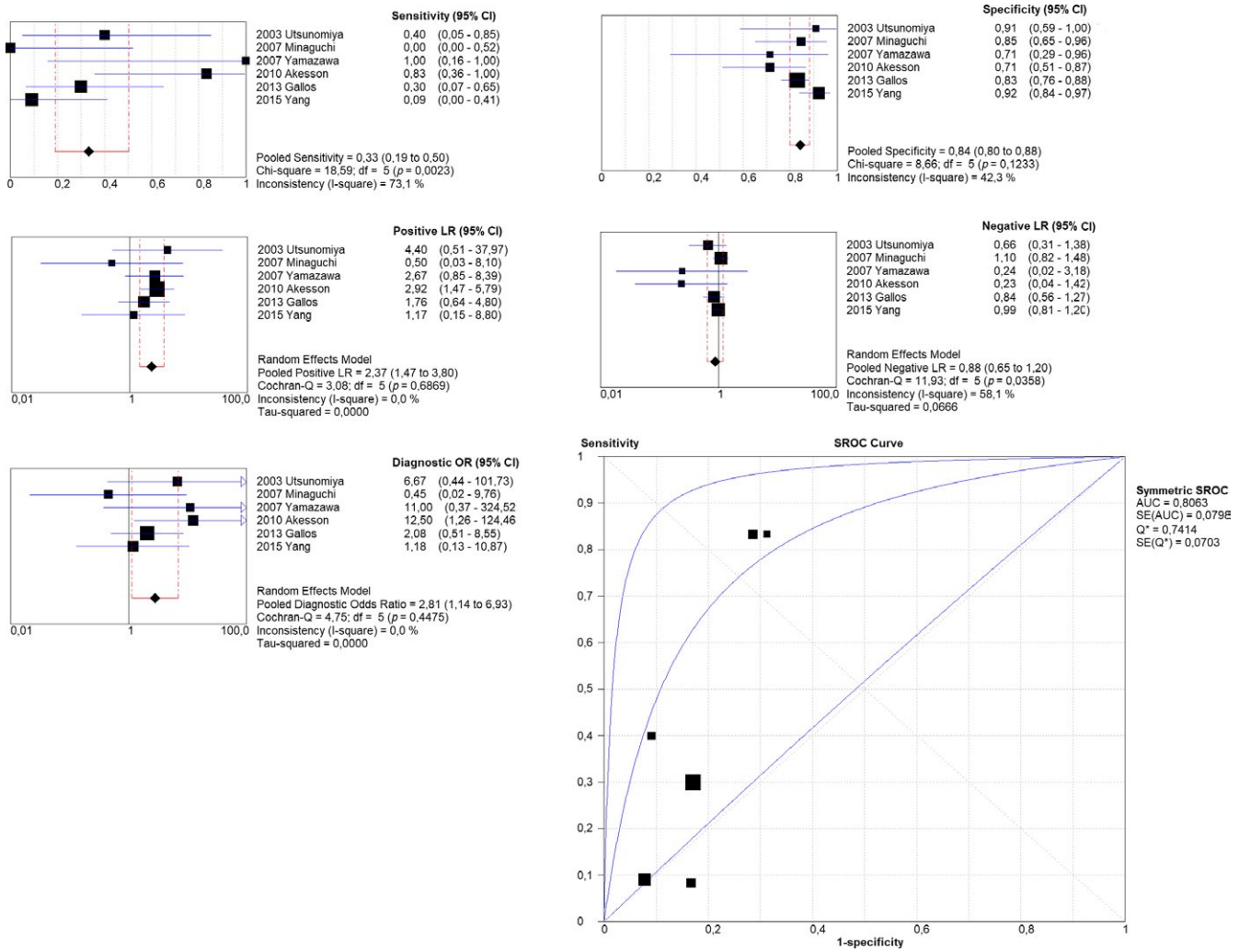


FIGURE 3 Plots reporting prognostic accuracy metrics for progesterone receptor [Colour figure can be viewed at wileyonlinelibrary.com]

CI .61-.91), respectively, with LR^+ and LR^- of .69 (95% CI .19-2.50) and 1.28 (95% CI .37-4.45), respectively, and a DOR of .53 (95% CI .05-5.94). Heterogeneity among studies was absent for DOR and LR^+ (I^2 = 0%), low for LR^- (42.0%), moderate for sensitivity (66.3%) and high for specificity (95.0%) (Figure 4a).

In the subgroup of patients treated with LNG-IUD, negative expression of ER was significantly predictive of poor response, with an RR of 7.35 (95% CI 2.98-18.16; $P < .0001$), without heterogeneity among studies (I^2 = 0%) (Figure 1b). Sensitivity and specificity were .50 (95% CI .25-.75) and .92 (95% CI .87-.97), respectively, with LR^+ and LR^- of 6.02 (95% CI 2.77-13.10) and .59 (95% CI .37-.94), respectively, and a DOR of 10.87 (95% CI 3.35-35.20). Heterogeneity among studies was absent for specificity, LR^+, LR^- and DOR (I^2 = 0%), and insignificant for sensitivity (I^2 = 7.6%) (Figure 4b).

In the subgroup of patients treated with oral progestins, PR was not significantly predictive of response, with an RR of 2.10 (95% CI .71-6.24; $P = .18$) and without heterogeneity (I^2 = 0%) (Figure 1e). Sensitivity and specificity were .33 (95% CI .10-.65) and .84 (95% CI .70-.93), respectively, with LR^+ and LR^- of 2.42 (95% CI .93-6.26) and .80 (95% CI .36-1.78), respectively, and a DOR of 3.18 (95% CI .47-21.51). Heterogeneity among studies was absent for specificity and LR^+ (I^2 = 0%), insignificant for DOR (I^2 = 16%), moderate for LR^- (I^2 = 56.6%) and high for sensitivity (I^2 = 76.6%) (Figure 5a).

In the subgroup of patients treated with LNG-IUD, negative expression of PR was significantly predictive of poor response, with an RR of 3.41 (95% CI 1.19-9.77; $P = .02$), with insignificant heterogeneity among studies (I^2 = 27%) (Figure 1e). Sensitivity and specificity were .50 (95% CI .25-.75) and .81 (95% CI .75-.87), respectively, with LR^+ and LR^- of 2.48 (95% CI 1.41-4.38) and .55 (95% CI .13-2.32), respectively, and a DOR of 4.04 (95% CI .72-22.57). Heterogeneity among studies was absent for LR^+ (I^2 = 0%), low for DOR (I^2 = 42.6%) and specificity (I^2 = 47.3%), moderate for LR^- (I^2 = 63.3%) and high for sensitivity (I^2 = 78.1%) (Figure 5b).

In the subgroup of patients with atypical EH and/or EEC, ER was not significantly predictive of response, with an RR of .84 (95% CI .29-2.48; $P = .75$) and without heterogeneity (I^2 = 0%) (Figure 1c). Sensitivity and specificity were .07 (95% CI .01-.23) and .90 (95% CI .84-.95), respectively, with LR^+ and LR^- of .81 (95% CI .29-2.26) and 1.00 (95% CI .89-1.12), respectively, and a DOR of .83 (95% CI .19-5.89). SROC analysis showed no predictive accuracy and

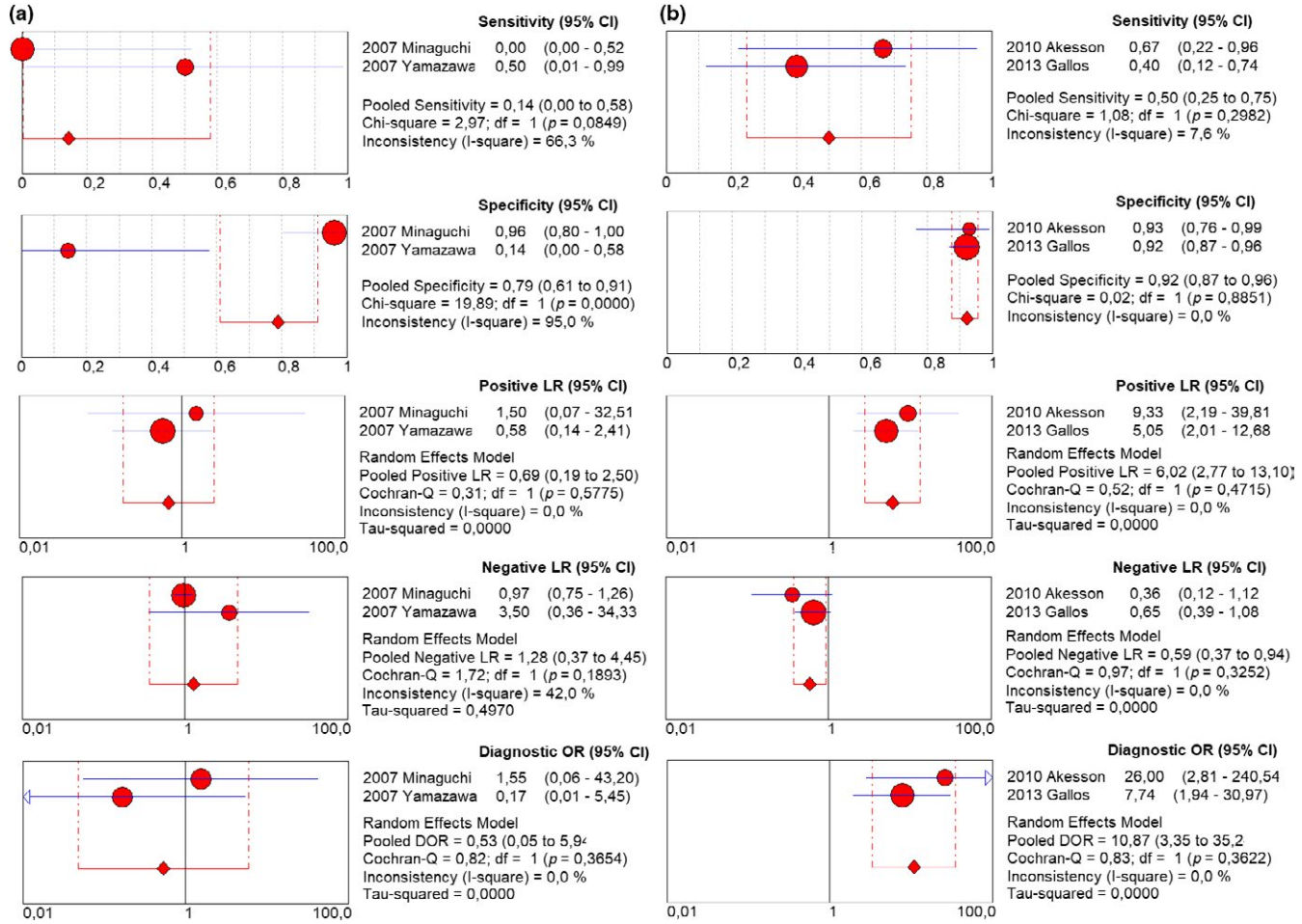


FIGURE 4 Forest plots reporting prognostic accuracy metrics for estrogen receptor in the subgroup treated with oral progestins (a) and LNG-IUD (b) [Colour figure can be viewed at wileyonlinelibrary.com]

no clinical usefulness, with an AUC of .304. Heterogeneity among studies was low for sensitivity (I^2 = 41%), high for specificity (I^2 = 88.2%), and absent for LR^+, LR^- and DOR (I^2 = 0%) (Supporting Information Figure S3).

In the subgroup of patients with atypical EH and/or EEC, PR was not significantly predictive of response, with an RR of 1.79 (95% CI .70-4.57; *P* = .22) and without heterogeneity (I^2 = 0%) (Figure 1f). Sensitivity and specificity were .22 (95% CI .07-.44) and .89 (95% CI .82-.94), respectively, with LR^+ and LR^- of 2.12 (95% CI .90-5.01) and .97 (95% CI .75-1.26), respectively, and a DOR of 2.18 (95% CI .85-8.61). SROC analysis showed moderate predictive accuracy and no actual clinical usefulness, with an AUC of .828. Heterogeneity among studies was moderate for sensitivity (I^2 = 71.8%), insignificant for specificity (I^2 = .2%), low for LR^- (I^2 = 35.5%) and absent for LR^+ and DOR (I^2 = 0%) (Supporting Information Figure S4).

Data regarding non-atypical EH were not separately extractable.

In the assessment of the risk of bias across studies for ER, although the funnel plot showed an evident asymmetry, the studies with higher accuracy were those showing the higher RR results (Figure 6a). Thus, publication bias was absent.

Regarding PR, the funnel plot showed a clear symmetry, thus excluding the possibility of a publication bias (Figure 6b).

4 | DISCUSSION

Our study showed that ER and PR expression in EH and EEC are predictive of response to LNG-IUD, whereas they do not have significant predictive value if oral progestins are administered.

ER and PR are nuclear receptors which play a crucial role in endometrial carcinogenesis. ER mediates the action of estrogens, promoting the proliferation of endometrium, with a physiological action in the proliferative phase of the menstrual cycle.[35] When the action of estrogens is not balanced by the action of progesterone, the endometrium goes through a phase of disordered proliferation (disordered proliferative endometrium), which leads to EH.[36] The increased and continued stimulation to proliferate predisposes to the development of genetic mutations with the emergence of a neoplastic clone (atypical EH/endometrioid intraepithelial neoplasia), initiating the carcinogenesis,[4,36] Most atypical EH and well differentiated EEC of endometrioid type, ("type 1 endometrial cancer" in

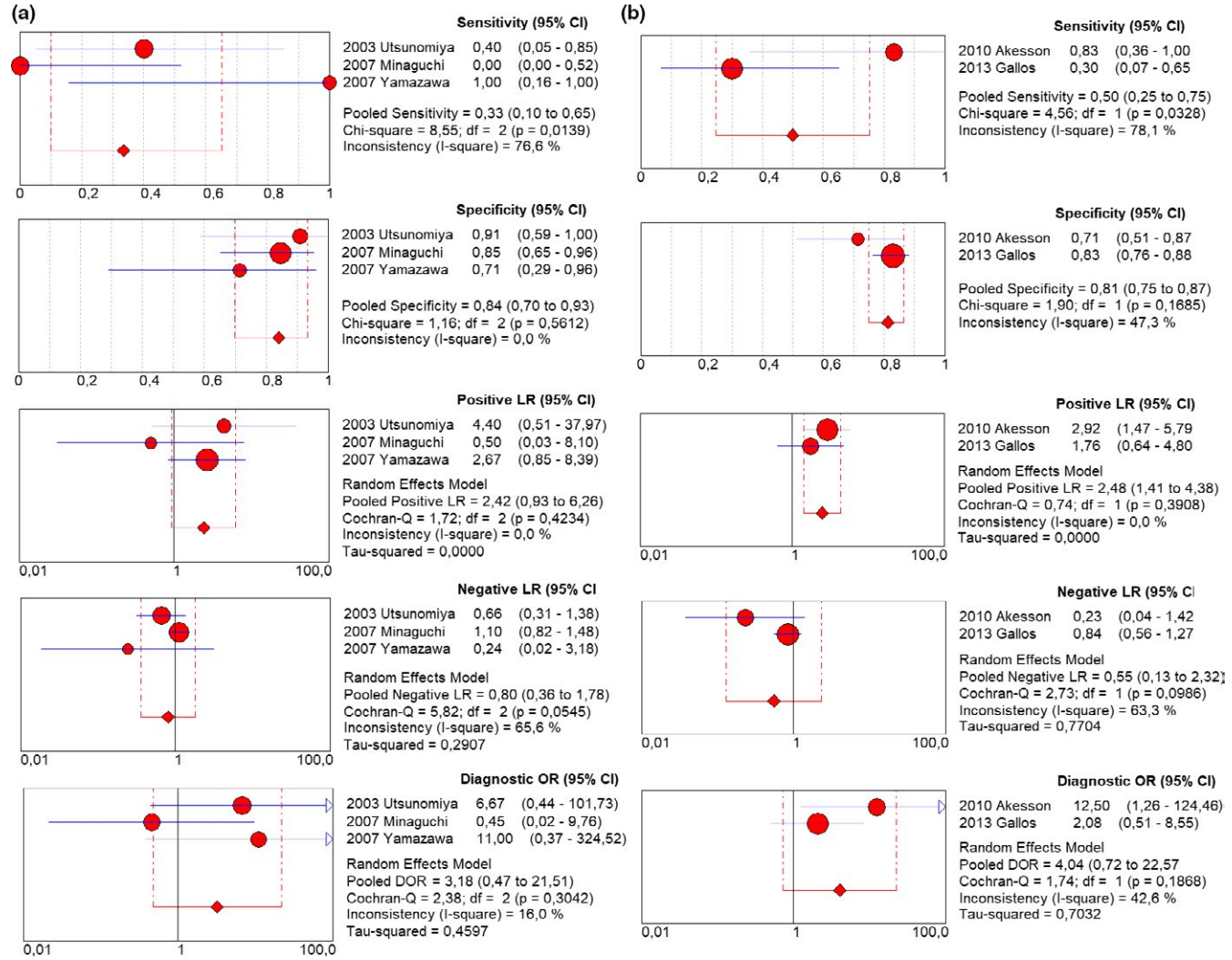


FIGURE 5 Forest plots reporting prognostic accuracy metrics for progesterone receptor in the subgroup treated with oral progestins (a) and LNG-IUD (b) [Colour figure can be viewed at wileyonlinelibrary.com]

the traditional classification) are still regulated by sexual hormones, with high expression of ER and PR.[2,3,22,36] In the 2013, the Cancer Genome Atlas Network identified four molecular-based prognostic categories of endometrial cancer (hypermutated; ultramutated; copy number low; copy number high); of these, the third one (classified as microsatellite stable, p53 wild type, with low somatic copy number) showed histological features superimposable to the "type 1" endometrial cancer, in most cases showing high expression of PR.[37]

Progestins have been widely used for the conservative treatment of EH and EEC, since they antagonize the growth-promoting action of estrogens, stimulating apoptosis in endometrial cells.[35,38]

Among the available progestin-based treatments, LNG-IUD showed better efficacy than oral progestin, as highlighted by several meta-analyses.[27-29] The intrauterine release of levonorgestrel provides a higher concentration of progestin at the level of the uterus, leading to marked changes in the endometrium, similar to the decidualization occurring during pregnancy.[17] On the other hand, the local action also limits the possible adverse effects due to the systemic action of the progestin.[6]

Nonetheless, there is great uncertainty in the literature regarding the predictive value of ER and PR. The 2016 ESMO-ESGO-ESTRO consensus conference on endometrial cancer discouraged the use of immunohistochemistry for PR in the selection of patients for conservative treatment, stating that about half of PR-null specimens still respond.[7] However, such a statement is based on only one study,[15] and several others showed different and conflicting results. While some studies advocate a high predictive utility of these receptors, other studies found no associations between their expression and the response to therapy. In our main analysis, we found that ER was not significantly predictive and had low accuracy, whereas PR was significantly predictive of response (RR), with moderate accuracy. These results might be in agreement with the scientific literature, which has given more importance to PR than ER. However, our analyses highlighted that the results were highly heterogeneous among studies and were strongly influenced by the administration route of progestins.

In the subgroup analysis, we observed that the predictive value of ER and PR was noticeably higher if LNG-IUD was used, whereas

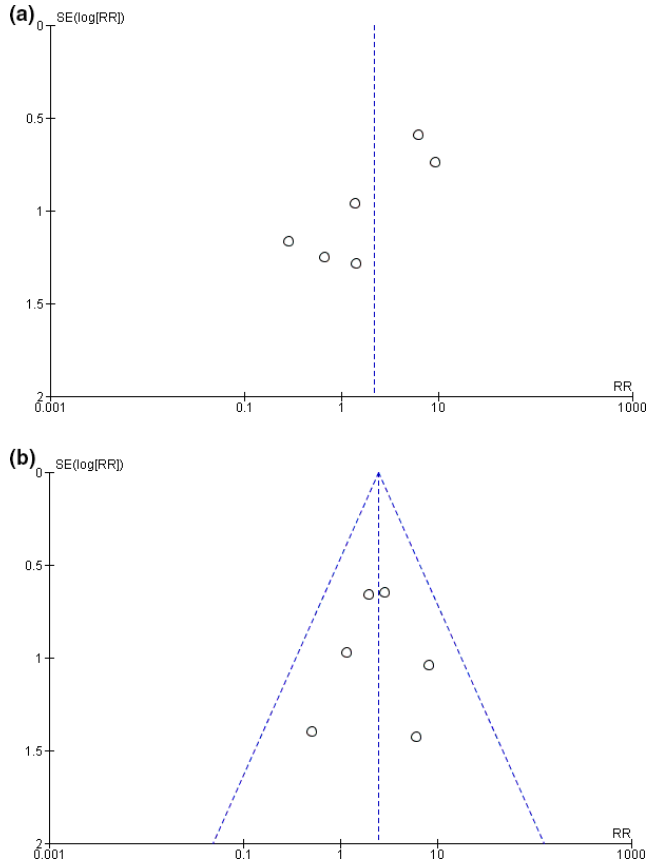


FIGURE 6 Funnel plots for the assessment of risk of bias across studies about estrogen receptor (a) and progesterone receptor (b) [Colour figure can be viewed at wileyonlinelibrary.com]

it was not significant in the case of oral administration of progestin. These results might be explained by the stronger local effect of LNG-IUD: if intrauterine administration activates a higher percentage of PR compared with the oral route, the difference in the response rate between PR-positive and PR-negative specimens can be expected to become higher.

In the subgroup of only atypical EH and EEC, the values of RR were not significant, with no accuracy for ER (AUC = .304) and moderate but still insufficient accuracy for PR (AUC = .828). These non-significant results may be explained by the fact that no studies in this subgroup exclusively considered LNG-IUD. Therefore, further studies are necessary in this regard.

However, even in the LNG-IUD subgroup, the predictive usefulness appeared insufficient, with a low sensitivity for both ER and PR (.50). This means that only half of resistant specimens have a loss of hormone receptor expression. In this regard, given the heterogeneity of the mutational background in endometrial neoplastic specimens,[37,39-41] it is possible that some genotypes are not responsive regardless of ER and PR expression. For example, a study by Zakhour et al[42] showed that mismatch repair-deficient atypical EH did not respond to therapy. Moreover, other clinic-pathological factors might affect the response in the resistant cases when ER and PR expression is maintained. On the other hand, specificity was high for both ER and PR, indicating that the vast majority of responsive cases express the hormonal receptors. While being an interesting result, in our opinion it still is not enough to be clinically useful in the decision-making between conservative treatment and hysterectomy, as supported by an AUC <.9, an LR^+ <10 and an LR^- >.1 for both ER and PR. In fact, progestin therapy is not the standard approach, and patients with atypical EH and EEC who choose to preserve fertility are closely followed with endometrial biopsies every 3-6 months, due to the risk of progression to invasive disease. A predictive marker that misses 50% of non-responsive patients and addresses hysterectomy in only 10%-20% of patients who would have responded definitely, appears inadequate, at least as a stand-alone marker.

In any case, further studies might obtain better results by assessing ER and PR isoforms (ERα, ERβ, PRA, PRB) separately and on large and more homogeneous samples.

To the best of our knowledge, this is the first systematic review and meta-analysis assessing the predictive value or ER and PR in the conservative treatment of EH and EEC, a long-standing controversial issue.

Although our study substantially confirms the guidance of ESMO-ESGO-ESTRO, it may provide a higher level of evidence to support this position, which was based on only one small study.

However, there are some limitations to our results. The small number of studies in the subgroup analysis did not allow us to calculate AUC on SROC according to the administration route of progestins, although we could use LR^+ and LR^- as surrogates.

Furthermore, it was impossible to extract data to assess the predictive value of ER and PR separately for neoplastic lesions (atypical EH, EEC) and reactive conditions (EH without atypia) in patients treated with LNG-IUD, which was the subgroup showing the best results.

5 | CONCLUSION

ER and PR expression in EH and EEC are predictive of response to LNG-IUD, but not to oral progestins. However, given the low sensitivity and the suboptimal specificity, they do not appear adequate as predictive markers if assessed alone. Further studies in this field may better assess separately the predictive value of ER and PR in each specific histological category (EH with and without atypia and EEC) and with regard to their receptor isoforms, and how it can be affected by the mutational background and the clinicopathological factors.

CONFLICT OF INTEREST

The authors have stated explicitly that there are no conflicts of interest in connection with this article.

ORCID

Antonio Raffone https://orcid.org/0000-0001-5443-2333
Antonio Travaglino https://orcid.org/0000-0003-4002-1618
Gabriele Saccone https://orcid.org/0000-0003-0078-2113

REFERENCES

1. Kurman RJ, Carcangiu ML, Herrington CS, Young RH, eds. *WHO Classification of Tumours of Female Reproductive Organs*, 4th ed. Lyon, France: IARC; 2014.
2. Sherman ME. Theories of endometrial carcinogenesis: a multidisciplinary approach. *Mod Pathol*. 2000;13:295-308.
3. Sanderson PA, Critchley HOD, Williams ARW, Arends MJ, Saunders PTK. New concepts for an old problem: the diagnosis of endometrial hyperplasia. *Hum Reprod Update*. 2017;23(2):232-254.
4. Mutter GL. Endometrial intraepithelial neoplasia (EIN): will it bring order to chaos? The Endometrial Collaborative Group. *Gynecol Oncol*. 2000;76(3):287-290.
5. Lacey JV, Sherman ME, Rush BB, et al. Absolute risk of endometrial carcinoma during 20-year follow-up among women with endometrial hyperplasia. *J Clin Oncol*. 2010;28:788-792.
6. Management of Endometrial Hyperplasia Green-top Guideline No. 67 RCOG/BSGE Joint Guideline | February 2016
7. Colombo N, Creutzberg C, Amant F, et al. ESMO-ESGO-ESTRO consensus conference on Endometrial Cancer: diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27:16-41.
8. Koskas M, Uzan J, Luton D, Rouzier R, Darai E. Prognostic factors of oncologic and reproductive outcomes in fertility-sparing management of endometrial atypical hyperplasia and adenocarcinoma: systematic review and meta-analysis. *Fertil Steril*. 2014;101:785-794.
9. Sato M, Arimoto T, Kawana K, et al. Measurement of endometrial thickness by transvaginal ultrasonography to predict pathological response to medroxyprogesterone acetate in patients with grade 1 endometrioid adenocarcinoma. *Mol Clin Oncol*. 2016;4:492-496.
10. Penner KR, Dorigo O, Aoyama C, et al. Predictors of resolution of complex atypical hyperplasia or grade 1 endometrial adenocarcinoma in premenopausal women treated with progestin therapy. *Gynecol Oncol*. 2012;124:542-548.
11. Travaglino A, Raffone A, Saccone G, et al. PTEN as a predictive marker of response to conservative treatment in endometrial hyperplasia and early endometrial cancer. A systematic review and meta-analysis. *Eur J Obstet Gynecol Reprod Biol*. 2018;231:104-110.
12. Utsunomiya H, Suzuki T, Ito K, et al. The correlation between the response to progestogen treatment and the expression of progesterone receptor B and 17beta-hydroxysteroid dehydrogenase type 2 in human endometrialcarcinoma. *Clin Endocrinol (Oxf)*. 2003;58(6):696-703.
13. Vereide AB, Kaino T, Sager G, Arnes M, Orbo A. Effect of levonorgestrel IUD and oral medroxyprogesterone acetate on glandular and stromal progesterone receptors (PRA and PRB), and estrogen receptors (ER-α and ER-β) in human endometrial hyperplasia. *Gynecol Oncol*. 2006;101:214-223.
14. Minaguchi T, Nakagawa S, Takazawa Y, et al. Combined phospho-Akt and PTEN expressions associated with post-treatment hysterectomy after conservative progestin therapy in complex atypical hyperplasia and stage Ia, G1 adenocarcinoma of the endometrium. *Cancer Lett*. 2007;248:112-122.
15. Yamazawa K, Hirai M, Fujito A, et al. Fertility-preserving treatment with progestin, and pathological criteria to predict responses, in young women with endometrial cancer. *Hum Reprod*. 2007;22:1953-1958.
16. Kashima H, Horiuchi A, Uchikawa J, et al. Up-regulation of nuclear receptor corepressor (NCoR) in progestin-induced growth suppression of endometrial hyperplasia and carcinoma. *Anticancer Res*. 2009;29:1023-1030.
17. Akesson E, Gallos ID, Ganesan R, Varma R, Gupta JK. Prognostic significance of estrogen and progesterone receptor expression in LNG-IUS (Mirena®) treatment of endometrial hyperplasia: an immunohistochemical study. *Acta Obstet Gynecol Scand*. 2010;89:393-398.
18. Orbo A, Arnes M, Pettersen I, Larsen K, Hanssen K, Moe B. Down-regulated progesterone receptor A and B coinciding with successful treatment of endometrial hyperplasia by the levonorgestrel impregnated intrauterine system. *Acta Obstet Gynecol Scand*. 2010;89:1438-1446.
19. Kamoi S, Ohaki Y, Mori O, Yamada T, Fukunaga M, Takeshita T. Determining best potential predictor during high-dose progestin therapy for early stage and well-differentiated endometrial adenocarcinoma using semiquantitative analysis based on image processing and immunohistochemistry. *J Nippon Med Sch*. 2011;78:84-95.
20. Upson K, Kimberley H, Reed SD, et al. Biomarkers of progestin therapy resistance and endometrial hyperplasia progression. *Am J Obstet Gynecol*. 2012;207:36.e1-36.e8.
21. Gallos ID, Devey J, Ganesan R, Gupta JK. Predictive ability of estrogen receptor (ER), progesterone receptor (PR), COX-2, Mlh1, and Bcl-2 expressions for regression and relapse of endometrial hyperplasia treated with LNG-IUS: a prospective cohort study. *Gynecol Oncol*. 2013;130:58-63.
22. Gunderson CC, Dutta S, Fader AN, et al. Pathologic features associated with resolution of complex atipica hyperplasia and grade 1 endometrial adenocarcinoma after progestin therapy. *Gynecol Oncol*. 2014;132:33-37.
23. Yang YF, Liao YY, Liu XL, Su SG, Li LZ, Peng NF. Prognostic factors of regression and relapse of complex atypical hyperplasia and well-differentiated endometrioid carcinoma with conservative treatment. *Gynecol Oncol*. 2015;139:419-423.
24. Reyes HD, Carlson MJ, Devor EJ, et al. Downregulation of FOXO1 mRNA levels predicts treatment failure in patients with endometrial pathology conservatively managed with progestin-containing intrauterine devices. *Gynecol Oncol*. 2016;140:152-160.
25. Sotiriadis A, Papatheodorou SI, Martins WP. Synthesizing evidence from diagnostic accuracy tests: the SEDATE guideline. *Ultrasound Obstet Gynecol*. 2016;47(3):386-395.
26. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
27. Gallos ID, Shehmar M, Thangaratinam S, Papapostolou TK, Coomarasamy A, Gupta JK. Oral progestogens vs levonorgestrel-releasing intrauterine system for endometrial hyperplasia: a systematic review and metaanalysis. *Am J Obstet Gynecol*. 2010;203:547.e1-547.e10.
28. Hashim HA, Ghaytay E, El Rakhawy M. Levonorgestrel-releasing intrauterine system vs oral progestins for non-atypical endometrial hyperplasia: a systematic review and metaanalysis of randomized trials. *Am J Obstet Gynecol*. 2015;213:469-478.
29. Yuk JS, Song JY, Lee JH, Park WI, Ahn HS, Kim HJ. Levonorgestrel-releasing intrauterine systems versus oral cyclic medroxyprogesterone acetate in endometrial hyperplasia therapy: a meta-analysis. *Ann Surg Oncol*. 2017;24:1322-1329.
30. Giampaolino P, Dispieziosardo A, Mollo A, et al. Hysteroscopic endometrial focal resection followed by levonorgestrel intrauterine device insertion as a fertility-sparing treatment of atypical endometrial hyperplasia and early endometrial cancer: a retrospective study. *J Minim Invasive Gynecol*. 2018. pii: S1553-4650(18)30347-9. https://doi.org/10.1016/j.jmig.2018.07.001. [Epub ahead of print]
31. Travaglino A, Raffone A, Saccone G, et al. Loss of B-cell lymphoma 2 immunohistochemical expression in endometrial hyperplasia: a specific marker of precancer and novel indication for treatment: a systematic review and meta-analysis. *Acta Obstet Gynecol Scand*. 2018;97(12):1415-1426.
32. Raffone A, Travaglino A, Saccone G, et al. Loss of PTEN expression as diagnostic marker of endometrial precancer: a systematic review and meta-analysis. *Acta Obstet Gynecol Scand*. 2019;98(3):275-286.

33. Travaglino A, Raffone A, Saccone G, et al. Endometrial hyperplasia and risk of coexistent cancer: WHO versus EIN criteria. *Histopathology*. 2018. https://doi.org/10.1111/his.13776.
34. Raffone A, Travaglino A, Saccone G, et al. PAX2 in endometrial carcinogenesis and in differential diagnosis of endometrial hyperplasia. A systematic review and meta-analysis of diagnostic accuracy. *Acta Obstet Gynecol Scand*. 2019;98:287-299.
35. Chandra V, Kim JJ, Benbrook DM, Dwivedi A, Rai R. Therapeutic options for management of endometrial hyperplasia. *J Gynecol Oncol*. 2016;27:e8.
36. Yang CH, Almomen A, Wee YS, Jarboe EA, Peterson CM, Janát-Amsbury MM. An estrogen-induced endometrial hyperplasia mouse model recapitulating human disease progression and genetic aberrations. *Cancer Med*. 2015;4(7):1039-1050.
37. Cancer Genome Atlas Research Network, Kandoth C, Schultz N, Cherniack AD, et al. Integrated genomic characterization of endometrial carcinoma. *Nature*. 2013;497(7447):67-73.
38. Vereide AB, Kaino T, Sager G, Ørbo A; Scottish Gynaecological Clinical Trials Group. Bcl-2, BAX, and apoptosis in endometrial hyperplasia after high dose gestagen therapy: a comparison of responses in patients treated with intrauterine levonorgestrel and systemic medroxyprogesterone. *Gynecol Oncol*. 2005;97(3):740-750.
39. Russo M, Broach J, Sheldon K, et al. Clonal evolution in paired endometrial intraepithelial neoplasia/atypical hyperplasia and endometrioid adenocarcinoma. *Hum Pathol*. 2017;67:69-77.
40. Travaglino A, Raffone A, Saccone G, et al. Immunohistochemical nuclear expression of β-catenin as a surrogate of CTNNB1 exon 3 mutation in endometrial cancer. *Am J Clin Pathol*. 2019. https://doi.org/10.1093/ajcp/aqy178. [Epub ahead of print]
41. Raffone A, Travaglino A, Saccone G, et al. Management of women with atypical polypoid adenomyoma of the uterus: a quantitative systematic review. *Acta Obstet Gynecol Scand*. 2019. https://doi.org/10.1111/aogs.13553. [Epub ahead of print]
42. Zakhour M, Cohen JG, Gibson A, et al. Abnormal mismatch repair and other clinicopathologic predictors of poor response to progestin treatment in young women with endometrial complex atypical hyperplasia and well-differentiated endometrial adenocarcinoma: a consecutive case series. *BJOG*. 2017;124(10):1576-1583.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Raffone A, Travaglino A, Saccone G, et al. Should progesterone and estrogens receptors be assessed for predicting the response to conservative treatment of endometrial hyperplasia and cancer? A systematic review and meta-analysis. *Acta Obstet Gynecol Scand*. 2019;00:1-12. https://doi.org/10.1111/aogs.13586