ESHRE guideline on Ovarian stimulation for IVF/ICSI

UPDATE 2025

European Society of Human Reproduction and Embryology

May 2025

ESHRE Guideline group on Ovarian stimulation for IVF/ICSI

REVIEW REPORT

www.eshre.eu/guidelines

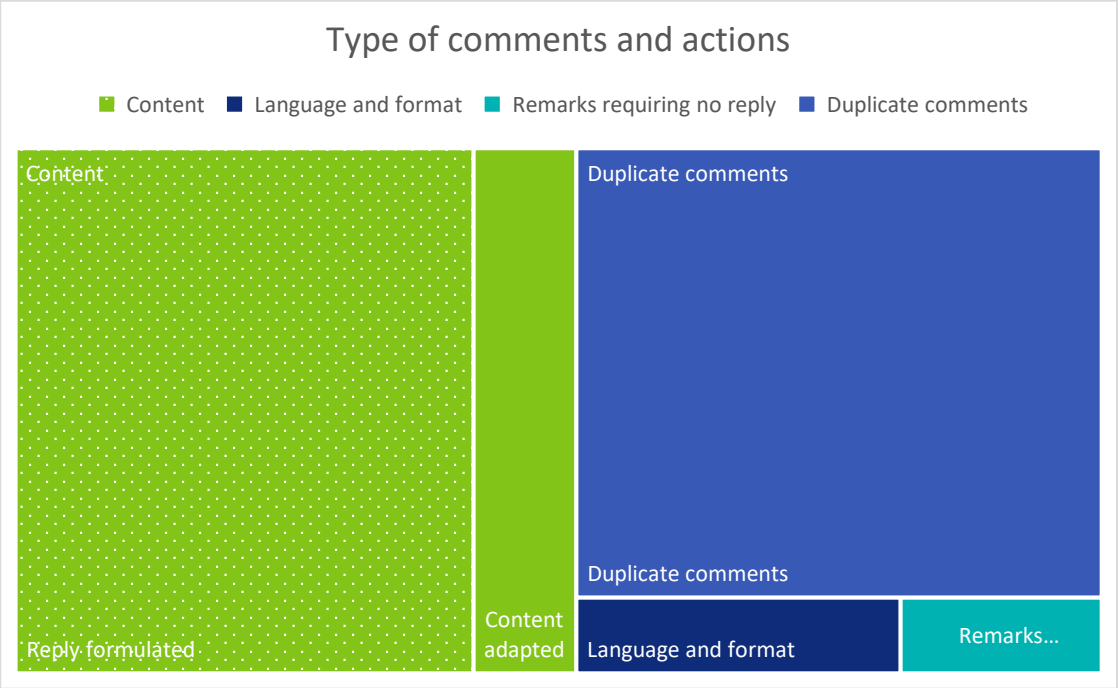
The draft of the ESHRE guideline on Ovarian Stimulation for IVF/ICSI update 2025 was published for public review for 6 weeks, between 5 May and 16 June 2025.

This report summarizes all reviewers, their comments and the reply of the guideline group and is published on the ESHRE website as supporting documentation to the guideline.

During the stakeholder review, a total of 486 comments were received from 156 reviewers.

The comments were focussed on the content of the guideline (452 comments), language and format (22 comments), or were remarks that did not require a reply (12 comments). All comments to the language and format were checked and corrected where relevant.

The comments to the content of the paper (n=452) were assessed by the working group and where relevant, adaptations were made in the paper (n=46; 18.3%). Adaptations included revisions and/or clarifications of the text, and amendments to the recommendations. For a number of comments, the working group considered them outside the scope of the paper or not appropriate/relevant (n=205; 81.7%).



Experts that participated in the stakeholder review

The list of representatives of professional organization, and of individual experts that provided comments to the guideline are summarized below.

Representatives of professional organisations

Organisation	Country	Representative
Gulsara Z. Eshimbetova	Uzbekistan	Association of reproductive medicine of Uzbekistan
Galina Grebennikova	Kazakhstan	Kazakhstan Association on Sexual and Reproductive Health (KMPA)
Assel Jaimbetova	Kazakhstan	Institute of Reproductive Health Almaty Kazakhstan
Monu Pattanayak	India	Shanti Memorial Hospital
Ulughbek Jabborov	Uzbekistan	Republican Perinatal Centre
Liudmyla Hutsikava	Republic of Belarus	Department of Obstetrics and Gynecology of Grodno State Medical University
Zaytuna Khamidullina	Kazakhstan	Federation of obstetrician-gynecologists of Astana city
Pavika Lal	India	Ganesh Shankar Vidyarthi Medical College
Hisham A. Arab	Saudi Arabia	Saudi Obstetrics and Gynecology Society
Farah Gari	Algeria	Department of gynecology and obstetrics, university hospital of Blida
Zeev Shoham **Ariel Weissman** **Raoul Orvieto**	Israel	IVF-Worldwide
Johannes Ott	Austria	Austrian association of Gynecology and Obstetrics (Working Group for Gynecologic Endocrinology and Reproductive Medicine
Ariane Germeyer **Martin Birkhäuser** **Bettina Böttcher** **Bruno Imthurn** **Alfred O Mueck** **Joseph Neulen** **Petra Stute** **Christian Thaler**	Germany	Zürcher Kreis working group

Inka Wiegratz **Ludwig Wildt**		
Alexander Katalinic **Maria Noftz** **Juan A Garcia-Velasco** **Lee P Shulman** **John N van den Anker** **Jerome F Strauss III**	Germany	International research group REASSURE:
José María Regalado Pedrajas	Spain	Onafiv, Fertilidad y Ginecología, S. L.
Emad Darwish	Egypt	Integrated Fertility Centre, Alexandria, Egypt
Hassan Sallam	Egypt	Alexandria Fertility and ART Centre International Representative Committee of the Royal College of Obstetricians and Gynaecologists in Egypt
Saghar Salehpour	Iran	Iranian Society of Reproductive medicine (ISRM)
Emre Goksan Pabuccu	Turkey	Association of Infertility Medicine and Surgery (UTCD), Turkey
Christophe Blockeel **Christos Venetis** **Biljana Popovic** **Ying Cheong** **Alexandra Freis**	Belgium Greece Serbia UK Norway	ESHRE SIG Reproductive Endocrinology
Yun Sun **Lei Jin** **Juanzi Shi** **Fenghua Liu** **Songying Zhang** **Cuilian Zhang** **Guimin Hao** **Jichun Tan** **Junhao Yan** **Qun Lv** **Jianqiao Liu** **Feiyang Diao** **Xiru Liu** **Yan Zhao** **Rong Li**	China	The Chinese Expert Review Panel for ESHRE OS Guideline
Veaceslav Mosin	Moldova	Alternativa Clinic
Jayesh Amin	India	Nova Wings Fertility Chains
	Germany	German Society of Reproductive Medicine

Alexandra Kohl Schwartz	Switzerland	Ager Switzerland, working group for gynaecological endocrinology and reproductive medicine
Adrija Kumar Datta **Stuart Campbell** **Geetta Nargund**	UK	International society for mild approaches for assisted reproduction (ISMAAR)

Individual experts

Reviewer	Country
Raj Mathur	UK
Natalia Pedachenko	Ukraine
K. K. Pandey	India
Nodira Ruzieva	Uzbekistan
Vyacheslav Lokshin	Kazakhstan
Anagani Manjula	India
Sujoy Dasgupta	India
Mita Aggarwal	India
Namita Kotia	India
Sridevi Nellimarla	India
Padmaja Veeramachaneni	India
Tetiana Tutchenko	Ukraine
Fei Gong	China
Sonia Naik	India
Raoul Orvieto	Israel
Biswajyoti Guha	India
Shikha Gupta	India
Suyesha Khanijao	India
Qinjie Tian	China
Ginny Gupta	India
Surinder Pal Singh Kochar	India
Alberto Revelli	Italy
Tamal Bhattacharyya	India
Mukesh Gupta	India
Sunita Arora	India
Ritu Joshi	India

Dubrovina Svetlana	Russia
Debankur Barman	India
Monu Pattanayak	India
Arnab Bhowmik	India
Bharat S.	India
Puja Kumari	India
Meeta Meeta	India
Ahmed Samy Abdelazim Saad	Egypt
Padmaja veeramachaneni	India
Ahmed Elsayed Hassan Hamed Elbohoty	UAE
Ulughbek Jabborov	Uzbekistan
Tian-Min Ye	China
Yan Gong	China
Jyothi G S	India
Nisha bhatnagar	India
Olena Yashyna	Ukraine
Priti Arora Dhamija	India
Isabel De Almeida	Brazil
Ritesh Sinha	India
Pavika Lal*	India
Manju Khemani	India
Poornima Durga	India
Farrukh Naheed	Pakistan
Feruza Gafurova	Uzbekistan
Geeta Khanna	India
Yun Sun	China
Elena Grudnitskaya	Republic of Belarus
Ayman Hany Ahmed	Egypt
Sandro C. Esteves	Brazil
Hassan Mostafa Gaafar	Egypt
Hisham A. Arab	Saudi Arabia
Yullia*	Ukraine
Diane De Neubourg	Belgium
Aniruddha Bhattacharjee	India

Semra Kahraman	Turkey
Yasser Orief	Egypt
Farah Gari	Algeria
Chengyan Deng	China
Xiu Luo	China
Sabirova Venera	Russia
Aleksandra Khramtsova	Russia
Juan-Enrique Schwarze **Shiv Gupta** **Susana Montenegro***	Germany
Xi Dong	China
Tapilskaya Natalia	Russia
Mohamed Ashraf Mohamed	Egypt
Yasser El Kassar	Egypt
Mitranovici Melinda Ildiko	Romania
Roberto Matorras	Spain
Dongzi Yang	China
Eduardo Correa Allende*	
Shamugiya Nato	Georgia
Umesh N Jindal	India
Madhu Shrivastav	India
Maneesha Jain	India
Anima Prasad	India
Aboubakr Mohamed Elnashar	Egypt
Willem Verpoest	The Netherlands
Guivarc'h Leveque Anne	France
Fang Xiong **Yinyang Bai**	China
Xi Xia	China
Sandeep Karunakaran	India
Sharda Jain	India
Kanad Dev Nayar	India
Kanad Dev Nayar	India
Manpreet Sharma	India
Miaoxin Chen	China

Andrii Berbets	Ukraine
Matthias Mueller*	Switzerland
Shaily Agarwal	India
Rishma Dhillon Pai	India
Rekha Rani	India
Amr Abdel Aziz Nadim	Egypt
Ayman Abo El Nour	Egypt
T. Ramani Devi	India
Pedro Augusto Araujo Monteleone	Brazil
Gavisova Alla	Russia
Monica Varma	India
Gatagazheva Aza Aslanovna	Russia
Tatyana Pestova	Russia
Kasi V Sellappan	India
Robert Fischer	Germany
Srilatha Gorthi	India
Surveen Ghumman	India
Alberto Vaiarelli	Italy
Ayman Oraif	Saudi Arabia
Kokkoni Kiose	Greece
Apostolos Tsironis	UK
Roberto de Azevedo Antunes	Brazil
Michael H. Dahan	Canada
Teraporn Vutyavanich	Thailand
Suresh Nair	Singapore
Nayana Patel	India
Sadiah Ahsan	Pakistan
Philippe Pinton*	Denmark
Alessandro Conforti **Robert Fisher** **Peter Humaidan** **Carlo Alviggi**	Italy Germany Denmark
Kastubh Kulkarni	India
Sonia Malik	India
Colin Howles	Switzerland

Karolina Palinska-Rudzka	UK
Vardanyan Rusudan	Armenia
Mohamed Bedis Chanoufi	Tunisia
BV Shobha	India
Himabindu Annamraju	India
Marianne Vendola	UK
Mekhala Dwarakanath B	India
Liudmila Stavinskaia	Republic of Moldova
Stefan Matik	North Macedonia
Cedrin Durnerin	France
Abdellatif Elkholy	Egypt

Reviewer comments and replies

NR	Reviewer	Page	Line	Comment	Action / Reply
INTRODUCTION					
7	Sujoy Dasgupta	9	251-262	"Due to the lack of universally accepted definitions of…."- In this guideline, follicle/ oocyte number more than 18 is used to defined "high responders" and less than 4 to define "low responders". Therefore, "normal responders" should develop 4-18 follicles/ oocytes during ovarian stimulation. This should be highlighted as the "optimum" response to ovarian stimulation or the "Goal" of ovarian stimulation. The paucity of robust evidence in this area should also be specified. One very recent individual patient data meta-analysis suggested that the live birth rate is optimized and the risk of OHSS is minimized if 8-14 oocytes are retrieved in fresh embryo transfer cycle using follitropin delta. Lobo R, Falahati A, Moley K, Pinborg A, Santos-Ribeiro S, Macklon NS, Jepsen IE. Oocyte yield and live birth rate after follitropin delta dosing and fresh embryo transfer: an individual patient data meta-analysis. Reprod Biomed Online. 2025 Feb;50(2):104451.	This statement is appreciated. At the same time, the response classes High and Low are clearly 'designed' to be avoided, as they both are linked to disadvantages regarding safety and/or lower efficacy. Aiming for the optimum response is therefore the implicit basis for all the efforts to forecast such classes of response with the challenge of trying to avoid them becoming reality and reaching the optimum range. The Lobo study was published after the final literature search and offers knowledge that has been broadly established in previous work.
7	Sujoy Dasgupta	11	318-323	"…next to the in vitro handling of gametes and embryos, and the embryo replacement procedure."- In addition, the improved efficiency of vitrification techniques of oocytes and embryos significantly improved the survival rates of oocytes/ embryos after freeze-thaw procedure.	This phrase was focussed on the role of Ovarian stimulation in the IVF/ICSI process. It is agreed that in this three step process many added procedures can be identified that have helped optimising the outcome, but here we limited to the basics.
79	Mitranovici Melinda Ildiko	12	348-350	Did you observe an increase in FSH levels after estradiol administration in elderly women?	The text explains that raising the FSH exposure by other procedures than injecting it directly,

NR	Reviewer	Page	Line	Comment	Action / Reply
					reduction in circulating oestradiol levels will do the same, although mostly with mild levels of response of the ovaries. Administering oestradiol itself will only suppress FSH levels, in older women too.
91	Willem Verpoest	12	367	Stating that obtaining only a few oocytes is an agonizing condition, is exaggerated; there are many bvariables that affect the prognosis. Agree that it is disappointing and produces a lower prognosis. Please change the word agonizing.	This may indeed be a too strong expression, however, from our current way of labelling the low response this term is perhaps not so much a misnomer. The text was adjusted.
91	Willem Verpoest	12	373	Typo: clinically useful strategies (not clinical); besides: ..are awaited; besides, not sure they are awaited, as it has never been established that producing extra oocytes in a low responder produces oocytes of sufficient quality	Thank you for the correction and suggestion, the text was adapted.
91	Willem Verpoest	13	397	Typo: potentially life-threatening condition (not: potential)	This was adapted
127	ESHRE SIG RE	12	347	The sentence reads “Only one compound is delivered in micrograms” – Both corifollitropin alfa and follitropin delta are delivered in mcgs.	The reviewers are correct, the text was adapted.
127	ESHRE SIG RE	13	400	“human chorion gonadotrophin” should be “human chorionic gonadotrophin”	Thank you for the correction.
137	Jayesh Amin	12	382	This is certainly much dependant on the Antral Follicle Count or AMH result.	The GDG agrees with the reviewer.
156	Adrija Kumar Datta Stuart Campbell Geetta Nargund	13	389	We appreciate that the GDG has recognized that there is unlikely to have a causal relationship between the oocyte number and cumulative live birth rate (cLBR). On that note, those studies that showed rising cLBR with increasing oocyte yield, have also showed that women with good prognostic factors e.g. younger age and higher ovarian reserve are more likely to yield more embryos and achieve a live birth (Chen, et al., 2015, Ji, et al., 2013,	I must agree on the point raised here. Still, if we believe in the good of many oocytes, we will strive for the many by advising higher FSH dosing. It appears however from the studies by the Ireland group that high FSH exposure relates to lower LBRs, even if we would discard the low responders. With this, the urge to refrain from

NR	Reviewer	Page	Line	Comment	Action / Reply
				Polyzos, et al., 2018) and cLBR rises with increasing ovarian reserve (Chen, et al., 2015). Point to note that mean starting and total dose of gonadotropin were lower in those who had a live birth than dose who did not have a livebirth. (Chen, et al., 2015, Polyzos, et al., 2018). Thus, "more" oocytes or live births are not a result of "higher" the stimulation dose; rather the inverse appears to exist (Chen, et al., 2015, Shaia, et al., 2020).	very high FSH exposure (> 300 IU per day) is strong. In the very end the lack of support for causality from number or FSH exposure and LBRs, implies that any response is good, with the exception of the high response.
156	Adrija Kumar Datta Stuart Campbell Geetta Nargund	13	390-393	Comment: Although systematic reviews of randomized controlled trials (RCTs) found higher oocyte yield with higher stimulation dose, all RCTs that compared the mean number or proportion of high-grade embryos between low (≤ 150 IU) and higher dose found no difference in poor as well as normal responder patients (Datta, et al., 2021). This may explain why LBR and cLBR were no different between low and high stimulation dose, though the data on cLBR is limited to only a few RCTs (Liu, et al., 2020, Oudshoorn, et al., 2017, van Tilborg, et al., 2017) and meta-analysis of those RCTs (Datta, et al., 2021). Rising cLBR with increasing oocyte yield is related to the prognostic factors while cLBRs seem to be no different when these prognostic factors (age, ovarian reserve etc) are matched within a RCT.	The GDG could not agree more. And we think we have tried to make that point sufficiently.
1. Ovarian response prediction					
106	Rishma Dhillon Pai	15	R1	For predicting high and low response to ovarian stimulation, use of either antral follicle count (AFC) AND or anti-Müllerian hormone (AMH) is recommended. (Add AND OR AMH IS RECOMMENDED)	The GDG does not agree to recommend both as this is not supported by the current evidence.
106	Rishma Dhillon Pai	24	470	EXTREME ovarian response (not extremely)	The reviewer is correct, this was adapted.

NR	Reviewer	Page	Line	Comment	Action / Reply
116	Gatagazheva Aza Aslanovna	29	R2 R3	"Basal LH not recommended" for ovarian response (Rec 2) and pregnancy prediction (Rec 3) ignores the critical role of mid-stimulation LH. Data show: Low mid-stimulation LH (< 4 IU/L) reduces cumulative live birth in GnRH-antagonist cycles (OR 0.76; 95 % CI 0.60–0.97; P = 0.014) . Premature LH rise in women ≥ 37 years lowers cumulative live birth and correlates with poor embryo potential . | Reword Rec 2–3: "Basal LH alone is insufficient; measure serum LH on Day 5–6 of stimulation to detect and manage: • Low LH (below physiological threshold) – risk of reduced cumulative live birth; • Premature/high LH (> 10 IU/L) – risk of premature luteinization and compromised embryo potential." | | 124 | Rec 80–81 (Part E: Monitoring) | Currently no LH measurement is specified under hormonal monitoring . Clinicians have no tool to identify cycles at risk of low or excessive LH. My IFFS 2025 case series confirms that targeted LH modulation with dydrogesterone improves outcomes in high-LH patients. | Add new recommendation: "During COS, measure serum LH on Day 5–6 and aim to maintain it within 1.2–10 IU/L by: • If LH < 1.2 IU/L – reduce GnRH-analogue dose by 30–50 % or add 75 IU r-LH; • If LH > 10 IU/L – administer dydrogesterone 10 mg BID, titrate to LH < 10 IU/L." | | 58–59 | Rec 25 (Part C: Pituitary suppression) | Rec 25 endorses "freeze-all with progestin" without dosage/timing details for LH control . Yet progestins (e.g., dydrogesterone) exert dose-dependent LH suppression in PPOS (Yu et al. 2018) and in my practice. | Amend Rec 25: "If a 'freeze-all' strategy is planned, consider progestin-primed ovarian stimulation (e.g., dydrogesterone 10–20 mg/day from Day 5) to achieve dose-dependent LH modulation; further prospective studies should define optimal regimens.	The GDG does not consider that there are enough data to support measurement of LH before or during ovarian stimulation.

NR	Reviewer	Page	Line	Comment	Action / Reply
39	Ahmed Elsayed Hassan Hamed Elbohoty	30	600	The prediction of ovarian response categories by age alone is not sufficiently reliable. While age alone is not a reliable predictor of ovarian response quantity, it remains a key determinant of oocyte, embryo quality and live birth potential, and should be considered alongside ovarian reserve markers (Broekmans FJ et al., Hum Reprod Update, 2009; Nelson SM et al., Hum Reprod, 2011).	In the justification it is explained that age alone has some, however insufficient, predictive value for ovarian response. In chapter 2 it is discussed that age alone is the strongest predictor of ongoing pregnancy.
127	ESHRE SIG RE	30	620	However, as it is later acknowledged "As all original studies have been performed using different assays or ranges for AFC and AMH, it is not possible to combine these data to calculate cut-offs for the prediction of a low or high response.". Therefore, what is the practical significance of this recommendation? Given the heterogeneity in the definitions of low and high ovarian response and the cut-offs used, and the lack of specific practical guidance, can this be a "strong" recommendation?	Although there are different assays, they all show the same effect. However, it is not possible to define a cut-off that is applicable to all assays.
132	The Chinese Expert Review Panel for ESHRE OS Guideline	15	R2 R3	Do not totally agree with the original statement. Suggestion: These factors are not recommended as independent predictors of ovarian response / pregnancy and live birth, which means NOT take any of them alone as a predictor. Because 1. These factors are still used as references in clinical practice, such as age. 2. Each factor cannot be used as an independent predictor is reasonable and acceptable for clinical practice.	In the justification it is explained that each of these factors alone may have some, however insufficient, or no predictive value for ovarian response.
7	Sujoy Dasgupta	31	625	The clinicians should keep it in mind that the correlation between AMH and AFC is poor, particularly in poor responders. Arvis P, Rongières C, Pirrello O, Lehert P. Reliability of AMH and AFC measurements and their correlation: a large multicenter study. J Assist Reprod Genet. 2022 May;39(5):1045-1053.	The GDG advices to use either AMH or AFC. The justification explains that the GDG did not study the effect of using both.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Again, both AMH and ultrasound should not be used to identify high responders who have polycystic ovarian morphology (PCOM) to avoid overdiagnosis. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin J, Redman LM, Boyle JA, Norman RJ, Mousa A, Joham AE. Recommendations From the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2023 Sep 18;108(10):2447-2469.	
128	Apostolos Tsironis	31	621	Possibly we need to change the wording of the recommendation: age, BMI should not be used "ALONE" for the prediction of response. The current wording indicates that both age and BMI should not be taken into account for response prediction which in my opinion is incorrect.	In the justification it is explained that age alone has some, however insufficient, predictive value for ovarian response. It also states that BMI alone is not a predictor for ovarian response.
2. Pregnancy prediction					
137	Jayesh Amin	/	/	In case of GnRH Agonist trigger, LH estimation can be recommended on day 2 of menses, on the day of trigger and 12 hours post trigger ; to predict the chances of sub-optimal response to trigger	This was deemed to be outside the scope of the current guideline update.
116	Gatagazheva Aza Aslanovna	15	R2 R3	"Basal LH not recommended" for ovarian response (Rec 2) and pregnancy prediction (Rec 3) ignores the critical role of mid-stimulation LH. Data show: Low mid-stimulation LH (< 4 IU/L) reduces cumulative live birth in GnRH-antagonist cycles (OR 0.76; 95 % CI 0.60–0.97; P = 0.014) . Premature LH rise in women ≥ 37 years lowers cumulative live birth and correlates with poor embryo potential . | Reword Rec 2–3: "Basal LH alone is insufficient; measure serum LH on Day 5–6 of stimulation to detect and manage: • Low LH (below physiological threshold) – risk of reduced cumulative live birth; • Premature/high LH (> 10 IU/L) – risk of premature luteinization and compromised embryo potential." | | 124 | Rec 80–81 (Part E: Monitoring) | Currently no LH measurement is	The GDG does not consider that there are enough data to support measurement of LH before or during ovarian stimulation.

NR	Reviewer	Page	Line	Comment	Action / Reply
				specified under hormonal monitoring . Clinicians have no tool to identify cycles at risk of low or excessive LH. My IFFS 2025 case series confirms that targeted LH modulation with dydrogesterone improves outcomes in high-LH patients. | Add new recommendation: "During COS, measure serum LH on Day 5–6 and aim to maintain it within 1.2–10 IU/L by: • If LH < 1.2 IU/L – reduce GnRH-analogue dose by 30–50 % or add 75 IU r-LH; • If LH > 10 IU/L – administer dydrogesterone 10 mg BID, titrate to LH < 10 IU/L." | | 58–59 | Rec 25 (Part C: Pituitary suppression) | Rec 25 endorses "freeze-all with progestin" without dosage/timing details for LH control . Yet progestins (e.g., dydrogesterone) exert dose-dependent LH suppression in PPOS (Yu et al. 2018) and in my practice. | Amend Rec 25: "If a 'freeze-all' strategy is planned, consider progestin-primed ovarian stimulation (e.g., dydrogesterone 10–20 mg/day from Day 5) to achieve dose-dependent LH modulation; further prospective studies should define optimal regimens.	
132	The Chinese Expert Review Panel for ESHRE OS Guideline	15	R2 R3	Do not totally agree with the original statement. Suggestion: These factors are not recommended as independent predictors of ovarian response / pregnancy and live birth, which means NOT take any of them alone as a predictor. Because 1. These factors are still used as references in clinical practice, such as age. 2. Each factor cannot be used as an independent predictor is reasonable and acceptable for clinical practice.	The aim of the guideline was to evaluate the evidence supporting the use of these variables in clinical practice.
132	The Chinese Expert Review	15	R4	Do not agree with taking BMI as a predictor for both pregnancy and live birth, and the strength may NOT be strong. Agree with that female age is a predictor of pregnancy and live birth.	The Guideline Development Group (GDG) does not concur with the assertion that evidence is lacking regarding BMI as a predictor of

NR	Reviewer	Page	Line	Comment	Action / Reply
	Panel for ESHRE OS Guideline			Because 1. Lack of evidence on demonstrating BMI as a predictor of live birth. 2. BMI is found negatively linearly correlated with live birth rate only when BMI is higher than 25 kg/m2 but the relationship is not clear when BMI is low.	pregnancy and live birth. Accordingly, the justification has been revised to reflect this position.
106	Rishma Dhillon Pai	34	739	AFC alone is not a predictor for the outcome pregnancy. (OUTCOME PREGNANCY iIS NOT CLEAR TERMINOLOGY)	The overall outcome of pregnancy pertains to the achievement of an ongoing or clinical pregnancy.
7	Sujoy Dasgupta	35	766	“AMH alone is not a predictor of the outcome pregnancy”. This sentence needs further explanation. One meta-analysis showed that low values of ovarian reserve tests (AMH and AFC) correlate with higher risk of miscarriage. Busnelli A, Somigliana E, Cirillo F, Levi-Setti PE. Is diminished ovarian reserve a risk factor for miscarriage? Results of a systematic review and meta-analysis. Hum Reprod Update. 2021 Oct 18;27(6):973-988. doi: 10.1093/humupd/dmab018. PMID: 34254138. Another very recent meta-analysis also confirmed that both low AMH and low AFC were associated with increased risk of miscarriage. However, subgroup analysis found that the risk was significantly higher only in women age below 35 years. Kasaven LS, Anson N, Jones BP, Odia R, Cordero J, Nagi JB , Theodorou E., Is ovarian reserve associated with increased risk of miscarriage? A systematic review and meta-analysis, Reprod BioMedicine Online (2025), doi: https://doi.org/10.1016/j.rbmo.2025.105041	The cited studies address the association between AMH levels and miscarriage, which represents an intermediate outcome. However, the PICO question explicitly concerns the prognostic value of AMH in predicting the overall outcome of pregnancy—namely, the achievement of an ongoing or clinical pregnancy—rather than its individual components. While miscarriage is undoubtedly a factor that can affect pregnancy outcomes, current evidence does not support AMH as a reliable overall predictor of pregnancy success following IVF.
7	Sujoy Dasgupta	36	815-816 831-832	I feel these two statements are contradictory and may increase the confusion among the readers. “Due to the low incidence it seems unnecessary to evaluate this research question for progesterone levels >1.6 ng/ml on cycle day 3.”	Available evidence suggests that elevated progesterone levels at the intended start of ovarian stimulation (typically cycle day 2) may be associated with reduced pregnancy rates.

NR	Reviewer	Page	Line	Comment	Action / Reply
				"Assessment of progesterone prior to initiation of stimulation on cycle day 2 in women undergoing ovarian stimulation with GnRH antagonist and gonadotrophins may be beneficial to identify cases".	Accordingly, measuring progesterone at this specific timepoint could help identify patients with a potentially lower likelihood of treatment success, particularly when levels exceed clinically relevant thresholds. However, the prevalence of elevated progesterone levels on cycle day 3 appears to be very low—for instance, 0.3% as reported by Faulisi et al. (2017). Given this low incidence, the guideline group concluded that specifically investigating progesterone levels >1.6 ng/mL on cycle day 3 holds limited clinical relevance and is therefore not justified.
1	Raj Mathur	38	894	It would be useful to have a recommendation on BMI level and its effect on live birth. This is potentially an 'actionable' area where the guidance could help clinicians deliver good evidence-based care. At present, it is hard to know what to make of the recommendation? Is it high BMI or low that we should be worried about, and what level should we ask patients to strive to reach? As a technical point, the recommendation concerning Age and BMI is not phrased like a recommendation, but more like a statement of evidence.	The statement is correct. It is mostly intended to inform the patient. The justification was adapted.
7	Sujoy Dasgupta	39	907	"The recommendation is not applicable to patients >39 years of age."- This sentence needs further explanations.	This refers to the age of the female participants included in the relevant studies, with the aim of enhancing the applicability and extrapolation of the results to broader clinical populations.
114	Monica Varma	39	905	Which blood test is required at initiation of stimulation cycle Day 2 ?	This refers to the assessment of oestradiol levels to rule out the presence of developing follicles

NR	Reviewer	Page	Line	Comment	Action / Reply
					or simple ovarian cysts at the intended start of the stimulation cycle.
3. Pre-treatment therapies					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	15	R8	Suggestion：Remove the specific time (12-28 days) for COCP pretreatment Because The clinical efficacy of COCP is uncertain when it is not used within the specific time	The GDG reviewed the comment. The day's range specified in the recommendation was based on the time range from the included RCTs. The GDG agrees that there is no evidence with regards to longer or shorter exposure. Therefore the GDG agreed to take out the specified time range.
91	Willem Verpoest	15	R9	There is insufficient evidence to state that a minimum COCP wash-out period of 5 days should be applied.	The GDG agreed to reformulate the GPP with more caution.
144	Karolina Palinska-Rudzka	15	R9	The recommendation for a fixed "minimum 5-day wash-out" after COCP use prior to fresh transfer appears to lack a clear physiological or trial-based explanation. Importantly, it does not specify the duration of COCP use, an aspect likely to be more clinically relevant than the wash-out interval itself. Recent studies have shown no detrimental effect on outcomes with shorter wash-out periods, provided total COCP exposure was limited: • Celik et al. (2019) and Rombauts et al. (2021) reported no adverse impact on pregnancy rates with 2–3 day or no wash-out, where COCP use did not exceed 24–28 days [8,9]. I would welcome clarification from the GDG as to the basis for the 5-day recommendation and whether a more flexible approach might be acceptable in clinical practice, particularly where scheduling or laboratory logistics are a factor, and duration of COCP is considered.	The GDG agreed to reformulate the GPP with more caution.
127	ESHRE SIG RE	41	980	A more recent systematic review and meta-analysis (including also a network meta-analysis) assessing estrogen pretreatment in the GnRH	The GDG has evaluated the Venetis 2023 network meta-analysis. Due to methodological

NR	Reviewer	Page	Line	Comment	Action / Reply
				antagonist protocol (Venetis et al., 2023, Human Reproduction Update, doi: 10.1093/humupd/dmac040) has been published (including the newer RCT published in 2022 by Fernandez-Prada).	shortcomings, this network meta-analysis could not be included in the evidence section. However, its results were added in a sentence in the justification.
127	ESHRE SIG RE	42	1018-1019	A more recent systematic review and meta-analysis (including also a network meta-analysis) assessing progestogen pretreatment in the GnRH antagonist protocol (Venetis et al., 2023, Human Reproduction Update, doi: 10.1093/humupd/dmac040) has been published.	The GDG has evaluated the Venetis 2023 network meta-analysis. Due to methodological shortcomings, this network meta-analysis could not be included in the evidence section. However, its results were added in a sentence in the justification.
37	Ahmed Samy Abdelazim Saad	43	1030	Also, can be used for synchronization of follicular growth	This is mentioned in the introduction to the chapter.
127	ESHRE SIG RE	43	1041-1043	A more recent systematic review and meta-analysis (including also a network meta-analysis) assessing OCP pretreatment in the GnRH antagonist protocol (Venetis et al., 2023, Human Reproduction Update, doi: 10.1093/humupd/dmac040) has been published (including the newer RCT published in 2022 by Fernandez-Prada).	The GDG has evaluated the Venetis 2023 network meta-analysis. Due to methodological shortcomings, this network meta-analysis could not be included in the evidence section. However, its results were added in a sentence in the justification.
127	ESHRE SIG RE	44	1062	A minimal wash out period of 5 days should be applied if COCP is used for programming cycle in the case of a fresh transfer. Can you please provide the evidence for this GPP? The meta-analysis by Venetis et al., 2023 showed that “There was a significantly lower OPR with an OCP-free interval of 4–5 days compared with no pretreatment (RR 0.81, 95% CI 0.66 to 0.99; I2 = 0%; 3 RCTs; n = 1020 participants; low-quality evidence)”	The GDG agreed to reformulate the GPP with more caution.

NR	Reviewer	Page	Line	Comment	Action / Reply
39	Ahmed Elsayed Hassan Hamed Elbohoty	45	1100	GnRH antagonist pre-treatment before ovarian stimulation in a delayed-start gonadotrophin protocol is probably not recommended. [2019] I agree that routine use is not supported by current evidence but I think that adding GnRH antagonist pre-treatment may be considered selectively to modify the stimulation start date or improve follicular synchronization in poor responders. Studies such as Cakmak et al., Hum Reprod., 2014 and Eftekhar et al., Iran J Reprod Med., 2019 explored this approach. Although live birth benefit has not been demonstrated, it may offer flexibility in cycle scheduling	The RCT by Eftekhar et al. is included in the body of evidence; the study by Cakmak et al. was not because of its retrospective nature. In addition, as stated in the introduction of the chapter, cycle scheduling is not within the scope of the guideline.
4. Ovarian stimulation protocols					
114	Monica Varma	16	R13	A reduced gonadotropin dose is probably recommended to decrease the risk of OHSS in predicted high responders. Is it specific for a Fresh Embryo Transfer?	Very high oocyte yield can be associated with early OHSS. In this context we believe that the recommendation is balanced to include also these cases
132	The Chinese Expert Review Panel for ESHRE OS Guideline	16	R18	Do not totally agree with the original statement that the use of modified natural cycle is probably not routinely recommended over conventional stimulation for low responders. Suggestion: Modified natural cycle can be considered in Low responders Because 1. Latest evidence shows similar pregnancy rates with controlled ovarian stimulation. 2. The evidence cited in the guideline was published too long ago.	Unfortunately, the reviewers have not provided references for the GDG to check.
156	Adrija Kumar Datta Stuart Campbell	16	R18 R19	We appreciate that the GDG recognizes the place for Natural Modified IVF in women with “very low” ovarian reserve. However, slightly different opinion about Modified Natural IVF stated in Recommendation 18 may confuse the reader unless the indications are more clearly specified.	We do not believe it is confusing since R18 states the MNC is probably not routinely recommended and R19 refers as a GPP to the group that it can be offered to: women with very low reserve

NR	Reviewer	Page	Line	Comment	Action / Reply
	Geetta Nargund				
132	The Chinese Expert Review Panel for ESHRE OS Guideline	16	R20 R21	Suggestion：Merge these two recommendations Because The recommended dosages for both recommendations are limited to below 300 IU	The GDG formulated two separate recommendations to emphasize that most low responders benefit from conventional gonadotropin dosing. However, they understand how the message of both recommendations may be similar, and have agreed to take out the first recommendation.
156	Adrija Kumar Datta Stuart Campbell Geetta Nargund	16	R20 R21	We appreciate that the GDG identified that a stimulation dose >225 IU/ day may not improve the pregnancy outcome. From multiple systematic reviews of RCTs comparing between 150 IU and higher stimulation dose (Datta, et al., 2021, Ngwenya, et al., 2024, Song, et al., 2016), we are now confident that a higher than a standard stimulation dose (150 IU/ day) may not improve the pregnancy outcomes in the poor responders. Having said that, stimulation dose up to 300 IU/ day stimulation in recommendation 21 appears somewhat contradictory unless its place is specified with quoted evidence.	The GDG formulated two separate recommendations to emphasize that most low responders benefit from conventional gonadotropin dosing. However, they understand how the message of both recommendations may be similar, and have agreed to take out the first recommendation.
156	Adrija Kumar Datta Stuart Campbell Geetta Nargund	48	1208	GDG may refer to ICMART/ WHO's glossary to define Modified Natural IVF which is: An IVF procedure in which one or more oocytes are collected from the ovaries during a spontaneous menstrual cycle. Drugs are administered with the sole purpose of blocking the spontaneous LH surge and/or inducing final oocyte maturation (Zegers-Hochschild, et al., 2017).	The reference was added to the text.
1	Raj Mathur	49	1243	In my opinion, this recommendation is too strong. For the reasons stated in the justification, it is hard to say that a 'reduced' dose of FSH in GnRH antagonist cycles is evidence-based to the extent that would justify such a recommendation. It may be true for agonist cycles, but the Antagonist	The second recommendation was adapted.

NR	Reviewer	Page	Line	Comment	Action / Reply
				should be preferred to the Agonist for this group of patients. A recent evidence-based guideline came to different conclusions for Agonist and Antagonist cycles Tsampras, N., Palinska-Rudzka, K., Alebrahim, Y., Craciunas, L., & Mathur, R. (2024). Prevention of ovarian hyperstimulation syndrome (OHSS): British Fertility Society policy and practice guideline. Human Fertility, 28(1). https://doi.org/10.1080/14647273.2024.2441827 Further, the recommendation that of a 'reduced' gonadotropin dose is difficult for clinicians to act upon, unless it is also stated what the dose is. This is not possible from the evidence, as the justification describes very clearly.	
37	Ahmed Samy Abdelazim Saad	49	1243	some high responders PCOS don't respond except with a step- down protocol. In such cases we begin ovarian stimulation with high dose of FSH then decrease the dose.	We agree with the comment, That is the reason why we used the term probably and also this is a conditional recommendation
126	Emre Goksan Pabuccu	49	1243	The GnRH antagonist protocol is recommended for predicted high responders. However, if GnRH agonist protocols are used, a reduced gonadotropin dose is recommended to decrease the risk of OHSS. [updated]à'..a reduced gonadotropin dose is recommended to decrease the risk of OHSS; however, a high number of follicles >10 mm in diameter may still confer a considerable OHSS risk."1	Unfortunately, the GDG does not understand the point the reviewer is trying to make.
127	ESHRE SIG RE	49	1243	A reduced gonadotropin dose is probably recommended to decrease the risk of OHSS in predicted high responders. [2025] What is considered "a reduced dose" for these patients? Can the GDG provide more practical guidance? (especially considering how q similar recommendation is phrased for normal and low responders)	Conventional dosing is 150-225 IU. In predicted high responders, a reduced gonadotropin dose (100 to <150 IU) is probably recommended, based on other patient characteristics, oocyte trigger and embryo transfer strategy.

NR	Reviewer	Page	Line	Comment	Action / Reply
137	Jayesh Amin	53	1353	If antral follicle count is 3 or less, minimal stimulation protocol by using 75 or 150 IU HMG with or without adding clomiphene citrate can be recommended.	This chapter refers to MNC that's why the recommendation refers to the MNC and not to minimal stimulation
37	Ahmed Samy Abdelazim Saad	53	1370	I would suggest rephrasing it to equally recommended	Although the evidence is of low quality it does not equally support the use of MNC over ovarian stimulation for all patients categories. Due to the lack of robust evidence we used the term probably and the recommendation was conditional . As a GPP we added that MNC can be used in women with very low ovarian reserve
37	Ahmed Samy Abdelazim Saad	53	1371	clinicians could choose to use a modified natural cycle. I would add to that: especially in a repeated stimulation cycle after a poor (or low) response in a conventional stimulation cycle	Although the GDG understands how the question came about, the PICO does not refer to this specific population.
138	Philippe Pinton*	54	20-21	We would also like to address a potential ambiguity between recommendations 20 and 21: • Recommendation 20: "A higher gonadotropin dose is probably not recommended over conventional (150–225 IU) for predicted low responders." • Recommendation 21: "A gonadotropin dose higher than 300 IU is not recommended for predicted low responders." While Recommendation 21 clearly sets an upper limit (≥300 IU) with a strong recommendation not to exceed it, Recommendation 20 seems to take a step back by advising against doses above 150–225 IU. This overlap on the same patient's subgroup may create confusion for clinicians: if doses up to 300 IU are deemed acceptable, why is there a separate caution against exceeding the 150–225 IU range? To improve clarity and consistency, we suggest merging or rephrasing	The GDG formulated two separate recommendations to emphasize that most low responders benefit from conventional gonadotropin dosing. However, they understand how the message of both recommendations may be similar, and have agreed to take out the first recommendation.

NR	Reviewer	Page	Line	Comment	Action / Reply
				these two recommendations to communicate a unified message that: • Confirms simply 300 IU as the upper limit not to be exceeded Such refinement would support more transparent, consistent, and actionable guidance for managing this challenging patient population.	
5. Pituitary suppression regimens					
106	Rishma Dhillon Pai	16	R23	The GnRH antagonist protocol is recommended over the GnRH agonist protocols given the comparable efficacy and higher safety in the general IVF/ICSI population. (IS GENERAL IVF POPULATION DEFINED OR IS NORMAL RESPONDER PREFERABLE?)	General population is a broader definition that includes normal responders, high responders and most of poor responders.
106	Rishma Dhillon Pai	16	R25	If freeze-all is planned, the use of progestin for pituitary suppression is probably equally recommended to GnRH analogues OR GnRh Antagonists (add GnRh antagonists)	Analogues include agonists and antagonists.
123	Alberto Vaiarelli	57	1473	Pituitary suppression regimes: The PPOS protocol has proven to be effective not only in conventional ovarian stimulation protocols when Fresh ET is not planned but also in unconventional stimulation settings (is Duostim) Retrospective data have demonstrated that its use does not negatively impact oocyte competence, embryonic competence, or ovarian sensitivity, as measured by the Follicle-Oocyte Index (FOI). To date, this is the first published study to investigate the combination of the PPOS protocol with Duostim. REF: A multicycle approach through DuoStim with a progestin-primed ovarian stimulation (PPOS) protocol: a valuable option in poor prognosis patients undergoing PGTA. Vaiarelli A, Pittana E, Cimadomo D, Ruffa A, Colamaria S, Argento C, Giuliani M, Petrone P, Fabozzi G, Innocenti F, Taggi M, Ata B, Rienzi L, Ubaldi FM. J Assist Reprod Genet. 2025 Jan;42(1):255-264. doi: 10.1007/s10815-024-	The GDG does not understand the point the reviewer is trying to make. In addition, only RCTs were considered for inclusion in the duostim section.

NR	Reviewer	Page	Line	Comment	Action / Reply
				03317-0. Epub 2024 Nov 13. PMID: 39538089	
39	Ahmed Elsayed Hassan Hamed Elbohoty	59	1545-1548	The guideline currently states that “a systematic review and meta-analysis, including 7 RCTs, compared fixed and flexible GnRH antagonist protocols (Venetis et al., 2023) showed no significant difference in ongoing pregnancy rate...” This is inaccurate and not consistent with the published findings. In fact, Venetis et al. (2023) reported a statistically significant difference favoring fixed protocols (RR 0.76 [95% CI: 0.62–0.94], p=0.02) based on analysis from 6 RCTs, not 7. We recommend correcting this statement to accurately reflect the published data.	The recommendation was adapted.
37	Ahmed Samy Abdelazim Saad	59	1550	As regards to the mentioned systematic review and meta-analysis by Venetis et., 2023: This is coated from the meta-analysis in their own words: “A flexible GnRH antagonist protocol resulted in a significantly lower ongoing pregnancy rate (OPR) compared with a fixed Day 5/6 protocol (relative risk (RR) 0.76, 95% CI 0.62 to 0.94, I2 = 0%; 6 RCTs; n = 907 participants; low certainty evidence) There was a lack of data regarding live birth when comparing the flexible and fixed GnRH antagonist protocols or cetrorelix and ganirelix. There was insufficient evidence of a difference between fixed/flexible or OCP pretreatment/no pretreatment interventions regarding other outcomes, such as ovarian hyperstimulation syndrome and miscarriage rates. So, we cannot conclude that both fixed and flexible are equally recommended. Also, by experience we may notice more premature ovulation or luteinization with flexible protocol.” So, it can be rephrased to: both flexible & fixed protocol can be used but	The GDG respectfully disagrees. The evidence shows that the ongoing pregnancy rate is significantly lower with the flexible protocol. Since data on live birth are still insufficient, ongoing pregnancy remains the best available estimate. It already accounts for early pregnancy loss and is therefore a reliable surrogate for live birth. Additionally, the comparison of OHSS rates between the two protocols has not been adequately addressed, but there is no biologically plausible reason to expect a difference in this outcome.

NR	Reviewer	Page	Line	Comment	Action / Reply
				no available strong evidence about a preference of one over the other. Or we may reword it to: The flexible GnRH antagonist protocol is probably as good as fixed GnRH antagonist protocol.	
39	Ahmed Elsayed Hassan Hamed Elbohoty	59	1550	The flexible and fixed GnRH antagonist protocol is probably equally recommended. [2025] The meta-analysis by Venetis et al. (2023) is based on a general IVF population and does not account for responder subgroups. I recommend specifying that in poor responders, the flexible protocol may offer clinical advantages, such as improved follicular recruitment, lower cancellation rates, and potentially better outcomes. This is supported by Lainas et al. (2008) and Yildiz et al. (2022), both of which evaluated these protocols in poor responder cohorts. Including this distinction would improve the applicability of the recommendation to individualized care.	The GDG agrees, the recommendation was adapted.
127	ESHRE SIG RE	59	1550	The flexible and fixed GnRH antagonist protocol is probably equally recommended. [2025] The systematic review and meta-analysis by Venetis et al. 2023, clearly states that “A flexible GnRH antagonist protocol resulted in a significantly lower OPR compared with a fixed Day 5/6 protocol (relative risk (RR) 0.76, 95% CI 0.62 to 0.94, I2 = 0%; 6 RCTs; n = 907 participants; low certainty evidence)”. This originates from the pairwise comparison. Please note that in none of these RCTs was a pre-treatment used in any arm. The GDG seems to have based their recommendation on the result of the network meta-analysis comparing the flexible no pretreatment protocol to the fixed no pretreatment protocol which showed a RR for ongoing pregnancy of: 0.85, 95% CI: 0.73-1.00).	The recommendation was adapted.

NR	Reviewer	Page	Line	Comment	Action / Reply
				However, it is customary for the pairwise comparisons to take precedence compared to the NMA, especially when formulating recommendations. Moreover, the RR of the NMA still implies a potentially negative effect although marginally not significant.	
127	ESHRE SIG RE	59	1559-1561	"Although there is high heterogeneity in RCTs comparing flexible to fixed GnRH antagonist protocols, results show that live birth and ongoing pregnancy rates are similar with a flexible GnRH antagonist protocol" The systematic review and meta-analysis by Venetis et al. 2023, states that "A flexible GnRH antagonist protocol resulted in a significantly lower OPR compared with a fixed Day 5/6 protocol (relative risk (RR) 0.76, 95% CI 0.62 to 0.94, I2 = 0%; 6 RCTs; n = 907 participants; low certainty evidence)". This shows a difference in ongoing pregnancy rates and also 0% statistical heterogeneity. It is not clear, therefore, how this statement is justified.	The recommendation was adapted.
127	ESHRE SIG RE	61	1624	If freeze-all is planned, the use of progestin for pituitary suppression is probably equally recommended to GnRH analogues. The GDG argues that PPOS has similar efficacy to GnRH analogues but has lower cost, is easier and more patient friendly. Would that potentially render PPOS a more attractive options (based on the principle of "simplicity" discussed in other parts of the document) and therefore warrant a more supportive statement?	The GDG believes that the statement "equally recommended" is the most appropriate, as both protocols are comparable in terms of efficacy and safety. Simplicity is a more subjective and debatable criterion, and we should be cautious and not to consider it decisive.
39	Ahmed Elsayed Hassan Hamed Elbohoty	61-62	1624-1643	If freeze-all is planned, the use of progestin for pituitary suppression is probably equally recommended to GnRH analogues. [updated] The guideline appropriately supports the use of progestin-based pituitary suppression (PPOS) in freeze-all cycles and summarizes the available evidence for safety and efficacy. However, it does not explicitly address cost-effectiveness, which is a relevant factor in clinical decision-making. I	While the group acknowledges the relevance of the comment, cost-effectiveness analyses are considered to fall outside the scope of this guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				recommend that the guideline acknowledge the economic advantages of PPOS, especially in fertility preservation and PGT cycles, where fresh transfer is not intended. Recent studies support this point: Shi et al. (2023) showed that among 163 fertility preservation patients, PPOS achieved comparable oocyte yield to GnRH antagonist protocols with significantly lower medication cost and fewer injections (Reprod Biol Endocrinol, PMID: 37265085). Stimpfel et al. (2024) demonstrated that in donor cycles, PPOS was more cost-effective while yielding comparable numbers of metaphase II oocytes. Incorporating this perspective would strengthen the clinical and economic rationale for recommending PPOS in freeze-all scenarios.	
44	Nisha bhatnagar	61	1637	Have used dydrogesterone for a very long time as oral antagonist in PPOS(progesterone primed ovarian stimulation),with no reported side effects	Individual experience and practice, while valuable is not a substitute for published clinical studies.
6. Types of gonadotropins and other ovarian stimulation drugs					
82	Eduardo Correa Allende*	/	/	A. Gonadotropins. Inclusion of EMA-approved follitropin alfa biosimilars, alongside other available follitropins (alfa, beta, and delta), as equally recommended options for ovarian stimulation.	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.
114	Monica Varma	/	/	Do we need to specify it as HP HMG or is it presumed ?	The comment lacks clarity regarding the specific section or statement to which it refers.
125	Kokkoni Kiose	/	/	I would like to respectfully draw the panel's attention to a notable omission in the updated section concerning the types of gonadotrophins and other ovarian stimulation agents. Since the publication of the 2020 ESHRE guideline, two systematic reviews (Chua et al., 2021, Kiose et al., 2025) have been published evaluating the clinical efficacy and safety associated with biosimilars follitropin alfa compared to the reference product. Despite their methodological rigor and clinical relevance, the	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.

NR	Reviewer	Page	Line	Comment	Action / Reply
				updated draft guideline does not address or reference these biosimilars. Importantly, both reviews suggest that treatment with biosimilars of follitropin alfa may be associated with inferior live birth rates compared to the originator product—an observation with direct clinical implications. Such an inclusion would not only ensure scientific completeness but also align with ESHRE's mission to provide guidelines that reflect the totality of the available evidence. Could the Guideline Development Group (GDG) clarify the rationale for this omission?	
127	ESHRE SIG RE	/	/	We could not find recommendations regarding the use of biosimilars of follitropin alfa when compared with the originator especially considering the recently published meta-analysis of RCTs by Kiose et al., 2025 (Human Reproduction) which suggest inferior live birth, ongoing and clinical pregnancy rates with the biosimilars compared to the originator. Please note that the same finding was also observed regarding the biosimilars that are currently being used in Europe.	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.
128	Apostolos Tsironis	/	/	There is no reference to the use of biosimilars in ovarian stimulation – it would be really important to hear the views of the guideline committee on the subject and whether they are recommended equally with other stimulation agents	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.
143	Colin Howles	/	/	ADD Recombinant FSH (rFSH) biosimilars vs reference recombinant FSH (follitropin alfa) Evidence A biosimilar is a biological medicine highly similar (in terms of structure, biological activity, efficacy, safety and immunogenicity profile) to another already approved biological medicine (the 'reference medicine') in the European Union (European Medicines Agency (EMA) website). Biosimilars are approved according to the same standards of pharmaceutical quality, safety and efficacy that apply to all biological medicines (EMA website). Biosimilars in EU are	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.

NR	Reviewer	Page	Line	Comment	Action / Reply
				assessed and registered by EMA after having satisfied the principle of structural and functional comparability (Wolff-Holz et al., 2019). Both the UK MHRA and a recent EMA reflection paper (EMA 2025) have asserted that efficacy trials are not an effective discriminating tool in biosimilar development. The European Union monitoring system had not identified any relevant difference in the efficacy or in the adverse effects between biosimilar medicines and their reference medicines (Correia-Pinheiro et al., 2021). Comparative clinical trials carried out for regulatory submission in EU, comparing the FSH biosimilar with the reference medicine have demonstrated equivalence in terms of the primary endpoint (no of oocytes) and no significant differences in secondary endpoints including immunogenicity profile, OHSS and live birth rates (Kaplan et al 2021; Rettenbacher et al 2015; Strowitzki et al 2016a,b). A recent meta-analysis (Kiose et al 2024) is at odds with the strict regulatory terminology defining the science of biosimilarity. In this meta-analysis only 2 out of 7 FSH products incorporated into the analysis are true FSH biosimilars. One of the products included doesn't claim to be an FSH biosimilar. Thus, claiming homogeneity of the meta-analysis is erroneous. Combining data from the 2 EU approved FSH biosimilars demonstrated no significant difference in live births. Justification Since the launch of the first biosimilar in 2006, the EMA has approved 110 biosimilars (and no biosimilar medicine has been withdrawn or suspended. Additional proof of similarity is reflected in the statement from EMA & Heads of Medicines Agencies confirming that biosimilar medicines approved in the EU are interchangeable with their reference medicine or with an equivalent biosimilar (EMA). In 2024, the French medicines agency (ANSM 2024) also recommended substitution of the reference FSH medicine with biosimilars. In February 2025 (Republique Francaise 2025), a law was passed in France documenting the conditions for biosimilar FSH substitution. The available evidence, from the highest regulatory authority in Europe as well as Australia supports the same	

NR	Reviewer	Page	Line	Comment	Action / Reply
				safety and efficacy and interchangeability of FSH biosimilars with the reference medicine. Recommendation Follitropin alfa biosimilars are equally safe and effective to the reference medicine (follitropin alfa) and beta for ovarian stimulation	
138	Philippe Pinton*	64-65	R26	The current statement indicates that the use of recombinant FSH (rFSH) and human menopausal gonadotropin (hMG) for ovarian stimulation is equally recommended. We respect the guidelines' data selection criteria, which primarily consider meta-analyses. However, we would like to highlight that, based on the meta-analysis published by Bordewijk et al. (2019), which included 28 RCTs (seven of which compared HP-hMG to rFSH and reported live birth rate (LBR)), a significantly higher LBR in fresh cycles was reported for hMG compared to rFSH. Regarding the cumulative live birth rate (CLBR), we acknowledge that it is the most accurate measure of effectiveness. However, CLBR data were not consistently reported at the time the included studies were conducted, a fact reflected by only three publications including this endpoint. Furthermore, the recommendation does not distinguish between highly purified human menopausal gonadotropin (HP-hMG) and traditional (non-HP) hMG. This differentiation is critical for both patient safety and treatment efficacy in assisted reproduction. Unlike conventional hMG, HP-hMG undergoes advanced chromatographic purification, resulting in a product with minimal protein contaminants and a consistent FSH:LH ratio. This higher purity translates into tangible clinical benefits: a direct comparative study (though not an RCT) by Ismail Aboul Foutouh et al. (2007), which followed 174 patients, demonstrated that HP-hMG yields a significantly higher number of MII oocytes with lower doses than non-HP hMG, without increasing adverse effects or altering stimulation characteristics. Parsanezhad et al. (2017), which was included as an RCT to support the equality between non-HP hMG and rFSH, was not powered to assess LBR outcomes (40 patients per arm). From a safety perspective, the reduction in urinary protein contaminants in HP-	Safety Advantage of hMG Over rFSH Witz et al., 2020 – Large randomized controlled trial (RCT) - Reported lower OHSS rates with HP-hMG in high responders. However, severe OHSS still occurred in both groups: - Severe OHSS with hMG: 2.6% - Severe OHSS with rFSH: 2.9% - Despite the use of GnRH agonists for final oocyte maturation in high-risk patients, OHSS was not completely prevented. In the small RCT by Figen Turkcapar et al. (2013) the difference in OHSS(undefined) was not significant, and since GnRH agonists were used for LH supression, no possibility of contemporary management of the risk for OHSS with GnRH agonist triggerinmg was possible Clinical Importance of HP-hMG vs. Non-HP hMG Bordewijk et al., 2019 – Meta-analysis of 28 RCTs - Slightly higher live birth rate in fresh cycles with HP-hMG vs. rFSH. This difference is already acknowledged in the

NR	Reviewer	Page	Line	Comment	Action / Reply
				hMG minimizes the risk of immune reactions and injection site complications, concerns that are more pronounced with less purified preparations. The enhanced batch-to-batch consistency of HP-hMG also ensures more predictable ovarian responses and dosing, reducing the risk of overstimulation and associated complications (Claudio Wolfenson 2005). In an RCT published by Keye et al. (2005), treatment with HP-hMG led to a seven-fold reduction in injection site reactions compared to treatment with a less purified hMG. Given these clear advantages in both efficacy and safety, the transition toward HP-hMG represents a natural evolution in clinical practice. Adopting HP-hMG as the standard for ovarian stimulation not only optimizes patient outcomes but also aligns with modern expectations for pharmaceutical safety and reliability. For these reasons, we would like to ask the committee to specify the recommendation to HP-hMG. In addition, within the guidelines' framework for recommendation drafting, safety was explicitly mentioned as an important endpoint, specifically regarding reducing the risk of developing OHSS, especially in high-risk populations. In the justification section, two RCTs were included: a large RCT published by Witz et al. (2020), which demonstrated a clear safety benefit for HP-hMG versus rFSH in predicted hyper-responders; and a small RCT by Figen Turkcapar et al. (2013), which, though not powered to show a significant effect on live birth rate (LBR) or OHSS, reported no OHSS in the hMG arms while OHSS occurred in 11.9% of the patient treated with rFSH. Given the two publications used by the GDG committee for justification, and based on the data, we suggest addressing that the potential benefit of HP-hMG is preferable to rFSH in terms of safety for hyper-responders.	guideline. However, the advantage in LBR was marginal and has not been confirmed in more recent trials such as Witz et al. (2020). Therefore, the clinical significance of this finding remains uncertain.
106	Rishma Dhillon Pai	17	R29	(IS GENERAL IVF POPULATION DEFINED OR IS NORMAL RESPONDER PREFERABLE?)	General population includes, but not exclusively, normal responders.
137	Jayesh Amin	17	R29	The combination of rFSH with HMG and rFSH alone are probably equally recommended for the general IVF population.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.

NR	Reviewer	Page	Line	Comment	Action / Reply
150	Mekhala Dwarakanath B	17	R29	The combination of rFSH with rLH and rFSH alone are probably equally recommended for the general IVF population. It may be more beneficial in general population with FSH/LH Polymorphisms	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
88	Aboubakr Mohamed Elnashar	17	R29 R30 R31	In general population, poor responders and above 35 years, no benefit of adding rec lh to rec fsh	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
132	The Chinese Expert Review Panel for ESHRE OS Guideline	17	R29-31	Do not totally agree with the original statement that the combination of rFSH with rLH and rFSH alone are probably equally recommended for the general IVF population/ low responders/women of advanced age (≥35 year). Because Some patients may benefit from the combination of rFSH with rLH	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
138	Philippe Pinton*	17	R29-31	The current statement suggests that rFSH alone and the combination of rFSH + rLH are "probably equally recommended" for the general IVF population, low responders, and women of advanced maternal age. However, as acknowledged in the justification section, the available evidence does not support a clinically meaningful benefit in live birth rate (LBR) when rLH is added to rFSH. This raises the question of why a recommendation was formulated in the absence of outcome-driven evidence—particularly given the principles typically underpinning guideline development. The Cochrane meta-analysis by Mochtar et al. (2017) concluded: "We found no clear evidence of a difference between rLH combined with rFSH and rFSH alone in rates of live birth or OHSS." The quality of evidence was rated very low for live births and low for OHSS. Similarly, the systematic review by Conforti et al. (2021) suggested an increase in clinical pregnancy and implantation rates in women aged 35–40 treated with rFSH + rLH, but did not assess live birth as a primary endpoint. Even the authors called for more focused RCTs with narrower age bands to confirm their findings. Moreover, two RCTs cited in the guideline justification—Lahoud et al. (2017) and	The recommendation that recombinant follicle-stimulating hormone (rFSH) combined with recombinant luteinizing hormone (rLH) is equally recommended as rFSH alone for the general IVF population is supported by evidence showing comparable live birth rates between the two approaches, as noted in the provided reference. Regarding the principle of preferring simpler treatments, the combination of rFSH and rLH does not significantly increase treatment complexity. Both rFSH alone and rFSH+rLH can be administered via a single subcutaneous injection, ensuring ease of use and patient convenience.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Humaidan et al. (2017)—also showed no significant difference in LBR between rFSH alone and the combination with rLH, whether in the general IVF population or in poor responders (per Bologna criteria). • Lahoud et al.: "The addition of rLH… did not improve live birth or clinical pregnancy rates. Results were not conclusive; further large RCTs are needed." • Humaidan et al.: "The primary and secondary endpoints were comparable between groups." Finally, the GDG appropriately emphasized the simplicity of ovarian stimulation protocols. When comparing compounds, dosages, or add-ons, preference was consistently given to simpler options—unless a clear and demonstrable benefit was shown. Given this, we would appreciate clarification on the rationale for recommending rLH + rFSH, particularly in the absence of consistent or high-quality evidence for improved live birth outcomes.	
25	Alberto Revelli	17	R30	Recommendation 30: The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. • The following publications reported that the association of r-FSH+r-LH brings some advantage for low responders: • Conforti et al. (2019) reported that in hypo responders, r-FSH:r-LH yielded more than a two-fold increase in the odds of achieving clinical pregnancy vs. r-FSH alone. • Ferraretti AP et al. (2004) A prospective RCT emphasized that adding r-hLH is more effective than increasing FSH dose for hypo-responders.	In this comment, the term hyporesponder is used interchangeably with low responder. However, the benefit referenced by the reviewer pertains to clinical pregnancy rate rather than live birth rate, for which the study by Conforti et al. did not demonstrate a significant difference. The study by Ferraretti et al. (2004), which was included in the meta-analysis by Conforti et al. (2019), had the following characteristics: 1. Both treatment arms received increased FSH dosing: Group A: FSH escalation up to 450 IU Group B: The same FSH escalation protocol plus recombinant LH (rLH) Therefore, the comparison was not between "FSH vs. FSH + LH" in a controlled manner, but rather between varying FSH starting doses based on female age

NR	Reviewer	Page	Line	Comment	Action / Reply
					(criteria not clearly reported), escalated up to 450 IU, and the same FSH escalation plus a variable dose (75–150 IU) of rLH. 2. Definition of hyporesponse: The study employed two distinct definitions for hyporesponders: Patients requiring increased or prolonged stimulation despite normal follicular recruitment Patients showing a plateau in follicular growth and estradiol levels between days 7–10 of stimulation 3. Lack of biomarker-based inclusion criteria: The study did not include selection based on objective biomarkers such as antral follicle count (AFC) or anti-Müllerian hormone (AMH); FSH starting doses varied by female age. 4. Absence of trial registration: The study was not registered in a clinical trial registry, making it impossible to verify whether outcomes were pre-specified or whether there were post hoc modifications to the study design, analytical approach, or outcome prioritization.
60	Sandro C. Esteves	17	R30	This statement suggests an equivalence between r-hFSH alone and r-hFSH:r-hLH in all low responders, without considering the distinct subgroup of hypo-responders—patients who exhibit an inadequate ovarian response to r-hFSH despite normal ovarian reserve markers. These patients differ from “low responders” and require specific attention. Several studies underscore the clinical benefit of r-hFSH:r-hLH in this context: • A systematic review and meta-analysis by Conforti et al. (2019) showed	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the

NR	Reviewer	Page	Line	Comment	Action / Reply
				a more than two-fold increase in clinical pregnancy odds for hypo-responders treated with r-hFSH:r-hLH compared to r-hFSH alone (OR 2.03; 95% CI: 1.27–3.25; p = 0.003). • A randomized controlled trial by Ferraretti et al. (2004) demonstrated that r-hLH supplementation was superior to simply increasing the FSH dose in hypo-responders, with significantly higher implantation and live birth rates (40.7% vs. 22%; p < 0.05). • These findings have been echoed in Delphi consensus statements advocating r-hLH supplementation in hypo-responders as a more effective strategy than monotherapy with r-hFSH. In light of this evidence, the guideline should consider differentiating low responders from hypo-responders and recommending that r-hFSH:r-hLH is probably the preferred option in patients with a previously suboptimal response to r-hFSH alone.	populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using

NR	Reviewer	Page	Line	Comment	Action / Reply
					contemporary definitions and live birth as a primary endpoint are required
64	Diane De Neubourg	17	R30	• Conforti et al. (2019) reported that in hypo responders, r-hFSH:r-hLH yielded more than a two-fold increase in the odds of achieving clinical pregnancy vs. r-hFSH alone. • Ferraretti AP et al. (2004) A prospective RCT emphasized that r-hLH is more effective than increasing FSH dose for hypo-responders. Therefore, the statement “probably equally recommended” underappreciates the current literature and it is suggested to be replaced by “rFSH with rLH can be recommended over rFSH alone in low responders”. References: • Conforti A, Esteves SC, Di Rella F, et al. The role of recombinant LH in women with hypo-response to controlled ovarian stimulation: a systematic review and meta-analysis. Reprod Biol Endocrinol. 2019;17(1):18. • Ferraretti AP, Gianaroli L, Magli MC, D'angelo A, Farfalli V, Montanaro N. Exogenous luteinizing hormone in controlled ovarian hyperstimulation for assisted reproduction techniques. Fertil Steril. 2004;82(6):1521-1526.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today’s hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline’s critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the

NR	Reviewer	Page	Line	Comment	Action / Reply
					analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
66	Semra Kahraman	17	R30	The guidelines address only low responders, without offering any recommendations for sub-optimal or hypo-responders, who actually constitute a significant proportion of patients in clinical practice. Both Conforti et al. (2019) and Ferraretti et al. (2004) highlighted the benefits of r-hLH in hypo-responders. Conforti et al. demonstrated that combining r-hFSH with r-hLH more than doubled the odds of achieving clinical pregnancy compared to r-hFSH alone, while Ferraretti et al., in a prospective RCT, showed that r-hLH was more effective than simply increasing the FSH dose. Furthermore, "Recombinant luteinizing hormone administration was found to potentially increase clinical pregnancy rates in patients aged 35–39 years in our study (Tayyar and Kahraman et al., 2019)."	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints.

NR	Reviewer	Page	Line	Comment	Action / Reply
					The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
69	Chengyan Deng	17	R30	Suggest the recommendation 30 could be justified as: The combination of FSH with LH and FSH alone are probably equally recommended for low responders. FSH and rLH cotreatment is probably over FSH alone in patients who are hypo responders to FSH alone.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which

NR	Reviewer	Page	Line	Comment	Action / Reply
				• The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 ($p = 0.003$), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH	forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically

NR	Reviewer	Page	Line	Comment	Action / Reply
				as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
70	Xiu Luo	17	R30	The recommendation 30 could be justified as: The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. Optimal benefit obtained in hypo-responders (POSEIDON group 1 and group 2) with combination of rFSH with rLH (GPP, RCT, meta analysis). There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in hypo-responders. Benefits of rFSH and rLH cotreatment in hypo-responders have been acknowledged by clinicians and achieved consensus in difference regions/countries1-6. A meta-analysis indicated that women with a previous hypo-response to exogenous FSH stimulation benefit from rFSH and rLH cotreatment as a means of increasing clinical pregnancy, implantation, and number of oocytes retrieved7. A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%8.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility.

NR	Reviewer	Page	Line	Comment	Action / Reply
				The RCT included 132 patients stratified according to POSEIDON classification group, patients were randomized into rFSH+rLH group and rFSH mono group during OS. The pregnancy rate was statistically higher in patients treated with rFSH + rLH compared to patients treated with rFSH (59.37% versus 34.30%, p≤ 0.05)9.	Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
75	Xi Dong	17	R30	Suggest the recommendation 30 could be justified as: The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. r-hFSH and r-hLH cotreatment is probably over r-hFSH alone in patients who are hypo responders to r-hFSH alone. • The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions.

NR	Reviewer	Page	Line	Comment	Action / Reply
				these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 ($p = 0.003$), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required

NR	Reviewer	Page	Line	Comment	Action / Reply
93	Fang Xiong Yinyang Bai	17	R30	The recommendation 30 could be justified as: The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. Optimal benefit obtained in hypo-responders (POSEIDON group 1 and group 2) with combination of rFSH with rLH (GPP, RCT, meta analysis). There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in hypo-responders. Benefits of rFSH and rLH cotreatment in hypo-responders have been acknowledged by clinicians and achieved consensus in difference regions/countries1-6. A meta-analysis indicated that women with a previous hypo-response to exogenous FSH stimulation benefit from rFSH and rLH cotreatment as a means of increasing clinical pregnancy, implantation, and number of oocytes retrieved7. A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%8. The RCT included 132 patients stratified according to POSEIDON classification group, patients were randomized into rFSH+rLH group and rFSH mono group during OS. The pregnancy rate was statistically higher in patients treated with rFSH + rLH compared to patients treated with rFSH (59.37% versus 34.30%, $p \leq 0.05$)9. Reference: 1. *Alviggi C, Humaidan P, Fischer R, et al. Reprod Biol Endocrinol. 2024;22(1):122. Published 2024 Oct 10. doi:10.1186/s12958-024-01291-x.* 2. *Raoul Orvieto, et al. Front Endocrinol (Lausanne). 2021 May 1012675670.* 3. *Barrenetxea G, Hernández C, Herrero J, et al. J Obstet Gynaecol, 2023, 43(1): 1-8.* 4. *Saghar Salehpour, et al. Front Reprod Health. 2024 May 9:6:1397446.*	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical

NR	Reviewer	Page	Line	Comment	Action / Reply
				5. *Rong Li, et al. Fertility and reproduction, 2024, 6 (3): 135-142.* 6. *Alviggi C, et al.Reprod Biol Endocrinol.2025 Mar 10;23(Suppl 1):38.* 7. *Conforti et al., Reprod Biol Endocrinol. 2019;17(1):18.* 8. *Ferraretti AP, et al. Fertil Steril. 2004;82(6):1521-1526. doi:10.1016/j.fertnstert.2004.06.041.* *M. Santonastaso, et al. ESHRE 2020, P-169.*	concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
94	Xi Xia	17	R30	The recommendation 30 could be justified as: The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. Optimal benefit obtained in hypo-responders (POSEIDON group 1 and group 2) with combination of rFSH with rLH (GPP, RCT, meta analysis). There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in hypo-responders. Benefits of rFSH and rLH cotreatment in hypo-responders have been acknowledged by clinicians and achieved consensus in difference regions/countries1-6. A meta-analysis indicated that women with a previous hypo-response to exogenous FSH stimulation benefit from rFSH and rLH cotreatment as a means of increasing clinical pregnancy, implantation, and number of oocytes retrieved7. A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield

NR	Reviewer	Page	Line	Comment	Action / Reply
				randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%8.	increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
95	Sandeep Karunakaran	17	R30	The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido

NR	Reviewer	Page	Line	Comment	Action / Reply
				However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a twofold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hyporesponsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for rhFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that

NR	Reviewer	Page	Line	Comment	Action / Reply
					current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
98	Kanad Dev Nayar	17	R30	Sub-optimal or hypo-responder subgroup benefit by the addition of rLH to rFSH for ovarian stimulation to optimize the results. • The guidelines do not provide any statements on patients with sub-optimal or hypo-responders (that form significant proportion of patients in clinical practice) but only on low responders. • Conforti et al. (2019) reported that in hypo responders, r-hFSH:r-hLH yielded more than a two-fold increase in the odds of achieving clinical pregnancy vs. r-hFSH alone. • Ferraretti AP et al. (2004) A prospective RCT emphasized that r-hLH is more effective than increasing FSH dose for hypo-responders.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both

NR	Reviewer	Page	Line	Comment	Action / Reply
					arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
102	Miaoxin Chen	17	R30	The justification for recommendation 30 can be framed as follows: The combination of rFSH with rLH and rFSH alone may be equally effective for low responders. However, for patients who are hypo-responders to rFSH alone, the co-treatment of r-hFSH and r-hLH is likely to be more beneficial than r-hFSH alone. • The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial

NR	Reviewer	Page	Line	Comment	Action / Reply
				than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
118	Saghar Salehpour	17	R30	The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically

NR	Reviewer	Page	Line	Comment	Action / Reply
				and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy 1 . • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment (p < 0.05). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22% 2 . • These findings have been widely recognized by reproductive medicine experts andreinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for	relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can

NR	Reviewer	Page	Line	Comment	Action / Reply
				patients with suboptimal response to r-hFSH alone 3-7 . • In light of above data, it will be more appropriate to probably recommend of r- hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
119	Kasi V Sellappan	17	R30	This draft does not specifically comment of patients who have had a sub optimal response to gonadotrophins – especially when there is growing evidence on supplementing rLH to rFSH which provides a better yield of oocytes (conforti et al 2019) rather than rFSH alone.. This point is also emphasized by Ferratreti AP el al (2004) where “rFSH being more effective when supplemented with rLH rather than just increasing the dose of rFSH alone”. So it would be prudent to amend this recommendation with an emphasis only on rFSH & rLH in hyporesponders / poor responders rather than only rFSH in such patients. The above management is also recognized and reinforced by the Delphi consensus.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today’s hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline’s critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without

NR	Reviewer	Page	Line	Comment	Action / Reply
					corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
120	Robert Fischer	17	R30	• The guidelines do not provide any statements on patients with hypo-responders but only on low responders. • Conforti et al. (2019) reported that in hypo responders, r-hFSH:r-hLH yielded more than a two-fold increase in the odds of achieving clinical pregnancy vs. r-hFSH alone. • Ferraretti AP et al. (2004) A prospective RCT emphasized that r-hLH is more effective than increasing FSH dose for hypo-responders. • From the above statement, it could be perceived that patient had low	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria

NR	Reviewer	Page	Line	Comment	Action / Reply
				or hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal

NR	Reviewer	Page	Line	Comment	Action / Reply
					recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
130	Michael H. Dahan	17	R30	• The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 ($p = 0.003$), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH,

NR	Reviewer	Page	Line	Comment	Action / Reply
				• These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence:

NR	Reviewer	Page	Line	Comment	Action / Reply
					Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
133	Suresh Nair	17	R30	• The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support

NR	Reviewer	Page	Line	Comment	Action / Reply
				a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1. • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment (p < 0.05). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH alone.	conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it

NR	Reviewer	Page	Line	Comment	Action / Reply
					inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
134	Nayana Patel	17	R30	• The guidelines do not provide any statements on patients with hypo-responders but only on low responders. Hyporesponders (Poseidon 1 and 2) are significant proportion of patients in clinical practice. • From the above statement, it could be perceived that patients with	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use.

NR	Reviewer	Page	Line	Comment	Action / Reply
				even patients with hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. However, several clinical studies among hypo responders, support that these patients could benefit more with combination of r-hFSH+r-hLH than with r-hFSH alone. The findings have been synthesized in a systematic review by Conforti et al. (2019), which included both randomized controlled trials and observational studies. The meta-analysis evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy1 • A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment (p < 0.05). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSH group was 22%2. • These findings have been widely recognized by reproductive medicine experts and reinforced through published Delphi consensus. The supporting publications and expert agreement advocate the use of r-hLH as a more effective alternative for patients with suboptimal response to r-hFSH alone3-7. • In light of above data, it will be more appropriate to probably	The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the

NR	Reviewer	Page	Line	Comment	Action / Reply
				recommend of r-hFSH+r-hLH over r-hFSH alone in patients who are hypo responders to r-hFSH	foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality

NR	Reviewer	Page	Line	Comment	Action / Reply
					Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
137	Jayesh Amin	17	R30	The combination of rFSH with HMG and rFSH alone are probably equally recommended for Low responders.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
139	Alessandro Conforti Robert Fisher Peter Humaidan Carlo Alviggi	17	R30	From the above statement, it could be perceived that patients with low ovarian response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. There are several issues related to these statements that are reported below. 1) Definition of low response is vague and not consistent with worldwide adopted criteria. According to this guideline, low response is defined as ≤ 3 follicles on the day of oocyte maturation trigger and/or ≤ 3 oocytes obtained after ovum pick up. This definition is not consistent either with ESHRE1 or Poseidon criteria2, that are so far the most widely adopted classification systems for poor ovarian response and low prognosis patients, respectively. Thus, we suggest considering both ESHRE and Poseidon criteria to avoid confusion among readers. 2) There is evidence that rFSH and rLH co-treatment could be useful in specific subgroup of low prognosis women, including those in whom suboptimal (4-9 eggs) or poor ovarian response (less equal 3 eggs) are obtained despite normal ovarian reserve. This profile, also defined as hypo-response, is characterized by ovarian resistance to FSH monotherapy. These findings have been synthesized in a systematic	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the

NR	Reviewer	Page	Line	Comment	Action / Reply
				review and meta-analysis conducted by Conforti et al. (2019)3. In detail, in hypo-responders, stimulation with rFSH + rLH resulted in a statistically significant increase in terms of both clinical pregnancy rate and the number oocytes retrieved, when compared to rFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy. Regarding live birth rate a prospective RCT demonstrated that this outcome is significantly higher in women rFSH rLH co-treatement than rFSH alone group (40.7% vs 22%, p < 0.05)4. 3) Specific subgroups of poor responders according to ESHRE classification could benefit from rFSH and rLH co-treatment in terms of live birth rate and cumulative live birth rate. The largest RCT5 conducted so far and a real-world analysis6 (9,787 cycles) demonstrated that rFSH/rLH co-treatment could significantly improve the live birth rate and the cumulative live birth rate in women displaying moderate or severe poor response according to Prosper score. In light of above data, it will be more appropriate to suggest rFSH+rLH over r-hFSH alone in women with moderate or severe poor response according to ESHRE criteria and in patients with unexpected poor or suboptimal response (hypo-responders, groups 1- 2 Poseidon).	number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant

NR	Reviewer	Page	Line	Comment	Action / Reply
					improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
149	Marianne Vendola	17	R30	This recommendation, as currently expressed, may lack clarity regarding its strength and implications. Although it is classified as conditional—recognizing the limited strength of the supporting evidence—it could potentially be misinterpreted as a firm recommendation. Recent studies have focused on distinguishing low-prognosis patients from poor responders, particularly supporting the benefit of combining r-hFSH with r-hLH over r-hFSH alone in hypo-responders. This evidence is	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago,

NR	Reviewer	Page	Line	Comment	Action / Reply
				synthesized in a systematic review by Conforti et al. (2019), which incorporated both randomized controlled trials and observational studies. The meta-analysis demonstrated a statistically significant increase in clinical pregnancy rates with ovarian stimulation using r-hFSH combined with r-hLH compared to r-hFSH alone. These findings have been widely acknowledged by experts in reproductive medicine and further reinforced through published Delphi consensus statements in 2024. The supporting publications and expert agreements collectively advocate for considering r-hLH as a more effective alternative for patients exhibiting suboptimal responses to r-hFSH monotherapy.	prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the

NR	Reviewer	Page	Line	Comment	Action / Reply
					methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of

NR	Reviewer	Page	Line	Comment	Action / Reply
					r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
150	Mekhala Dwarakanath B	17	R30	The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. In a prospective study done at our centre(n=579, unpublished data) we observed that a fixed dose combination of r-FSH; r-LH in a ratio of 2;1 resulted in higher live birth rate in comparison to standard of care with r-FSH in POSEIDON Group 1,2 & 3 attaining statistical significance. Use of r-fsh +r-lh increased the number of oocytes and Blastocysts but did not increase LBR	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
154	Cedrin Durnerin	17	R30	• The guidelines do not provide any statements on patients with sub-optimal or hypo-responders (that form significant proportion of patients in clinical practice Poseidon group 1 b and 2 b) but only on low responders. • Conforti et al. (2019) reported that in hypo responders, r-hFSH:r-hLH yielded more than a two-fold increase in the odds of achieving clinical pregnancy vs. r-hFSH alone. • Ferraretti AP et al. (2004) De placido et al (2004) Two prospective RCT emphasized that r-hLH is more effective than increasing FSH dose for hypo-responders.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to

NR	Reviewer	Page	Line	Comment	Action / Reply
					interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed

NR	Reviewer	Page	Line	Comment	Action / Reply
					higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
119	Kasi V Sellappan	17	R30 R31	This draft does on fully reflect on current evidence especially when it comes to discussing management of patients in the Poseidon 1 & 2 group, who form a significant proportion of patients presenting in clinic. Statements in recommendation 30 & 31 could mislead clinicians that	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.

NR	Reviewer	Page	Line	Comment	Action / Reply
				both rFSH:RLH and only rFSH are equally effective as that's contradicted by current evidence.	
122	Surveen Ghumman	17	R30 R31	The following may be noted 1. The RCT (Lahoud et al., 2017) quoted in line 1830 as evidence is for general population with a decreased level of LH. No subgroup analysis of low/hyporesponder was done. The population was only those with low LH whereas LH values can be high in specific gene mutations or polymorphism in LH or LH receptor genes. This important population was not studied in this RCT. These would be hyporesponders in normal ovarian reserve women and may require extra LH during stimulation 2. A systematic review by Conforti et al. (2021),quoted in line 1838 of guideline document as evidence included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate Twelve studies were identified.. In women aged between 35 and 40 years, r-hFSH/r-hLH co-treatment was associated with higher clinical pregnancy rates (OR 1.45, CI 95% 1.05–2.00, I2 = 0%, P = 0.03) and implantation rates (OR 1.49, CI 95% 1.10–2.01, I2 = 13%, P = 0.01) versus r-hFSH monotherapy. Fewer oocytes were retrieved in r-hFSH/r-hLHtreated patients than in r-hFSH-treated patients both in women aged ≥35 years (WMD -0.82 CI 95% -1.40 to − 0.24, I2 = 88%, P = 0.005) and in those aged between 35 and 40 years (WMD - 1.03, CI − 1.89 to − 0.17, I2 = 0%, P = 0.02). Conclusion of the systemic review: Although more oocytes were retrieved in patients who underwent r-hFSH monotherapy, this meta-analysis suggests that r-hFSH/r-hLH co-treatment improves clinical pregnancy and implantation rates in women between 35 and 40 years of age undergoing ovarian stimulation for assisted reproduction technology. Besides evidence quoted in this guideline there has been evidence to support addition of r LH to rFSH in poor/low /hypo responders. Both RCT and metanalysis with subgroup analysis on poor/low responders have shown an advantage with addition of rLH specially in the group of hyporesponders and women >35 years as cited below	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the

NR	Reviewer	Page	Line	Comment	Action / Reply
				1. Conforti et al. (2019) performed a systematic review and meta-analysis to compare the effects of combined r LH and rFSH over rFSH monotherapy in hyporesponders. They synthesized data from four RCTs and one observational study. Improvement in CPR was greater with combined rLH and rFSH therapy than with rFSH monotherapy (RR 2.03, 95% CI 1.27–3.25, I2=0%, four studies). Similar effects were observed in a subgroup with only RCTs (RR 2.02, 95% CI 1.18–3.45, I2=0%, three RCTs). The implantation rate too was better in the combined rLH and rFSH therapy group (OR: 2.62, 95% CI 1.37–4.99, five studies) and in the subgroup of RCTs (OR 2.58, 95% CI 1.09–6.07). Analysis of RCTs indicated that more oocytes were retrieved in the combined rLH and rFSH group than in the rFSH monotherapy group (MD 2.90, 95% CI 1.88–3.92). 2. Alviggi et al. (2018) systematically reviewed literature on rLH supplementation in six groups of patients. A meta-analysis was not performed. Women with adequate ovarian reserve findings had an unexpected hyporesponse to rFSH monotherapy, and women aged 36–39 years seemed to benefit from this supplementation. The first group, with hyporesponse to rFSH monotherapy and a normal ovarian reserve, included 848 patients from four RCTs. The authors concluded that addition of rLH would be beneficial than continuing rFSH with the same or an increased dosage. However, the inclusion criteria and outcome parameters differed across studies. In the second group with women 36–39 years of age, 10 RCTs (2901 patients) involving agonist and antagonist protocols were analysed. The authors concluded that rhLH exerted a beneficial effect on the implantation rate. No effect on pregnancy rate was observed. Further, no significant effect was observed among women >40 years receiving an agonist or an antagonist regimen. In a 3. Cochrane review by Mochtar et al. (2017), eight of 36 RCTs included poor responders. On subgroup analysis of low responders for livebirth outcomes, one RCT by Ferraretti et al. (2014) was identified with an OR of 9.33 and 95 % CI of 1.03, 84.2. On subgroup analysis of the ongoing pregnancy outcomes based on ovarian response, three RCTs, namely by Ferraretti et al. (2004), de Placido et al. (2005), and Ruvolo et al. (2007) were identified. These compared 143 (rLH +	analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. 1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations.

NR	Reviewer	Page	Line	Comment	Action / Reply
				rFSH) and133 (rFSH alone) patients, with an OR of 2.06 and a 95 % CI of 1.2, 3.53 favouring the rLH + rFSH group. There was little or no difference in cancellation rates between the rLH + rFSH and rFSH groups due to a low response (OR 0.77, 95% CI 0.54–1.10; n=2251; 11 studies; I2=16%, lowquality evidence). The evidence suggests that if the risk of cancellation due to low response following treatment with rFSH alone is 7%, it would be between 4% and 7% on using rLH +rFSH. 4. In a systematic review with meta-analysis by Lehert et al. (2014), data from 43 studies (40 RCTs, 6443 patients) comparing the outcomes of rFSH and rFSH + rLH were included. Of them, 12 studies had a cohort of poor responders. In these, rLH was started on day 1 of stimulation in three studies and mid-cycle in five studies; four articles had no mention of the timing of initiation. This study was graded as having low confidence based on AMSTAR-2 criteria. No significant results were observed in the per protocol population (RR 1.29, 95% CI 0.96–1.73). Significantly higher CPRs were observed with recombinant human FSH (r-hFSH) + r-hLH than with r-hFSH alone in the overall population (RR 1.09; 95% CI 1.01–1.18) and poor responders (n=1179; RR 1.30; 95% CI 1.01–1.67; ITT population); the observed difference was more pronounced in poor responders. 5. In an RCT by Humaidan et al. (2017), the patients were randomised into two groups administered a 2:1 combination of r-hFSH/r-hLH (n=477) and rhFSH (n=462). In the ITT population, the mean (standard deviation) number of retrieved oocytes primary endpoint) (3.3 [2.7]) in the r-hFSH/r-hLH group was not significantly different from that in the r-hFSH group (3.6 [2.82]). The biochemical pregnancy rate, OPR, and LBR did not differ significantly between the groups. A post hoc logistic regression analysis considering baseline characteristics indicated that the incidence of total pregnancy outcome failure (defined as the combination of preclinical miscarriage, clinical miscarriage [early + late] and ectopic pregnancy) was lower in the 2:1 r-hFSH/r-hLH group (6.7%) than in the r-hFSH group (12.4%) with an OR of 0.52 (95% CI 0.33, 0.82; p=0.005).	3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required

NR	Reviewer	Page	Line	Comment	Action / Reply
				Ovarian stimulation is undoubtedly one of the most complex challenging procedures in reproductive endocrinology. To simplify it is an excellent idea - as long as we are not missing an important detail.	
25	Alberto Revelli	17	R31	Recommendation 31: The combination of rFSH with rLH and rFSH alone are probably equally recommended for women of advanced age (≥35 years). • Ovarian stimulation outcomes worsen with age due to ovarian aging, reduced LH bioactivity, and impaired steroidogenesis. • Bosch et al. (2021) highlighted functional LH/FSH deficiency in women ≥35, reducing ART outcomes. • Conforti et al. (2021) published a systematic review demonstrating that in women aged 35–40 years, r-FSH:r-LH resulted in significantly higher implantation and clinical pregnancy rates compared to r-FSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). • Bielfeld et al. (2023) German IVF Registry (Real world data) also confirmed in a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher in the r-FSH:r-LH group compared to the r-FSH alone group.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study represents a large retrospective observational cohort based on the German IVF registry. While it offers valuable insights into routine clinical practice, real-world data of this nature cannot establish causal relationships due to the potential for residual confounding, even when adjustments for baseline characteristics are attempted. → Implication: Observational studies are informative and hypothesis-generating but cannot substitute for randomized controlled trials (RCTs) in the development of clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Rationale This publication is a narrative review discussing hypothesized physiological mechanisms—such as functional LH deficiency in older women. Although it presents a compelling biological rationale, it does not provide new clinical trial data or comparative outcome-based evidence. → Implication: Mechanistic plausibility is valuable for generating hypotheses but is insufficient, in isolation, to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This is the only cited analysis based on randomized controlled trials and was considered during guideline

NR	Reviewer	Page	Line	Comment	Action / Reply
					development. However, two key limitations must be highlighted: The primary outcome of interest for guideline formulation was live birth rate (LBR). The Conforti meta-analysis did not demonstrate a significant difference in LBR, either overall in women >35 years or in the subgroup aged 35–40 years (a subgroup defined arbitrarily). The observed benefits in implantation and clinical pregnancy rates do not necessarily translate into improvements in live birth outcomes—a distinction that is essential in evidence grading and recommendation formulation. → Implication: Although the Conforti meta-analysis provides valuable data, its lack of demonstrated benefit in live birth limits its influence on guideline recommendations, which prioritize LBR as the most clinically meaningful endpoint.
60	Sandro C. Esteves	17	R31	This recommendation does not adequately account for age-related pathophysiology affecting ovarian stimulation outcomes: • Advanced maternal age is associated with reduced LH bioactivity, impaired LH receptor function, and diminished intra-ovarian steroidogenesis, which together contribute to suboptimal ovarian responses and poorer ART outcomes. • Bosch et al. (2021) described women ≥35 years as having a functional LH/FSH deficiency, emphasizing the need for LH supplementation. • A systematic review of RCTs by Conforti et al. (2021) demonstrated that r-hFSH:r-hLH significantly improved both implantation (OR 1.49; p = 0.01) and clinical pregnancy rates (OR 1.45; p = 0.03) in women aged 35–40 years.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and

NR	Reviewer	Page	Line	Comment	Action / Reply
				• A large observational study using the German IVF Registry confirmed a 16% increase in both clinical pregnancy and live birth rates with r-hFSH:r-hLH compared to r-hFSH alone in women aged 35–40 with normal ovarian reserve. These findings, along with expert consensus publications, argue against treating r-hFSH:r-hLH and r-hFSH alone as clinically interchangeable in this population. Patients of advanced maternal age often seek to maximize outcomes in their first ART attempt; an individualized, evidence-driven approach is therefore essential. In this regard, please consider changing the statement to that r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.

NR	Reviewer	Page	Line	Comment	Action / Reply
64	Diane De Neubourg	17	R31	• Ovarian stimulation outcomes worsen with age due to ovarian aging, reduced LH bioactivity, and impaired steroidogenesis. Bosch et al. (2021) highlighted functional LH/FSH deficiency in women ≥35, reducing ART outcomes. • Conforti et al. (2021) conducted a systematic review demonstrating that in women aged 35–40 years, r-hFSH:r-hLH resulted in significantly higher implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). • Bielfeld et al. (2023) German IVF Registry also confirmed in a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group. Therefore, the statement “probably equally recommended” underappreciates the current literature and it is suggested to be replaced by “rFSH with rLH can be recommended over rFSH alone for women of advanced age (>=35 years)”. References • Bosch E, Alviggi C, Lispi M, et al. Reduced FSH and LH action: implications for medically assisted reproduction. Hum Reprod. 2021;36(6):1469-1480. • Conforti A, Esteves SC, Humaidan P, et al. Recombinant human luteinizing hormone co-treatment in ovarian stimulation for assisted reproductive technology in women of advanced reproductive age: a systematic review and meta-analysis of randomized controlled trials. Reprod Biol Endocrinol. 2021;19(1):91	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not

NR	Reviewer	Page	Line	Comment	Action / Reply
				• Bielfeld AP, Schwarze JE, Verpillat P, et al. Effectiveness of recombinant human follicle-stimulating hormone (r-hFSH): recombinant human luteinizing hormone versus r-hFSH alone in assisted reproductive technology treatment cycles among women aged 35-40 years: A German database study. Best Pract Res Clin Obstet Gynaecol. 2023;89:102350.	show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
66	Semra Kahraman	17	R31	Advancing age negatively affects ovarian stimulation outcomes, largely as a result of ovarian aging, diminished luteinizing hormone (LH) bioactivity, and disrupted steroidogenesis. According to Bosch et al. (2021), a functional deficiency in LH and FSH among women aged ≥35 years contributes to reduced success rates in assisted reproductive technology (ART). Conforti et al. (2021), in a systematic review, demonstrated that in women aged 35–40 years, the combination of r-hFSH and r-hLH significantly improved implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). The same topic was also discussed in one of our group's studies. Karlıkaya et al (2003) Comparison of three different protocols in the group of advanced maternal age with diminished ovarian reserve. 59th Annual Meeting of the American Society for Reproductive Medicine, October 11-15, 2003, San Antonio, TX, USA	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-

NR	Reviewer	Page	Line	Comment	Action / Reply
					based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
69	Chengyan Deng	17	R31	Suggest the recommendation 31 could be justified as: The combination of FSH with LH is probably recommended for women of advanced age patients especially aged 35-40. • It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing1, reduction of follicle endocrine	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual

NR	Reviewer	Page	Line	Comment	Action / Reply
				milieu2, reduction in LH bioactivity3 and LH receptor activity4 & decreased steroidogenesis5. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes6. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)7. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group8. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone9-11. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient	confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading.

NR	Reviewer	Page	Line	Comment	Action / Reply
				dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	→ Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
70	Xiu Luo	17	R31	The recommendation 31 could be justified as: The combination of rFSH with rLH is probably recommended for women of advanced age (≥35 year), especially aged 35-40. (GPP, meta analysis, real world study). Advanced reproductive age is one of the factors for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes1.There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in women aged 35-40 years old. Benefits of rFSH and rLH cotreatment in advanced reproductive age patients have been endorsed by expert consensus/review in different regions/countries2-6. Meta analysis (Conforti et al. 2021): In women aged between 35 and 40 years, r-hFSH/r-hLH co-treatment was associated with higher clinical pregnancy rates (OR 1.45, CI 95% 1.05–2.00, I2 = 0%, P = 0.03) and implantation rates (OR 1.49, CI 95% 1.10–2.01, I2 = 13%, P = 0.01) versus r-hFSH monotherapy7. The results of real world studies (Bielfeld AP, et al. 2024) indicated that women aged 35-40 years old in combination of rFSH with rLH group achieved significant higher clinical pregnancy rate and live birth rate8.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized

NR	Reviewer	Page	Line	Comment	Action / Reply
					controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
75	Xi Dong	17	R31	Suggest the recommendation 31 could be justified as: The combination of rFSH with rLH is probably recommended for women of advanced age patients especially aged 35-40. • It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing1, reduction of follicle endocrine milieu2, reduction in LH bioactivity3 and LH receptor activity4 & decreased steroidogenesis5. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes6. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion

NR	Reviewer	Page	Line	Comment	Action / Reply
				systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)7. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group8. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone9-12. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
93	Fang Xiong Yinyang Bai	17	R31	The recommendation 31 could be justified as: The combination of rFSH with rLH is probably recommended for women of advanced age (≥35	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis)

NR	Reviewer	Page	Line	Comment	Action / Reply
				year), especially aged 35-40. (GPP, meta analysis, real world study). Advanced reproductive age is one of the factors for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes1.There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in women aged 35-40 years old. Benefits of rFSH and rLH cotreatment in advanced reproductive age patients have been endorsed by expert consensus/review in different regions/countries2-6. Meta analysis (Conforti et al. 2021): In women aged between 35 and 40 years, r-hFSH/r-hLH co-treatment was associated with higher clinical pregnancy rates (OR 1.45, CI 95% 1.05–2.00, I2 = 0%, P = 0.03) and implantation rates (OR 1.49, CI 95% 1.10–2.01, I2 = 13%, P = 0.01) versus r-hFSH monotherapy7. The results of real world studies (Bielfeld AP, et al. 2024) indicated that women aged 35-40 years old in combination of rFSH with rLH group achieved significant higher clinical pregnancy rate and live birth rate8.	This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40

NR	Reviewer	Page	Line	Comment	Action / Reply
					years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
94	Xi Xia	17	R31	The recommendation 31 could be justified as: The combination of rFSH with rLH is probably recommended for women of advanced age (≥35 year), especially aged 35-40. (GPP, meta analysis, real world study). Advanced reproductive age is one of the factors for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes1.There is evidence to suggest compared to r-hFSH alone, the combination of r-hFSH with r-hLH improves pregnancy outcomes in women aged 35-40 years old. Benefits of rFSH and rLH cotreatment in advanced reproductive age patients have been endorsed by expert consensus/review in different regions/countries2-6. Meta analysis (Conforti et al. 2021): In women aged between 35 and 40 years, r-hFSH/r-hLH co-treatment was associated with higher clinical pregnancy rates (OR 1.45, CI 95% 1.05–2.00, I2 = 0%, P = 0.03) and implantation rates (OR 1.49, CI 95% 1.10–2.01, I2 = 13%, P = 0.01) versus r-hFSH monotherapy7. The results of real world studies (Bielfeld AP, et al. 2024) indicated that women aged 35-40 years old in combination of rFSH with rLH group achieved significant higher clinical pregnancy rate and live birth rate8.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important

NR	Reviewer	Page	Line	Comment	Action / Reply
					for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
95	Sandeep Karunakaran	17	R31	• It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing8, reduction of follicle endocrine milieu9, reduction in LH bioactivity10 and LH receptor activity11 & decreased steroidogenesis12. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes13.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics.

NR	Reviewer	Page	Line	Comment	Action / Reply
				• Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)14. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group15. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone5-7. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention rhFSH: r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	→ Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a

NR	Reviewer	Page	Line	Comment	Action / Reply
					benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
98	Kanad Dev Nayar	17	R31	Recent studies suggest a statistically significant increase in clinical pregnancy rate and live birth rate with the addition of rLH to rFSH in patients of advanced age (35-40 years) undergoing ovarian stimulation. • Ovarian stimulation outcomes worsen with age due to ovarian aging, reduced LH bioactivity, and impaired steroidogenesis. • Bosch et al. (2021) highlighted functional LH/FSH deficiency in women ≥35, reducing ART outcomes. • Conforti et al. (2021) conducted a systematic review demonstrating that in women aged 35–40 years, r-hFSH:r-hLH resulted in significantly higher implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). • Bielfeld et al. (2023) German IVF Registry also confirmed in a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However,

NR	Reviewer	Page	Line	Comment	Action / Reply
					two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
102	Miaoxin Chen	17	R31	Recommendation 31 can be justified by stating that the combination of rFSH with rLH is likely advisable for women of advanced reproductive age, particularly those between 35 and 40 years old. • It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing1, reduction of follicle endocrine milieu2, reduction in LH bioactivity3 and LH receptor activity4 & decreased steroidogenesis5. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes6. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g.,

NR	Reviewer	Page	Line	Comment	Action / Reply
				systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)7. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group8. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone9-12. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
118	Saghar Salehpour	17	R31	• It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing 8 , reduction of follicle endocrine milieu 9 ,	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it

NR	Reviewer	Page	Line	Comment	Action / Reply
				reduction in LH bioactivity 10 and LH receptor activity 11 & decreased steroidogenesis 12 . • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes 13 . • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03) 14 . • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group 15 . • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone 5-7 . • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based	provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and

NR	Reviewer	Page	Line	Comment	Action / Reply
				treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
119	Kasi V Sellappan	17	R31	Ovarian reserves reduce with age, It also a known fact that LH bio activity also reduces with age. Reduced functional FSH/LH activity results in ART outcomes not being optimal (Bosch et al., 2021).Current evidence reveals that using rFsh along with r hLH results in a statistically significant clinical pregnancy rates (conforti et al; 2021; implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03)in women aged 35 – 40. This was also supported by real world data (Bielfield et al; 2023; German IVF registry) in patients between 35 – 40 years of age having a normal ovarian reserve resulting in a statiscally significant 16% increase in live birth rates when using rFSH:rLH compared to rFSH alone. So, the inference here would be that rfsh along with rlh would yield better outcomes that rfsh alone. Hence an amendment should be considered in this statement as well. The above data management is also recognized and reinforced by the Delphi consensus.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations.

NR	Reviewer	Page	Line	Comment	Action / Reply
					3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
120	Robert Fischer	17	R31	• Ovarian stimulation outcomes worsen with age due to ovarian aging, reduced LH bioactivity, and impaired steroidogenesis. • Bosch et al. (2021) highlighted functional LH/FSH deficiency in women ≥35, reducing ART outcomes. • Conforti et al. (2021) conducted a systematic review demonstrating that in women aged 35–40 years, r-hFSH:r-hLH resulted in significantly higher implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). • Bielfeld et al. (2023) German IVF Registry also confirmed in a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials

NR	Reviewer	Page	Line	Comment	Action / Reply
				16% in the r-hFSH:r-hLH group compared to the r-hFSH group. • It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing8, reduction of follicle endocrine milieu9, reduction in LH bioactivity10 and LH receptor activity11 & decreased steroidogenesis12. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes13. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)14. • This finding was also observed in an observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group15. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone5-7. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably	(RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact

NR	Reviewer	Page	Line	Comment	Action / Reply
				recommended to be prescribed to advanced maternal age patients especially aged 35-40.	on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
130	Michael H. Dahan	17	R31	• It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing8, reduction of follicle endocrine milieu9, reduction in LH bioactivity10 and LH receptor activity11 & decreased steroidogenesis12. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes13. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)14. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group15. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized:

NR	Reviewer	Page	Line	Comment	Action / Reply
				hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone5-7. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	• The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
132	The Chinese Expert Review Panel for ESHRE OS Guideline	17	R31	Suggestion： Update the strength and the quality of evidence to GPP Based on the Delphi consensus, and enhance the operability of clinical practice.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
133	Suresh Nair	17	R31	• It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing8, reduction of follicle endocrine milieu9, reduction in LH bioactivity10 and LH receptor activity11 & decreased steroidogenesis12. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes13.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics.

NR	Reviewer	Page	Line	Comment	Action / Reply
				• Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)14. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group15. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone5-7. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	→ Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a

NR	Reviewer	Page	Line	Comment	Action / Reply
					benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
134	Nayana Patel	17	R31	• It is well established that with age increases, risk of leading to poor ovarian stimulation outcomes also increase. This is due to several reasons, including ovarian ageing8, reduction of follicle endocrine milieu9, reduction in LH bioactivity10 and LH receptor activity11 & decreased steroidogenesis12. • In 2021 Bosch publication, advanced reproductive age is one of the factor for functional LH & FSH deficiency and this would in turn reduce ART clinical outcomes13. • Several clinical studies have proven the effect of r-hFSH:r-hLH over r-hFSH in patients older than 35. The findings have been synthesized in a systematic review by Conforti et al. (2021), which included only RCTs. The meta-analysis evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03)14. • This finding was also observed in a large observational study using data registered by the German IVF Registry (Deutsche IVF Register). In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and

NR	Reviewer	Page	Line	Comment	Action / Reply
				were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group15. • These clinical findings have also been broadly endorsed in recent Delphi consensus publications. Expert panels reaffirmed the benefits of r-hFSH:r-hLH in women of advanced reproductive age, highlighting its role in improving clinical outcomes compared to r-hFSH alone5-7. • Patients hope to achieve the best possible outcome in their first stimulation cycle. Clinicians should therefore adopt evidence-based treatment strategies to minimize the time to live birth and reduce patient dropout rates. In this regard, we suggest ESHRE guideline should mention r-hFSH:r-hLH treatment is probably recommended to be prescribed to advanced maternal age patients especially aged 35-40.	Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
137	Jayesh Amin	17	R31	The combination of rFSH with HMG and rFSH alone are probably equally recommended for women of advanced reproductive age (≥35 year).	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
139	Alessandro Conforti Robert Fisher Peter	17	R31	Advanced reproductive age presents different segments of prognosis with the worst one above 40 years old. Indeed, embryo euploidy rates, which are the most important factor governing cumulative and live birth following ART, are remarkably higher in women aged 35–39 years than in	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical

NR	Reviewer	Page	Line	Comment	Action / Reply
	Humaidan Carlo Alviggi			those above the age of 40 years. Considering all RCTs1-5 performed in women between 35-40 years, rFSH and rLH co-treatement showed a significantly higher clinical pregnancy rate than rFSH alone 6. This finding was also confirmed in a large observational study (4238 women in each treatment group) based on data of the German IVF Registry. In a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the rFSH + rLH group compared to the rFSH group7 In this regard, we suggest ESHRE guideline should mention rFSH and rLH co-treatment is suggested in advanced maternal age patients especially aged 35-40 .	practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate

NR	Reviewer	Page	Line	Comment	Action / Reply
					into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
149	Marianne Vendola	17	R31	Similarly, this recommendation should be re phrased It is not strongly recommend in fact because there are evidence out of the benefit of patients over 35 to add rLH to rec FSH during COS. It is well-recognized that ovarian stimulation outcomes tend to decline with age due to factors such as ovarian aging, decreased LH bioactivity, and impaired steroidogenesis. Bosch et al. (2021) highlighted functional LH/FSH deficiency in women aged 35 and above, which can negatively impact ART success. A systematic review by Conforti et al. (2021) demonstrated that in women aged 35–40 years, treatment with r-hFSH combined with r-hLH resulted in significantly higher implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). Furthermore, the German IVF Registry study by Bielfeld et al. (2023), involving a balanced population of women aged 35–40 with normal ovarian reserve, confirmed that both clinical pregnancy and live birth rates were approximately 16% higher in the group receiving r-hFSH combined with r-hLH versus r-hFSH alone. Given this evidence, it would be advisable to rephrase this	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence.

NR	Reviewer	Page	Line	Comment	Action / Reply
				recommendation to provide clearer guidance on its conditional nature and the specific patient populations it pertains to.	→ Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
150	Mekhala Dwarakanath B	17	R31	Addition of r- LH with r- FSH may be considered	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.

NR	Reviewer	Page	Line	Comment	Action / Reply
154	Cedrin Durnerin	17	R31	• Ovarian stimulation outcomes worsen with age due to ovarian aging, reduced LH bioactivity, and impaired steroidogenesis. • Bosch et al. (2021) highlighted functional LH/FSH deficiency in women ≥35, reducing ART outcomes. • Conforti et al. (2021) conducted a systematic review demonstrating that in women aged 35–40 years, r-hFSH:r-hLH resulted in significantly higher implantation and clinical pregnancy rates compared to r-hFSH alone (implantation: OR 1.49, P = 0.01; clinical pregnancy: OR 1.45, P = 0.03). • Bielfeld et al. (2023) German IVF Registry also confirmed in a balanced patient population aged 35–40 years with normal ovarian reserve (5–14 oocytes), both clinical pregnancy rate and live birth rate were statistically significantly higher by 16% in the r-hFSH:r-hLH group compared to the r-hFSH group.	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not

NR	Reviewer	Page	Line	Comment	Action / Reply
					show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
88	Aboubakr Mohamed Elnashar	17	R34	Follitropin delta results in higher cumulative live birth rates compared with follitropin alfa/beta	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
138	Philippe Pinton*	17	R34	The guideline states that follitropin delta and alfa/beta are “probably equally recommended” (conditional). This recommendation appears based on the body of evidence generated from the development program for follitropin delta, including three large Phase III RCTs: 1. ESTHER-1 – Nyboe Andersen et al., Fertil Steril. 2017;107(2):387–396.e4 2. STORK – Ishihara et al., Reprod Biomed Online. 2021;42(5):909–918 3. GRAPE – Qiao et al., Hum Reprod. 2021;36(9):2452–2462 These trials—each comparing follitropin delta to either alfa or beta—demonstrated non-inferiority in primary outcomes (ongoing pregnancy, oocyte yield) and comparable live birth rates across treatment arms. Notably, they consistently showed a lower incidence of extreme ovarian response and fewer	The recommendation was changed to strong in the updated guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				cycle cancellations with follitropin delta. The meta-analysis' by Palomba et al; Nelson et al; and Komiya et al; all published in 2024 further confirmed these findings: similar efficacy with a reduced OHSS risk, particularly in fresh GnRH-antagonist cycles. We appreciated the GDG's recognition of different clinical approaches, particularly the contrast between fixed and individualized dosing. Follitropin delta is the only gonadotropin approved with an algorithm-based, individualized dosing strategy using AMH and body weight. This approach provides a more tailored and potentially safer stimulation, especially in IVF-naïve patients, those undergoing fresh transfers, or individuals at high risk for OHSS—aligning with the guideline's aim to promote simpler, safer protocols. Given this context, we would respectfully request further clarification: • What is the rationale for assigning only a conditional recommendation to follitropin delta, despite robust evidence from multiple RCTs and regulatory approval grounded in its individualized and safety-oriented approach? • The GDG's justification referencing "different follitropin medications" appears to rely on pooled data from alfa and beta comparators. This may unintentionally overlook the strength and consistency of findings from individual RCTs specifically designed to evaluate follitropin delta. • Might it be worth considering whether the clinical value of algorithm-based dosing with follitropin delta could be more explicitly acknowledged—particularly in the context of first-cycle IVF patients or those at increased risk of over-response in fresh embryo transfer settings? • We fully recognize and support the GDG's emphasis on simplicity. However, in this case, individualization represents a data-driven advancement toward precision medicine—enhancing both safety and therapeutic efficiency, rather than adding unnecessary complexity. In light of this, we kindly suggest reconsidering the strength of the current recommendation and exploring the possibility of including a dedicated statement that highlights the role of individualized, algorithm-based protocols in advancing ovarian stimulation practices in specific settings.	

NR	Reviewer	Page	Line	Comment	Action / Reply
144	Karolina Palinska-Rudzka	17	R34	The draft recommendation that follitropin delta and alfa/beta are "probably equally recommended" may not fully reflect the available safety data. While efficacy outcomes are indeed comparable, several large studies and recent analyses have consistently demonstrated lower rates of OHSS with follitropin delta when used according to its AMH- and weight-based dosing algorithm: • The ESTHER-1, STORK, and GRAPE trials each reported reduced early OHSS rates and fewer cycle cancellations with follitropin delta compared to conventional rFSH. • In patients with AMH >35 pmol/L, Višnová et al. (2021) observed a significantly lower rate of OHSS or need for preventive intervention with delta (7.7%) versus alfa (26.7%) [3]. • Most recently, Nelson et al. (2024) published an individual participant data meta-analysis of randomised controlled trials, showing that individualized dosing with follitropin delta improves both live birth outcomes and safety [4]. In light of this evidence, the current phrasing may understate the safety benefits associated with algorithm-based follitropin delta use in high responders. These findings underpin the BFS recommendation that algorithm-based follitropin delta should be considered in high responders. Framing it as "probably equal" may unintentionally downplay these clinically relevant safety advantages.	The recommendation was changed to strong in the updated guideline. As explained in the guideline the relevant RCTs 1929 included two interventions: a) different follitropin medications, and b) individualised versus fixed dosing. Therefore, it is uncertain that the effect on OHSS rate is due to the gonadotropin or the dosing regimen.
73	Juan-Enrique Schwarze Shiv Gupta Susana Montenegro*	17	R36	The current recommendation is extremely misleading especially given the fact that between June 2007 (when marketing authorization was granted for r-hLH:r-hLH) and the end of September 2024, more than 2.5 million cycles have been performed with r-hFSH:r-hLH worldwide without any higher risk of safety compared to other gonadotropins. The current recommendation is solely based on a single small study: Urinary hMG (Menopur) versus recombinant hFSH plus recombinant hLH (Pergoveris) in	Stakeholders correctly note that the only published RCT directly comparing rFSH + rLH with hMG is Pacchiarotti et al. (2010). Importantly, this study does not demonstrate clinical equivalence between these regimens but instead highlights that no conclusive evidence

NR	Reviewer	Page	Line	Comment	Action / Reply
				IVF: a multicenter, prospective, randomized controlled trial" Pacchiarotti et al., Fertility and Sterility, 2010. A thorough methodological assessment reveals that the study suffers from critical design limitations, severe underpowering, lack of transparency in randomization, flawed statistical procedures, and legal constraints that invalidate the comparability of stimulation outcomes. As such, its findings cannot be relied upon as robust evidence for clinical equivalence between these two gonadotropin protocols. Below are detailed report of the limitations and flaws in the study: 1. Lack of Clear Research Question and Estimand - No primary endpoint, clear research question, or specific estimand defined. - Multiple outcomes treated as co-primary without correction for multiplicity. - No hypothesis or MCID declared; clinical interpretation impossible. - Power only to detect ≥25 percentage point difference; observed difference ~1%. 2. Underpowered to Detect Clinically Meaningful Differences - Study is statistically incapable of confirming equivalence or non-inferiority. - Equating non-significance with 'no difference' is incorrect. 3. Inadequate Randomization Transparency - No details on allocation concealment, sequence generation, or baseline characteristics. - High risk of bias in the randomization process domain (Cochrane RoB 2). 4. Implausible Clinical Outcomes and Questionable Internal Validity - Oocyte yield and MII proportions far below ESHRE KPIs. - Raises questions on lab performance, protocol adherence, or reporting integrity. 5. Incomplete Reporting of Embryological and Clinical Data - No details on embryo morphology, fertilization, or cumulative outcomes. - Lack of linkage between embryo quality and clinical outcomes. 6. Flawed Statistical Practice - 16+ tests without applied correction for multiplicity.	exists to favor one approach over the other. Given its methodological limitations—including a small sample size, lack of clear endpoints, and the impact of Italian law (Law 40/2004) on stimulation protocols—the study cannot be considered a robust basis for strong recommendations. For these reasons the guideline appropriately issued a conditional, safety-based recommendation, limited to GnRH agonist protocols.

NR	Reviewer	Page	Line	Comment	Action / Reply
				- Mention of Bonferroni/Sidak correction unsubstantiated by adjusted values. 7. Legal Constraints Compromising Trial Validity - Italian Law 40/2004 biased the downstream outcomes of stimulation. - Conditioning on number of oocytes/embryos transferred introduces structural bias. 8. Risk of Bias (Cochrane RoB 2 Summary) Domain | Risk of Bias Randomization Process | High Deviations from Intended Interventions | Some Concerns Missing Outcome Data | Low Measurement of the Outcome | Some Concerns Selection of the Reported Result | High **Overall Judgment** | **High Risk** Final Appraisal Given the combination of unclear hypotheses, poor reporting, flawed statistical analysis, and severe structural bias introduced by national legislation, this trial fails to meet the methodological standards required for reliable evidence synthesis. It is inappropriate to use this study as a sole reference to this recommendation and given the long safety data, we suggest removing this recommendation from the guideline.	
75	Xi Dong	17	R36	The recommendation 36 could be justified as: The use of recombinant LH + recombinant FSH (rFSH+rLH) for ovarian stimulation is probably recommended over hMG. • The recommendation 36 is based on the cancelation rate of one RCT included 122 patients underwent GnRH agonist. However, another larger sample real world study (n=999) showed significant higher cancelation rate in hMG group compare with rLH+rFSH group in GnRH agonist cycles. The clinical pregnancy per started cycle was higher in the rFSH+rLH group1. • A prospective randomized cohort study included 94 patients aged 38-	Stakeholders correctly note that the only published RCT directly comparing rFSH + rLH with hMG is Pacchiarotti et al. (2010). Importantly, this study does not demonstrate clinical equivalence between these regimens but instead highlights that no conclusive evidence exists to favor one approach over the other. Given its methodological limitations—including a small sample size, lack of clear endpoints, and the impact of Italian law (Law 40/2004) on

NR	Reviewer	Page	Line	Comment	Action / Reply
				40 years old undergoing GnRH agonist downregulation. 58 patients received rFSH + HMG and 36 received rFSH + rLH, no hyperstimulation syndrome occurred in either group. Significantly more MII oocytes and pregnancy rate were seen in the group treated with rFSH + rLH than rFSH + HMG2. • A cohort study enrolled patients (n=122) who underwent IVF stimulation with a long GnRH agonist protocol and received FSH plus r-LH or hMG. rFSH+ rLH group has statistically high number of oocytes and embryo, pregnancy rate compared with hMG group. Ovarian hyperstimulation syndrome was avoided in all cases3. References: 1. *Conforti A, Esteves SC, Di Rella F, et al. The role of recombinant LH in women with hypo-response to controlled ovarian stimulation: a systematic review and meta-analysis [published correction appears in Reprod Biol Endocrinol. 2019 Mar 14;17(1):31. doi: 10.1186/s12958-019-0475-x.]. Reprod Biol Endocrinol. 2019;17(1):18. Published 2019 Feb 6. doi:10.1186/s12958-019-0460-4* 2. *Ferraretti AP, Gianaroli L, Magli MC, D'angelo A, Farfalli V, Montanaro N. Exogenous luteinizing hormone in controlled ovarian hyperstimulation for assisted reproduction techniques. Fertil Steril. 2004;82(6):1521-1526. doi:10.1016/j.fertnstert.2004.06.041* 3. *Alviggi C, Humaidan P, Fischer R, et al. Patients with low prognosis in ART: a Delphi consensus to identify potential clinical implications and measure the impact of POSEIDON criteria. Reprod Biol Endocrinol. 2024;22(1):122. Published 2024 Oct 10. doi:10.1186/s12958-024-01291-x* 4. *Orvieto R, Venetis CA, Fatemi HM, et al. Clinical recommendations for the use of LH in controlled ovarian stimulation: a consensus paper. Front Endocrinol (Lausanne). 2021;12:675670. doi:10.3389/fendo.2021.675670*	stimulation protocols—the study cannot be considered a robust basis for strong recommendations. For these reasons the guideline appropriately issued a conditional, safety-based recommendation, limited to GnRH agonist protocols.

NR	Reviewer	Page	Line	Comment	Action / Reply
				5. *Barrenetxea G, Hernández C, Herrero J, et al. Use of gonadotropins in ovarian stimulation in Spain: Delphi consensus. J Obstet Gynaecol. 2023;43(1):2174692. doi:10.1080/01443615. 2023.2174692* 6. *Salehpour S, Aleyasin A, Moini A, et al. Luteinizing hormone supplementation in controlled ovarian stimulation: the Iran Delphi consensus. Front Reprod Health. 2024;6:1397446. doi:10.3389/frph.2024.1397446* *Li R, Wang Y; China Expert Suggestions Group on the definition of in-vitro fertilization success. How to define in-vitro fertilization success: a Delphi consensus among China experts. Fertil Reprod. 2024;6(3):135-142. doi:10.1142/S2661318224500191*	
102	Miaoxin Chen	17	R36	Recommendation 36 can be justified by stating that the use of recombinant LH combined with recombinant FSH (rFSH + rLH) for ovarian stimulation is likely preferred over hMG. • The recommendation 36 is based on the cancelation rate of one RCT included 122 patients underwent GnRH agonist. However, another larger sample real world study (n=999) showed significant higher cancelation rate in hMG group compare with rLH+rFSH group in GnRH agonist cycles. The clinical pregnancy per started cycle was higher in the rFSH+rLH group1. • A prospective randomized cohort study included 94 patients aged 38-40 years old undergoing GnRH agonist downregulation. 58 patients received rFSH + HMG and 36 received rFSH + rLH, no hyperstimulation syndrome occurred in either group. Significantly more MII oocytes and pregnancy rate were seen in the group treated with rFSH + rLH than rFSH + HMG2. • A cohort study enrolled patients (n=122) who underwent IVF stimulation with a long GnRH agonist protocol and received FSH plus r-LH or hMG. rFSH+ rLH group has statistically high number of oocytes and	Stakeholders correctly note that the only published RCT directly comparing rFSH + rLH with hMG is Pacchiarotti et al. (2010). Importantly, this study does not demonstrate clinical equivalence between these regimens but instead highlights that no conclusive evidence exists to favor one approach over the other. Given its methodological limitations—including a small sample size, lack of clear endpoints, and the impact of Italian law (Law 40/2004) on stimulation protocols—the study cannot be considered a robust basis for strong recommendations. For these reasons the guideline appropriately issued a conditional, safety-based recommendation, limited to GnRH agonist protocols.

NR	Reviewer	Page	Line	Comment	Action / Reply
				embryo, pregnancy rate compared with hMG group. Ovarian hyperstimulation syndrome was avoided in all cases3.	
106	Rishma Dhillon Pai	17	R38	ITS ALMOST ALWAYS LETROZOLE FOLLOWED BY GONADOTRPHINS AND NOT GONADOTROPHINS FOLLOWED BY LETROZOLE	The recommendation was deleted.
137	Jayesh Amin	17	R39	The addition of letrozole to gonadotropins in stimulation protocols for predicted high responders is probably not recommended except in cases of cancer patients, where letrozole along with gonadotropins can be recommended in Hormone- receptor positive cancer patients.	This is addressed in the chapter on fertility preservation.
132	The Chinese Expert Review Panel for ESHRE OS Guideline	17	R39-41	Do not totally agree with the statement that the addition of letrozole to gonadotropins in stimulation protocols for predicted high / normal / low responders is probably not recommended. Because Some patients, such as patients with high risk of OHSS or patients with estrogen-sensitive tumors, may benefit from adding letrozole.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
39	Ahmed Elsayed Hassan Hamed Elbohoty	64	1738	The guideline does not currently address the cost-effectiveness of different gonadotrophin preparations (e.g., rFSH vs. hMG), despite robust evidence showing comparable clinical outcomes in many patient populations. Given the significant cost implications for patients and healthcare systems, I recommend the inclusion of a brief summary of the economic literature comparing gonadotrophin types. Recombinant components is more expensive than human menopausal gonadotrophins and Most RCTs and meta-analyses suggest comparable live birth and pregnancy rates between them. Guidance on Cost-effectiveness would be of high clinical importance.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
39	Ahmed Elsayed Hassan Hamed Elbohoty	64	1738	It does not provide specific recommendations for the use of particular gonadotrophins in unique clinical situations such as hypopituitarism or hypogonadotropic hypogonadism.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.

NR	Reviewer	Page	Line	Comment	Action / Reply
				The guideline currently does not address the use of specific gonadotropin preparations in defined clinical contexts. It would be helpful to include a statement regarding conditions such as hypogonadotropic hypogonadism or hypopituitarism, where protocols using FSH alone may be insufficient. In these patients, the absence of endogenous LH necessitates the inclusion of exogenous LH activity, either through human menopausal gonadotropin (hMG) or the combination of recombinant FSH with recombinant LH, to ensure appropriate follicular development and steroidogenesis.	
19	Shikha Gupta	64	1763	Use of rFSH +HMG for COS is equally recommended 1-Includes studies with many heterogenous group of patients 2.success rate used here has been - CBR(Clinical birt rate) Currently with improved freezing and thawing technologies, better index for measuring success outcome should be CLBR (Cumulative live birth rate)	Heterogeneity in patient groups is expected and not considered unusual. The relevant outcome is indeed the cumulative live birth rate
126	Emre Goksan Pabuccu	64	1763	The use of recombinant FSH (rFSH) and human menopausal gonadotropin (hMG) for ovarian stimulation is equally recommended. [2019]à“The use of recombinant FSH (rFSH) and human menopausal gonadotropin (hMG) for ovarian stimulation is equally recommended. [2019] Based on current evidence, although the number of oocytes retrieved is higher in the rFSH group, pregnancy and live birth rates have been found to be similar between the two groups.”2	The guideline group issued a strong recommendation based on a relevant systematic review and meta-analysis that included data from 3,397 women. Additional studies addressing the specific PICO question were published after the completion of the systematic review and were therefore not included in the pooled analysis. Nonetheless, the findings of these subsequent studies are consistent with the conclusions of the Cochrane review.
127	ESHRE SIG RE	64	1763	The use of recombinant FSH (rFSH) and human menopausal gonadotropin (hMG) for ovarian stimulation is equally recommended. [2019] The GDG seems to be basing this strong recommendation on a meta-	While we acknowledge the reviewer's methodological concerns, we believe the evidence base is more robust than suggested and supports the current recommendation for

NR	Reviewer	Page	Line	Comment	Action / Reply
				analysis published in 2019 which includes only 3 RCT with cumulative live birth rates. The effect size is RR: 0.91 (95% CI: 0.80-1.04) which is very similar to the effect size regarding live birth rates RR: 0.88 (95% CI:0.78-0.99). Could the non-significant finding in CLBR be a type II error? Is this evidence sufficient to exclude a potential difference in efficacy and to formulate a strong recommendation?	several reasons: Regarding cumulative live birth rates (CLBR): Beyond the 2019 meta-analysis of 3 RCTs (n=2,109 women) showing no significant difference between rFSH and hMG (RR: 0.91, 95% CI: 0.80-1.04), this finding has been corroborated by a subsequent individual RCT that also demonstrated no significant difference in cumulative live birth rates. Regarding live birth rates: The evidence base extends beyond the original meta-analysis. Three additional RCTs have been published, consistently showing no statistically significant differences in live birth rates between hMG and rFSH: Parsanezhad et al. (2017): 27.5% vs. 40% (hMG vs. rFSH) Turkcapar et al. (2013): 23.1% vs. 35.7% (hMG vs. rFSH) Witz et al. (2020): 52.2% vs. 48.7% (hMG vs. rFSH)
128	Apostolos Tsironis	64	1763	On the RCT by Witz et al, the risk of OHSS is significantly higher with r-FSH vs h-MG therefore this should be reflected in our recommendation, for example: in predicted high responders the risk of OHSS is higher with r-FSH vs HMG.	While Witz et al. reported a statistically significant difference in overall OHSS rates, the study showed no difference in severe OHSS rates between hMG and rFSH groups (2.6% vs. 2.9%, respectively; total severe OHSS rate: 2.7%). Severe OHSS represents the clinically most concerning outcome, and the equivalent

NR	Reviewer	Page	Line	Comment	Action / Reply
					rates between treatments suggest that any difference in milder forms of OHSS may not translate to clinically meaningful safety concerns.
125	Kokkoni Kiose	65	1782-1788 1810-1816	While the recommendation asserts that both agents (rFSH and pFSH) can be equally used, the accompanying justification raises concerns that do not appear to support such a conclusion with the level of certainty typically required for a strong recommendation. Specifically, the justification notes that in GnRH agonist protocols, the use of pFSH is not preferable to rFSH, and that in GnRH antagonist protocols, there is insufficient evidence to draw conclusions regarding the comparative effectiveness of the two preparations. This evidence base seems misaligned with the issuance of a strong recommendation implying clinical equivalence across all protocols and patient groups.	The evidence presented in the updated guideline supports the formulation of such a recommendation specifically in the context of GnRH agonist protocols.
98	Kanad Dev Nayar	66-67	1820-1858	These recommendations may not fully reflect the latest scientific evidence of optimal treatment options for patients with r-hFSH:r-hLH treatment. Patients deserve the best chance of success in their first cycle.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
80	Roberto Matorras	67	1832-1846	The ESHRE draft states In a sub-analysis of the meta-analysis, a small RCT in poor responders showed a beneficial effect of rLH 1833 pre-treatment to rFSH on live birth rate (OR 9.33, 95% CI 1.03-84.20, 43 women) (Ferraretti et al., 2014, 1834 Mochtar et al., 2017). However, a large RCT (939 women), more recent than the meta-analysis, 1835 reported no effect of rLH addition to rFSH in Bologna poor responders on live birth rate (10.6% (49/462) 1836 vs. 11.7% (56/477)) (Humaidan et al., 2017). In this trial, only one event of mild early OHSS occurred in 1837 the rFSH+rLH group. However, in a more recent meta-analysis (Conforti et al, 2019), it was concluded that Significantly higher clinical pregnancy rates (odds ratio: 2.03, P = 0.003), implantation rates (odds ratio: 2.62, P = 0.004) and number of oocytes retrieved	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions.

NR	Reviewer	Page	Line	Comment	Action / Reply
				(weight mean differences: 1.98, P = 0.03) were observed in hypo-responders supplemented with recombinant LH versus hypo-responders who underwent FSH monotherapy Thus, in our opinion the statement should be The combination of LH and FSH is recommended over the use of LH alone in hyporesponders women.	Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required

NR	Reviewer	Page	Line	Comment	Action / Reply
80	Roberto Matorras	67	1838-1847	We would like to make the following comments. The guideline refers to Conforti et al meta-analysis (2021), and concludes that in advanced age women (> 35 years) the combination of rFSH with rLH and rFSH alone alone are probably equally recommended for women of advanced age (≥35 year). However the aforementioned meta-analysis it was concluded that in women aged between 35 and 40 years, r-hFSH/r-hLH co-treatment was associated with higher clinical pregnancy rates (OR 1.45, CI 95% 1.05-2.00, I2 = 0%, P = 0.03) and implantation rates (OR 1.49, CI 95% 1.10-2.01, I2 = 13%, P = 0.01) versus r-hFSH monotherapy. Thus in our opinion the recommendation should be split in two. The combination of LH and FSH is recommended over the use of LH alone in women aged 35 to 40 years. The combination of rFSH with rLH and rFSH alone are probably equally recommended for women of advanced age (> 40 year)	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not

NR	Reviewer	Page	Line	Comment	Action / Reply
					show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
126	Emre Goksan Pabuccu	67	1845	The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. [updated]à We recommend that future guidelines address sub-optimal or hypo-responders, a distinct group not currently covered. RCTs suggest that r-hFSH:r-hLH may be more effective than r-hFSH alone or higher FSH doses in improving outcomes in this population. 3,4	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield

NR	Reviewer	Page	Line	Comment	Action / Reply
					increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
127	ESHRE SIG RE	67	1845	The combination of rFSH with rLH and rFSH alone are probably equally recommended for the general IVF population. [updated] This is justified as follows “According to the best available evidence, the combination of rFSH with rLH results in similar live birth rates compared to rFSH alone.” If the addition of rLH does not confer any	The recommendation that recombinant follicle-stimulating hormone (rFSH) combined with recombinant luteinizing hormone (rLH) is equally recommended as rFSH alone for the general IVF population is supported by evidence showing comparable live birth rates between the two

NR	Reviewer	Page	Line	Comment	Action / Reply
				benefit, based on the principle of preferring simpler forms of treatment (lines 1856-1858), how is this recommendation justified?	approaches, as noted in the provided reference. Regarding the principle of preferring simpler treatments, the combination of rFSH and rLH does not significantly increase treatment complexity. Both rFSH alone and rFSH+rLH can be administered via a single subcutaneous injection, ensuring ease of use and patient convenience.
125	Kokkoni Kiose	67	1845-1858	I would like to provide a comment regarding the conditional recommendation stating that recombinant FSH (rFSH) alone and rFSH combined with recombinant LH (rLH) are equally recommended for ovarian stimulation. According to the justification provided, current evidence demonstrates no clear benefit in terms of live birth rate from the addition of rLH to rFSH in the general IVF population. In this context, it is unclear how the conclusion of equal recommendation is derived when the addition of rLH appears to confer no additional efficacy and may increase cost and complexity of treatment.	The GDG considered the principle of simplicity in its deliberations; therefore, the recommendation was framed as a conditional recommendation, using the phrasing “probably recommended.”
19	Shikha Gupta	67	1846	The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. Several meta analysis over part decade have suggested Rlh+FSH as compared to rFSH alone for cpr and CBR is poor responder especially advanced maternal age So recommendation may be changed.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
127	ESHRE SIG RE	67	1846	The combination of rFSH with rLH and rFSH alone are probably equally recommended for low responders. [updated] This is justified as follows “Current evidence from a large RCT in low responders indicated no beneficial effect of the combination of rFSH with rLH and rFSH alone on live birth rate.” If the addition of rLH	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes

NR	Reviewer	Page	Line	Comment	Action / Reply
				does not confer any benefit, based on the principle of preferring simpler forms of treatment (lines 1856-1858), how is this recommendation justified? It could be perceived that patient had low or hypo response in previous cycle with r-hFSH alone, both options of r-hFSH+r-hLH or r-hFSH alone are equally effective. The systematic review by Conforti et al. (2019), evaluating the clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone. The pooled odds ratio was 2.03, with a 95% confidence interval (CI) of 1.27 to 3.25 (p = 0.003), indicating more than a two-fold increase in the odds of achieving clinical pregnancy with the combination therapy. A prospective RCTs had shown r-hLH is more effective than increasing the dose of r-hFSH in patients with an initial inadequate ovarian response to r-hFSH alone. Women showing hypo-responsiveness to r-hFSH were randomized to receive an increased dose of r-hFSH, or the combination of r-hLH 75–150 IU and an increased dose of r-hFSH. Implantation rates and pregnancy rates were higher in those with r-hFSH + r-hLH treatment ($p < 0.05$). The LBR for r-hFSH + r-hLH group was 40.7% whereas the r-hFSHgroup was 22% (Ferraretti et l., 2004).	only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited.

NR	Reviewer	Page	Line	Comment	Action / Reply
					In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are requiredThe GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
128	Apostolos Tsironis	67	1846	I respectfully disagree with the recommendation as I think we have evidence to suggest that the combination of FSH+LH may result in improved outcomes in poor responders: 1. Lehert 2014: the combination of FSH+LH resulted in higher oocyte numbers and clinical pregnancy rates in poor responders as defined by the Bologna criteria. 2. A post hoc analysis on ESPART trial(Humaidan 2017): wonen with moderate to severe POR have significantly less early pregnancy failures when received FSH+LH compared to FSH alone 3. Comforti 2019: syst review and metanalysis on unexpected low responders – the use of FSH+LH is associated with higher number of eggs, implantation and clinical pregnancy rates. 4. Ferraneti 2014 – Already included Furthermore, the study used as reference (Lahoud 2017) defines low responders based on day6 LH levels on a long protocol which is largely irrelevant in the current practice.	We thank the reviewers for highlighting the role of r-hLH in hypo-responders. While the topic is clinically relevant, the current evidence base does not support a guideline-level recommendation for its routine use. The meta-analysis by Conforti et al. (2019), which forms the basis of much of the cited support, includes only two RCTs—Ferraretti et al. (2004) and De Placido et al. (2005)—conducted nearly two decades ago, prior to the development of the POSEIDON criteria and current stratification methods. As a result, the populations studied are not clearly translatable to today's hypo-responder definitions. Importantly, the primary outcome in these trials was clinical pregnancy, not live birth. Only one trial reported live birth rates, and with a small sample size (N=104), the evidence is insufficient to support conclusions on the guideline's critical endpoints. The reported clinical benefits are also difficult to interpret biologically. Although total oocyte yield increased modestly, there was no difference in the number of mature (MII) oocytes. Without

NR	Reviewer	Page	Line	Comment	Action / Reply
					corresponding improvements in embryo number, quality, or fertilization rate, the observed increase in pregnancy rates lacks mechanistic plausibility. Both trials also introduced confounding through cointerventions: one used escalating FSH in both arms but added r-hLH only to one group; the other replaced increased FSH dosing with r-hLH, complicating attribution of effects. Additionally, the analysis of implantation rate as a binary outcome rather than a per-embryo metric raises statistical concerns. Finally, while Delphi consensus statements can provide expert guidance, they rely on the quality of the supporting evidence. In this case, the foundational data are outdated and methodologically limited. In addition, While we acknowledge the reviewer's interest in optimizing treatment for poor responders, the cited evidence does not meet the methodological standards required for modifying our recommendation: Primary Evidence Assessment: Mochtar et al. (2017) - Cochrane Review: This systematic review of 36 RCTs identified only one relevant study (Ferraretti et al., 2014) for poor responders. Critically, this trial evaluated LH pretreatment rather than co-treatment, making it inadequate to address the clinical question. Lehert et al. (2014): While this meta-analysis showed higher clinical pregnancy rates in poor responders (RR 1.30), it demonstrated no significant improvement in live birth rate - the most clinically

NR	Reviewer	Page	Line	Comment	Action / Reply
					relevant endpoint for patients and the primary outcome measure for our recommendations. Humaidan et al. (2017) - ESPART Trial: The original pre-specified analysis showed similar live birth rates between groups, which is the appropriate basis for guideline recommendations. Methodological Concerns with Post Hoc Evidence: Post hoc analyses, while scientifically interesting, cannot directly inform clinical practice guidelines because they: Are exploratory and not pre-specified Carry high risk of bias and type I error Do not meet GRADE framework standards for evidence quality Risk misleading clinicians without replication in prospective RCTs In light of these considerations, we believe that current evidence is insufficient to support the use of r-hLH in hypo-responders as a formal recommendation. Further prospective studies using contemporary definitions and live birth as a primary endpoint are required
25	Alberto Revelli	67	1846-1847	The published evidence indicates that the addition of r-LH to r-FSH is likely to improve IVF results in low responders and women above 35 years, so the association of both gonadotropins should probably be recommended for those subcategories of IVF patients	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
126	Emre Goksan Pabuccu	67	1847	The combination of rFSH with rLH and rFSH alone are probably equally recommended for women of advanced age (≥35 year). [updated] àWe suggest that future updates reconsider the recommendation for women of advanced age (≥35 years), as age-related decline in LH activity and	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical

NR	Reviewer	Page	Line	Comment	Action / Reply
				impaired steroidogenesis may impact outcomes. Evidence from Bosch et al. (2021), Conforti et al. (2021), and Bielfeld et al. (2023) indicates that r-hFSH:r-hLH combination therapy yields significantly better clinical pregnancy and live birth rates than r-hFSH alone in this age group. 5-7	practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate

NR	Reviewer	Page	Line	Comment	Action / Reply
					into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint.
127	ESHRE SIG RE	67	1847	The combination of rFSH with rLH and rFSH alone are probably equally recommended for women of advanced age (≥35 year). [updated] This is justified as follows "Similarly, a systematic review and meta-analysis focussing on women of advanced age (≥35 years) found no evidence of a benefit of adding rLH to ovarian stimulation with rFSH (Conforti et al., 2021)." If the addition of rLH does not confer any benefit, based on the principle of preferring simpler forms of treatment (lines 1856-1858), how is this recommendation justified? A meta-analysis by Conforti et al. (2021), which included only RCTs evaluating both implantation rate and clinical pregnancy rate demonstrated a statistically significant benefit in favor of ovarian stimulation with r-hFSH + r-hLH, compared to r-hFSH alone in women >35 -40 years of age. The pooled odds ratio was 1.49 for implantation rate, with a 95% confidence interval (CI) of 1.10 to 2.01 (p = 0.01) and a pooled odds ratio was 1.45 for clinical pregnancy rate, with a 95% confidence interval (CI) of 1.05 to 2.00 (p = 0.03).	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and

NR	Reviewer	Page	Line	Comment	Action / Reply
					Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
128	Apostolos Tsironis	67	1847	I have the belief that women over the age of 35y might benefit from the combination of FSH +LH 1. Comforti 2021: combination of FSH+LH results in significantly higher implantation and clinical pregnancy rates in women between 35-40y (as well as higher oocyte numbers) 2. Bosch 2021: combination of FSH+LH results in significantly higher implantation and clinical pregnancy rates in women between 36-39y	1. Bielfeld et al., 2023 – Real-World Data (Registry Analysis) This study is a large retrospective observational cohort, based on the German IVF registry. While it provides valuable insights into routine clinical practice, such real-world data cannot establish causal relationships due to the potential for residual confounding, despite attempts to adjust for baseline

NR	Reviewer	Page	Line	Comment	Action / Reply
				The 2 trials that included Live birth data were: a. Vuong et al 2015: this trial was open label which may have introduced bias and the authors admit in the abstract that the study was likely to be underpowered b. Matorras et al 2009: this trial was underpowered to detect differences in pregnancy outcomes and the authors mention that the addition of LH actually is associated with improved outcomes (not statistically significant) which should be considered clinically relevant Finally, real world data and registry studies (Bielfeld et al 2024) indicate a positive impact of the use of combination FSH+LH in women of advanced reproductive age.	characteristics. → Implication: Observational studies are informative but cannot replace randomized controlled trials (RCTs) when developing clinical recommendations. 2. Bosch et al., 2021 – Narrative Review and Mechanistic Discussion This publication is a narrative review, focused on hypothesized physiological mechanisms (e.g., functional LH deficiency in advanced age). While it offers a compelling biological rationale, it does not include new clinical trial data or provide outcome-based comparative evidence. → Implication: Mechanistic plausibility is important for hypothesis generation but is insufficient alone to support clinical recommendations. 3. Conforti et al., 2021 – Systematic Review and Meta-analysis of RCTs This was the only meta-analysis of randomized controlled trials cited, and it was taken into account during the guideline development process. However, two key points must be emphasized: • The primary endpoint in guideline development was live birth rate (LBR). The Conforti analysis did not show a significant difference in LBR either overall in women >35 years or in the subgroup aged 35–40 years, which was an arbitrary defined subgroup. • The pooled benefits reported in implantation and clinical pregnancy rates do not necessarily translate into live birth improvements, and this distinction is critical in evidence grading. → Implication: While the Conforti meta-analysis

NR	Reviewer	Page	Line	Comment	Action / Reply
					contributes useful data, it did not demonstrate a benefit in live birth outcomes, which limits its impact on guideline recommendations that prioritize LBR as the clinically most meaningful endpoint. The GDG took the priniciple of simplicity into account, therefore the recommendation was formulated as a conditional recommendation, with the wording "probably recommended".
37	Ahmed Samy Abdelazim Saad	68	1886	It's better to change it to equally recommended or there is no strong evidence than not recommended over. It gives more safe area to use any of them according to availability & cost and patient and doctor's satisfaction to use rec FSH alone or Combined rec FSH & LH in some patients or at the end of stimulation or from the start. The evidence as mentioned in the guideline from only a handful studies, so we cannot withdraw solid conclusion from that.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
128	Apostolos Tsironis	68	1886	The recommendation to avoid this widely used combination is based on 2 trials 1. Shu et al 2019: a study including patient with average age 28 and BMI 22 – NO DIFERENCE WOULD BE EXPECTED ANYWAY 2. Qiu et al 2023: a study on Poseidon 4 patients where no standardization in protocols took place and also the majority of embryos transferred was day 3. It is probably better to state that there is not enough evidence to recommend a specific approach.	A conditional recommendation was formulated based on the presence of only a handful studies suggesting that i adding hMG either in the beginning of the stimulation with rFSH or after a rFSH stimulation period of 5-8 days, does not create any benefits in patients using either the GnRH agonist or antagonist pituitary suppression protocol.
82	Eduardo Correa Allende*	70	1923	A biosimilar is a biological medicine highly similar to another biological medicine already approved in the EU (the so-called 'reference medicine')1. In the EU, two rFSH alfa biosimilars have been approved, showing no significant differences in pharmacokinetics, immunogenicity, or safety compared to their originators. Multiple studies across Spain,	This section has been introduced in the guideline, however, due to disagreement in the GDG, no recommendation was formulated.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Germany, Austria, and France have confirmed their comparable effectiveness and safety in ovarian stimulation protocols, with pregnancy and oocyte retrieval rates aligning with national ART registries. Adverse event rates, including OHSS, were low. A large French study also found no significant difference in cumulative live birth rates between biosimilars and originators. Endorsed by European authorities, the EMA affirms biosimilars' clinical equivalence, highlighting their safe interchangeability without the need for additional switch studies, based on more than a decade of evidence and experience1,2. References available upon request.	
39	Ahmed Elsayed Hassan Hamed Elbohoty	71	1952	The use of recombinant LH (rLH)+recombinant FSH (rFSH+LH) for ovarian stimulation is probably not recommended over human menopausal gonadotropin (hMG) in GnRH agonist protocols with regards to safety. [2019] I am not convinced by the statement that rFSH+rLH is "probably not recommended over hMG" with regard to safety. While some studies showed no consistent safety concern associated with rLH use. Moreover, subgroups such as older women, poor responders, or hypo-responders may benefit from rFSH+rLH. I suggest rephrasing this statement to reflect the lack of superiority, rather than suggesting inferiority of rLH.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
128	Apostolos Tsironis	71	1952	The wording of the recommendation should probably be different, for example if FSH+LH is used on a long protocol consider adjusting the dose to avoid OHSS due to higher potency of Rfsh compared to HMG.(Pacchiarotti et al are using long protocol and the same dose of HMG and Pergoveris – 225iu- with no mention of baseline ovarian reserve markers)	Stakeholders correctly note that the only published RCT directly comparing rFSH + rLH with hMG is Pacchiarotti et al. (2010). Importantly, this study does not demonstrate clinical equivalence between these regimens but instead highlights that no conclusive evidence exists to favor one approach over the other.

NR	Reviewer	Page	Line	Comment	Action / Reply
				When the comparison between FSH and HMG was described OHSS is not even mentioned in the recommendation (see my previous comments)	Given its methodological limitations—including a small sample size, lack of clear endpoints, and the impact of Italian law (Law 40/2004) on stimulation protocols—the study cannot be considered a robust basis for strong recommendations. For these reasons the guideline appropriately issued a conditional, safety-based recommendation, limited to GnRH agonist protocols.
37	Ahmed Samy Abdelazim Saad	73	2031	I agree with the recommendation, but many recent highly purified HMG now has mainly HCG than LH in their composition and it was mentioned in this guideline before that it's equally effective with FSH alone or rec FSH. So if needed to be mentioned, specify that point alone then clarify it as low doses of HCG alone not in HMG vial	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
153	Stefan Matik	75	2114	It might be added the combination of letrozole with gonadotropins could be considered in the setting of fertility preservation in women with estrogen-sensitive diseases (e.g. breast cancer patients) – according to the recommendations from the 2020 Female fertility preservation guidelines from the ESHRE Female fertility preservation guideline development group	This is discussed in the fertility preservation chapter
37	Ahmed Samy Abdelazim Saad	75	2115	I would recommend it. It will decrease the dose of FSH, it will be much safer	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
37	Ahmed Samy Abdelazim Saad	75	2117	I would recommend its use to use the patient own FSH & LH. We permit using modified natural cycle in low responders, but we are against the use of letrozole in addition to FSH, this is a contradiction. Let's make it equally recommended at least instead of not recommended.	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.

NR	Reviewer	Page	Line	Comment	Action / Reply
39	Ahmed Elsayed Hassan Hamed Elbohoty	75	2117	The guideline currently does not address the use of letrozole as an adjunct for ovarian stimulation in specific clinical contexts where estrogen suppression is desirable, such as: § Women with estrogen-sensitive malignancies, including current or prior hormone receptor-positive breast cancer § Patients with endometriosis, where minimizing estrogen may reduce lesion activity § Women at increased risk of thromboembolism, where high estradiol levels could exacerbate vascular risk Letrozole-based protocols are widely used in fertility preservation and in minimal stimulation IVF strategies. They reduce systemic estradiol levels without compromising oocyte yield and are endorsed by multiple international societies in these specific scenarios (Oktay K et al., Fertil Steril. 2005; Azim AA et al., J Clin Oncol. 2008; Kim JY et al., Clin Exp Reprod Med. 2014; Oktay K et al., J Clin Oncol. 2018). I strongly recommend that the guideline include a brief, evidence-based statement acknowledging the use of letrozole in ovarian stimulation for these special populations, to reflect current clinical practice and improve the applicability of the guideline.	With regard to safety, recent large-scale analyses—such as the study by Tulandi et al. (2015)—have shown no increased risk of congenital anomalies associated with letrozole compared to clomiphene citrate or natural conception. Although manufacturer warnings remain in place due to initial preclinical concerns, these have not been substantiated by current human data and have been superseded by multiple clinical guidelines, including the ESHRE 2020 guideline on fertility preservation. Nevertheless, given that letrozole is used off-label for ovulation induction, appropriate informed consent remains essential
1	Raj Mathur	76	2124	Is it still true to say that there are concerns about the teratogenicity of letrozole? This seems a bit excessive, especially without any further context	With regard to safety, recent large-scale analyses—such as the study by Tulandi et al. (2015)—have shown no increased risk of congenital anomalies associated with letrozole compared to clomiphene citrate or natural conception. Although manufacturer warnings remain in place due to initial preclinical concerns, these have not been substantiated by current human data and have been superseded

NR	Reviewer	Page	Line	Comment	Action / Reply
					by multiple clinical guidelines, including the ESHRE 2020 guideline on fertility preservation. Nevertheless, given that letrozole is used off-label for ovulation induction, appropriate informed consent remains essential
19	Shikha Gupta	76	2124	The addition of letrozole to gonadotropins in stimulation protocols for predicted high responders is probably not recommended. Letrozole used: concern has been raised about tertatogenecity which is unjustified since it has been endorsed as first line treatment in PCOD patients so., this line should be removed	With regard to safety, recent large-scale analyses—such as the study by Tulandi et al. (2015)—have shown no increased risk of congenital anomalies associated with letrozole compared to clomiphene citrate or natural conception. Although manufacturer warnings remain in place due to initial preclinical concerns, these have not been substantiated by current human data and have been superseded by multiple clinical guidelines, including the ESHRE 2020 guideline on fertility preservation. Nevertheless, given that letrozole is used off-label for ovulation induction, appropriate informed consent remains essential
7. Adjustment of gonadotropin dose					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	18	R45	Do not totally agree with the statement that adjustment (increase or decrease) of the gonadotrophin dose in the mid-stimulation phase during ovarian stimulation is probably not recommended Because Reducing the gonadotrophin dose during the late follicular phase can help prevent the progesterone elevation.	This is not supported by the current evidence. An RCT, cited in the evidence section, has shown that "a step-down approach of daily 12.5 IU rec-FSH did not achieve a significantly reduced progesterone level on the day of HCG trigger." Lawrence et al., 2021

NR	Reviewer	Page	Line	Comment	Action / Reply
88	Aboubakr Mohamed Elnashar	18	R46	Patient characteristics should be defined	In the justification it is explained that ovarian reserve testing, patient preferences etc should be used to determine the appropriate gonadotropin starting dose.
7	Sujoy Dasgupta	84	2455	Regarding dose adjustment during the stimulation-phase in poor responders, one retrospective study found that adjusting the dose of gonadotropins leads to comparable outcomes similar to the fixed-dose group. Aslan K, Kasapoğlu I, Mesut C, Gurbuz TB, Çakır C, Avcı B, Uncu G. The Effect of the Gonadotropin Dose Increment During Controlled Ovarian Hyperstimulation on Live Birth Rates of POSEIDON Group 3-4 Patients. Uludağ Tıp Derg. October 2024;50(2):203-208.	There is no reason to consider the cited retrospective study in the presence of RCTs.
1	Raj Mathur	84	2457	This recommendation is non-specific in the extreme. Would the GDG consider transferring the first sentence from the justification to the recommendation?	The GDG has discussed the formulation of this GPP thoroughly, and this formulation covers the message they intended.
8. Adjunct therapies					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	18	R47	The patient profile needs to be more specific, for instance, without insulin resistance, rather than just general PCOS patients.	Thank you for the comment. Most included studies involving women with PCOS followed the Rotterdam criteria which is the standard definition for PCOS globally. There is currently no international consensus for sub-classification of women with PCOS based on insulin resistance. There are no studies.
132	The Chinese Expert Review Panel for ESHRE OS Guideline	18	R49	Do not agree with the statement that Use of adjuvant growth hormone before and/or during ovarian stimulation is not recommended for low responders. Because The evidence cited in the guideline indicates that low responders with	The GDG has reviewed the evidence again and has decided to change the recommendation from strong to conditional against.

NR	Reviewer	Page	Line	Comment	Action / Reply
				growth hormone therapy achieve better clinical outcome compared to the control group, in terms of more MII numbers (MD 1.63, 95% CI 1.13-2.13, 11 RCTs), higher clinical pregnancy rates (OR 1.92, 95% CI 1.51-2.43, 19 RCTs, 1763 women), and higher live birth rates (OR 1.80, 95% CI 1.22-2.64, 9 RCTs, 945 women).	
137	Jayesh Amin	18	R50	Adjuvant growth hormone during ovarian stimulation can be considered when PCOS is associated with co-morbid conditions and hyporesponse to ovarian stimulation is there. Adjuvant growth hormone during ovarian stimulation can be considered in PCOS patients where even after using recombinant FSH alone or recombinant FSH along with recombinant LH; hyporesponse is there. Adjuvant growth hormone during ovarian stimulation can be considered in PCOS patients (with or without doing polymorphism test for LH and FSH receptors) where even after using recombinant FSH alone or recombinant FSH along with recombinant LH; hyporesponse is there.	Thank you for the comment. Our recommendation is based on the updated search for the PICO question and rigorous evidence synthesis process (as set out in the ESHRE manual). Please refer to the details of evidence synthesis for the PICO question and justification which explains the rationale for the recommendation.
119	Kasi V Sellappan	18	R51	The use of testosterone in poor responders in also a controversial topic as this cannot be completely ruled out as not recommended. Mireia Gonzalez-comadran et al;2012 in their meta-analysis showed statistically significant increases in live birth rates and reduced usage of gonadotrophins. This increase in live birth rates were also acknowledged by another meta analysis by Marco Noventa et al 2019. Cochrane database review also acknowledges the possibility of as increase in live birth rates with pre treatment with testosterone. Hence this should be prescribed on a case by case basis as opposed to completely ruling it out.	The GDG has reviewed the evidence again and has decided to change the recommendation from strong to conditional against.
1	Raj Mathur	86	2499	Would the GDG consider assessing adjuvant Co-enzyme Q10 role in poor responders? It is widely used by patients and clinicians are often asked about it.	Thank you for the suggestion. Adjuvant Co-enzyme Q10 was not addressed in the current guideline. The GDG will consider including with future updates. We would like to relay that

NR	Reviewer	Page	Line	Comment	Action / Reply
					adjuvant anti-oxidants (including Co-enzyme Q10) were addressed in the Unexplained Infertility Guideline, although this was not in the context of ovarian stimulation or low responders.
7	Sujoy Dasgupta	87	2533	Recommendation. In GnRH agonist protocol, metformin can be considered to reduce the risk of OHSS and miscarriage. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin J, Redman LM, Boyle JA, Norman RJ, Mousa A, Joham AE. Recommendations From the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2023 Sep 18;108(10):2447-2469.	Thank you for the comment. The GDG does not recommend GnRH agonist protocol for high responders. Our recommendation is based on the updated search for the PICO question and rigorous evidence synthesis process (as set out in the ESHRE manual). Please refer to the details of evidence synthesis for the PICO question and justification which explains the rationale for the recommendation.
7	Sujoy Dasgupta	88	2579	The word “probably” should be added, given some evidences on beneficial effects of growth hormone in low responders. Additionally, the WHO suggested that growth hormone can be considered in poor responders. Farquhar C, Marjoribanks J, Brown J, Fauser BCJM, Lethaby A, Mourad S, Rebar R, Showell M, van der Poel S. Management of ovarian stimulation for IVF: narrative review of evidence provided for World Health Organization guidance. Reprod Biomed Online. 2017 Jul;35(1):3-16.	The GDG has reviewed the evidence again and has decided to change the recommendation from strong to conditional against.
7	Sujoy Dasgupta	89	2622	The word “probably” should be added, given some evidences on beneficial effects of testosterone in low responders. Additionally, the WHO suggested that testosterone can be considered in poor responders. Farquhar C, Marjoribanks J, Brown J, Fauser BCJM, Lethaby A, Mourad S, Rebar R, Showell M, van der Poel S. Management of ovarian stimulation	The GDG has reviewed the evidence again and has decided to change the recommendation from strong to conditional against.

NR	Reviewer	Page	Line	Comment	Action / Reply
				for IVF: narrative review of evidence provided for World Health Organization guidance. Reprod Biomed Online. 2017 Jul;35(1):3-16.	
119	Kasi V Sellappan	89	2622	Testosterone pre treatment has added a little hope to women with poor reserves. There is some evidence pointing towards use of testosterone pretreatment which results in increased live birth rates in women with poor ovarian reserves. Therefore such pre treatment cannot be absolutely ruled out depriving the possibility of a live birth in that group of women.	The GDG has reviewed the evidence again and has decided to change the recommendation from strong to conditional against.
7	Sujoy Dasgupta	90	2643	The word "probably" should be added, given some evidences on beneficial effects of DHEA in low responders. Additionally, the WHO suggested that DHEA can be considered in poor responders. Farquhar C, Marjoribanks J, Brown J, Fauser BCJM, Lethaby A, Mourad S, Rebar R, Showell M, van der Poel S. Management of ovarian stimulation for IVF: narrative review of evidence provided for World Health Organization guidance. Reprod Biomed Online. 2017 Jul;35(1):3-16.	Thank you for the comment. Our recommendation is based on the updated search for the PICO question and rigorous evidence synthesis process (as set out in the ESHRE manual). Please refer to the details of evidence synthesis for the PICO question and justification which explains the rationale for the recommendation.
9. Non-conventional start of ovarian stimulation					
123	Alberto Vaiarelli	/	/	Thank you for integrating the section referring to the double stimulation protocol, which should be considered as one of the strategies within the multicycle approach. In fact, the DuoStim protocol, alongside oocyte and embryo accumulation, represents an additional option within the multicycle strategy for specific subgroups of patients with poor prognosis, characterized by advanced maternal age, low ovarian response, and reduced oocyte and embryo quality. Several aspects should be taken into account when discussing these	The GDG respectfully disagrees. This has not been shown in an RCT. In addition, these outcomes are not within the scope of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				strategies: • the potential reduction of treatment discontinuation; • the opportunity for strategic family planning in poor prognosis patients, where reproductive chances decrease over time; • a potentially better cost-effectiveness compared to two conventional stimulations, as it may reduce drop-out rates and increase the cumulative live birth rate; • new potential indications in candidates undergoing PGT-A and PGTSR. Although further investigations are warranted, these considerations are supported by growing evidence suggesting that DuoStim, when applied following an appropriate protocol, may represent a valuable option within the multicycle strategy framework.	
123	Alberto Vaiarelli	97-101	/	I suggest to add the conclusion of these metanalysis: The conclusions of three meta-analyses of published studies suggest that unconventional stimulation protocols offer comparable outcomes in terms of oocyte biological competence and reproductive results when compared to conventional cycles. However, these approaches may provide increased flexibility and improved IVF treatment efficiency by reducing the time to obtain competence embryos. Preliminary studies also indicate that this multi-cycle strategy may decrease treatment discontinuation, shorten the overall time in treatment, and improve the cost-effectiveness of IVF treatments in terms of live births and family planning outcomes. These putative benefit should be confirmed in future study. REF: How effective are the non-conventional ovarian stimulation protocols in ART? A systematic review and metaanalysis. Glujovsky D, Pesce R, Miguens M, Sueldo CE, Lattes K, Ciapponi A.J Assist Reprod	The conclusion of the meta-analysis is in accordance with the recommendation. However, the outcomes proposed are not within scope of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Genet. 2020 Dec;37(12):2913-2928. doi: 10.1007/s10815-020-01966-5. Epub 2020 Nov 21.PMID: 33219862 What is the true place of a double stimulation and double oocyte retrieval in the same cycle for patients diagnosed with poor ovarian reserve? A systematic review including a metaanalytical approach. Sfakianoudis K, Pantos K, Grigoriadis S, Rapani A, Maziotis E, Tsioulou P, Giannelou P, Kontogeorgi A, Pantou A, Vlahos N, Koutsilieris M, Simopoulou M.J Assist Reprod Genet. 2020 Jan;37(1):181- 204. doi: 10.1007/s10815-019-01638-z. Epub 2019 Dec 3.PMID: 31797242 The impact of Duostim protocol on pregnancy outcomes in infertile patients: A meta-analysis comparing single and double conventional stimulation cycles. Zeng Y, Liu W, Luo Y, Luo B, Zhu L, Yang Z, Feng K, Li D, Chen SA, Li X. J Assist Reprod Genet. 2024 Dec;41(12):3455-3466. doi: 10.1007/s10815-024-03304-5. Epub 2024 Nov 27. PMID: 39601990	
123	Alberto Vaiarelli	97-101	2951	I suggest to add this issues regarding this approach: Reduction of treatment discontinuation: A multicycle approach that exploits the recruitment of multiple follicular waves within the same ovarian cycle can reduce the rate of treatment discontinuation, which is particularly relevant in this patient population: - Second stimulation in the same ovarian cycle: an option to fully-personalize the treatment in poor prognosis patients undergoing PGT-A.Vaiarelli A, Cimadomo D, Gennarelli G, Guido M, Alviggi C, Conforti A, Livi C, Revelli A, Colamaria S, Argento C, Giuliani M, De Angelis C, Matteo M, Canosa S, D'Alfonso A, Cimadomo V, Rienzi L, Ubaldi FM.J Assist Reprod Genet. 2022 Mar;39(3):663-673. doi: 10.1007/s10815-022-02409-z. Epub 2022 Feb 7. PMID: 35128583	The GDG respectfully disagrees. This has not been shown in an RCT. In addition, these outcomes are not within the scope of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				- Luteal phase after conventional stimulation in the same ovarian cycle might improve the management of poor responder patients fulfilling the Bologna criteria: a case series.Vaiarelli A, Cimadomo D, Conforti A, Schimberni M, Giuliani M, D'Alessandro P, Colamaria S, Alviggi C, Rienzi L, Ubaldi FM.Fertil Steril. 2020 Jan;113(1):121-130. PMID31837743	
123	Alberto Vaiarelli	97-101	2960	I would add another potential benefits of this approach for couples undergoing PGT-A or PGT-SR when a limited number of embryos is available, as DuoStim allows the collection of a higher number of oocytes and embryo available for the biopsy in a shorter period of time. REF A multi-cycle approach via DuoStim is beneficial to treat couples indicated to PGT-M plus PGT-A. A propensity score matching-based case series. Vaiarelli A, Cimadomo D, Blancafort C, Trabucco E, Alviggi E, Vallefuoco R, Livi C, Benini F, Canosa S, Llácer J, Ruffa A, Borini A, Capalbo A, Rienzi L, Gennarelli G, Maria Ubaldi F. Eur J Obstet Gynecol Reprod Biol. 2024 Dec;303:272-278. doi: 10.1016/j.ejogrb.2024.11.003. Epub 2024 Nov 4. PMID: 39509926	The suggested study is not an RCT. There are several RCTs available on the topic, therefore cohort studies will not be considered to inform this recommendation.
123	Alberto Vaiarelli	97-101	2960	I suggest to add these issues: Family planning considerations in poor prognosis patients: In women with poor reproductive prognosis due to advanced maternal age or diminished ovarian reserve, the multicycle approach with DuoStim allows the accumulation of multiple embryos within a short time frame. This is particularly important when the reproductive plan includes the desire to have more than one child, as having surplus frozen embryos after achieving a first pregnancy becomes crucial. REF: Second stimulation in the same ovarian cycle: an option to	These are not RCTs. In addition, the recommendation already states that duostim can be used with the intention to accumulate oocytes or embryos.

NR	Reviewer	Page	Line	Comment	Action / Reply
				fully-personalize the treatment in poor prognosis patients undergoing PGT-A. Vaiarelli A, Cimadomo D, Gennarelli G, Guido M, Alviggi C, Conforti A, Livi C, Revelli A, Colamaria S, Argento C, Giuliani M, De Angelis C, Matteo M, Canosa S, D'Alfonso A, Cimadomo V, Rienzi L, Ubaldi FM.J Assist Reprod Genet. 2022 Mar;39(3):663-673. doi: 10.1007/s10815- 022-02409-z. Epub 2022 Feb 7.PMID: 35128583 Cost-effectiveness: The cost per live birth of the DuoStim protocol appears to be superior compared to two conventional stimulation in a private setting, primarily due to the reduction in cycle discontinuation risk and to high number of embryo available for ET. Second stimulation in the same ovarian cycle: an option to fully-personalize the treatment in poor prognosis patients undergoing PGT-A. Vaiarelli A, Cimadomo D, Gennarelli G, Guido M, Alviggi C, Conforti A, Livi C, Revelli A, Colamaria S, Argento C, Giuliani M, De Angelis C, Matteo M, Canosa S, D'Alfonso A, Cimadomo V, Rienzi L, Ubaldi FM.J Assist Reprod Genet. 2022 Mar;39(3):663-673. doi: 10.1007/s10815- 022-02409-z. Epub 2022 Feb 7.PMID: 35128583	
123	Alberto Vaiarelli	97-101	2972	SAFETY: Biological safety and reproductive and perinatal outcomes. Available evidence has demonstrated the safety of this protocol in terms of oocyte competence, implantation potential and reproductive outcomes, especially after embryo transfer of euploid embryos obtained from non-conventional phases of stimulation. REF: The euploid blastocysts obtained after luteal phase stimulation show the same clinical, obstetric and perinatal outcomes as follicular phase stimulationderived ones: a multicenter study. Vaiarelli A, Cimadomo D, Alviggi E, Sansone A, Trabucco E, Dusi L, Buffo L, Barnocchi N, Fiorini F, Colamaria S, Giuliani M, Argento C, Rienzi L, Ubaldi FM. Hum	The GDG agrees, this is why the recommendation states "could be used".

NR	Reviewer	Page	Line	Comment	Action / Reply
				Reprod. 2020 Nov 1;35(11):2598-2608. doi: 10.1093/humrep/deaa203. PMID: 32951051	
7	Sujoy Dasgupta	100	2974	"Double stimulation can be used with the intention to accumulate oocytes or embryos when fresh transfer is not planned" and when there is no possibility of natural conception	The GPP was adapted to include the risk of concurrent spontaneous conception.
39	Ahmed Elsayed Hassan Hamed Elbohoty	100	2974	Double stimulation can be used with the intention to accumulate oocytes or embryos when fresh transfer is not planned. [updated] While the guideline acknowledges that double stimulation may be used to accumulate oocytes or embryos when fresh transfer is not planned, I suggest explicitly noting that this strategy may be particularly beneficial in poor responders, especially those ≥35 years old. In these patients, embryo pooling through DuoStim can help overcome low yield per cycle and reduce time to treatment completion, especially when euploid embryos are desired. Studies by Ubaldi et al. (2016), Vaiarelli et al. (2018), and Cimadomo et al. (2020) support the safety, feasibility, and potential benefit of DuoStim in this population.	The GDG respectfully disagrees. This has not been shown in an RCT.
19	Shikha Gupta	100	2977	Dual stimulation in clinical research setting. Poor responders doing double stimulation in same cycle has shown to increase the availability of euploid embryos fit for transfer .	The availability of euploid embryos is not more than after 2 conventional cycles.
123	Alberto Vaiarelli	97-101	2980	Although this study was conducted as a randomized controlled trial (RCT), several methodological limitations and potential sources of bias have been identified, as also highlighted in the commentary published in Human Reproduction. In light of these limitations, the data and the conclusions drawn from the study should be interpreted with particular caution. What protocol should not be adopted, and which patients should not be suggested double stimulation in the same ovarian cycle? A	The GDG does not understand the objection to the Boudry trial in relation to the sentence in the justification it is cited in.

NR	Reviewer	Page	Line	Comment	Action / Reply
				randomized controlled trial answers. Alviggi C, Yarali H, Cimadomo D, Rienzi L, Ubaldi FM, Vaiarelli A. Hum Reprod. 2024 Aug 1;39(8):1860-1861. doi: 10.1093/humrep/deae141. PMID: 38908018	
126	Emre Goksan Pabuccu	97	2880	Random-start ovarian stimulation could be used when a fresh transfer is not intended and there is no possibility of natural conception. [Updated] Random-start ovarian stimulation could be used when a fresh transfer is not intended; nonetheless, the potential for spontaneous conception should always be considered.8	Random-start ovarian stimulation could be used when a fresh transfer is not intended; nonetheless, the risk of OHSS in case of concurrent spontaneous conception should always be discussed with the patient
153	Stefan Matik	97-98	2881 2935	For the random-start ovarian stimulation and luteal start ovarian stimulation, although it is stated that the they could be used when there is no possibility of natural conception, it might be wise to add a recommendation for blood or urine pregnancy test before their start, unless the medical indications for undergoing fertility treatment exclude the possibility of natural conception	The GDG has discussed the need to add this to the recommendation, however, has decided to refrain from it. In most cases, the pregnancy would not be advanced enough to detect is by a blood or urine hCG test.
10. Ovarian stimulation for fertility preservation					
91	Willem Verpoest	19	R67	For final oocyte maturation in elective oocyte cryopreservation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist trigger is advised.: there is no evidence to suggest this, please omit from your recommendation until proper studies have been performed; it also contradicts rec 78 and 79	The GPP was adapted to reflect a preference for GnRH agonist trigger.
144	Karolina Palinska-Rudzka	19	R67	The current draft recommends hCG as the preferred trigger unless the patient is at risk of OHSS. In the context of fertility preservation, particularly for patients about to start systemic cancer therapy or pelvic/abdominal surgery, this hierarchy may not fully capture the clinical value of a GnRH agonist trigger. Avoiding OHSS is essential in this population, as even mild or moderate OHSS may delay time-critical cancer treatment. Ovarian response in	The GPP was adapted to reflect a preference for GnRH agonist trigger.

NR	Reviewer	Page	Line	Comment	Action / Reply
				young oncology patients can be unpredictable, and conventional OHSS risk markers are not always reliable. Relevant data include: • Massarotti et al. (2023) reported no OHSS following a long-acting GnRH agonist trigger in oncology patients, with comparable MII oocyte yield to hCG and timely treatment commencement within five days [1]. • A Cochrane review by Youssef et al. (2014) showed an 85% reduction in moderate/severe OHSS with GnRH agonist trigger versus hCG (OR 0.15, 95% CI 0.05–0.41) in antagonist cycles, highly relevant where no fresh transfer is planned [2]. In view of the safety evidence, the practical considerations in caring for patients with cancer, and the need to minimise the burden of ovarian stimulation, it may be worth reconsidering whether the GnRH agonist trigger should be framed as the default approach for fertility preservation in oncology, rather than being reserved only for those with clear OHSS risk.	
132	The Chinese Expert Review Panel for ESHRE OS Guideline	19	R70	Do not agree with preferring hCG triggering in elective oocyte cryopreservation. Suggestion: Non-GnRH-agonist protocol, such as GnRH-antagonist protocol, are widely used in fertility preservation, then GnRH-agonist trigger could be an alternative.	The GPP was adapted to reflect a preference for GnRH agonist trigger.
79	Mitranovici Melinda Ildiko	102	3052-3054	Fertility preservation is of great interest	The GDG agrees with the reviewer.
127	ESHRE SIG RE	107	3240	For ovarian stimulation in women seeking fertility preservation for medical reasons the GnRH antagonist protocol is probably recommended. [2019] Considering previous statements on PPOS (lower cost, easy, patient friendly, similar efficacy), shouldn't there be a discussion here regarding	A sentence was added to the justification to discuss the potential use of PPOS in fertility preservation.

NR	Reviewer	Page	Line	Comment	Action / Reply
				the potential use of PPOS (similar to the elective cryopreservation section)? Of course, this should discuss that in hormone sensitive cancers, PPOS should likely be avoided and the GnRH antagonist preferred (GPP).	
7	Sujoy Dasgupta	107	3241	"In ovarian stimulation for fertility preservation in oestrogen sensitive diseases the concomitant use of anti-oestrogen therapy, such as letrozole or tamoxifen, can be considered" to reduce oestrogen but not for the purpose of improving the outcome.	The GDG agrees with the reviewer, a sentence was added to the justification.
39	Ahmed Elsayed Hassan Hamed Elbohoty	107	3241	In ovarian stimulation for fertility preservation in oestrogen sensitive diseases the concomitant use of anti-oestrogen therapy, such as letrozole or tamoxifen, can be considered. [2019] I suggest rephrasing this as a more affirmative, evidence-based recommendation, specifically endorsing letrozole-based protocols as the current standard of care for oestrogen sensitive diseases. Letrozole is extensively validated in this context and is endorsed by ASCO, ASRM, and the International Breast Cancer Study Group for fertility preservation. It allows ovarian stimulation with controlled estrogen exposure, and multiple studies have shown no increase in cancer recurrence risk when used appropriately. Tamoxifen-based protocols are less commonly used and less well-studied in this setting. Letrozole should be highlighted as the preferred agent.	There is a lack of evidence on the long-term cancer outcomes of anti-oestrogen use during OS. Therefore, some caution was taken into account.
7	Sujoy Dasgupta	108	3286	The word "Probably" should be added, given the lack of strong evidence.	The GPP was adapted to reflect a preference for GnRH agonist trigger.
39	Ahmed Elsayed Hassan Hamed Elbohoty	108	3286	For final oocyte maturation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist triggering is advised. [2025] I suggest revising this statement. In freeze-all protocols—including	The GPP was adapted to reflect a preference for GnRH agonist trigger.

NR	Reviewer	Page	Line	Comment	Action / Reply
				fertility preservation, donor cycles, and PGT—GnRH agonist (GnRHa) trigger should be considered the preferred option, not only for patients at risk of OHSS. Multiple studies have demonstrated that GnRHa trigger yields comparable or superior outcomes in terms of mature oocyte yield, fertilization rates, and blastocyst development, with no adverse effect on embryo euploidy or competence.	
127	ESHRE SIG RE	108	3286	Trigger for final oocyte maturation in case of fertility preservation For final oocyte maturation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist triggering is advised. Justification: not clear to us why GnRH agonist should not be preferred, is it possible to add references?	The GPP was adapted to reflect a preference for GnRH agonist trigger.
140	German Society of Reproductive Medicine	108	3286	For final oocyte maturation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist triggering is advised. [2025] (Recommendation) The literature and study evidence do not support a preference for hCG trigger, as it is clearly stated that the number of mature oocytes is comparable between hCG and GnRH agonist trigger.	The GPP was adapted to reflect a preference for GnRH agonist trigger.
1	Raj Mathur	111	3356	Is it not more accurate to say that 'HCG and GnRH agonist are equally preferred' unless there is a risk of OHSS?	The GPP was adapted to reflect a preference for GnRH agonist trigger.
39	Ahmed Elsayed Hassan Hamed Elbohoty	111	3356	For final oocyte maturation in elective oocyte cryopreservation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist trigger is advised. [2025] Relevant evidence includes: Haas et al., 2019; Babayev et al., 2017; Youssef et al., 2016 (Cochrane Review); Humaidan et al., 2010. However, I acknowledge that GnRHa trigger may be unsuitable in rare clinical contexts such as hypogonadotropic hypogonadism, prolonged use of combined hormonal contraception (CHC), or cases with	The GPP was adapted to reflect a preference for GnRH agonist trigger.

NR	Reviewer	Page	Line	Comment	Action / Reply
				low/undetectable LH levels. In such situations, hCG (or dual trigger) may be more appropriate. Reflecting this in the guideline would promote safer, individualized practice in freeze-all cycles.	
92	Guivarc'h Leveque Anne	111	3356	triggering should be done with agonist each time it's possible (exclusion of LH defiency) As the efficiency of agonist is as good as HCG ,agonist should be preferred as the discomfort is less important with agonist	The GPP was adapted to reflect a preference for GnRH agonist trigger.
127	ESHRE SIG RE	111	3356	Trigger for final oocyte maturation in elective oocyte cryopreservation …hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist trigger is advised. Justification: hCG and GnRH agonist for final oocyte maturation result in similar numbers of mature oocytes. Given the similar number of oocytes, it is not clear why GnRH agonist should not be preferred	The GPP was adapted to reflect a preference for GnRH agonist trigger.
140	German Society of Reproductive Medicine	111	3356	For final oocyte maturation in elective oocyte cryopreservation, hCG is preferred, unless the patient is at risk of early OHSS, in which case GnRH agonist trigger is advised. [2025] (Recommendation) The literature and study evidence do not support a preference for hCG trigger, as it is clearly stated that the number of mature oocytes is comparable between hCG and GnRH agonist trigger.	The GPP was adapted to reflect a preference for GnRH agonist trigger.
11. Ovarian stimulation for oocyte donation					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	20	R78-79	Suggestion：Merge these two recommendations	The recommendations 79 was separately formulated in order to refer to the cases in which GnRH agonist protocol is used in oocyte donation in which only hCG can be used. This is why we state "The use of a hCG trigger is not routinely recommended in oocyte donation cycles" (referring to the exception of the cases

NR	Reviewer	Page	Line	Comment	Action / Reply
					in which agonist protocol is used. The GDG however clearly does not recommend the use of GnRH agonist in oocyte donors "A GnRH agonist protocol is not recommended in oocyte donors. [2025]"
140	German Society of Reproductive Medicine	116	3564	The use of any type of contraception (hormonal, non-hormonal, oral, vaginal or intrauterine) before or during ovarian stimulation is not a contraindication in oocyte donors.[2025] (Recommendation and GPP) While this is true for progestogens only, the continued use of combined hormonal contraception during ovarian stimulation does not appear to be plausible. This needs to be revised.	The GDG agrees with the reviewer. The GPP was adapted and split into two parts.
7	Sujoy Dasgupta	120	3688	".....also obtain an optimal number of oocytes"- There is paucity of data on what is the optimum number of oocytes in such cases.	While the GDG understands that a specific range for the optimal number of oocytes could be helpful, the term "optimal" reflects an individual balance between efficacy and safety. Therefore the GDG refrained from defining the optimal number of oocytes.
12. Hormonal assessment during ovarian stimulation					
98	Willem Verpoest	/	/	I believe it is useful to review evidence of the effect of elevated progesterone in the late follicular phase on implantation and embryo quality; see Venetis CA, Kolibianakis EM, Bosdou JK, Tarlatzis BC. Progesterone elevation and probability of pregnancy after IVF: a systematic review and meta-analysis of over 60 000 cycles. Hum Reprod Update. 2013 Sep-Oct;19(5):433-57. doi: 10.1093/humupd/dmt014. Epub 2013 Jul 4. PMID: 23827986.	This topic is addressed in a separate section of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
147	Jayesh Amin	/	/	In case of GnRH Agonist trigger, LH estimation can be recommended on day 2 of menses, on the day of trigger and 12 hours post trigger ; to predict the chances of sub-optimal response to trigger	This was considered to fall outside the scope of the current guideline update.
142	The Chinese Expert Review Panel for ESHRE OS Guideline	20	R80-81	Do not totally agree with the statement that the addition of oestradiol or a hormonal panel consisting of a combination of oestradiol, progesterone and LH measurements to ultrasound monitoring is probably not recommended. Because 1. It is widely used in clinical practice 2. High responders may prevent OHSS risks by testing for E2 3. Testing LH and P may prevent premature ovulation	There is no doubt that monitoring LH, estradiol, and progesterone levels offers additional information on follicular development and endometrial status, complementing ultrasound, as the reviewer correctly notes. However, this was not the focus of the current guideline question. Rather, the question was whether this additional biochemical information leads to improved safety and efficacy. At present, the available evidence does not support such a benefit.
21	Shikha Gupta	125	3843	Monitoring: addition of Estradiol measurement to USG monitoring is probably not recommended In OHSS risk management-elevated or rapidly rising serum estrogen levels are used as factor of risk OHSS	The key question in this context is not whether serum oestradiol (E2) levels are associated with the risk of OHSS—this association is well established. Rather, the question is whether adding E2 measurements to ultrasound monitoring during ovarian stimulation improves clinical efficacy and safety. In this regard, the recommendation provided in the current guideline is justified. However, it is not definitive, as the number of patients included in the available studies is currently insufficient and the overall quality of evidence is low. Therefore, this recommendation remains subject to revision should future randomized controlled

NR	Reviewer	Page	Line	Comment	Action / Reply
					trials demonstrate a clear benefit of incorporating E2 assessment alongside ultrasound evaluation.
126	Gatagazheva Aza Aslanovna	125	3858	The draft omits any recommendation for measuring serum LH in the follicular phase (e.g. Day 5–6 of stimulation) and provides no actionable algorithms when LH is too low or too high. Key data include: 1. Low mid-stimulation LH (< 4 IU/L) in GnRH-antagonist cycles independently reduces cumulative live birth (OR 0.76; 95 % CI 0.60–0.97; P = 0.014) . 2. A premature LH rise in women ≥ 37 years lowers cumulative live birth and correlates with compromised embryo potential . 3. Dydrogesterone (10 mg BID) in PPOS prevents LH surges without reducing oocyte yield or embryo quality (RCT, n = 516) . 4. My IFFS 2025 case series shows that dose-dependent LH suppression with dydrogesterone can be titrated to an optimal range, improving oocyte competence in high-LH patients (unpublished). Without integrating LH monitoring and modulation (GnRH analogues, progestins, r-LH) tailored to each patient's profile—and striving for physiological LH/FSH ratios—we risk suboptimal oocyte yield, repeat full-cycle stimulations, and unnecessary patient burden. |	Unfortunately, it is not possible to provide a response, as the comment does not include the relevant citations needed to support its claims.
21	Shikha Gupta	125	3867	The addition of a hormonal panel consisting of a combination of oestradiol, progesterone and LH measurements to ultrasound monitoring is probably not recommended. Serum Serum progesterone level elevation on the day of HCG is linked to poor pregnancy rate after ET (Systemic review) Monitoring of progesterone may compliment USG monitoring during ovarian stimulation	This section of the guideline examined whether monitoring ovarian stimulation with both ultrasound and progesterone assessment improves safety and efficacy compared to ultrasound alone. At present, no recommendation can be made due to the absence of relevant clinical trials.
13. Endometrial thickness					

NR	Reviewer	Page	Line	Comment	Action / Reply
91	Willem Verpoest			I believe it is useful to review evidence of the effect of elevated progesterone in the late follicular phase on implantation and embryo quality; see Venetis CA, Kolibianakis EM, Bosdou JK, Tarlatzis BC. Progesterone elevation and probability of pregnancy after IVF: a systematic review and meta-analysis of over 60 000 cycles. Hum Reprod Update. 2013 Sep-Oct;19(5):433-57. doi: 10.1093/humupd/dmt014. Epub 2013 Jul 4. PMID: 23827986.	The published literature on hormonal levels during ovarian stimulation is reviewed in another chapter of the document, and the present chapter focuses on the independent predictive value of EMT. In addition, the study by Griesinger et al. reported that the independent contribution of EMT (assessed on day of embryo transfer) to live birth likelihood is small and may result from (undetermined) confounding factors. If EMT indeed is an independent factor affecting outcome, this finding implies that at a baseline live birth rate of 20% an increase of 2 mm in EMT should result in an increase of the live birth rate of ~1.6% (Griesinger, et al., 2018).
7	Sujoy Dasgupta	129	3958	According to the Canadian Fertility and Andrology Society, in case of endometrial thickness less than 8 mm, fresh embryo transfer should be avoided and elective cryopreservation of all the embryos should be considered. Liu KE, Hartman M, Hartman A. Management of thin endometrium in assisted reproduction: a clinical practice guideline from the Canadian Fertility and Andrology Society. Reprod Biomed Online. 2019 Jul;39(1):49-62.	The recommendations in the Canadian guideline are supported by very low-quality evidence. The GDG doesn't think it is appropriate to formulate such a recommendation on the available evidence. Instead, the GDG formulated a GPP in the 2019 version of the guideline with regards to the importance of counselling, which still stands.
79	Mitranovici Melinda Ildiko	129	3967	How about personalized treatment? In older women under stimulation the implantation window happens earlier and endometrial measurement could be helpful. Further investigation should be necessary?	The GDG would like to draw the reviewers' attention to the GPP the GDG formulated: "the guideline group suggests performing a single measurement of the endometrium during ultrasound assessment on the day of triggering

NR	Reviewer	Page	Line	Comment	Action / Reply
					or oocyte pick-up to counsel patients on potentially lower pregnancy chance".
14. Criteria for final oocyte maturation					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	20	R85	The timing of triggering should be also considered according to the different characteristics of patients, ovarian stimulation protocols, and embryo transfer strategies. For example, women of advanced age or with poor ovarian reserve/response may consider triggering earlier.	The recommendation already acknowledges that the decision on triggering is multifactorial and includes variables that reflect patient characteristics and protocol type, especially the size of the growing follicle cohort, but also duration of stimulation, and prior cycle experience. These factors inherently reflect patient age, ovarian reserve, and stimulation protocol, as their interaction becomes apparent only during the course of stimulation. Therefore, we deliberately refrain from including a priori predictors (e.g. age or AMH) in the recommendation, as these are not absolute determinants of follicular development or functional maturity. Instead, we emphasize factors that are observable and actionable at the time of decision-making, supplemented by overarching considerations such as organizational and financial aspects. This ensures clinical flexibility and individualized care without prescribing specific timing rules for subgroups. However, we acknowledge that embryo transfer strategy (e.g. fresh transfer vs. freeze-all; SET vs. DET; blastocyst vs. cleavage) is

NR	Reviewer	Page	Line	Comment	Action / Reply
					an overarching aspect yet missed and have adopted the recommendation accordingly.
101	José María Regalado Pedrajas	131	4036	I recommend that hormonal testing be excluded from the general criteria for determining the optimal trigger day. This is based on consistent findings across multiple sections indicating that hormonal parameters are not predictive and are therefore not recommended for this purpose. Instead, the general recommendations should prioritize the following criteria: leading follicles size, economic and organisational criteria and the experience of the previous cycles.	For consistency with Chapter 15 ("It is probably recommended to measure serum progesterone levels on the day of final oocyte maturation in cycles aimed for a fresh embryo transfer"), the inclusion of “hormonal data” among the multiple factors informing trigger timing is warranted. Its use in this context is neither prescriptive nor exclusionary, but rather complementary to sonographic assessment. In particular, serum estradiol levels may be considered in the context of OHSS risk assessment in high responders in selected cases, and to support physiological coherence, i.e. assessing the functional maturity of the follicular cohort—especially in cases with poor response or mono-follicular development. Therefore, a flexible, individualized approach that allows for the use of hormonal data in selected scenarios appears more appropriate than its categorical exclusion.
101	José María Regalado Pedrajas	131	4036 4047	I find some contradictions between the general recommendations and the specific management of hormonal analysis for determining the trigger’s day.	For consistency with Chapter 15 ("It is probably recommended to measure serum progesterone levels on the day of final oocyte maturation in cycles aimed for a fresh embryo transfer"), the inclusion of “hormonal data” among the multiple factors informing trigger timing is

NR	Reviewer	Page	Line	Comment	Action / Reply
					warranted. Its use in this context is neither prescriptive nor exclusionary, but rather complementary to sonographic assessment. In particular, serum estradiol levels may be considered in the context of OHSS risk assessment in high responders in selected cases, and to support physiological coherence, i.e. assessing the functional maturity of the follicular cohort—especially in cases with poor response or mono-follicular development. Therefore, a flexible, individualized approach that allows for the use of hormonal data in selected scenarios appears more appropriate than its categorical exclusion.
15. Hormonal assessment on the day of final oocyte maturation					
88	Aboubakr Mohamed Elnashar	21	R89	What is definition of high progesterone level on day of final oocyte maturation?	While different serum progesterone levels have been investigated in the available studies, the negative effect has been observed when serum progesterone level was >0.8 ng/ml on the day of trigger. The available studies suggest a dose response relationship, i.e., the higher the progesterone level the lower the chance of ongoing pregnancy and live birth rates. However, the decision to continue with fresh transfer or cancelling it depends on other factors, precluding a recommendation for cancellation of fresh transfer at a particular progesterone level.

NR	Reviewer	Page	Line	Comment	Action / Reply
127	ESHRE SIG RE	135	4112	It is probably recommended to measure serum progesterone levels on the day of final oocyte maturation in cycles aimed for a fresh embryo transfer. [2025] Should the GDG provide more practical guidance on how to use/ incorporate these measurements in the clinical management of patients? Please consider the recently published IPD meta-analysis by the INFORM network suggesting that the serum P4 cut-off of 1.2 ng/mL should be used to potentially convert to a freeze-all.	The INFORM IPD meta-analysis has been published only as an abstract so it cannot be used to inform the guideline at the moment. Moreover, negative effect has been observed when serum progesterone level was >0.8 ng/ml on the day of trigger.
39	Ahmed Elsayed Hassan Hamed Elbohoty	135	4113	If serum progesterone levels are high, the patient should be counselled about potentially lower ongoing pregnancy/live birth rates. The decision to defer embryo transfer should include other factors (number of oocytes, number of embryos, and embryo quality). [2025] I suggest that the guideline include a specific serum progesterone threshold—>1.5 ng/mL—to aid in clinical decision-making regarding freeze-all strategies. This value is consistently associated with reduced implantation and live birth rates in fresh embryo transfer cycles, as demonstrated in multiple high-quality studies and meta-analyses. Specifying this cutoff would enhance clarity and support real-time counseling and treatment planning. (Bosch et al., 2010; Venetis et al., 2013)	The available studies suggest a dose response relationship, i.e., the higher the progesterone level the lower the chance of ongoing pregnancy and live birth rates. It is not possible to mention a single threshold, particularly 1.5 ng/ml as similar effect has also been observed with lower serum progesterone levels on the day of trigger. While different serum progesterone levels have been investiagetd in the available studies, the negative effect has been observed when serum progesterone level was >0.8 ng/ml on the day of trigger.
126	Emre Goksan Pabuccu	137	4203	It is not recommended to measure serum LH levels on the day of HCG trigger in ovarian stimulation cycles aimed for a fresh embryo transferà“It is not recommended to measure serum LH levels on the day of HCG trigger in ovarian stimulation cycles aimed for a fresh embryo transfer. However, if a GnRH agonist trigger is planned followed by a fresh	As mentioned in the following question, serum LH levels on the day of trigger are not discriminatory for response to a GnRH agonist. Patients with certain characteristics, which render them at risk of inadequate response can be identified at the start of stimulation cycle.

NR	Reviewer	Page	Line	Comment	Action / Reply
				transfer, caution is warranted due to the risk of empty follicle syndrome.9	The question does not address LH levels the day after a GnRH agonist trigger.
16. Criteria for cycle cancellation					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	21	R95	Do not agree with the strength of this recommendation is strong. Because 1. The risk of OHSS can be reduced by other methods, such as reducing hCG dosage or freezing all embryos. 2. It would be a great psychological and economic burden on patients, who are actually not a small proportion, if cancelling final oocyte maturation.	The recommendation is to use GnRH antagonist in high responders. This avoids the need to cancel the cycle due to high response.
144	Karolina Palinska-Rudzka	21	R95	The draft currently recommends that in antagonist cycles with ≥19 follicles ≥11 mm, cancellation of final oocyte maturation should be "primarily" considered. In practice, cancellation is typically a last resort. A high follicle number may include many small follicles (<14 mm), which individually pose lower OHSS risk. The use of an appropriate preventative strategies often allows safe continuation. Evidence supports the efficacy of alternative preventive measures: • Cabergoline: Randomised controlled trials (Alvarez et al., 2007; Papanikolaou et al., 2009) and a Cochrane review (Tang et al., 2021) demonstrated that dopamine agonists reduce the incidence of moderate/severe OHSS [5–7]. • Elective segmentation (freeze-all): A widely adopted measure to mitigate the risk of late OHSS. These strategies are already established in practice and form part of the BFS OHSS guideline. Given this, it may be helpful for the guideline to highlight these approaches as the means of managing (unexpected) high responders undergoing agonist cycle, with cancellation reserved for	Thank you for your comments. In fact all the steps could be undertaken for anatgonist cycles but in this section the guideline refers to situations in which agonist cycle would be applied.

NR	Reviewer	Page	Line	Comment	Action / Reply
				selected situations rather than applied routinely to all with ≥19 follicles ≥11 mm .	
39	Ahmed Elsayed Hassan Hamed Elbohoty	142	4387	A low response to ovarian stimulation alone is not a reason to cancel a cycle. [2019] I support the recommendation that a low follicular response should not automatically result in cycle cancellation. However, this section could be enhanced by including tailored strategies based on follicular cohort dynamics, especially in poor responders. Consider the following management pathways: § Presence of one or 2 dominant follicles with several small antral follicles (<10 mm): It may be clinically beneficial to trigger with a GnRH agonist (GnRHa) without proceeding to egg retrieval, and initiate luteal phase stimulation (LPS) 5–7 days later to rescue the second cohort aiming to get more oocytes.(Kuang Y et al., Reprod Biol Endocrinol. 2014;12:108) § If only a few large follicles are present and no visible second cohort: Proceeding with oocyte retrieval may still be worthwhile to avoid missing a valuable opportunity, especially in older patients or those with limited reserve. (Vaiarelli A et al., Front Endocrinol (Lausanne). 2018;9:317) § We may proceed also with trigger and oocyte retrieval then luteal phase stimulation and another trigger (Double OPU for embryo pooling). This approach may be particularly helpful for POSEIDON Group IV patients (>35 years, low AFC/AMH).(Massarotti C et al., Ther Adv Reprod Health. 2020;14:2633494120971535; Ubaldi FM et al., Fertil Steril. 2016;105(6):1488–1495).	Thank you for your comment. All options are possible but in this section we focused only on cycle cancelation criteria. For an unexpected low responder, the GDG recommends the physician to counsel patients individually regarding pregnancy prospects and the decision to continue this cycle. The exact description of possibilities are besides this section could be taken into account in the future
17. Triggering of final oocyte maturation					

NR	Reviewer	Page	Line	Comment	Action / Reply
137	Jayesh Amin	22	103	The addition of a GnRH agonist to hCG as a dual trigger for final oocyte maturation can be recommended for low responders for optimum yield of oocytes.	The GDG sees your concern. However, it does not seem to be the case according to the available evidence.
91	Willem Verpoest	22	Rec/	Following rec 102 and 103, this is an unexpected statement; if dual trigger is probably not recommended, rec / in the next line should be deleted	This was adjusted, it was added that this was a conclusion.
132	The Chinese Expert Review Panel for ESHRE OS Guideline	22	R101-103, 117	Do not totally agree with the statement that the dual trigger for final oocyte maturation is probably not recommended for normal/low/high responders. Because According to the recommendation, there will be no suitable population for dual trigger.	The GDG sees your concern. However, it seems to be the case according to the available evidence.
48	Isabel De Almeida	147	4544	I'd like to know about how to handle the empty follicle syndrome in the next ovarian stimulation. I mean, which is the best protocol in this case	The GDG sees your concern. However, the available evidence is limited in order to provide solid conclusions
39	Ahmed Elsayed Hassan Hamed Elbohoty	148	4560	If the GnRH agonist trigger with triptorelin is applied, dosages ranging of 0.1-0.4 mg can be chosen. I suggest adding a clarification to guide clinicians in selecting the appropriate triptorelin dose based on clinical context. For example, 0.3–0.4 mg may be considered in patients with high BMI, hypothalamic suppression (e.g., prolonged CHC use), or a history of suboptimal LH response. In such cases, clinicians should also consider a dual trigger (GnRH agonist + low-dose hCG), or monitor serum LH levels ~12 hours post-trigger to confirm an adequate surge and reduce the risk of failed maturation (Humaidan et al., 2005; Kolibianakis et al., 2012; Youssef et al., 2016).	The GDG sees your concern. However, the available evidence is limited by its observational status to provide solid conclusions
17	Raoul Orvieto	149	4582	"concept of dual and dual trigger" Should be "concept of dual and double trigger"	Thank you for pointing this out, this was corrected.

NR	Reviewer	Page	Line	Comment	Action / Reply
17	Raoul Orvieto	149	4584	For better clarification, please add “In clinical practice, a Dual trigger is generally used to improve outcomes in predicted normal responders, while a Double trigger is typically reserved for patients with abnormal final follicular maturation (Orvieto R. Triggering final follicular maturation for IVF cycles. Reprod Biol Endocrinol. 2025;23(Suppl 1):12.).”	Thank you for comment. The text was adapted.
17	Raoul Orvieto	149	4588	Beebeejaun et al meta-analysis suffers from major methodological flaws. Inclusion of methodologically inconsistent studies, lack of adherence to established definitions and misclassification of studies for comparison that have led to inaccurate conclusions. Letter to the Editor was accepted for publication in F&S	The GDG sees your concern.Thank you for pointing this out
17	Raoul Orvieto	149	4595	Does the sample size is sufficient??? For LBR comparison	The GDG sees your concern.Thank you for pointing this out
17	Raoul Orvieto	149	4598	In Singh et al study: " the number of MII oocytes (7.82 vs. 5.92, p=0.003) and day-3 grade-1 embryos (4.24 vs. 1.8, p<0.001) and consequently, number of embryos cryopreserved (2.68 vs. 0.94, p<0.001) were significantly higher in the dual trigger group". Therefore, Cumulative PR would be probably higher. Moreover:" clinical pregnancy rates between the two groups (21% vs. 19.6%, p=0.770) were comparable". Do these figures are acceptable? or indicate a low quality outcome/programe?	The GDG sees your concern.Thank you for pointing this out
17	Raoul Orvieto	149	4606	Published in a very low quality Journal IF 0.519 Q3	The GDG sees your concern.Thank you for pointing this out
17	Raoul Orvieto	149	4611-2	Again, does the sample size is sufficient? More than 50% increase	The GDG sees your concern.Thank you for pointing this out
17	Raoul Orvieto	150	4619	Keskin et al- very problematic study. The patients in the Dual were Posiedon group 4, while in the hCG- group 3 (Older, lower AMH and more previous IVF attempts) Moreover, 36.3 and 39.2% LBR pre OPU or ET, respectively, in Poseidon	Thank you for comment. The text was adapted.

NR	Reviewer	Page	Line	Comment	Action / Reply
				group 3/4 that were triggered with hCG (control group). These figures are too good to be true!!!	
1	Raj Mathur	149	4622	The GDG provides a good overview of the evidence on dual trigger in normal responders. It seems to me that this suggests that dual trigger can be 'equally recommended' to HCG trigger in normal responders.	The GDG sees your concern. However, it seems to be the case according to the available evidence.
17	Raoul Orvieto	150	4622	Please refer to "Orvieto R. Triggering final follicular maturation for IVF cycles. Reprod Biol Endocrinol. 2025;23(Suppl 1):12."	Only systematic reviews with meta-analysis were considered for inclusion in the evidence section.
134	Nayana Patel	150	4622-4623	Adding a GnRH agonist dual trigger should be considered for low responders, as well as certain normal responders with a diachronous cohort on the day of trigger.	The GDG sees your concern. However, it seems to be the case according to the available evidence.
37	Ahmed Samy Abdelazim Saad	150	4623	Better to be equally recommended or may be used in cases of failure after failed ovum pickup in a previous cycle in the form of low retrieved numbers or immature oocytes.	The GDG sees your concern. However, it does not seem to be the case according to the available evidence.
18. Luteal phase support					
100	Reassure group	/	/	Chapter 18 – Luteal Phase Support raises significant concerns regarding methodological consistency and impartiality, particularly in its treatment of dydrogesterone relative to other progestins. Dydrogesterone—supported by multiple RCTs, IPD meta-analyses, and the largest evidence base—receives only a conditional recommendation, despite comparable efficacy and comparable safety in clinical data. Moreover, pharmacovigilance (PV) data are selectively applied to dydrogesterone, without similar scrutiny of other agents, and disproportionality-based, hypothesis-generating PV signals are prioritized over higher-quality RCT data, contradicting ESHRE's methodological standards. The overall tone downplays positive evidence on dydrogesterone while emphasizing low-certainty or speculative findings, resulting in a biased and unbalanced representation.	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stands with the 2025 update of the guideline. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in pharmacovigilance, teratology, and perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	
100	Reassure group	/	/	A review of the full draft guideline and annexed literature search reveals that pharmacovigilance (PV) data were neither included in the search strategy described in the methods section nor systematically reported for any intervention other than dydrogesterone. The sole use of a single PV study (Henry et al.) in the context of dydrogesterone, without comparable evaluation for other agents used in ART, suggests that this reference was selectively introduced rather than identified through a predefined, systematic approach. This selective inclusion carries negative implications for the methodological transparency, consistency, and neutrality of the guideline and may undermine confidence in how safety signals are evaluated across interventions. Suggested revision: Provide a comprehensive and systematic literature search and safety evaluation of all drugs and interventions assessed.	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.
100	Reassure group	/	/	A major concern with the current guideline draft is the selective inclusion of evidence regarding fetal safety. The DEBC cohort, conducted in a general obstetric population, is the only observational dataset cited in support of concerns about dydrogesterone. However, other relevant evidence from non-ART populations is systematically omitted, particularly data suggesting a favorable safety profile. For example, the guideline does not reference a recent network meta-analysis (DOI: 10.1002/ijgo.15987), which found that oral dydrogesterone had the highest probability of being the safest intervention in terms of congenital	The suggested meta-analysis was published after the final literature search for the guideline. All studies reporting on safety of dydrogesterone, included in the literature search, were reviewed and included in the justification section of the guideline. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				anomaly risk. The SUCRA ranking indicated oral dydrogesterone (82%) outperformed oral progesterone (67%), placebo or no treatment (47%), and vaginal progesterone (4%). These findings should be presented and balanced against other evidence. Omitting positive safety data from non-ART studies while emphasizing a single cohort which is likewise not an ART study (i.e. the DEBC study in which dosages predominantly used are not LPS dosages) with methodological limitations introduces bias and undermines the evidence balance expected in guideline development. Suggested revision: For the guideline to remain balanced, neutral, and methodologically consistent, safety data from all available sources and for all relevant interventions must be systematically searched, transparently evaluated, and consistently integrated. This includes data from pharmacovigilance databases, observational cohorts, randomized trials, and meta-analyses—both favorable and unfavorable. Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in pharmacovigilance, teratology, and perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	
104	Matthias Mueller*	/	/	The recommendations for luteal phase support are clearly divided into a section for progesterone (all different routes of administration) and dydrogesterone. However, the evidence used as backbone for the validity of progesterone use in luteal phase support is the Cochrane meta-analysis from 2015 (van der Linden et al.) whereas 2 of the used 5 RCTs have been conducted with dydrogesterone. The dydrogesterone treatment consequently contributed to the conclusion of the Cochrane	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands. Had the reviewer provided references of more recent studies comparing progesterone to placebo, the GDG could have reviewed them.

NR	Reviewer	Page	Line	Comment	Action / Reply
				meta-analysis which is valid for all progestins. In the study selection used to evaluate the efficacy and safety of the different treatment regimens the most recent study is from 1996. Cochrane meta-analysis also remains unchanged since 2015. For example, the Cochrane meta-analysis is still including studies using Hydroxyprogesterone-acetate that has been recommended for suspension by EMA in 2024) (Details on suggested studies included in specific comments). Given the significant evolvement of ART and especially luteal phase support protocols, more recent studies should be taken into consideration to reflect the actual clinical practice and time-relevant recommendations given.	
104	Matthias Mueller*	/	/	As per the guideline group explanation, each individually concluded recommendation is based on efficacy and safety evaluation. In the area of reproductive medicine, the safety aspect can be divided into two: the maternal and the fetal safety. Apart from the section on progestogen use in LPS, there is very little focus on the aspect of fetal safety after certain treatment interventions throughout the guideline due to the complexity and the ethical difficulty of its assessment in high-quality trials (Katalinic et al. 2024). It is important to give perspective on the overall context and data available at the time. In the late 1960-1970s all sex steroids were suspected to increase the evidence of congenital malformations which led to an FDA warning label in pregestational and contraceptive drugs. However, in 1999 the FDA cleared this warning from all packages, and later, in 2005, through a comprehensive review by Brent et al. 2005,(Nongenital malformations following exposure to pregestational drugs: The last chapter of an erroneous allegation - Brent - 2005 - Birth Defects Research Part A: Clinical and Molecular Teratology - Wiley Online Library)	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				it has been concluded that sex steroids taken during pregnancy do not increase the incidence of congenital malformations.	
104	Matthias Mueller*	/	/	For the safety aspect, evidence focusing on maternal and fetal safety must be assessed equally to formulate a recommendation for treatment. Especially in the first trimester treatment, where progestogens are used to safeguard pregnancy, the occurrence of pregnancy complications (vaginal bleeding, HDP, infections, nausea etc.) potentially impacting the live birth or pregnancy progression are equally important. Maternal safety is equally critical, as it significantly influences pregnancy progression and live birth outcomes. In the context of assisted reproductive technologies (ART), it also plays a pivotal role in determining patient retention and the continuation of treatment cycles.	The GDG agrees with the reviewer.
104	Matthias Mueller*	/	/	Chapter 18 presents notable concerns regarding the consistency of its methodology, the weighting of evidence, and the impartiality of its conclusions—particularly in its comparative treatment of dydrogesterone and other progestins. Unsubstantiated Endorsement of Vaginal Progesterone: Vaginal progesterone monotherapy is granted a strong recommendation based on low-certainty evidence due to absence of any placebo-controlled randomized trials supporting its use for luteal phase support (LPS). In contrast, dydrogesterone—backed by the highest level of clinical evidence including multiple RCTs, meta-analyses, and individual participant data (IPD) studies—receives only a conditional recommendation. This is despite high-quality data showing dydrogesterone's non-inferiority to micronized vaginal progesterone (MVP) in terms of both ongoing pregnancy and live birth rate (Tournaye et al. 2017, Griesinger et al. 2018) Notably, an individual participant data (IPD) meta-analysis further established dydrogesterone's superiority over MVP in achieving these outcomes, while maintaining a comparable safety profile (Griesinger et al., 2020). When it comes to long-term safety, expert consensus acknowledges that current evidence is limited. However, the only systematic review conducted on this topic	The GDG has not formulated a recommendation specifically in favour of vaginal progesterone. Furthermore, the incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a potential increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				(Katalinic et al., 2025) reported a 10% lower incidence of fetal malformations with dydrogesterone compared to MVP—though this difference did not reach statistical significance. Taken together, the available data on both efficacy and safety support a balanced and equally strong recommendation for the use of either progesterone or dydrogesterone in clinical practice. Uneven Application of Pharmacovigilance Data: The chapter includes pharmacovigilance research only to dydrogesterone, while other agents have not been studies explicitly for safety. This selective focus lacks justification within an evidence-based framework. PV signals, which are exploratory and not designed to establish causality, should not be treated as equivalent to robust RCT findings. Deviation from Evidence-Based Standards: The selective use of disproportionality-based PV analyses on dydrogesterone, despite the availability of superior RCT data, represents a departure from both the GRADE approach and ESHRE's own methodological principles. While other sections of the guideline appropriately prioritize high-level evidence and exclude lower-tier studies when stronger data are available, this standard is inconsistently applied in Chapter 18. Skewed Narrative and Omission of Key Findings: The chapter minimizes or omits favorable data on dydrogesterone—such as lower malformation rates, better patient tolerability, and higher patient preference—while placing undue emphasis on less definitive or negative findings. This imbalance raises concerns about the objectivity and fairness of the presentation.	
109	Amr Abdel Aziz Nadim	/	/	In the draft prepared by the European Society of Human Reproduction and Embryology (ESHRE) as a 2025 update of the guideline Ovarian Stimulation for IVF/ICSI , the section on luteal phase support includes a safety footnote regarding dydrogesterone that deviates from the guideline's stated methodology and evidence-based medicine (EBM) principles. This critique outlines inconsistencies in the guideline's approach to evaluating dydrogesterone safety, highlights methodological concerns.	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic

NR	Reviewer	Page	Line	Comment	Action / Reply
				1. Inconsistent Application of Evidence Hierarchy The guideline's methodology (page 179, line 5419) describes an iterative literature search prioritizing systematic reviews and meta-analyses, followed by randomized controlled trials (RCTs), cohort studies, and case reports, in accordance with the EBM hierarchy. However, the safety conclusion for dydrogesterone relies heavily on a pharmacovigilance study (Henry et al.), which is not considered clinical evidence in the EBM framework, as reporting odds ratios do not constitute clinical outcomes. This study is given disproportionate weight compared to higher-quality clinical trials and meta-analyses, undermining the guideline's methodological rigor. § Recommendation: The Guideline Development Group (GDG) should consistently apply the EBM hierarchy by prioritizing high-quality clinical evidence (e.g., RCTs and meta-analyses) over pharmacovigilance studies, which are hypothesis-generating and cannot establish causality. A transparent justification for deviations from the stated methodology should be provided. 2. Methodological Flaws in Cited Pharmacovigilance Study The safety footnote for dydrogesterone references pharmacovigilance reports, specifically citing Henry et al. This study is limited by exposure, temporal, and selection biases, as well as underreporting of birth defects in the VigiBase database. Pharmacovigilance studies are inherently designed for signal detection, not for establishing causal relationships, yet the guideline emphasizes these findings without sufficient context. § Recommendation: Clarify the limitations of pharmacovigilance studies in the guideline and avoid speculative language that may amplify concerns without robust evidence. The Guideline Development Group should ensure that safety conclusions are grounded in high-quality clinical data rather than hypothesis-generating studies. 3. Omission of Relevant Safety Studies The guideline acknowledges a lack of long-term offspring health studies for dydrogesterone but omits existing studies questioning the safety of comparator progestins (e.g., Carmichael et al., 2005). Additionally, studies reporting lower	systematic review, and where the EBM-pyramid does not apply.

NR	Reviewer	Page	Line	Comment	Action / Reply
				tolerability of progesterone in pregnancy (e.g., Bespalova et al., 2021; Astrankantseva et al., 2021; Zhao et al., 2025) and adverse events in progesterone-treated women are not addressed. This selective inclusion creates an unbalanced representation of safety data across progestins. § Recommendation: Conduct a comprehensive review of all relevant safety studies for dydrogesterone and comparator progestins to ensure equitable evaluation. The guideline should transparently report both positive and negative findings to provide a balanced perspective. 4. Inconsistent Safety Footnotes Across Treatment Recommendations Dydrogesterone is the only progestin—and the only drug in the guideline—singled out with a safety footnote, potentially heightening concern without adequate context or consensus. Although the guideline states that no consensus was reached, yet the inclusion of the safety footnote risks misleading clinicians by implying a unique safety concern not applied to other treatments. § Recommendation: Apply consistent criteria for safety statements across all treatments to avoid bias. Remove or revise the dydrogesterone-specific footnote unless supported by robust, consensus-driven evidence, ensuring alignment with the guideline's purpose of providing clear, evidence-based guidance. 5. Established Safety Profile of Dydrogesterone Dydrogesterone has a well-documented safety profile, supported by 65 years of use by more than 140 million women, including over 20 million in the first trimester of pregnancy. The two existing studies on fetal safety after dydrogesterone use are included in the guideline but are overshadowed by the pharmacovigilance study's speculative conclusions. Recent pharmacovigilance data do not outweigh the robust clinical evidence supporting dydrogesterone's safety and efficacy. § Recommendation: Acknowledge the extensive clinical evidence and long-term use supporting dydrogesterone's safety in the guideline. Ensure that conclusions reflect the weight of this evidence rather than emphasizing methodologically weaker studies. 6. The Need for Clear and Reliable Clinical Guidance	

NR	Reviewer	Page	Line	Comment	Action / Reply
				The guideline's speculative language and inconsistent methodology in this section risk undermining its purpose of providing clear, evidence-based guidance for physicians in daily practice. The safety footnote for dydrogesterone may create uncertainty, potentially limiting access to a trusted and effective option for luteal phase support. § Recommendation: Revise the guideline to provide clear, reliable, and consistent guidance that supports clinicians' decision-making. Please avoid speculative language and ensure that safety statements are based on high-quality evidence to maintain trust in dydrogesterone as a viable treatment option. Conclusion The ESHRE 2025 Update guideline's safety footnote for dydrogesterone deviates from the guideline's stated methodology, overemphasizes a methodologically flawed pharmacovigilance study, and inconsistently applies evidence standards compared to other treatments. To enhance credibility and utility, the Guideline Development Group should: 1.Consider removing the footnote entirely unless supported by consensus-driven, high-quality evidence. 2. Adhere to the EBM hierarchy by prioritizing high-quality clinical evidence. 3. Transparently address the limitations of pharmacovigilance studies and Correct inaccuracies in referencing pharmacovigilance data. 4. Comprehensively review safety data for all progestins and apply consistent criteria for safety statements across different treatment modalities. This would mitigate unintended concern, and uphold the guideline's purpose of providing clear, reliable guidance for clinical practice. 5. Acknowledge dydrogesterone's established safety profile. By addressing these issues, the ESHRE guideline can better serve as a reliable resource for physicians, ensuring continued access to effective and trusted options for luteal phase support.	

NR	Reviewer	Page	Line	Comment	Action / Reply
140	German Society of Reproductive Medicine	/	/	Chapter 18 – Luteal Phase Support raises significant concerns regarding methodological consistency and impartiality, particularly in its treatment of dydrogesterone relative to other progestins. Dydrogesterone—supported by multiple RCTs, IPD meta-analyses, and the largest evidence base—receives only a conditional recommendation, despite comparable efficacy and comparable safety in clinical data. Moreover, pharmacovigilance (PV) data are selectively applied to dydrogesterone, without similar scrutiny of other agents, and disproportionality-based, hypothesis-generating PV signals are prioritized over higher-quality RCT data, contradicting ESHRE's methodological standards. The overall tone downplays positive evidence on dydrogesterone while emphasizing low-certainty or speculative findings, resulting in a biased and unbalanced representation. Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in pharmacovigilance, teratology, and perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic systematic review.
140	German Society of Reproductive Medicine	/	/	The current guideline cites pharmacovigilance and observational safety data exclusively in relation to dydrogesterone, but omits comparable data on natural progesterone, despite its widespread use in ART. This omission undermines the neutrality, balance, and methodological consistency expected of a clinical guideline. Notably, several studies from non-dydrogesterone settings—particularly in the U.S., where dydrogesterone is not available—have raised concerns about congenital anomalies potentially associated with natural progesterone exposure in ART contexts. For example:	Only studies published between 31 October 2018 and 2 February 2025 were considered for inclusion in the guideline update. Both suggested studies included al PROGESTINS, which also includes dydrogesterone, and because no specific compound or route is identified, mentioning of these studies in the justification would be non-actionable.

NR	Reviewer	Page	Line	Comment	Action / Reply
				· Carmichael et al. (2005) reported an association between maternal progestin intake and hypospadias in a large U.S. population-based study (Arch Pediatr Adolesc Med. 159(10):957–962). While no specific compound or route was identified, dydrogesterone was not among the available treatments. · Silver et al. (1999) observed an increased risk of hypospadias in IVF pregnancies, noting that the only recognized difference between cases and controls was maternal progesterone administration. Again, this observation occurred in a context where dydrogesterone was not in use. By excluding such data, the guideline creates the impression that dydrogesterone uniquely raises safety concerns, while other progestins are implicitly regarded as safer. This asymmetry in evidence presentation may mislead clinicians and does not reflect the available literature. Moreover, if the Henry et al. pharmacovigilance study is considered sufficient to support a conditional recommendation against dydrogesterone, then—to maintain methodological parity—similar standards should be applied to progesterone for which comparable safety concerns exist. Failure to do so introduces a non-bipartisan application of evidence standards, which compromises both the internal consistency and the perceived impartiality of the guideline.	
140	German Society of Reproductive Medicine	/	/	A review of the full draft guideline and annexed literature search reveals that pharmacovigilance (PV) data were neither included in the search strategy described in the methods section nor systematically reported for any intervention other than dydrogest dydrogesterone. The sole use of a single PV study (Henry et al.) in the context of dydrogesterone, without comparable evaluation for other agents used in ART, suggests that this reference was selectively introduced rather than identified through a predefined, systematic approach. This selective inclusion carries negative	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				implications for the methodological transparency, consistency, and neutrality of the guideline and may undermine confidence in how safety signals are evaluated across interventions. Suggested revision: Provide a comprehensive and systematic literature search and safety evaluation of all drugs and interventions assessed.	
140	German Society of Reproductive Medicine	/	/	A major concern with the current guideline draft is the selective inclusion of evidence regarding fetal safety. The DEBC cohort, conducted in a general obstetric population, is the only observational dataset cited in support of concerns about dydrogesterone. However, other relevant evidence from non-ART populations is systematically omitted, particularly data suggesting a favorable safety profile. For example, the guideline does not reference a recent network meta-analysis (DOI: 10.1002/ijgo.15987), which found that oral dydrogesterone had the highest probability of being the safest intervention in terms of congenital anomaly risk. The SUCRA ranking indicated oral dydrogesterone (82%) outperformed oral progesterone (67%), placebo or no treatment (47%), and vaginal progesterone (4%). These findings should be presented and balanced against other evidence. Omitting positive safety data from non-ART studies while emphasizing a single cohort which is likewise not an ART study (i.e. the DEBC study in which dosages predominantly used are not LPS dosages) with methodological limitations introduces bias and undermines the evidence balance expected in guideline development. Suggested revision: For the guideline to remain balanced, neutral, and methodologically consistent, safety data from all available sources and for all relevant interventions must be systematically searched, transparently evaluated, and consistently integrated. This includes data from pharmacovigilance databases, observational cohorts, randomized trials, and meta-analyses—both favorable and unfavorable. Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in pharmacovigilance, teratology, and	The suggested meta-analysis was published after the final literature search for the guideline. All studies reporting on safety of dydrogesterone, included in the literature search, were reviewed and included in the justification section of the guideline. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	
142	Sonia Malik	/	/	Pharmacovigilance study and review should not be taken as evidence for the competence of the drug or its safety. Many studies taken for dydrogesterone are methodologically flawed or very old and should not have been considered in this data.	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
151	Alexandra Kohl Schwartz	/	/	Chapter 18 – Luteal Phase Support raises significant concerns regarding methodological consistency and impartiality, particularly in its treatment of dydrogesterone relative to other progestins. Dydrogesterone—supported by multiple RCTs, IPD meta-analyses, and the largest evidence base—receives only a conditional recommendation, despite comparable efficacy and comparable safety in clinical data. Moreover, pharmacovigilance (PV) data are selectively applied to dydrogesterone, without similar scrutiny of other agents, and disproportionality-based, hypothesis-generating PV signals are prioritized over higher-quality RCT data, contradicting ESHRE's methodological standards. The overall tone downplays positive evidence on dydrogesterone while emphasizing low-certainty or speculative findings, resulting in a biased and unbalanced representation. Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic systematic review.

NR	Reviewer	Page	Line	Comment	Action / Reply
				pharmacovigilance, teratology, and perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	
151	Alexandra Kohl Schwartz	/	/	The inclusion of pharmacovigilance remarks only for dydrogesterone, while omitting similar reporting for other agents used in ovarian stimulation and IVF, introduces a clear inconsistency and bias. Other ART drugs have PV data, often more substantial (https://oehha.ca.gov/proposition-65/crnr/notice-intent-list-chemical-formally-required-be-labeled-or-identified-mechanism-clomiphene-citrate), for example on clomifen the recent studies in humans DOIs: 10.1080/14740338.2024.2358972, 10.1158/1055-9965.EPI-16-0809 , 10.1158/1055-9965.EPI-16-0880, 10.1186/s13048-022-01084-z, 10.1155/2018/7191704 and animals 10.1210/endocr/bqae047. Without systematic consideration of safety signals across all options, this selective emphasis may misrepresent relative risks within the ART field overall and undermine the guideline's neutrality. This selective inclusion introduces a notable imbalance in how safety is presented across treatment options and may inadvertently stigmatize dydrogesterone in the absence of consistent or stronger evidence compared to alternatives. Such asymmetry risks undermining the scientific neutrality and credibility of the guideline. Suggested revision: (a) provide a comprehensive, structured and balanced overview of available maternal and fetal safety data for all relevant drugs and interventions in the context of OS and ART; or (b) omit pharmacovigilance remarks altogether unless they are based on systematic and consistent evidence review with clear implications for clinical decision-making. This aligns with the methodological standards of GRADE, GIN, and the	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Institute of Medicine, which emphasize consistency, transparency, and fair comparison across all alternatives.	
151	Alexandra Kohl Schwartz	/	/	The current sentence wrongly implies that safety concerns apply uniquely or with special emphasis to dydrogesterone, rather than to all substances used in early pregnancy and ART. This selective negative framing is methodologically inappropriate and inconsistent with basic principles of reproductive pharmacology and teratology, where fetal safety is universally relevant, regardless of whether the drug is synthetic, natural, oral, or vaginal. Moreover, comparator drugs (e.g., micronised progesterone) exhibit strong variation in intrauterine exposure due to differences in route of administration. Vaginal progesterone may deliver high local endometrial concentrations with extremely high early fetal exposure, while IM or oral routes result in greater systemic exposure and impact on placenta/placental transfer. Importantly, natural progesterone is chemically unstable, rapidly metabolized, and yields a variety of metabolites—some of which have been shown to exhibit anti-mitotic or anti-proliferative activity (e.g., 5β-dihydroprogesterone) in vitro and in animal studies (e.g. doi.org/10.1186/s41936-021-00212-3, doi.org/10.1038/s41598-020-78976-x). These aspects are not accounted for in the current guideline but are essential for interpreting potential class effects or differences across preparations. Suggested revision: Avoid negative framing. As is, the statement is problematic for two reasons: 1. It implies that being orally active and structurally different uniquely triggers safety concerns. 2. It lacks comparative context — safety is relevant for all progestins in early pregnancy, not just dydrogesterone. Suggested sentence: Dydrogesterone is a retrosteroid with distinct	Had a studies reporting concerns been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline. It has been shown that dydrogesterone has differt binding properties compared to natural progesterone. This is one of the hypothesis underlying the mechanism of congenital malformations with dydrogesterone.

NR	Reviewer	Page	Line	Comment	Action / Reply
				stereochemical and structural features compared to natural progesterone, resulting in a unique metabolic and pharmacokinetic profile.	
151	Alexandra Kohl Schwartz	/	/	The current guideline cites pharmacovigilance and observational safety data exclusively in relation to dydrogesterone, but omits comparable data on natural progesterone, despite its widespread use in ART. This omission undermines the neutrality, balance, and methodological consistency expected of a clinical guideline. Notably, several studies from non-dydrogesterone settings—particularly in the U.S., where dydrogesterone is not available—have raised concerns about congenital anomalies potentially associated with natural progesterone exposure in ART contexts. For example: · Carmichael et al. (2005) reported an association between maternal progestin intake and hypospadias in a large U.S. population-based study (Arch Pediatr Adolesc Med. 159(10):957–962). While no specific compound or route was identified, dydrogesterone was not among the available treatments. · Silver et al. (1999) observed an increased risk of hypospadias in IVF pregnancies, noting that the only recognized difference between cases and controls was maternal progesterone administration. Again, this observation occurred in a context where dydrogesterone was not in use. By excluding such data, the guideline creates the impression that dydrogesterone uniquely raises safety concerns, while other progestins are implicitly regarded as safer. This asymmetry in evidence presentation may mislead clinicians and does not reflect the available literature. Moreover, if the Henry et al. pharmacovigilance study is considered sufficient to support a conditional recommendation against dydrogesterone, then—to maintain methodological parity—similar	Only studies published between 31 October 2018 and 2 February 2025 were considered for inclusion in the guideline update. Both suggested studies included all PROGESTINS, which also includes dydrogesterone, and because no specific compound or route is identified, mentioning of these studies in the justification would be non-actionable.

NR	Reviewer	Page	Line	Comment	Action / Reply
				standards should be applied to progesterone for which comparable safety concerns exist. Failure to do so introduces a non-bipartisan application of evidence standards, which compromises both the internal consistency and the perceived impartiality of the guideline.	
151	Alexandra Kohl Schwartz	/	/	A review of the full draft guideline and annexed literature search reveals that pharmacovigilance (PV) data were neither included in the search strategy described in the methods section nor systematically reported for any intervention other than dydrogest dydrogesterone. The sole use of a single PV study (Henry et al.) in the context of dydrogesterone, without comparable evaluation for other agents used in ART, suggests that this reference was selectively introduced rather than identified through a predefined, systematic approach. This selective inclusion carries negative implications for the methodological transparency, consistency, and neutrality of the guideline and may undermine confidence in how safety signals are evaluated across interventions. Suggested revision: Provide a comprehensive and systematic literature search and safety evaluation of all drugs and interventions assessed.	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.
155	Abdellatif Elkholy	/	/	The draft GL on luteal phase support in ART doesn't seem to apply evidence-based principles consistently—especially when it comes to how dydrogesterone is assessed compared to other progestins. This raises some important concerns about how fair and balanced the recommendations really are. Here are a few key points to highlight: The guideline mentions a pharmacovigilance study by Henry et al. (2025) to raise concerns about the safety of dydrogesterone, especially the risk of birth defects. But it's important to understand that these kinds of studies are meant to spot possible safety signals and not to prove that one thing causes another. And this particular study has several major issues: Data problems: Reports in databases like VigiBase were incomplete. For example, the study reported a 0.7% rate of birth defects, which is actually lower	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the jusification, where the Li et al and the Henry et al studies as well as the SR by Katalinic et al. are cited and where the EBM-pyramid does not apply. Furthermore, it is stated in the justification: "It needs to be pointed out here that the observed

NR	Reviewer	Page	Line	Comment	Action / Reply
				than what's normally expected in ART pregnancies (which is usually 2–3%). This suggests the data may be unreliable or underreported. Missing context: The study doesn't compare its results to known risks in ART, and it doesn't account for key factors like the mother's health or the type of IVF treatment used, both of which can affect outcomes. Risk calculations can be misleading: The study uses something called a reporting odds ratio (ROR) to estimate risk. But this method can be confusing and even misleading, especially if the comparison group has lots of side effects in the mother, which might hide or distort the results in the baby. These numbers are often misunderstood, even by non-experts. We don't know how many patients took the drug: One big limitation of this type of study is that we don't know how many people were actually exposed to dydrogesterone or progesterone. Without that, it's impossible to figure out how common or rare any side effects really are. Stronger studies don't show the same risk: Most importantly, these findings haven't been backed up by better-quality studies. In fact, a recent meta-analysis by Katalinic et al. (2024) found no increased risk of birth defects with dydrogesterone. On the other side, Katalinic et al. (2022, 2024): a scoping review and meta-analysis followed by systematic review and meta-analysis, looked at all the available evidence and found no link between dydrogesterone and birth defects. LOTUS I & II trials: These large, well-run studies found that oral dydrogesterone works just as well as vaginal progesterone for luteal phase support, with a similar safety record. Griesinger et al. (2020): This study even showed that dydrogesterone led to higher pregnancy rates compared to vaginal progesterone. So, this kind of strong evidence should carry the most weight when developing clinical guidelines. However, with all the strong evidence supporting dydrogesterone, it only gets a conditional recommendation, while other progestins, some with less evidence, are strongly recommended. That feels inconsistent and a bit unfair.	relations from these two studies cannot be translated into a conclusion on causality".

NR	Reviewer	Page	Line	Comment	Action / Reply
				And what's even more concerning is that dydrogesterone is the only progestin—and actually the only drug in the entire guideline—that has a safety warning, based on pharmacovigilance data, even though the limitations of this data have already been pointed out. None of the other drugs have similar warnings. In conclusion, it's important to stick to the evidence hierarchy—well-designed clinical trials and systematic reviews, like the LOTUS studies and those by Katalinic, should carry more weight than early-signal studies like Henry et al., which have known limitations. Safety concerns should be based on strong, reliable evidence and applied fairly across all medications. The selective safety note on dydrogesterone is not consistent with this approach and should either be removed or properly explained in context highlighting that higher quality evidence supports dydrogesterone safety.	
100	Reassure group	/	/	The current sentence wrongly implies that safety concerns apply uniquely or with special emphasis to dydrogesterone, rather than to all substances used in early pregnancy and ART. This selective negative framing is methodologically inappropriate and inconsistent with basic principles of reproductive pharmacology and teratology, where fetal safety is universally relevant, regardless of whether the drug is synthetic, natural, oral, or vaginal. Moreover, comparator drugs (e.g., micronised progesterone) exhibit strong variation in intrauterine exposure due to differences in route of administration. Vaginal progesterone may deliver high local endometrial concentrations with extremely high early fetal exposure, while IM or oral routes result in greater systemic exposure and impact on placenta/placental transfer. Importantly, natural progesterone is chemically unstable, rapidly metabolized, and yields a variety of metabolites—some of which have been shown to exhibit anti-mitotic or anti-proliferative activity (e.g., 5β-dihydroprogesterone) in vitro and in animal studies (e.g. doi.org/10.1186/s41936-021-00212-3,	The Henry et al. pharmacovigilance study and the Li et al. DEBC cohort study were included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				doi.org/10.1038/s41598-020-78976-x). These aspects are not accounted for in the current guideline but are essential for interpreting potential class effects or differences across preparations. Suggested revision: Avoid negative framing. As is, the statement is problematic for two reasons: 1. It implies that being orally active and structurally different uniquely triggers safety concerns. 2. It lacks comparative context — safety is relevant for all progestins in early pregnancy, not just dydrogesterone. Suggested sentence: Dydrogesterone is a retrosteroid with distinct stereochemical and structural features compared to natural progesterone, resulting in a unique metabolic and pharmacokinetic profile.	
67	Yasser Orief	/	/	1. Despite strong evidence—including randomized controlled trials (RCTs) and meta-analyses—demonstrating superior efficacy and comparable safety, dydrogesterone is only granted a conditional recommendation. 2. Micronized vaginal progesterone (MVP) receives a strong recommendation for luteal phase support (LPS), despite the absence of placebo-controlled RCTs supporting its efficacy. 3. While the guideline claims to prioritize systematic reviews and RCTs, disproportionate emphasis is placed on a lower-tier pharmacovigilance study focused on dydrogesterone 4. According to the guideline's own methodology (page 179, line 5419), literature selection follows an evidence hierarchy—starting with systematic reviews and meta-analyses, followed by RCTs, and then observational studies. Nevertheless, lower-quality studies (e.g., Atarieh et al. for efficacy and Henry et al. for safety) are given undue prominence over high-quality clinical trials and meta-analyses	The GDG has no doubts about the efficacy of dydrogesterone for LPS. The RCT by Atarieh et al. 2024 was taken out of the guideline because it has been retracted. However, when formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies as well as the SR by Katalinic et al. are cited and where the EBM-pyramid does not apply. Furthermore, it is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be

NR	Reviewer	Page	Line	Comment	Action / Reply
				5. There is selective inclusion of RCTs and exclusion of more recent, robust studies. Outdated or methodologically weak trials (e.g., Atarieh 2024, Kupferminc 1999) are considered, while newer, high-quality studies on dydrogesterone post-2017 are ignored—failing to reflect current ART protocols. 6. In the progesterone section, four out of five recommendations are based on 'Good Practice Points' (GPP), with no supporting evidence cited. Meanwhile, despite two Phase 3 trials for dydrogesterone, this is not deemed adequate for a strong recommendation 7. The 2015 Cochrane meta-analysis cited in support of progesterone also included dydrogesterone trials. Nevertheless, the guideline separates them without clear justification.	translated into a conclusion on causality". GPP's are formulated when supporting evidence is missing or insufficient. The 2015 Cochrane review was replaced by the individual RCTs comparing micronised progesterone to placebo/no intervention.
100	Reassure group	/	/	The current guideline cites pharmacovigilance and observational safety data exclusively in relation to dydrogesterone, but omits comparable data on natural progesterone, despite its widespread use in ART. This omission undermines the neutrality, balance, and methodological consistency expected of a clinical guideline. Notably, several studies from non-dydrogesterone settings—particularly in the U.S., where dydrogesterone is not available—have raised concerns about congenital anomalies potentially associated with natural progesterone exposure in ART contexts. For example: • Carmichael et al. (2005) reported an association between maternal progestin intake and hypospadias in a large U.S. population-based study (Arch Pediatr Adolesc Med. 159(10):957–962). While no specific compound or route was identified, dydrogesterone was not among the available treatments. • Silver et al. (1999) observed an increased risk of hypospadias in IVF pregnancies, noting that the only recognized difference between cases	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline. Only studies published between 31 October 2018 and 2 February 2025 were considered for inclusion in the guideline update. Both suggested studies included al PROGESTINS, which also includes dydrogesterone, and because no specific compound or route is identified, mentioning of these studies in the justification would be non-actionable.

NR	Reviewer	Page	Line	Comment	Action / Reply
				and controls was maternal progesterone administration. Again, this observation occurred in a context where dydrogesterone was not in use. By excluding such data, the guideline creates the impression that dydrogesterone uniquely raises safety concerns, while other progestins are implicitly regarded as safer. This asymmetry in evidence presentation may mislead clinicians and does not reflect the available literature. Moreover, if the Henry et al. pharmacovigilance study is considered sufficient to support a conditional recommendation against dydrogesterone, then—to maintain methodological parity—similar standards should be applied to progesterone for which comparable safety concerns exist. Failure to do so introduces a non-bipartisan application of evidence standards, which compromises both the internal consistency and the perceived impartiality of the guideline.	
100	Reassure group	22	R104	The current strong recommendation for “progesterone” after IVF/ICSI is not supported by direct, high-quality evidence as summarized in the Cochrane 2015 review. Critically, not a single placebo- or no-treatment-controlled RCT has evaluated vaginal progesterone monotherapy for live birth rate, which is the most widely used route. The cited Cochrane meta-analysis includes studies using dydrogesterone—accounting for approximately 85% of the weight in the analysis—as well as combination regimens (e.g., vaginal progesterone + estradiol or intramuscular progesterone), rendering the evidence indirect and likely not applicable to the predominant clinical practice due to relevant pharmacokinetic differences between administration routes. According to GRADE and ESHRE’s own methodological principles, a strong recommendation based on absent or only indirect and low-certainty evidence is inappropriate, unless accompanied by an explicit and compelling rationale (e.g. strong patient values, ethical imperatives etc.).	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands. The setting of IUI, where the need of luteal support is still under discussion, is not comparable to IVF, where a new RCT comparing vaginal progesterone to placebo would be unethical given the current knowledge of LPS.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Such justification is currently lacking, and instead, inferences are made that are not supported by the literature. Furthermore, this important limitation is obscured in the guideline text. Instead of clearly presenting the evidence base and the absence of sufficient direct evidence, it is simply stated that 'the evidence clearly supports the use of progestins in the luteal phase' in the justification text. Suggested revision: The recommendation should therefore be revised to reflect the limited, indirect and heterogeneous evidence base. A call for high-quality RCTs is necessary. Unlike the statement that such trials are not going to happen, such trials are indeed feasible, as demonstrated by the ongoing RCT comparing vaginal progesterone with placebo in women undergoing ovarian stimulation and IUI (NTR NL24508).	
100	Reassure group	22	104	Reliance on outdated evidence: The current recommendation for progesterone use in luteal phase support (LPS) after IVF/ICSI fails to reflect the current state of evidence. It continues to rely heavily on the outdated Cochrane meta-analysis by van der Linden et al. (2015), which is over a decade old, does not include large recent phase III trials and is limited to direct comparisons, while omitting the far more comprehensive and recent network meta-analysis by Kastoras et al. (2024) (DOI: 10.1038/s41598-024-64804-z), published in Nature Scientific Reports. This study includes 76 RCTs, 22 interventions, and over 26,000 participants, and uses advanced methodology that allows comparisons across different routes and formulations within a single analysis. The failure to include this analysis obscures a critical fact: there is substantial uncertainty and heterogeneity in the efficacy profiles of the various progesterone formulations (oral, vaginal, subcutaneous, intramuscular with or without Agonist or hCG or E2). This is also supported by another earlier large systematic review and network meta-analysis (https://doi.org/10.1186/s12958-021-00782-5). Pertinent to progestin LPS: the Kastoras network meta-analysis	The systematic review by Kastoras et al., 2024 has been reviewed by the GDG during the evidence synthesis process. Unfortunately, they do not report which studies were included for each of the outcomes and each of the interventions, nor did they report the number of events in the study and control groups. Therefore, the GDG was unable to evaluate the quality of this systematic review and the study was excluded, as specified in annex 7.

NR	Reviewer	Page	Line	Comment	Action / Reply
				reports that oral dydrogesterone is significantly more effective for live birth rate than vaginal progesterone, while subcutaneous progesterone in the dosages tested is significantly less effective (Figure 5B). These findings may have direct implications for clinical care, yet they are neither addressed nor acknowledged in the current draft. Issuing a strong, undifferentiated recommendation for “progesterone” by any route or type or dose based on outdated and methodologically weak analyses thus misrepresents the evidence and ignores unresolved clinical uncertainty. This contradicts core GRADE principles, which require that strong recommendations be based on high-certainty, consistent, and directly applicable evidence—none of which is met here. Instead, the draft applies a one-size-fits-all recommendation, despite well founded uncertainty that not all progesterone routes of administration, formulations and dosages may be equivalent. Suggested revision: The systematic review of Kastoras et al. 2024 with network meta-analysis should be incorporated into the evidence base. If not revised, the guideline risks giving misleading clinical direction and confidence based on outdated, incomplete, and oversimplified interpretation of the available evidence. Based on ESHREs own rules: ‘Recognise that different choices will be appropriate for individual patients and that you must help each patient arrive at a management decision consistent with his or her values and preferences ‘) a conditional recommendation should be considered.	
140	German Society of Reproductive Medicine	22	104	The Cochrane meta-analysis cited (van der Linden et al., 2015) reports a benefit of progestins versus placebo/no treatment for luteal phase support (OR 1.77, 95% CI 1.09–2.86). However, approximately 85% of the statistical weight in that meta-analysis derives from studies investigating dydrogesterone. Given that this guideline includes a separate chapter specifically addressing dydrogesterone, it is methodologically inconsistent and misleading to present this pooled analysis under the general heading of "progesterone" without clarification. Likewise, HMG	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.

NR	Reviewer	Page	Line	Comment	Action / Reply
				RCTs are not presented under a rFSH chapter and findings from Antagonist trials are not presented under Agonist trials. Suggested revision: To ensure logical structure and consistency, the DYD-containing RCTs (active comparator vs. nil) from the Cochrane review should be presented in the DYD chapter, and section 18.1 should reflect the evidence (active comparator vs. nil/placebo) relevant to progesterone (e.g. oral, MVP, i.m., s.c.). This is a necessity of the principle of internal consistency in evidence presentation.	
140	German Society of Reproductive Medicine	22	104	Reliance on outdated evidence: The current recommendation for progesterone use in luteal phase support (LPS) after IVF/ICSI fails to reflect the current state of evidence. It continues to rely heavily on the outdated Cochrane meta-analysis by van der Linden et al. (2015), which is over a decade old, does not include large recent phase III trials and is limited to direct comparisons, while omitting the far more comprehensive and recent network meta-analysis by Kastoras et al. (2024) (DOI: 10.1038/s41598-024-64804-z), published in Nature Scientific Reports. This study includes 76 RCTs, 22 interventions, and over 26,000 participants, and uses advanced methodology that allows comparisons across different routes and formulations within a single analysis. The failure to include this analysis obscures a critical fact: there is substantial uncertainty and heterogeneity in the efficacy profiles of the various progesterone formulations (oral, vaginal, subcutaneous, intramuscular with or without Agonist or hCG or E2). This is also supported by another earlier large systematic review and network meta-analysis (https://doi.org/10.1186/s12958-021-00782-5). Pertinent to progestin LPS: the Kastoras network meta-analysis reports that oral	The systematic review by Kastoras et al., 2024 has been reviewed by the GDG during the evidence synthesis process. Unfortunately, they do not report which studies were included for each of the outcomes and each of the interventions, nor did they report the number of events in the study and control groups. Therefore, the GDG was unable to evaluate the quality of this systematic review and the study was excluded, as specified in annex 7.

NR	Reviewer	Page	Line	Comment	Action / Reply
				dydrogesterone is significantly more effective for live birth rate than vaginal progesterone, while subcutaneous progesterone in the dosages tested is significantly less effective (Figure 5B). These findings may have direct implications for clinical care, yet they are neither addressed nor acknowledged in the current draft. Issuing a strong, undifferentiated recommendation for “progesterone” by any route or type or dose based on outdated and methodologically weak analyses thus misrepresents the evidence and ignores unresolved clinical uncertainty. This contradicts core GRADE principles, which require that strong recommendations be based on high-certainty, consistent, and directly applicable evidence—none of which is met here. Instead, the draft applies a one-size-fits-all recommendation, despite well founded uncertainty that not all progesterone routes of administration, formulations and dosages may be equivalent. Suggested revision: The systematic review of Kastoras et al. 2024 with network meta-analysis should be incorporated into the evidence base. If not revised, the guideline risks giving misleading clinical direction and confidence based on outdated, incomplete, and oversimplified interpretation of the available evidence. Based on ESHREs own rules: ‘Recognise that different choices will be appropriate for individual patients and that you must help each patient arrive at a management decision consistent with his or her values and preferences ‘) a conditional recommendation should be considered.	
140	German Society of Reproductive Medicine	22	104	The current strong recommendation for “progesterone” after IVF/ICSI is not supported by direct, high-quality evidence as summarized in the Cochrane 2015 review. Critically, not a single placebo- or no-treatment-controlled RCT has evaluated vaginal progesterone monotherapy for live	The recommendation states "Progesterone is recommended for LPS", and proceeds with a GPP that all administration routes can be used. Even though the evidence is old and low quality,

NR	Reviewer	Page	Line	Comment	Action / Reply
				birth rate, which is the most widely used route. The cited Cochrane meta-analysis includes studies using dydrogesterone—accounting for approximately 85% of the weight in the analysis—as well as combination regimens (e.g., vaginal progesterone + estradiol or intramuscular progesterone), rendering the evidence indirect and likely not applicable to the predominant clinical practice due to relevant pharmacokinetic differences between administration routes. According to GRADE and ESHRE's own methodological principles, a strong recommendation based on absent or only indirect and low-certainty evidence is inappropriate, unless accompanied by an explicit and compelling rationale (e.g. strong patient values, ethical imperatives etc.). Such justification is currently lacking, and instead, inferences are made that are not supported by the literature. Furthermore, this important limitation is obscured in the guideline text. Instead of clearly presenting the evidence base and the absence of sufficient direct evidence, it is simply stated that 'the evidence clearly supports the use of progestins in the luteal phase' in the justification text. Suggested revision: The recommendation should therefore be revised to reflect the limited, indirect and heterogeneous evidence base. A call for high-quality RCTs is necessary. Unlike the statement that such trials are not going to happen, such trials are indeed feasible, as demonstrated by the ongoing RCT comparing vaginal progesterone with placebo in women undergoing ovarian stimulation and IUI (NTR NL24508).	there will be no new RCTs comparing the use of progesterone to placebo. Nevertheless, the GDG has no doubt that support of the luteal phase is needed. Therefore, a strong recommendation is warranted.
151	Alexandra Kohl Schwartz	22	104	The Cochrane meta-analysis cited (van der Linden et al., 2015) reports a benefit of progestins versus placebo/no treatment for luteal phase support (OR 1.77, 95% CI 1.09–2.86). However, approximately 85% of the statistical weight in that meta-analysis derives from studies investigating	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.

NR	Reviewer	Page	Line	Comment	Action / Reply
				dydrogesterone. Given that this guideline includes a separate chapter specifically addressing dydrogesterone, it is methodologically inconsistent and misleading to present this pooled analysis under the general heading of "progesterone" without clarification. Likewise, HMG RCTs are not presented under a rFSH chapter and findings from Antagonist trials are not presented under Agonist trials. Suggested revision: To ensure logical structure and consistency, the DYD-containing RCTs (active comparator vs. nil) from the Cochrane review should be presented in the DYD chapter, and section 18.1 should reflect the evidence (active comparator vs. nil/placebo) relevant to progesterone (e.g. oral, MVP, i.m., s.c.). This is a necessity of the principle of internal consistency in evidence presentation.	
151	Alexandra Kohl Schwartz	22	104	The current strong recommendation for “progesterone” after IVF/ICSI is not supported by direct, high-quality evidence as summarized in the Cochrane 2015 review. Critically, not a single placebo- or no-treatment-controlled RCT has evaluated vaginal progesterone monotherapy for live birth rate, which is the most widely used route. The cited Cochrane meta-analysis includes studies using dydrogesterone—accounting for approximately 85% of the weight in the analysis—as well as combination regimens (e.g., vaginal progesterone + estradiol or intramuscular progesterone), rendering the evidence indirect and likely not applicable to the predominant clinical practice due to relevant pharmacokinetic differences between administration routes. According to GRADE and ESHRE’s own methodological principles, a strong recommendation based on absent or only indirect and low-certainty evidence is inappropriate, unless accompanied by an explicit and compelling rationale (e.g. strong patient values, ethical imperatives etc.).	The recommendation states "Progesterone is recommended for LPS", and proceeds with a GPP that all administration routes can be used. Even though the evidence is old and low quality, there will be no new RCTs comparing the use of progesterone to placebo. Nevertheless, the GDG has no doubt that support of the luteal phase is needed. Therefore, a strong recommendation is warranted.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Such justification is currently lacking, and instead, inferences are made that are not supported by the literature. Furthermore, this important limitation is obscured in the guideline text. Instead of clearly presenting the evidence base and the absence of sufficient direct evidence, it is simply stated that 'the evidence clearly supports the use of progestins in the luteal phase' in the justification text. Suggested revision: The recommendation should therefore be revised to reflect the limited, indirect and heterogeneous evidence base. A call for high-quality RCTs is necessary. Unlike the statement that such trials are not going to happen, such trials are indeed feasible, as demonstrated by the ongoing RCT comparing vaginal progesterone with placebo in women undergoing ovarian stimulation and IUI (NTR NL24508).	
151	Alexandra Kohl Schwartz	22	104	Reliance on outdated evidence: The current recommendation for progesterone use in luteal phase support (LPS) after IVF/ICSI fails to reflect the current state of evidence. It continues to rely heavily on the outdated Cochrane meta-analysis by van der Linden et al. (2015), which is over a decade old, does not include large recent phase III trials and is limited to direct comparisons, while omitting the far more comprehensive and recent network meta-analysis by Kastoras et al. (2024) (DOI: 10.1038/s41598-024-64804-z), published in Nature Scientific Reports. This study includes 76 RCTs, 22 interventions, and over 26,000 participants, and uses advanced methodology that allows comparisons across different routes and formulations within a single analysis. The failure to include this analysis obscures a critical fact: there is substantial uncertainty and heterogeneity in the efficacy profiles of the various progesterone formulations (oral, vaginal, subcutaneous, intramuscular with or without Agonist or hCG or E2). This is also supported by another earlier large systematic review and network meta-analysis (https://doi.org/10.1186/s12958-021-00782-5). Pertinent to progestin LPS: the Kastoras network meta-analysis	The systematic review by Kastoras et al., 2024 has been reviewed by the GDG during the evidence synthesis process. Unfortunately, they do not report which studies were included for each of the outcomes and each of the interventions, nor did they report the number of events in the study and control groups. Therefore, the GDG was unable to evaluate the quality of this systematic review and the study was excluded, as specified in annex 7.

NR	Reviewer	Page	Line	Comment	Action / Reply
				reports that oral dydrogesterone is significantly more effective for live birth rate than vaginal progesterone, while subcutaneous progesterone in the dosages tested is significantly less effective (Figure 5B). These findings may have direct implications for clinical care, yet they are neither addressed nor acknowledged in the current draft. Issuing a strong, undifferentiated recommendation for “progesterone” by any route or type or dose based on outdated and methodologically weak analyses thus misrepresents the evidence and ignores unresolved clinical uncertainty. This contradicts core GRADE principles, which require that strong recommendations be based on high-certainty, consistent, and directly applicable evidence—none of which is met here. Instead, the draft applies a one-size-fits-all recommendation, despite well founded uncertainty that not all progesterone routes of administration, formulations and dosages may be equivalent. Suggested revision: The systematic review of Kastoras et al. 2024 with network meta-analysis should be incorporated into the evidence base. If not revised, the guideline risks giving misleading clinical direction and confidence based on outdated, incomplete, and oversimplified interpretation of the available evidence. Based on ESHREs own rules: ‘Recognise that different choices will be appropriate for individual patients and that you must help each patient arrive at a management decision consistent with his or her values and preferences ‘) a conditional recommendation should be considered.	
140	German Society of Reproductive Medicine	22	105	Inappropriate use of a GPP: Labeling the statement that “any of the previously mentioned administration routes (non-oral) for natural progesterone as luteal phase support can be used” as a Good Practice Point (GPP) is inappropriate and misleading. In reality, the relative efficacy of different progesterone formulations, dosages, regimens and administration routes in IVF/ICSI luteal phase support is uncertain, under-researched, and based on evidence of predominantly low quality.	According to the ESHRE manual for guideline development "A good practice point or GPP is written by the GDG to support the recommendations". This is exactly how this GPP on the administration routes was intended.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Suggested revision: The statement does not meet the criteria for a GPP by GRADE standards and should instead be presented as a conditional recommendation, with clear acknowledgment of the current evidence gaps and uncertainty in the justification text.	
151	Alexandra Kohl Schwartz	22	105	Inappropriate use of a GPP: Labeling the statement that “any of the previously mentioned administration routes (non-oral) for natural progesterone as luteal phase support can be used” as a Good Practice Point (GPP) is inappropriate and misleading. In reality, the relative efficacy of different progesterone formulations, dosages, regimens and administration routes in IVF/ICSI luteal phase support is uncertain, under-researched, and based on evidence of predominantly low quality. Suggested revision: The statement does not meet the criteria for a GPP by GRADE standards and should instead be presented as a conditional recommendation, with clear acknowledgment of the current evidence gaps and uncertainty in the justification text.	According to the ESHRE manual for guideline development "A good practice point or GPP is written by the GDG to support the recommendations". This is exactly how this GPP on the administration routes was intended.
2	Natalia Pedachenko	22	109	Dydrogesterone should be strongly recommended for LPS in IVF. Signals from pharmacovigilance reports require confirmation through higher levels of evidence. Recent several meta-analyses: by Katalinic et al. (2024), Katalinic et al. (2022), Griesinger et al. (2020) represent the most up-to-date and highest levels of evidence that did not reveal an additional risk of congenital anomalies with the use of dydrogesterone compared to other progestogens.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
3	K. K. Pandey	22	109	This is an excellent and highly effective drug formulation. Highly recommended Dydrogesterone	Individual experience and practice, while valuable is not a substitute for published clinical studies.
4	Nodira Ruzieva	22	109	Dydrogesterone should have a strong recommendation for LPS, since the conclusion of the pharmacovigilance study is in many ways contradictory to the extensive real-world clinical practice data on the use of dydrogesterone.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
5	Vyacheslav Lokshin	22	109	Dydrogesterone should be with strong recommendation. This position is supported by multiple high-quality studies. Recent meta-analyses have not confirmed the safety concerns (A.Katalinic,2024) and local study in Russia and Kazakhstan -IRIS also demonstrated high efficacy and safety.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
8	Gulsara Z. Eshimbetova	22	109	Based on data of below Dydrogesterone should be strongly recommended for lutein phase support in IVF. 1. Evidence-based medicine prioritizes meta-analyses and systematic reviews as the highest forms of evidence. Signals from case-non-case studies comparable to case-control studies require confirmation by higher levels of evidence. The recent meta-analysis by Katalinic et al. (2024) represents the most up-to-date and highest levels of evidence, which found no additional risk of congenital anomalies with dydrogesterone compared with other progestogens. Alexander Katalinic 1, Maria R Noftz 1, Juan A Garcia-Velasco 2 3, Lee P Shulman 4, John N van den Anker 5 6, Jerome F Strauss Iii 7, No additional risk of congenital anomalies after first-trimester dydrogesterone use: a systematic review and meta-analysis, PMID: 38344249 PMCID: PMC10859181 DOI: 10.1093/hropen/hoae004. 2. Lotus I Double-blind, double-dummy, multicenter, randomized controlled	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. As such the use of Dydrogesterone cannot be strongly supported.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Phase III trial, 1031 patients, 30mg dydrogesterone vs. 600mg MVP Capsule, 974 ET patients, IVF or ICSI, SET or DET - 38 centers in 7 countries: Austria, Belgium, Germany, Finland, Israel, Russia, and Spain 3. Lotus II Open, multicenter, randomized controlled Phase III trial, 1034 patients, 30mg dydrogesterone vs. 8% MPV Gel, 90mg, 980 ET patients, IVF or ICSI, SET or DET - 37 centers in 10 countries/regions: Australia, Belgium, China, Germany, Hong Kong, India, Russia, Singapore, Thailand, and Ukraine According to the Meta-analysis of individual patient data from LOTUS I and LOTUS II: Study Program for Fresh Cycles, which is one of the most methodologically robust registration studies on luteal phase support in IVF cycles, Dydrogesterone compared to MVP in IVF procedures is associated with significantly higher rates of ongoing pregnancy/1000 women 381 DYD vs 341 MVP and live birth 344 DYD vs 312 MVP. (Griesinger G, Blockeel C, Kahler E, et al. Dydrogesterone as an oral alternative to vaginal progesterone for IVF luteal phase support: A systematic review and individual participant data meta-analysis. PLoS One. 2020;15(11):e0241044. doi:10.1371/journal.pone.0241044). Additionaly, we want to say that in the Republic of Uzbekistan, Dydrogesterone has been widely used to support the luteal phase and to treat habitual miscarriage for more than 20 years and since 2019 in the treatment of infertility using IVF. During this period of using Dydrogesterone, we have not seen cases of congenital fetal abnormalities due to its use.	
10	Galina Grebennikova	22	109	Dydrogesterone should be strongly recommended for LPS. Safety concerns raised by pharmacovigilance signals have not been substantiated by recent meta-analytical data (Katalinic et al.2024) Within the evidence-based medicine hierarchy, systematic reviews and meta-analyses constitute the most robust sources of clinical data. Safety of Dydrogesterone is supported by multiple studies, Griesinger (2020), Ott (2021), and Katalinic (2022), which consistently confirm the	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation

NR	Reviewer	Page	Line	Comment	Action / Reply
				safety and efficacy of dydrogesterone in the context of luteal phase support.	rate. As such the use of Dydrogesterone cannot be strongly supported.
12	Sridevi Nellimarla	22	109	Vaginal route of progesterone is most preferred,yet adding dydrogesterone as added support in patients who have poor absorption due to inadequate knowledge of administration or due to overt vaginitis. In my experience of using dydrogesterone during hundreds of pregnancies,there has not been any anomaly specifically associated with its usage.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
12	Sridevi Nellimarla	22	109	Luteal Phase Support (LPS) It is noteworthy that the van der Linden (2015) Cochrane meta-analysis, which forms a cornerstone of the evidentiary support for progesterone's role in Luteal Phase Support, incorporated data from two randomized controlled trials (out of five) investigating dydrogesterone. Consequently, the therapeutic effects observed with dydrogesterone have substantively contributed to the meta-analysis's overall conclusion	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.
13	Padmaja Veeramachaneni	22	109	In india , when no other drug was available in the 80s for endometriosis , one of the most popular regimes was use of dydrogesterone 20 mg /day . numerous pregnancies were reported , who continues it as luteal phase support , to prevent miscarriage , to decease prêt erm labor , and no major abnormalities were reported . Pioneers like dr BN chakaborthy had followed up these pregnancies and live births . orally active dydrogeterone is the top prefernace of women .To now say that conditional approval , makes it difficult to recommend a drug with proven efficacy . So this statement should be made only if the proof is based on good quality studies .	The GDG can only rely on published data, in the setting of luteal phase support. The GDG has no doubts about the efficacy of dydrogesterone for LPS. However, the GDG has concerns about potentially higher rates of congenital malformations, therefore, the recommendation is conditional.
18	Biswajyoti Guha	22	109	Dydrogesteron is useful in pregnancy up to 12 weeks and I have been using this molecule for the last 20yrs without any major side effect	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
20	Assel Jaimbetova	22	109	Dydrogesterone must have a strong recommendation for LPS. According to the principles of evidence-based medicine, systematic reviews and meta-analyses represent the highest level of clinical evidence. A recent meta-analysis Katalinic 2024 demonstrated no association between dydrogesterone use and an increased risk of congenital anomalies or other birth defects, thereby reinforcing its safety profile.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
20	Assel Jaimbetova	22	109	Researchers from our Institute of Reproductive Medicine conducted the IRIS study, and the results showed high efficacy of dydrogesterone in supporting the luteal phase, as well as safety for both the mother and the fetus.	The GDG was unable to find a publication in English language related to the IRIS trial.
21	Suyesha Khanijao	22	109	IN MY PRACTICE OF THAT 15 YEARS , I HAVE USED DYDROGESTONE FOR INDICATIONS LIKE RECURRENT PREGNANCY LOSS , LUTEAL SUPPORT IN IU AND IVF CYCLES, TILL DATE I HAVE NOT COME ACROSS ANY ISSUES WITH SIDE EFFECT AS CONGENITAL ANOMALY IN FETUSES OR NEONATES. THE COMPLIANCE IS VERY GOOD AS ITS EASY TO TAKE .	Individual experience and practice, while valuable is not a substitute for published clinical studies.
23	Ginny Gupta	22	109	We in India have been using Dydrogesterone for luteal phase support and also in early pregnancy for more than a decade. We have not encountered any adverse effects in neonates born to these mothers. We recommend the use of Dydrogesterone in luteal phase support and in early pregnancy.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
26	Tamal Bhattacharyya	22	109	Dydrogesterone works well and is generally easy to tolerate—it's a reliable choicerelated. Highly suggested for its effectiveness.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
27	Mukesh Gupta	22	109	We have been using dydrogesterone in such conditions and personally never found any such incidence of increase in Hypospadias ans CHD.	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
28	Sunita Arora	22	109	I HAVE BEEN USING DYDROGESTERONE SALT FOR LAST 23 YEARS AND HAVE NOT COME ACROSS ANY INCREASED CASES OF MALFORMATIONS. iTS A GOOD MOLECULE AND EASY TO USE BECAUSE OF ORAL FORMULATIONS AVAILABLE.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
45	Olena Yashyna	22	109	The highest level of evidence for real clinical practice is provided by systematic meta-analyses and systematic reviews. Individual case data require verification and detailed analysis. Years of experience demonstrate the safety of using dydrogesterone in IVF cycles and stimulation cycles (Katalinic, 2024; Griesinger, 2020). This supports the rationale for prescribing dydrogesterone in future clinical practice.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.
46	Liudmyla Hutsikava	22	109	Teratogenicity is associated with a recognizable syndrome or group of malformations. Most non-genital structures do not contain specific sex steroid receptors, especially during early organogenesis when the heart and limbs are developing; therefore, tissues without receptors for progestational drugs are highly unlikely to have an isolated specific response to a drug when no other tissue or organ is responding (Brent RL., 2005). Evidence-based medicine prioritizes meta-analyses and systematic reviews as the highest forms of evidence. Signals from case-not-case studies comparable to case-control studies require confirmation through higher levels of evidence. A recent meta-analysis by Katalinic et al. (2024) represents the most up-to-date and highest levels of evidence that did not reveal an additional risk of congenital anomalies with the use of dydrogesterone compared to other progestogens. Dydrogesterone should be strongly recommended for LPS in IVF.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.
49	Zaytuna Khamidullina	22	109	Suggest upgrading this to a strong recommendation for oral dydrogesterone in luteal phase support. The finding of lower live birth	The RCT by Atarieh et al. 2024 was taken out of the guideline because it has been retracted.

NR	Reviewer	Page	Line	Comment	Action / Reply
				rates in Atarieh et al. (2024) contrasts with results from larger RCTs (LOTUS I and II, >2000 participants), which demonstrated non-inferiority to vaginal progesterone. Considering the weight of evidence and the practical advantages of oral administration, a stronger recommendation is justified. Recommend rephrasing the statement regarding pharmacovigilance concerns. Updated meta-analyses (Katalinic 2022, 2024) and long-term safety data have not confirmed a causal association between dydrogesterone and congenital malformations. Current wording may be misleading and does not fully reflect the totality of available safety data. Recommend revising the conclusion of this section. While recent pharmacovigilance signals are mentioned, they are not supported by prospective data or established causality. Updated meta-analyses (Katalinic et al., 2022; 2024) and post-marketing experience over decades support the safety of dydrogesterone in early pregnancy. The justification text currently overemphasizes non-conclusive signals and does not adequately reflect the weight of high-quality controlled studies. A more balanced interpretation is advised.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analyisis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up an increase in congenital malformation rate. One of the components to determine a recommendation is the balance between benefit and harms. Since the GDG does not want to disregard the signals from the Henry 2025 and Li 2024 studies, the recommendation cannot be strong.
55	Feruza Gafurova	22	109	Dydrogesterone should be strongly recommended for LPS in IVF. Based on sources below: 1)A Phase III randomized controlled trial of oral dydrogesterone versus intravaginal progesterone gel for luteal phase support in in vitro fertilization (Lotus II): results from the Chinese mainland subpopulation. Dong-Zi Yang, Georg Griesinger et all. 2)Oral dydrogesterone versus intravaginal micronized progesterone gel for luteal phase support in IVF: a randomized clinical trial. Georg Griesinger et all. – These studies demonstrates that oral dydrogesterone is a viable alternative to MVP, due to its comparable efficacy and tolerability profiles. Owing to its patient-friendly oral administration route, dydrogesterone may replace MVP as the	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.

NR	Reviewer	Page	Line	Comment	Action / Reply
				standard of care for luteal phase support in fresh-cycle IVF. 3)Dydrogesterone as an oral alternative to vaginal progesterone for IVF luteal phase support: A systematic review and individual participant data meta-analysis. Georg Griesinger et all. - Influence of significant predictor variables (including treatment) on ongoing pregnancy and live birth. 4) Role of Dydrogesterone for Luteal Phase Support in Assisted Reproduction/ Ameet Patki - Dydrogesterone has a good safety profile and is well tolerated. Its efficacy has been evaluated in several clinical studies and demonstrated to be non-inferior to micronized vaginal progesterone in large-scale clinical trials. Oral dydrogesterone may potentially become a preferred drug for luteal phase support in millions of women undergoing IVF. 5) Oral dydrogesterone along with vaginal micronized progesterone supplementation for luteal phase support in IVF patients, and its impact on pregnancy and live birth rates: a prospective randomized trial. Leonardo Rinaldi et all. - No differences were observed between the two groups in terms of pregnancy rate (Group A 34,9% vs. Group B 35,7%), live birth rate (Group A 30,6% vs. Group B 29,2%), miscarriage rate (Group A 12% vs. Group B 18%) and implantation rate (Group A 18,6% vs. Group B 17,1%). 6) Is oral dydrogesterone equivalent to vaginal micronized progesterone for luteal phase support in women receiving oocyte donation? Margaux Lorillon et all. – analysed 372 oocyte donation cycles with embryo transfer, Conclusion: Oral dydrogesterone seems to be a good alternative to vaginal micronized progesterone for LPS treatment during an artificial cycle, especially in combination with a weekly injection of intramuscular progesterone in the course of oocyte donation.	
62	Hisham A. Arab	22	R109	1. The effectiveness of dydrogesterone in Luteal Phase Support has been demonstrated in numerous RCTs and high-quality studies over the past decade, including LOTUS I and II, as well as others. 2. It was also found to be non-inferior to Micronized Vaginal Progesterone in terms of pregnancy rate at 12 weeks of gestation, with similar safety profiles in both mother and child.	The GDG has no doubts about the efficacy of dydrogesterone for LPS. However, since the GDG has concerns about potential congenital malformations, the recommendation is conditional. The incidence of congenital

NR	Reviewer	Page	Line	Comment	Action / Reply
				3. Concerning its safety, a. First, the above studies have shown that when it was used during LPS and the first trimester, when teratogenicity is of the highest concern, it did not prove to be teratogenic. b. Second, the pharmacovigilance studies that raised this concern were counting all cases, regardless of the genuineness of the data acquisition or how these anomalies were diagnosed. For instance, the 2008 Fertility and Sterility publication demonstrated the unfounded correlation between dydrogesterone and hypospadias, highlighting the variable nature of its presentation. There have also been some retracted studies due to their poor design and inappropriate data compilation on congenital anomalies, which should not be considered in such assessments. c. Third, we all know that pharmacovigilance reporting does not imply correlation or causation; therefore, such reporting should not be a cause for concern or overemphasized. d. Fourth, the safety of dydrogesterone has been continuously evaluated in our practice in Saudi Arabia since the introduction of the National Miscarriage Guidelines in 2014, which was endorsed by the Saudi Society of Obstetrics and Gynecology and updated in 2019 (Arab H et al. International Journal of Women's Health 2019:11 589–596). It has always been recommended that dydrogesterone be used to manage both threatened and recurrent miscarriages during the first trimester, and we have never witnessed a tendency to increase any form of congenital heart disease, hypospadias, or other anomalies above the general incidence globally. e. Fifth, our observation is well supported by the most recent systematic reviews and meta-analysis of the Reassure studies I & II (Katalinic 2022 & 2024). f. Sixth, based on my personal experience and work with Dydrogesterone for over a decade, both locally and globally, I believe that Dydrogesterone is safe and should be strongly recommended as one of the progestins for Luteal Phase Support.	malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
63	Yullia*	22	R109	In my practice, I have been using Duphaston during pregnancy for many years. No developmental defects were detected in newborns.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
91	Willem Verpoest	22	R109	In the light of the published pharmacovigilance papers as well as crucial historical medical errors, it is impossible to state that dydrogesterone is probably recommended for luteal phase support, hence this guideline should be deleted	The GDG has no doubts about the efficacy of dydrogesterone for LPS. However, since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
100	Reassure group	22	R109	This statement is alarmistic, too vague to be informative and borders on being trivial. Pharmacovigilance reports are an expected outcome of spontaneous reporting systems for any widely used drug, and their presence alone does not constitute a meaningful or actionable safety signal. The phrase “cannot necessarily be translated into a conclusion on causality” is awkward and imprecise; it is trivial that no causal inference can be drawn from such data. As currently written, the sentence acknowledges uncertainty about relevance but still implies a concern—without providing any context regarding the nature, frequency, or consistency of the reported malformations or how this reflects with other explicit safety evidence for the drug and within the field of ART. This creates a misleading impression of risk and may overstate what is, at most, a very speculative association. Selective inclusion of such remarks prominently in the summary of recommendations risks distorting the comparative safety profile of drugs used in ART and undermines the guideline’s neutrality and credibility. The current statement incorrectly and in an alarmistic way combines the terms ‘increased risk’ and ‘association’ in the context of a pharmacovigilance disproportionality analysis. This implies a risk-based association, which is misleading. Disproportionality analyses assess patterns in adverse event reporting—not incidence, risk, or causality. Apparent disproportionality can result from numerous artefacts, including notoriety bias, stimulated reporting, underreporting of other adverse events, or changes in reporting practices over	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				time. Moreover, practice patterns such as national licensing status, indication-specific use, and differences in surveillance intensity across countries can all distort PV signal strength. Disproportionality analyses on PV data are a hypothesis-generating tool for early signal detection. The statement is thus inappropriate. The use of the plural reports implies to the reader that there are at least two, or possibly multiple, consistent pharmacovigilance (PV) signals, which is not the case. PV is a multifaceted discipline encompassing a wide range of data sources and methods beyond the spontaneous report-based disproportionality analysis of Henry et al. that the statement refers to. These include observational studies (e.g., cohort and case–control designs), clinical trial data, registries, in vitro and in vivo (animal) studies, mechanistic toxicology, and structured causality assessments. The disproportionality analysis mentioned in the guideline (Henry et al.) represents only a minor and exploratory element within this broader framework. It is a hypothesis-generating tool for early signal detection and, on its own, cannot establish risk. Presenting it without specifying the nature of the study (i.e. early signal detection) and outcome (i.e. ROR) and proper context is therefore misleading. Furthermore, dydrogesterone has been in widespread clinical use for decades, and cumulative PV data—together with robust clinical evidence—have not demonstrated a consistent association with congenital anomalies. Isolating a single disproportionality analysis, without referencing the totality of evidence, risks overstating the signal and may compromise the neutrality and scientific impartiality of the guideline. It is also important to distinguish formal pharmacovigilance from observational pharmacoepidemiology. The DEBC pharmacoepidemiology study (Li et al., 2024), cited later in the document, is a prospective pregnancy cohort with several methodological limitations. It does not involve spontaneous reporting or signal detection, and describing it as PV data conflates distinct approaches and exaggerates the availability of PV evidence. Suggested revision: Remove this safety alert from the guideline summary.	

NR	Reviewer	Page	Line	Comment	Action / Reply
100	Reassure group	22	R109	The inclusion of pharmacovigilance remarks only for dydrogesterone, while omitting similar reporting for other agents used in ovarian stimulation and IVF, introduces a clear inconsistency and bias. Other ART drugs have PV data, often more substantial (https://oehha.ca.gov/proposition-65/crnr/notice-intent-list-chemical-formally-required-be-labeled-or-identified-mechanism-clomiphene-citrate), for example on clomifen the recent studies in humans DOIs: 10.1080/14740338.2024.2358972, 10.1158/1055-9965.EPI-16-0809 , 10.1158/1055-9965.EPI-16-0880, 10.1186/s13048-022-01084-z, 10.1155/2018/7191704 and animals 10.1210/endocr/bqae047. Without systematic consideration of safety signals across all options, this selective emphasis may misrepresent relative risks within the ART field overall and undermine the guideline's neutrality. This selective inclusion introduces a notable imbalance in how safety is presented across treatment options and may inadvertently stigmatize dydrogesterone in the absence of consistent or stronger evidence compared to alternatives. Such asymmetry risks undermining the scientific neutrality and credibility of the guideline. Suggested revision: (a) provide a comprehensive, structured and balanced overview of available maternal and fetal safety data for all relevant drugs and interventions in the context of OS and ART; or (b) omit pharmacovigilance remarks altogether unless they are based on systematic and consistent evidence review with clear implications for clinical decision-making. This aligns with the methodological standards of GRADE, GIN, and the Institute of Medicine, which emphasize consistency, transparency, and fair comparison across all alternatives.	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
103	Andrii Berbets	22	109	As an ObGyn, I prescribe dydrogesterone to my patients, including pregnant women, for many years. According to my clinical experience, this is a safe and effective medication. If used properly following medical indications and dosages, dydrogesterone causes no serious adverse events both to mother and child. Therefore, it deserves to be included into appropriate guidelines.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
107	Emad Darwish	22	109	Dydrogesterone receives only a conditional recommendation despite being supported by robust evidence, including randomized controlled trials and meta-analyses, demonstrating superior efficacy and comparable safety to other options. Griesinger G et al. Hum Reprod.2018. Griesinger G,et al, PloS One. 2020. Katalinic A,etal Hum Reprod Open 2024 Misrepresentation of Dydrogesterone Safety: Dydrogesterone possesses a well-documented safety profile established over 65 years of global use (>147 million women, including >20 million pregnancies) and reinforced by high-quality clinical trials. The guideline's inclusion of a cautionary safety footnote, based solely on a methodologically limited pharmacovigilance study (Henry et al.), deviates from evidence-based medicine principles and introduces unwarranted concern. Recommendation: The dydrogesterone safety footnote should be removed to align with the guideline's stated evidence hierarchy, prioritizing high-quality clinical studies (LOTUS I, LOTUS II,Griesenger Et.al 2020, Katalinic Et.al 2022,Katlinic Et.al 2024)and extensive real-world experience over hypothesis-generating pharmacovigilance data. The established safety profile warrants clear acknowledgment. In conclusion, the draft guideline's assessment of dydrogesterone exhibits methodological disparities inconsistent with its stated evidence-	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic systematic review.

NR	Reviewer	Page	Line	Comment	Action / Reply
				based approach. To optimize scientific rigor, transparency, and clinical relevance, the GDG is urged to: . Remove the dydrogesterone safety footnote to accurately reflect its extensive safety profile and evidence hierarchy.	
108	Rekha Rani	22	109	Dydrogesterone is been around for more than 60years, cases of CHD or Hypospadias attributing to Dydrogesterone alone can be misleading. The reported congenital anomaly incidence of 0.7% in Henry et al pharmacovigilance study is substantially lower than the recognized baseline incidence in natural pregnancies (approximately 2%), a discrepancy strongly suggestive of considerable underreporting of adverse events within the database. More longitudinal, RCT studies are needed to establish the association. Such misleading statements solely based on pharmacovigilance study can create unnecessary fear among practitioners and patients.	The GDG could not agree more that larger studies are necessary. However, in 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.
110	Ayman Abo El Nour	22	109	A High-Level Evidence-Based Perspective; - Clinical Trial Data: In the Phase III LOTUS I and LOTUS II studies, newborns exposed to dydrogesterone were followed up to 30 days post-delivery. No significant treatment-related adverse events linked to congenital, familial, or genetic disorders were observed. - Robust Evidence Supports Safety: A comprehensive scoping review and meta-analysis, representing the highest level of evidence, confirms no causal association between first-trimester dydrogesterone use and fetal abnormalities (Katalinic et al., 2024). This systematic review, preceded by a scoping review to ensure methodological rigor, validates the safety of dydrogesterone for both mother and child during early pregnancy. - Established Safety Profile: Meta-analysis reinforces dydrogesterone's favorable safety profile, aligning with its long-standing clinical use ;Griesinger et.al. 2020 & Katalinic et.al 2022 and Katalinic et.al 2024. Limitations of Pharmacovigilance (PV) Studies: Low-Tier Evidence: PV studies, including disproportionality analyses, are initial	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				tools for signal detection but cannot establish causation due to inherent limitations: - Data Quality Issues: Incomplete data, underreporting, and duplicate records compromise reliability. - Biases: Selection bias, time-related bias, and geographical reporting variations distort findings. - Not an Endpoint: PV signals require validation through higher-quality evidence, such as systematic reviews or clinical trials. - Misleading Data in PV Studies: - The reported birth defect prevalence (0.7%) in (Henry et.al 2025) is significantly lower than the globally recognized 3-5% rate for assisted reproductive technology (ART) pregnancies, indicating underreporting and undermining data significance. - The Reporting Odds Ratio (ROR) value, while not uncommon in PV analyses, has been misused in some cases to create unwarranted fear, particularly through sensationalized titles and inappropriate head-to-head drug comparisons. Ethical Concerns and Patient Impact: - Misuse of PV Data: The continued reliance on poor-quality PV data, coupled with misleading interpretations, risks creating unnecessary fear among patients and clinicians. This may unjustly deny women access to dydrogesterone, a valuable treatment option. - Unethical Fear Creation: Unlike other essential drugs where similar ROR values do not trigger alarm, dydrogesterone has been unfairly targeted, amplifying patient-level concerns without robust evidence. Conclusion: High-quality evidence from systematic reviews and meta-analyses strongly and clearly supports the safety and efficacy of dydrogesterone in early pregnancy. In contrast, PV studies, while useful for generating hypotheses, are low-tier evidence prone to biases and misinterpretation. Their findings should not overshadow the robust safety profile of dydrogesterone established through rigorous research.	

NR	Reviewer	Page	Line	Comment	Action / Reply
				Misusing PV data to incite fear is unethical and detrimental to patient care. Clinical guidelines should prioritize high-quality evidence to ensure informed decision-making and avoid unnecessary alarm that could misguide the Healthcare Professional decision and harm patient care.	
111	T. Ramani Devi	22	109	Comparable efficacy to that of NMP. Due to immune modulatory effect is preferable to NMP in RIF/recurrent miscarriage patients or it can be combined with NMP.	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, a strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
113	Gavisova Alla	22	109	As noted in the recommendations, dydrogesterone is a retro progesterone with a high affinity for progesterone receptors. According to scientific data, dydrogesterone reduces the production of key inflammatory cytokines (COX-2, IL-6, IL-8 and TNF α)1, which may be useful for its use in patients with a history of endometriosis to support the luteal phase. At the same time, in the presence of endometriosis, patients often develop resistance to progesterone, which can limit the effectiveness of progesterone. A study by Bezhenar V, 2023, showed that dydrogesterone is effective even in patients with progesterone sensitization2. In addition, in the study by Nazarenko T., 2025, where I was a co-investigator, support of the luteal phase with dydrogesterone in patients with endometriosis showed comparable efficacy along with other factors of infertility. Also, this study showed a favorable profile of the efficacy and safety of dydrogestene, which allows it to be used as an alternative to vaginal micronized progesterone3.	Patients with endometriosis is considered outside the scope of this guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				I propose to consider the possibility of revising the recommendation of dydrogesterone is recommended for luteal phase support after IVF/ICS	
114	Monica Varma	22	R109	Out of the 122 recommendations in only this recommendation an explanation has been added which is a pharmacovigilance report. Is it that significant to reflect clinical experience with dydrogesterone ?	The GDG has deliberately added the statement under the recommendation box.
114	Monica Varma	22	R109	May be this recommendation can be divided in to two parts regarding efficacy of dydrogesterone and its safety as luteal phase support. A Conditional recommendation has been given emphasizing the pharmacovigilance reports. This may be a hindrance for the continued use of dydrogesterone as a very good option for luteal phase support and in an early pregnancy for specific indications. On one hand we know that even with the use of natural progesterone as luteal phase support the incidence of hypospadias is increased in IVF patients. RI Silver et al 1999, AJ Macnab et al 1991. It may not only be because of the type of progesterone but also because of the increased serum levels of progesterone much more that it is in a normal pregnancy (KJ Siemienowicz et al 2020). As in patient cohorts using natural progesterone as luteal phase support in IVF the evidence is there for increased risk of hypospadias- the risk could be because of infertility, IVF, natural progesterone, increased progesterone serum levels compared to a normal pregnancy ? Same is the case with dydrogesterone but it being a newer molecule in comparison to natural progesterone, a congenital malformation occurring would be highlighted. As mentioned in line 4817 long term offspring health studies are currently lacking for natural progesterone, may be the recommendation 109 can be divided specifically for fetal safety for both dydrogesterone	When formulating recommendations, in addition to the evidence cited in the evidence section, benefits vs harms need to be considered. Therefore, it would not make sense to formulate 2 recommendations, one on efficacy and one on safety.

NR	Reviewer	Page	Line	Comment	Action / Reply
				and natural progesterone as long term offspring health studies are lacking for both and effect on the fetus as congenital malformations is equally questionable for both molecules specifically for the exact cause – effect relationship.	
115	Hassan Sallam	22	109	Why is the recommendation" conditional" when various RCTs reported its efficacy compared to other luteal support agents (including the LOTUS studies)?	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
136	Veaceslav Mosin	22	109	Among patients who conceived, pregnancy outcomes—including live birth rates around 1000—have been favorable, and no drug-related side effects have been reported.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
136	Veaceslav Mosin	22	109	Based on our clinical practice and experience, no adverse effects or tolerability issues have been observed in patients receiving dydrogesterone for luteal phase support in IVF protocols. The treatment has been well tolerated across all age groups. Among patients who conceived, pregnancy outcomes—including live birth rates around 1000—have been favorable, and no drug-related side effects have been reported.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
140	German Society of Reproductive Medicine	22	109	This statement is alarmistic, too vague to be informative and borders on being trivial. Pharmacovigilance reports are an expected outcome of spontaneous reporting systems for any widely used drug, and their presence alone does not constitute a meaningful or actionable safety signal. The phrase "cannot necessarily be translated into a conclusion on causality" is awkward and imprecise; it is trivial that no causal inference can be drawn from such data. As currently written, the sentence	The GDG does not agree with the reviewer. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				acknowledges uncertainty about relevance but still implies a concern—without providing any context regarding the nature, frequency, or consistency of the reported malformations or how this reflects with other explicit safety evidence for the drug and within the field of ART. This creates a misleading impression of risk and may overstate what is, at most, a very speculative association. Selective inclusion of such remarks prominently in the summary of recommendations risks distorting the comparative safety profile of drugs used in ART and undermines the guideline's neutrality and credibility. Suggested revision: Remove this safety alert from the guideline summary.	
140	German Society of Reproductive Medicine	22	109	The current statement incorrectly and in an alarmistic way combines the terms 'increased risk' and 'association' in the context of a pharmacovigilance disproportionality analysis. This implies a risk-based association, which is misleading. Disproportionality analyses assess patterns in adverse event reporting—not incidence, risk, or causality. Apparent disproportionality can result from numerous artefacts, including notoriety bias, stimulated reporting, underreporting of other adverse events, or changes in reporting practices over time. Moreover, practice patterns such as national licensing status, indication-specific use, and differences in surveillance intensity across countries can all distort PV signal strength. Disproportionality analyses on PV data are a hypothesis-generating tool for early signal detection. The statement is thus inappropriate. Suggested revision: Remove this safety alert from the guideline summary.	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.
140	German Society of Reproductive Medicine	22	109	The use of the plural reports implies to the reader that there are at least two, or possibly multiple, consistent pharmacovigilance (PV) signals, which is not the case. PV is a multifaceted discipline encompassing a wide range of data sources and methods beyond the spontaneous report-	The sentence in the justification was corrected. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic

NR	Reviewer	Page	Line	Comment	Action / Reply
				based disproportionality analysis of Henry et al. that the statement refers to. These include observational studies (e.g., cohort and case–control designs), clinical trial data, registries, in vitro and in vivo (animal) studies, mechanistic toxicology, and structured causality assessments. The disproportionality analysis relations cannot necessarily be translated into a conclusion on causality.” mentioned in the guideline (Henry et al.) represents only a minor and exploratory element within this broader framework. It is a hypothesis-generating tool for early signal detection and, on its own, cannot establish risk. Presenting it without specifying the nature of the study (i.e. early signal detection) and outcome (i.e. ROR) and proper context is therefore misleading. Furthermore, dydrogesterone has been in widespread clinical use for decades, and cumulative PV data—together with robust clinical evidence—have not demonstrated a consistent association with congenital anomalies. Isolating a single disproportionality analysis, without referencing the totality of evidence, risks overstating the signal and may compromise the neutrality and scientific impartiality of the guideline. It is also important to distinguish formal pharmacovigilance from observational pharmacoepidemiology. The DEBC pharmacoepidemiology study (Li et al., 2024), cited later in the document, is a prospective pregnancy cohort with several methodological limitations. It does not involve spontaneous reporting or signal detection, and describing it as PV data conflates distinct approaches and exaggerates the availability of PV evidence. Suggested revision: Remove this safety alert from the guideline summary.	nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
140	German Society of Reproductive Medicine	22	109	The inclusion of pharmacovigilance remarks only for dydrogesterone, while omitting similar reporting for other agents used in ovarian stimulation and IVF, introduces a clear inconsistency and bias. Other ART drugs have PV data, often more substantial (https://oehha.ca.gov/proposition-65/crnr/notice-intent-list-chemical-formally-required-be-labeled-or-identified-mechanism-clomiphene-citrate), for example on clomifen the recent studies in humans DOIs: 10.1080/14740338.2024.2358972, 10.1158/1055-9965.EPI-16-0809 , 10.1158/1055-9965.EPI-16-0880, 10.1186/s13048-022-01084-z, 10.1155/2018/7191704 and animals 10.1210/endocr/bqae047. Without systematic consideration of safety signals across all options, this selective emphasis may misrepresent relative risks within the ART field overall and undermine the guideline's neutrality. This selective inclusion introduces a notable imbalance in how safety is presented across treatment options and may inadvertently stigmatize dydrogesterone in the absence of consistent or stronger evidence compared to alternatives. Such asymmetry risks undermining the scientific neutrality and credibility of the guideline. Suggested revision: (a) provide a comprehensive, structured and balanced overview of available maternal and fetal safety data for all relevant drugs and interventions in the context of OS and ART; or (b) omit pharmacovigilance remarks altogether unless they are based on systematic and consistent evidence review with clear implications for clinical decision-making. This aligns with the methodological standards of GRADE, GIN, and the Institute of Medicine, which emphasize consistency, transparency, and fair comparison across all alternatives.	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
140	German Society of Reproductive Medicine	22	109	The current sentence wrongly implies that safety concerns apply uniquely or with special emphasis to dydrogesterone, rather than to all substances used in early pregnancy and ART. This selective negative framing is methodologically inappropriate and inconsistent with basic principles of reproductive pharmacology and teratology, where fetal safety is universally relevant, regardless of whether the drug is synthetic, natural, oral, or vaginal. Moreover, comparator drugs (e.g., micronised progesterone) exhibit strong variation in intrauterine exposure due to differences in route of administration. Vaginal progesterone may deliver high local endometrial concentrations with extremely high early fetal exposure, while IM or oral routes result in greater systemic exposure and impact on placenta/placental transfer. Importantly, natural progesterone is chemically unstable, rapidly metabolized, and yields a variety of metabolites—some of which have been shown to exhibit anti-mitotic or anti-proliferative activity (e.g., 5β-dihydroprogesterone) in vitro and in animal studies (e.g. doi.org/10.1186/s41936-021-00212-3, doi.org/10.1038/s41598-020-78976-x). These aspects are not accounted for in the current guideline but are essential for interpreting potential class effects or differences across preparations. Suggested revision: Avoid negative framing. As is, the statement is problematic for two reasons: 1. It implies that being orally active and structurally different uniquely triggers safety concerns. 2. It lacks comparative context — safety is relevant for all progestins in early pregnancy, not just dydrogesterone. Suggested sentence: Dydrogesterone is a retrosteroid with distinct stereochemical and structural features compared to natural	Had a studies reporting concerns been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline. It has been shown that dydrogesterone has differt binding properties compared to natural progesterone. This is one of the hypothesis underlying the mechanism of congenital malformations with dydrogesterone.

NR	Reviewer	Page	Line	Comment	Action / Reply
				progesterone, resulting in a unique metabolic and pharmacokinetic profile.	
141	Kastubh Kulkarni	22	109	Our Experience , regarding the safety of using the drug dychogesterone for luteal phase support in heating patients of infertility is as follows: We have used the drug , in the recommended dose and found NO INCREASE IN THE CONGENITAL ABNORMALITIES IN THE NEW BORN in the past 25 year. We feel it has a beneficial role, as immunomodulator in patients of RPL.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
145	Vardanyan Rusudan	22	109	Dydrogesterone should be strongly recommended for LPS in IVF. Dydrogesterone has shown comparable or superior outcomes to other progestogens without additional safety concerns, justifying its strong recommendation for the use in LPS protocols.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
146	Mohamed Bedis Chanoufi	22	109	As a Professor of Gynecology and Obstetrics and Head of Department, I would like to share my clinical perspective on the use of dydrogesterone in pregnancy care. In my extensive experience, dydrogesterone has consistently demonstrated a favorable safety profile when used for luteal phase support and early pregnancy management. I have not encountered any adverse fetal outcomes directly attributable to its use. This observation is consistent with the broader clinical literature, which has not established a causal link between dydrogesterone and congenital anomalies. While recent pharmacovigilance data from WHO's VigiBase have raised concerns, such findings should be interpreted with caution due to limitations such as reporting bias and lack of contextual clinical data. These signals, while important for ongoing monitoring, do not confirm causation.	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
				Given its decades-long use, well-documented tolerability, and patient-friendly oral administration, dydrogesterone remains a reliable option in pregnancy support. I support continued vigilance and further prospective studies, but based on current evidence and clinical practice, I remain confident in its safety when used appropriately.	
147	BV Shobha	22	109	Dydrogesterone is in the Indian market for >6 decades. Indian Drs prescribe it in conditions like Recurrent pregnancy loss, Threatened miscarriage, Luteal phase support in ART, Endometriosis, Preterm, Abnormal uterine bleeding etc. Such long history of molecule with established safety data in pregnancy (LOTUS I & II trails) recent meta-analysis (Katalinic 2022 & 2024) proves that Dydrogesterone isn't alone attributed for CHD or Hypospadias cases. More studies are needed to establish the association. Henry et al 2025 study solely should not be considered for laying the recommendations.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality".
152	Liudmila Stavinskaia	22	109	In RM, dydrogesterone has been widely used in obstetric and gynecological practice since 2004. This includes support for the luteal phase in natural ovulation stimulation cycles or reproductive technology cycles, demonstrating high efficacy, a high number of developing pregnancies, and no adverse effects on either the mother or the fetus. Our practical recommendations are based on publications Barbosa M.W. et alt., Mirza F.G. et alt., Chakravarty B.N et alt., Patki A. et alt., Saharkhiz N et alt., Tomic V et alt., G. Sukhikh, I. Baranov et alt., and, foremost, on the LOTUS I and LOTUS II studies by Tournaye H et alt.	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, a strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
152	Liudmila Stavinskaia	22	109	In randomized clinical trials (LOTUS I and LOTUS II) comparing the efficacy, safety, and tolerability of oral dydrogesterone and vaginal micronized	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, the

NR	Reviewer	Page	Line	Comment	Action / Reply
				progesterone for luteal phase support in in vitro fertilization cycles, the following was confirmed: - In the study population, the pregnancy rate at 12 weeks of gestation (10 weeks of pregnancy) was 37.6% and 33.1% (LOTUS I) and 36.7% and 34.7% (LOTUS II) for oral dydrogesterone and vaginal micronized progesterone, respectively. , respectively. The difference in pregnancy rates between the two groups was 4.7 (95% CI, −1.2; 10.6) (LOTUS I) and 2.0 (95% CI, −4.0; 8.0) (LOTUS II); - in the safety evaluation sample, which included 1,029 patients (LOTUS I) and 1,030 patients (LOTUS II) who received at least one dose of the study drug, the incidence of the most common adverse reactions was comparable in both treatment groups. Due to the nature of the patient population/indication studied, a certain number of early abortions/miscarriages are expected, especially before 12 weeks of gestation (10 weeks of pregnancy), as the expected pregnancy rate at this time point is approximately 35%. The safety profile observed in both LOTUS studies is consistent with expectations, given the well-established safety profile of dydrogesterone and the patient population/indication. According to Lotus 1, adverse events that developed during treatment of the mother and fetus/newborn were analyzed according to system organ class and reported for the dydrogesterone and MPC groups as follows: reproductive system and breast disorders (21.8% and 18.4%; vaginal bleeding was the most common AEFI overall for both treatment groups – 11.6% and 9.2%); gastrointestinal disorders (19.1% and 17.2%); disorders of the nervous system (7.7% and 8.2%); disorders of the vascular system (3.5% in both groups), including peripheral embolism and thrombosis (0.2% in both groups). Data on the safety of newborns collected during delivery were similar in the dydrogesterone and MPC groups, with most infants born without any abnormalities on physical examination (93.4% and 92.4% in the overall study population and 100.0% and 97.1% in the Russian population, respectively). The number of fetuses/newborns who developed at least one serious NPS was similar in both groups: 4.2% in the dydrogesterone group and 5.7% in the MPC	incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				group. In total, there were 11 fetuses/newborns with NANFL associated with congenital, hereditary, or genetic disorders, with some having more than one condition (5 in the dydrogesterone group and 6 in the MPC group). In the Russian population, there were two fetuses with trisomy 21, one in each treatment group. Ultimately, among Russian patients in the Lotus I study, mothers did not report any health problems in their infants during visit 11 (a telephone call 6 months after delivery to assess the safety and well-being of the infant(s))(according: Tournaye H., Sukhikh G.T., Kahler E., Griesinger G. A phase III randomized controlled trial comparing the efficacy, safety and tolerability of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization. Hum. Reprod. 2017; 32(5): 1019-27; G. Sukhikh, I. Baranov, G. Melnichenko et alt. Lotus I: A Phase III randomized controlled trial of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization, with focus on the Russian subpopulation, https://dx.doi.org/10.18565/aig.2017.7.75-95)	
88	Aboubakr Mohamed Elnashar	23	R112	This is against all evidence	The GDG is unsure about the biological rationale and the safety of using GnRH agonist for LPS. Hence the recommendation against the use of GnRH agonist for LPS.
124	Ayman Oraif	153	4721	OUTDATED META ANALYSIS : THE 2015 COCHRANE REVIEW USED TO SUPPORT PROGESTERONE ALSO INCLUDED DYDROGESTERONE TRIAL YET THE GUIDELINE SEPERATES THEM UNJUSTIFIABLY	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.
100	Reassure group	153	4721-4723	The Cochrane meta-analysis cited (van der Linden et al., 2015) reports a benefit of progestins versus placebo/no treatment for luteal phase support (OR 1.77, 95% CI 1.09–2.86). However, approximately 85% of the statistical weight in that meta-analysis derives from studies investigating dydrogesterone. Given that this guideline includes a separate chapter specifically addressing dydrogesterone, it is methodologically	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.

NR	Reviewer	Page	Line	Comment	Action / Reply
				inconsistent and misleading to present this pooled analysis under the general heading of "progesterone" without clarification. Likewise, HMG RCTs are not presented under a rFSH chapter and findings from Antagonist trials are not presented under Agonist trials. Suggested revision: To ensure logical structure and consistency, the DYD-containing RCTs (active comparator vs. nil) from the Cochrane review should be presented in the DYD chapter, and section 18.1 should reflect the evidence (active comparator vs. nil/placebo) relevant to progesterone (e.g. oral, MVP, i.m., s.c.). This is a necessity of the principle of internal consistency in evidence presentation.	
140	German Society of Reproductive Medicine	153	4721-4723	The Cochrane meta-analysis cited (van der Linden et al., 2015) reports a benefit of progestins versus placebo/no treatment for luteal phase support (OR 1.77, 95% CI 1.09–2.86). However, approximately 85% of the statistical weight in that meta-analysis derives from studies investigating dydrogesterone. Given that this guideline includes a separate chapter specifically addressing dydrogesterone, it is methodologically inconsistent and misleading to present this pooled analysis under the general heading of "progesterone" without clarification. Likewise, HMG RCTs are not presented under a rFSH chapter and findings from Antagonist trials are not presented under Agonist trials. Suggested revision: To ensure logical structure and consistency, the DYD-containing RCTs (active comparator vs. nil) from the Cochrane review should be presented in the DYD chapter, and section 18.1 should reflect the evidence (active comparator vs. nil/placebo) relevant to progesterone (e.g. oral, MVP, i.m., s.c.). This is a necessity of the principle of internal consistency in evidence presentation.	The systematic review by Van der Linden was taken out and replaced by the individual RCTs comparing progesterone to placebo. However, the recommendation still stands.

NR	Reviewer	Page	Line	Comment	Action / Reply
115	Hassan Sallam	154	4768	In the early cessation section of the guidelines, two RCTs reporting the increased incidence of vaginal bleeding with early cessation of luteal support were not included. Although no statistical difference in the miscarriage rate was found in both studies, vaginal bleeding in the pregnant ART patients is very alarming and psychologically taxing. I suggest including this to alert physicians who opt for early cessation of luteal support [Kohls et al. Fertil Steril 2012 Oct;98(4):858-62] and [Aboulghar et al. Hum Reprod 2008 Apr;23(4):857-62]	The studies mentioned by the reviewer are included in the systematic review by Watters et al., 2020, which is cited in the guideline. A sentence was added to the justification on possible bleeding after early cessation of progesterone for LPS.
124	Ayman Oraif	155	4797	PROGESTERONE HAS A STRONG RECOMMENDATION IN THE GUIDELINES DESPITE WEAK CLINICAL EVIDENCE	A strong recommendation indicates that the GDG is confident that most patients would benefit from the intervention.
100	Reassure group	155	4798	Inappropriate use of a GPP: Labeling the statement that “any of the previously mentioned administration routes (non-oral) for natural progesterone as luteal phase support can be used” as a Good Practice Point (GPP) is inappropriate and misleading. In reality, the relative efficacy of different progesterone formulations, dosages, regimens and administration routes in IVF/ICSI luteal phase support is uncertain, under-researched, and based on evidence of predominantly low quality. Suggested revision: The statement does not meet the criteria for a GPP by GRADE standards and should instead be presented as a conditional recommendation, with clear acknowledgment of the current evidence gaps and uncertainty in the justification text.	According to the ESHRE manual for guideline development "A good practice point or GPP is written by the GDG to support the recommendations". This is exactly how this GPP on the administration routes was intended.
7	Sujoy Dasgupta	155	4801	Progesterone support can probably be discontinued after a positive pregnancy test in a fresh embryo transfer cycle	The GDG has considered changing the recommendation. However, they have refrained from doing so because this would contradict the

NR	Reviewer	Page	Line	Comment	Action / Reply
					advice in the SPC's. The evidence investigating early cessation is also not strong, with only 3 RCTs and 830 women for the outcome of live birth. A sentence was added to the justification.
148	Himabindu Annamraju	156	4819	Barbosa MWP, Valadares NPB, Barbosa ACP, Amaral AS, Iglesias JR, Nastri CO, Martins WP, Nakagawa HM. Oral dydrogesterone vs. vaginal progesterone capsules for luteal-phase support in women undergoing embryo transfer: a systematic review and meta-analysis. JBRA Assist Reprod. 2018 Jun 1;22(2):148-156. doi: 10.5935/1518-0557.20180018. PMID: 29488367; PMCID: PMC5982562. Result summary: Good quality evidence indicates that oral dydrogesterone provided at least similar results than vaginal progesterone capsules on live birth/ongoing pregnancy (RR=1.08, 95%CI=0.92-1.26, I2=29%, 8 RCTs, 3,386 women) and clinical pregnancy rates (RR 1.10, 95% CI 0.95 to 1.27; I2=43%; 9 RCTs; 4,061 women). Additionally, moderate quality evidence suggests there is no relevant difference on miscarriage rates (RR=0.92, 95%CI=0.68-1.26, I2=6%, 8 RCTs, 988 clinical pregnancies; the quality of the evidence was downgraded because of imprecision). Based on this, Dydrogesterone can be recommended for luteal phase support, as stated in the previous ESCHRE guideline.	The systematic review by Barbosa was replaced by the more recent systematic review by Griesinger et al., 2020
16	Sonia Naik	156	4829-4836	Concerning Kupferminc et al. (1990) RCT (Page 156, Lines 4829-4836): The randomized controlled trial by Kupferminc et al. (1990), which assessed dydrogesterone against a placebo, demonstrates restricted relevance to current practices of luteal phase support (LPS) within assisted reproductive technology (ART). This limitation arises principally because the trial was conducted prior to dydrogesterone's recognized	The study showed no difference between hCG, dydrogesterone and placebo. Though the setting was similar in dosing three times a day, we cannot really just discard all the studies not showing the superiority of dydrogesterone.

NR	Reviewer	Page	Line	Comment	Action / Reply
				use for ART-associated LPS and, critically, did not employ ovarian suppression regimens. Consequently, the lack of a statistically significant difference between the study arms might be explained by the overriding influence of endogenous progesterone secretion, which could have obscured an underlying luteal phase insufficiency.	
47	Priti Arora Dhamija	156	4829-4836	Regarding Atarieh et al. (2024) RCT: The robustness of the primary outcome and the statistical interpretations presented in the Atarieh et al. (2024) randomized controlled trial concerning dydrogesterone are considerably compromised due to critical shortcomings in its methodology. A pivotal identified weakness is a significant disparity between the study's source data and the derived odds ratios, thereby challenging the dependability of its conclusions.	The RCT by Atarieh et al. 2024 was taken out of the guideline because it has been retracted.
51	Pavika Lal*	156	4829-4836	I am presently employed as Associate Professor in the Department of Obstetrics and Gynaecology, GSVM Kanpur India. My clinical as well academic experience spans a period of approximately 18 years. Since I am at a tertiary care centre I have a privilege to deal with many patients who are being IVF treated or have been given ovulation stimulation protocols with Gonadotropins . Such patients need progesterone support during the first 9-10 weeks of positive pregnancy test. Here are the following salient features which I acknowledge in favour of Dydrogestrone over other progesterone - 1. Better safety profile and well tolerated when given for luteal phase support as compared to vaginal progesterone(Vaginal progesterone causes side effects like vaginal irritation discharge thus seems inconvenient to patient) 2. Inter individual variability in progestrone levels is more as compared to Dydrogestrone. 3. I have seen higher patient satisfaction, acceptibility and good	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
				tolerability with Dydrogesterone over other progesterones. 4. Patients prefer oral routes over other routes due to socio cultural barriers leading to better patient compliance. 5. Since most of the patients are being delivered at our centre, I have not seen any congenital anomaly in the fetus whose mothers have been exposed to Dydrogesterone in their first trimester.	
53	Poornima Durga	156	4829-4836	for the last ten years of my practice ,using Dydrogesterone as luteal phase support without any adverse effects on the fetus . More research –as in MULTI CENTRIC STUDES, RANDOMISED CONTROL TRAILS are strongly recommended to establish an association of Dydrogesterone with any CHD.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
56	Geeta Khanna	156	4829-4836	I don’t comply with the findings of this study and its adverse events	Individual experience and practice, while valuable is not a substitute for published clinical studies.
57	Yun Sun	156	4829-4836	To ensure the authority and impartiality of the ESHRE guidelines, pharmacovigilance studies should be carefully included. As a European authoritative guide, the ESHRE guidelines have always been committed to a fair and objective approach. In the draft version of the 2025 update, a French pharmacovigilance report was cited. This report is mainly used for detecting pharmacovigilance signals and cannot be used as a basis for any safety conclusions. Due to the limited information collected in pharmacovigilance, VigiBase reports cannot be regarded as samples from the patient population like in clinical trials or observational studies. Confounding factors related to diseases or other sources of bias cannot be ruled out, which may lead to misinterpretation of the results. The literature in question was cited multiple times in this guideline, with the limitations and shortcomings of the pharmacovigilance article being	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.

NR	Reviewer	Page	Line	Comment	Action / Reply
				overlooked. It is recommended that citations be made with caution to ensure the fairness of the guideline.	
124	Ayman Oraif	156	4834	STUDIES CITED ARE FROM THE 1990S IGNORING MODERN ART PROTOCOLS AND NEWER RCT'S	As far as the GDG is aware, all RCTs comparing the use of dydrogesterone with placebo or micronised progesterone were included in the guideline
71	Sabirova Venera	156	4834-4836	A study conducted at IVF Scandinavia to evaluate the efficacy of progestogens for luteal phase support in fresh embryo transfer cycles demonstrated the efficacy and favorable safety profile of dydrogesterone in luteal phase support in ART. Comparative analysis of didroestrogesterone and progesterone showed comparable efficacy and safety when used in ART programs. Clinical pregnancy and live birth rates with dydrogesterone as LP support in ART programs were 25/67 and 19/67 (37.3% and 28.4%, respectively), with MVP – 27/79 and 19/79 (34.2% and 24.1%, respectively). In the study showed a favorable safety profile of both gestagens: in the micronized progesterone group, one child was diagnosed with a malformation, syndactyly of the second and third toes. In the dydrogesterone group, one term stillbirth was registered due to chronic fetoplacental insufficiency and intrauterine growth restriction (IUGR). However, the incidence of side effects and adverse reactions was significantly lower in the group of women receiving dydrogesterone compared to micronized progesterone (3/67 (4.5%) and 11/79 (13.9%) cases, respectively, p <0.05)1. In view of the accumulated scientific data on the comparable efficacy and safety of dydrogesterone compared to progesterone, I propose to assign them the same strength of recommendation. References: 1) Sabirova V.L., Kurbatina M.M., Minnullina F.F., Filyushina A.V. Comparative analysis of the efficacy and safety of gestagens for luteal phase support in fresh IVF/ICSI cycles with single embryo transfer. Vopr. ginekol. akus. perinatol.	Only studies published in English language can be considered for inclusion in the guideline. In addition, the incidence of congenital malformations in IVF/ICSI is very low. Therefore conclusions cannot be based on studies with such small study populations.

NR	Reviewer	Page	Line	Comment	Action / Reply
				(Gynecology, Obstetrics and Perinatology). 2023; 22(6): 28–35. (In Russian). DOI: 10.20953/ 1726-1678-2023-6-28-35 2) Nazarenko T.A., Pestova T.I., Lokshin V.N., Dzhusubalieva T.M., Serov V.N., Baranov I.I., Bezhenar V.F., Gavisova A.A., Gorodnova E.A., Dolgushina N.V., Kalugina A.S., Kvashnina E.V., Kogan I.Yu., Koloda Yu.A., Korsak V.S., Krasnopolskaya K.V., Molchanova I.V., Sabirova V.L., Tapilskaya N.I., Sukhikh G.T. Predictors of pregnancy rate in assisted reproductive technologies: results of the IRIS observational program in the population of Russia and Kazakhstan. Akusherstvo i Ginekologiya/Obstetrics and Gynecology. 2025; 3: 144-158 (in Russian) https://dx.doi.org/10.18565/aig.2025.82	
72	Aleksandra Khramtsova	156	4834-4836	We have long-term positive experience using dydrogesterone in IVF/ICSI programs in both fresh and thawed embryo transfer cycles. The results of retrospective study involving 390 patients shows comparable efficacy of micronized progesterone and dydrogesterone for LPS in IVF programs with native embryo transfer in a stimulated cycle. However, it should be noted that pregnancy-induced hypertension with/without proteinuria occurred statistically significantly more often in patients taking MVP for post-transfer support - 24/247 (9.7%) compared to patients in the dydrogesterone group - 6/143 (4.2%) (OR = 2.457; 95% CI 0.98–6.164; p = 0.045). According to the results obtained in the study, it can be concluded that the use of dydrogesterone helped to reduce late pregnancy complications such as pre-eclampsia1. Due to its high selectivity for progesterone receptors, dydrogesterone has some advantages over micronized progesterone in luteal phase support in patients with PCOS during FET in hormone replacement therapy (HRT) cycles. The results of a study that included 105 live births after an ART program showed patients with PCOS after FET who take micronized progesterone for LPS medicine compared to dydrogesterone have a higher risk of GDM and insulin therapy OR=0.143 [CI:0.017—1.209] (p=0.042)2. The accumulated experience and scientific data show that the use of dydrogesterone for LPS is an effective and safety method of therapy. In view of the above data, I suggest to revise the recommendation for dydrogesterone for	Individual experience and practice, while valuable is not a substitute for published clinical studies. With regards to the clinical studies: as per the ESHRE manual on guideline development, only studies written in English are considered to be included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				luteal phase support the same as for progesterone (strong recommendation). References: 1) Khramtsova A.Yu., Bashmakova N.V., Semenov Yu.A., Karibaeva Sh.K., Melkozerova O.A. Comparative analysis of pregnancy outcomes after embryo transfer in a stimulated in vitro fertilisation cycle depending on the progestogen type used for post-transfer support. Akusherstvo i Ginekologiya/Obstetrics and Gynecology. 2024; (10): (in Russian) https://dx.doi.org/10.18565/aig.2024.236 2) Khramtsova AYu, Dankova IV, Deryabina EG. Features of the course of induced pregnancy during the transfer of a thawed embryo in women with polycystic ovary syndrome. Russian Journal of Human Reproduction. 2025;31(1):54–62. (In Russ.). https://doi.org/10.17116/repro20253101154	
74	Zeev Shoham Ariel Weissman Raoul Orvieto	156	4834-4836	The updated 2025 ESHRE Ovarian Stimulation Guideline uses GRADE principles to evaluate scientific evidence quality in a clear hierarchy. Systematic reviews and meta-analyses of randomized controlled trials (RCTs) rank highest, followed by individual RCTs, which are prioritized for treatment questions due to their robust design that minimizes bias. Observational studies like cohort studies rank lower and are used mainly when RCTs are unavailable, as they're considered low-quality evidence due to confounding and selection bias. The conditional recommendation for dydrogesterone in the updated guideline (Section 18.2) does not align with the evidence hierarchy outlined by ESHRE's own methodology. This decision appears to overemphasize the results of two low- to moderate-quality observational studies (Henry et al., 2025; Li et al., 2024) while disregarding the substantial body of high-certainty evidence from randomized controlled trials (RCTs) and meta-analyses. Specifically: • Henry et al. (2025) conducted a pharmacovigilance disproportionality analysis using WHO VigiBase data, which the authors themselves noted is inherently limited by underreporting, confounding, and lack of control over key variables such as maternal age, infertility diagnosis, and ART type. • Li et al. (2024) performed a cohort analysis from a Chinese national birth	The GDG has no doubts about the efficacy of dydrogesterone for LPS. However, when formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the jusification, where the Li et al and the Henry et al studies as well as the SR by Katalinic et al. are cited and where the EBM-pyramid does not apply. Furthermore, it is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality".

NR	Reviewer	Page	Line	Comment	Action / Reply
				registry and found a statistically significant but modest association (adjusted RR 1.13, 95% CI 1.06–1.21) between dydrogesterone exposure and congenital malformations. Importantly, the authors explicitly state that their findings cannot be interpreted as causal and recommend further analysis. By contrast, high-quality RCTs and a robust meta-analysis by Katalinic et al. (2024) provide stronger and more reliable evidence: • Katalinic et al. (2024) performed a systematic review and meta-analysis including 6 RCTs and 3 observational studies (n = 5070, live births = 2680) and found no increased risk of congenital anomalies with dydrogesterone (RR 0.92; 95% CI 0.55–1.55), with a congenital anomaly rate of 2.5%—well within the EUROCAT baseline of 2.15%. • Griesinger et al. (2020) and Tournaye et al. (2017) demonstrated superior live birth and ongoing pregnancy rates with dydrogesterone over vaginal progesterone. In an IPD meta-analysis, dydrogesterone was associated with a significantly higher live birth rate (OR 1.28; 95% CI 1.04–1.57) [Griesinger et al., 2020]. These findings are derived from ICH-GCP compliant studies, offering regulatory-grade data that directly contradict the impression of safety uncertainty presented in the guideline. Suggested Reconsideration: Upgrade the recommendation for dydrogesterone to “strong” based on: • Evidence from RCTs and IPD meta-analyses showing superior efficacy. • No demonstrated increase in congenital anomalies based on high-quality data. • Better patient compliance and preference due to oral route. • In line with GRADE’s mandate to base strong recommendations on high-certainty evidence.	
76	Tapilskaya Natalia	156	4834-4836	Based on the evidence-based medicine pyramid, when choosing a therapy, I rely on data from meta-analyses, systematic reviews, randomized clinical trials, and real-life clinical practice studies. Among progestogens for LPS in IVF cycles, dydrogesterone is a suitable	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, a strong recommendation would indicate that the GDG is confident that most patients would

NR	Reviewer	Page	Line	Comment	Action / Reply
				alternative to micronized vaginal progesterone. RCTs by Tournaye et al., 20171, Griesinger et al., 20182 and meta-analyses by Griesinger et al., 20203, Katalinic et al., 20244 have shown the efficacy and safety of dydrogesterone in ART programs. The results of the prospective study by Nazarenko et al. 2025 fully correlate with the known results of RCTs and meta-analyses regarding the efficacy and safety of dydrogesterone for LPS in routine practice. The confirmed clinical pregnancy rate was 36.7% (37.2% in the Russian population), the live birth rate was 30.1% (30.7% in the Russian population), which was slightly higher compared to the data of the Russian ART registry for the same period. The results of the study showed a high level of satisfaction with oral dydrogesterone therapy, as well as a favorable safety profile for the fetus. Malformations and anomalies were detected in 3.2% of newborns, which correlates with population data, scientific data and real clinical practice5. I suggest adding comment with new data on the effectiveness and safety of using dydrogesterone in LPS (Nazarenko T., 2025). However, some low-certainty case-non-case study data show different results from RCT data, which may mislead health care professionals6. I suggest to revise this paragraph in favor of more substantial evidence of dydrogesterone safety and to revise the recommendation of dydrogesterone: Dydrogesterone is recommended for luteal phase support after IVF/ICSI (strong+) – the same as progesterone.	benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
77	Mohamed Ashraf Mohamed	156	4834-4836	The current body of high-quality evidence strongly supports the safety profile of oral dydrogesterone for luteal phase support in assisted reproductive technology (ART) and early pregnancy. This was documented by several well-designed studies (randomized controlled trials (RCTs), systematic reviews and meta-analysis) that emphasized both drug efficacy and safety	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes

NR	Reviewer	Page	Line	Comment	Action / Reply
				Regarding efficacy: The LOTUS I (Tournaye et al., 2017) and LOTUS II (Griesinger et al., 2018) phase III RCTs depicted that oral dydrogesterone is non-inferior to micronized vaginal progesterone (MVP) capsules and gel, respectively, for luteal phase support. There were comparable clinical pregnancy rates at 12 weeks gestation for both modalities of treatment Even more, Griesinger et al. (2020) performed a systematic review and individual participant data meta-analysis, which concluded that oral dydrogesterone was associated with significantly higher chance of ongoing pregnancy at 12 weeks of gestation (OR 1.32; 95% CI 1.08 to 1.61; P = 0.0075) and live birth (OR 1.28; 95% CI 1.04 to 1.57; P = 0.0214) compared to MVP Regarding safety: The same study by Griesinger et al. (2020) reported comparable safety outcomes compared to MVP Later on, Katalinic et al. (2022) performed a rigorous scoping review and meta-analysis including six RCTs. They reported no causal association between first-trimester dydrogesterone exposure and the risk of fetal abnormalities (RR 0.96; 95% CI 0.57–1.62). Again, Katalinic et al. (2024) conducted a systematic literature review and reached the same conclusion that dydrogesterone use is not associated with an increased risk of congenital anomalies beyond the baseline incidence attributable to environmental and genetic factors. This study represents the most robust and current level of evidence on this topic and resolved prior uncertainty emerging from low-quality and retracted publications. However, In contrast, Henry et al. (2025) conducted a pharmacovigilance analysis with a higher rate of disproportionate reporting of congenital anomalies associated with dydrogesterone (ROR: 5.4; 95% CI: 3.7–7.9) Yet, they denoted that their results should be taken with caution and their study has also limitations Indeed such findings warrant consideration, yet, pharmacovigilance data inherently suffer from several critical limitations: -Susceptibility to under-reporting and reporting bias being inherent to pharmacovigilance systems, which can significantly distort signal detection and impedes the measurement of the incidence of adverse drug reactions -Inability to establish causality, given the retrospective and spontaneous nature of adverse event reporting.	1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality".

NR	Reviewer	Page	Line	Comment	Action / Reply
				-Lack of clinical and biological context, which limits the interpretability and generalizability of such findings. - The lack of a true exposure denominator in the PV dataset impairs the ability to contextualize the number of reported cases relative to actual dydrogesterone use, introducing significant uncertainty into the interpretation of risk. As emphasized by Katalinic et al. (2024), in response to emerging concerns, pharmacovigilance data may serve as an initial signal for potential safety issues; however, its evidentiary value is inherently limited due to factors such as incomplete reporting, lack of standardized methodologies, and absence of adjustment for confounding risk factors. Consequently, there remains a critical need for high-quality, methodologically robust studies to ensure reliable safety assessments. We should be reminded that certain previous studies reported increased risk of congenital malformations with the use of dydrogesterone and were of low quality and were later retracted. This denotes that before a conclusion is drawn, further assessment by high quality studies is needed Accordingly, the inclusion of a footnote referencing the Henry et al. (2025) study in the current ESHRE guideline draft may inadvertently overemphasize a low-tier evidence source. This risks creating disproportionate concern among clinicians and patients, particularly in light of the extensive data from RCTs and meta-analyses that support dydrogesterone's safety. To ensure scientific rigor and uphold guideline credibility, it is recommended that the draft: Prioritize high-quality, peer-reviewed evidence derived from controlled and prospective settings. Contextualize lower-level evidence, such as pharmacovigilance signals, within their methodological constraints. Reconsider the inclusion of the aforementioned footnote, as its presence—absent appropriate qualification—may compromise the clarity and integrity of the guideline's messaging. In conclusion, up till now, the consistent and reproducible findings from robust clinical trials and systematic reviews strongly affirm dydrogesterone's efficacy and safety in ART and early pregnancy. The pharmacovigilance signal, while not to be dismissed, does not warrant equal evidentiary weight in clinical guideline development. Its current presentation risks misinterpretation and could undermine confidence in a clinically valuable therapeutic agent.	

NR	Reviewer	Page	Line	Comment	Action / Reply
78	Yasser El Kassar	156	4834-4836	Despite robust evidence from randomized controlled trials and meta-analyses LOTUS I (Tournaye et al., 2017) and LOTUS II (Griesinger et al., 2018) and Griesinger et al. (2020) affirming the superior efficacy and comparable safety of dydrogesterone, its designation as a conditional recommendation in clinical guidelines is unjustified. The inclusion of a cautionary note regarding dydrogesterone's safety, grounded in low-quality evidence, is inappropriate, particularly considering the retraction of earlier studies that erroneously suggested safety concerns. Pharmacovigilance (PV) studies, while useful for post-marketing surveillance, do not meet the rigorous standards of evidence-based medicine. These studies lack the methodological robustness of clinical trials, provide no clinical outcome data, and rely on reporting odds ratios (RORs), which are prone to reporting biases and lack a defined denominator population. Consequently, PV data cannot establish causality and should not independently influence clinical recommendations. In clinical practice, dydrogesterone is highly effective, well-tolerated, and offers superior patient convenience compared to alternatives such as vaginal pessaries or intramuscular injections, owing to its excellent oral bioavailability. Given the high-quality evidence supporting its efficacy and safety, dydrogesterone warrants a strong recommendation in clinical guidelines to reflect its therapeutic value accurately.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analyisis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up an increase in congenital malformation rate. A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
19	Shikha Gupta	156	4837	Dydrogesterone supported by robust studies including RCTs and meta analyses have shown efficacy and similar safety but only receive recommendation.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has

NR	Reviewer	Page	Line	Comment	Action / Reply
				A pharmacovigilance study cannot be ranked on the EBM –pyramid as it is not considered to be clinical entity (reporting odds ratio is not a clinical outcome).	concerns about potential congenital malformations, the recommendation is conditional.
39	Ahmed Elsayed Hassan Hamed Elbohoty	156	4837	The guideline does not mention any comment on synthetic progestins I strongly recommend that the guideline explicitly state that synthetic progestins such as medroxyprogesterone acetate, norethisterone, and levonorgestrel should not be used for luteal phase support in IVF. These compounds are unsupported by randomized controlled trials, lack evidence for efficacy in luteal support, may negatively affect endometrial receptivity and may result in congenital malformation and androgenic effects (Cicinelli E et al., 2003; Griesinger G et al., 2012; Practice Committee of the ASRM., 2015)	The GDG only considered regimens or medications currently in use for LPS.
44	Nisha bhatnagar	156	4837	Dydrogesterone is good for luteal phase support Have used it for last 25 years ,with excellent results Have not come across any malformations or defects coz of use of dydrogesterone (none of my patients have reported)	Individual experience and practice, while valuable is not a substitute for published clinical studies.
50	Ritesh Sinha	156	4837	Ensures favorable impacts and increases effectiveness. Shows effective treatment results.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
52	Manju Khemani	156	4837	I have been using Dydrogesteron for more than 20 years. Initially, in cases of recurrent abortion, than for IUi cases. I have not kept a record but I don't recall a single patient who has delivered a baby with hypospadias or a congenital heart defect over so many years.I have delivered many patients who have got ART done and I have delivered their babies but never come across any patients with these defects. A Pharmacopoeia vigilance study should not be taken as a final verdict. Maybe people did not report in the progesterone group. This will send	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
				the wrong message to doctors using dydrogesterone without any confirmation.	
58	Elena Grudnitskaya	156	4837	In light of the current evidence base (H. Tournaye, et al. 2017; G. Griesinger et al. 2018; G. Griesinger et al. 2020; Y. Atzmon et al. 2020; LN Vuong et al. 2021; J. Metello et al. 2022) dydrogesterone should be considered a first-line agent and strongly recommended for luteal phase support in assisted reproductive technologies. Direct comparative randomized trials have not revealed any negative impact of dydrogesterone on the development of congenital malformations in children, including congenital heart defects (H. Tournaye, et al. 2017; G. Griesinger et al. 2018)	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional. The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.
65	Aniruddha Bhattacharjee	156	4837	Dydrogesterone offers beneficial results and Generates strong performance. Strongly suggested and Ensures strong therapeutic impact	Individual experience and practice, while valuable is not a substitute for published clinical studies.
68	Farah Gari	156	4837	I use Dydrogestrone for luteal phase support in my daily practice and I haven't any adverse effect like congenital malformation.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
81	Dongzi Yang	156	4837	As an investigator involved in a randomized, open-label, multicenter international study on dydrogesterone, the findings from this research (Yang, D. Z., et al., 2020) demonstrated that within the Chinese mainland subpopulation, the incidence of congenital, familial, and genetic disorders in the fetal/neonatal population was numerically lower in the oral dydrogesterone group (4.6% [4/87]) compared to the micronized vaginal progesterone gel group (12.7% [9/71]). However, this observed	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies

NR	Reviewer	Page	Line	Comment	Action / Reply
				difference was not statistically significant and was not considered clinically relevant. RECOMMENDATION 1: Clinical guidelines should prioritize referencing high-quality clinical evidence.	can pick up a potential increase in congenital malformation rate.
85	Madhu Shrivastav	156	4837	Has good results and I am using it since long.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
86	Maneesha Jain	156	4837	I have never experienced any adverse effects with dydrogesetrone in my practice I am using this Molecule since starting from practice	Individual experience and practice, while valuable is not a substitute for published clinical studies.
87	Anima Prasad	156	4837	It is a wonder drug and I have faced no issues with it.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
97	Kanad Dev Nayar	157	4837	I have been using dydrogesterone in my clinical practice for more than 20 years. There has not been any incident of congenital anomalies in my practice. I administer dydrogesterone to my patients for luteal phase support, it is a cornerstone in my clinical practice since 2 decades.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
99	Manpreet Sharma	157	4837	Has good results and I am using it since long. I am prescribing it for progestrone supports in RPL and TM is found to very effective,	Individual experience and practice, while valuable is not a substitute for published clinical studies.
96	Sharda Jain	157	4837	I have been using dydrogesterone for luteal phase support for decades, and it continues to be a cornerstone in my clinical practice. Its targeted progestogenic action, excellent oral bioavailability, and lack of androgenic or glucocorticoid side effects make it an ideal choice for women requiring hormonal support—whether in cases of luteal insufficiency, infertility, or assisted reproductive techniques. Over the years, I have found it to be well-tolerated by my patients, with minimal to no side effects, which greatly improves adherence and comfort during	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
				treatment. The consistency of results, coupled with a strong safety profile, has made dydrogesterone a trusted and effective therapeutic option in the management of women's reproductive health. I truly believe in the molecule and its safety profile.	
89	Johannes Ott	156	4837	In our humble opinion, dydrogesterone should be recommended as an effective option for luteal phase support. High-quality randomized studies and meta-analyses demonstrated a superior live birth rate compared to vaginal progesterone with a comparable safety profile [Griesinger et al 2020 Plos One]. Recent data from pharmacovigilance databases indicated a possible association with malformations, although the causality of this is unclear due to methodological limitations and further clarification is required. The decision for use should therefore be made in the context of shared decision making, taking into account the proven benefit and the current data situation. We also take the liberty asking why is only dydrogesterone is subjected to this intensive safety screening by pharmacovigilance data in the guideline draft. Why did the authors chose to not include such data about other progestins, e.g. MPA in the PPOS protocol, or clomiphene citrate, for which there have also been discussions about possible teratogenic effects for decades? This approach in the draft guideline seems methodologically inconsistent.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.
111	T. Ramani Devi	156	4837	Dydrogesterone – Non-inferior to NMP in LPD	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, a strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital

NR	Reviewer	Page	Line	Comment	Action / Reply
					malformations, the recommendation is conditional.
124	Ayman Oraif	156	4837	DYDROGESTERONE HAS A CONDITIONAL RECOMMENDATION IN THE GUIDELINES DESPITE STRONG CLINICAL EVIDENCE	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
129	Roberto de Azevedo Antunes	156	4837	Why aren't there any clear recommendations stated regarding the dosage of dydrogesterone for LPS. The data is only addressed on line 4820 of the guideline, but it should be presented as a topic like the one regarding the dosages of progesterone for LPS.	The GDG has considered your comment, however, there is not enough evidence from RCTs to make a dose comparison.
54	Farrukh Naheed	156	4837-40	Recommendation regarding "Dydrogesterone" for LPS(2019) was taken as "Conditional". I have reviewed many studies a mentioned in your guidelines i.e. Kata linic et al, Henry et al and Li et al. None of the study proves any superiority for MVP over oral dydrogesterone. So conclusively, pharmaco vigilance based upon small meta analysis study unable to generate any consensus but can be individualized depending upon demographic features as well as the compliance of patient who underwent ART in order to achieve safety outcome. ART Practices (South Asian Countries) like Pakistan, India and Bangladesh had difficult views for MVP VS Oral Dydrogesterone due to socio-economic status of population and religious cultural adaptations, so the compliance for MVP is unable to prove superior over oral dydrogesterone. Many ART centers have their own experiences and none of them has adopted similar methodology for LPS either with MVP or Dydrogesterone. In term of live birth and Congenital Malformation, similar results of <5% with adverse outcome of pregnancy noted with both MVP and dydrogesterone. The	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.

NR	Reviewer	Page	Line	Comment	Action / Reply
				standard protocol followed in ESHRE guideline development hopefully does not qualify only pharmaco vigilance meta analysis studies which required further explanation regarding the efficacy and safety of oral dydrogesterone. So it should not remain as questionable and doubtful for ORAL DYDROGESTERONE. So, I requested you to kindly review that comment (conditional) for oral dydrogesterone which creates unnecessary confusions that was scientifically ruled out in LOTUS II study already.	
59	Ayman Hany Ahmed	156	4837-4840	The proposed ESHRE guideline includes the statement: "There are pharmacovigilance reports of association between dydrogesterone exposure and increased risk of congenital malformations, although the observed relations cannot necessarily be translated into a conclusion on causality.". The guideline group justified the conditional recommendation for dydrogesterone in luteal phase support (LPS) partly by "reflecting concerns about potential safety signals from recent pharmacovigilance data.". This pharmacovigilance study has several weaknesses that justify reassessing the cautionary statement 1. Inherent Limitations of Spontaneous Reporting Spontaneous reports are prone to under-reporting and external influences (like media or policies), which can introduce bias and affect data accuracy. 2. Inability to Establish Causality Disproportionality analysis estimates reporting risk (not actual risk), so it cannot prove a causal link between the drug and adverse events. 3. Susceptibility to Confounding Factors Patient-related factors like infertility and other exposures may confound results, making it hard to isolate the drug's specific effect. 4. Lower Level of Evidence for Risk Assessment Pharmacovigilance studies are mainly for detecting signals and generating hypotheses, not for providing strong evidence like RCTs. 5. Contradictory Evidence from Higher-Level Studies: The sources also present	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				findings from a systematic review and meta-analysis of published literature , including RCTs and observational studies, specifically assessing the risk of congenital anomalies with dydrogesterone exposure in the first trimester. This meta-analysis of RCTs found no statistically significant increased risk (RR 0.92, 95% CI 0.55; 1.55). Even when including observational studies, the overall risk remained non-significant (RR 1.11, 95% CI 0.73; 1.68). This study, which represents the "highest current level of evidence" for this question, provides "clear reassurance" that dydrogesterone is not a "relevant additional risk factor" for congenital anomalies. The authors of this meta-analysis explicitly state that pharmacovigilance data should primarily serve as a "sign giver" because of its limited evidence level. So, Kindly remove the conditional recommendation and the specific cautionary sentence about pharmacovigilance reports and the increased risk of congenital malformations. Alternatively, modify the sentence to clarify that pharmacovigilance data raised a signal of disproportionate reporting , but rigorous clinical studies have not confirmed an increased risk. Reiterate that evidence from clinical trials, including a recent meta-analysis, found no relevant additional risk of congenital anomalies, thus providing reassurance based on stronger evidence. The core message is that while signal detection is important for triggering further research, a clinical guideline aimed at providing evidence-based recommendations for practice should primarily rely on study designs capable of assessing actual risk and causality, like RCTs and systematic reviews thereof, which in this case, offer a reassuring safety profile for dydrogesterone regarding congenital anomalies.	
61	Hassan Mostafa Gaafar	156	4837-4840	Despite substantial evidence, including multiple randomized controlled trials (RCTs) and meta-analyses demonstrating superior efficacy and comparable safety, dydrogesterone has been assigned only a conditional recommendation. Conversely, micronized vaginal progesterone (MVP) receives a strong recommendation for luteal phase support (LPS) despite the absence of placebo-	The RCT by Atarieh et al. 2024 was taken out of the guideline because it has been retracted. The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to

NR	Reviewer	Page	Line	Comment	Action / Reply
				controlled RCTs supporting its efficacy. The inclusion of a footnote referencing safety concerns directly beneath a conclusive recommendation is a methodological anomaly, unprecedented in both this guideline and other ESHRE documents. This approach deviates from the stated methodological framework outlined by the guideline development group (GDG). Available safety data for alternative progestins have not been comprehensively assessed, indicating selective consideration of the evidence base. Although the guideline asserts a preference for systematic reviews and RCTs, it affords disproportionate weight to a lower-tier pharmacovigilance study concerning dydrogesterone—evidence that lies outside the conventional hierarchy of clinical evidence Pharmacovigilance analyses, such as reporting odds ratios (RORs), are not recognized as clinical evidence and cannot be positioned within the evidence-based medicine (EBM) pyramid. RORs serve only as hypothesis-generating signals and do not constitute proof of causality or clinical effect The methodological framework described on page 179 (line 5419) stipulates a hierarchical, iterative search strategy beginning with systematic reviews and RCTs. Despite this, lower-tier studies (e.g., Atarieh et al., Henry et al.) are given undue prominence over high-quality meta-analyses and large-scale clinical trials The guideline references a large retrospective cohort study by Li et al. (2024), which includes a mixed obstetric population, to question dydrogesterone's safety in ART. However, several robust safety studies in both ART and general pregnancy contexts are overlooked. Li et al. evaluated 18 pharmaceutical agents during pregnancy, including progesterone. If the findings are deemed applicable to dydrogesterone, they must be equally considered for progesterone The selection of RCTs is inconsistent. Methodologically outdated or flawed studies (e.g., Kupferminc 1999, Atarieh 2024) are included, whereas high-quality, contemporary trials supporting dydrogesterone—especially those conducted post-2017 reflecting modern ART practices—are excluded.	detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. Statements below the recommendations boxes have been included in this same guideline before. When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited and where the EBM-pyramid does not apply. Furthermore, it is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality".

NR	Reviewer	Page	Line	Comment	Action / Reply
				Of five recommendations concerning progesterone, four are based solely on 'Good Practice Points' (GPP) without accompanying evidence. In contrast, dydrogesterone—despite being evaluated in two Phase 3 trials—is not granted a strong recommendation. The 2015 Cochrane review, cited in support of progesterone, includes studies involving dydrogesterone. Nonetheless, the guideline separates the conclusions in a manner that lacks justification. Basing a contemporary guideline on data derived predominantly from early 2000s studies—when ART protocols differed substantially—fails to reflect current clinical practice. The guideline's safety conclusions regarding dydrogesterone rest on two studies with significant limitations: Henry et al. is subject to temporal ambiguity, exposure misclassification, and selection bias . VigiBase data are not designed for hypothesis testing, let alone causal inference. Li et al. lacks sufficient methodological clarity and rigor, rendering its findings unreliable for informing clinical guidelines. Several studies report reduced maternal tolerability with progesterone (e.g., Bespalova et al. 2021; Astrankantseva et al. 2021; Zhao et al. 2025), yet data on adverse events in progesterone-treated women are excluded from the analysis. Additional safety concerns raised in the literature regarding other progestogens (e.g., Carmichael et al. 2005) are not reflected in the guideline, despite the stated absence of long-term data on offspring health. Although the two available studies on fetal safety following dydrogesterone use are included, yet broader safety literature is not reviewed with equivalent depth for comparator agents. The term "reporting OR" (used in relation to the pharmacovigilance study) may be misleading. While it visually resembles 'odds ratio,' RORs are fundamentally different, offering insight into reporting frequencies rather than actual clinical risk. The assertion that dydrogesterone warrants additional safety scrutiny due to its	

NR	Reviewer	Page	Line	Comment	Action / Reply
				synthetic nature overlooks its origin and synthesis, which parallel that of natural progesterone. Dydrogesterone is the only progestin—and the only pharmaceutical agent within the entire guideline—accompanied by a safety footnote, creating a disproportionate perception of risk in the absence of a consistent evidentiary rationale While the guideline acknowledges that a causal link between dydrogesterone and congenital malformations has not been established, it nevertheless introduces pharmacovigilance data using speculative and cautionary language, which undermines the clarity and objectivity expected of evidence-based guidance. Although the guideline explicitly states that no consensus was reached among the GDG members, it still includes a cautionary footnote that may unduly amplify concern in the absence of unified agreement	
81	Dongzi Yang	156	4838	Guideline Acknowledges the lack of proven causality but also drawing a confusing safety footnote.	The safety footnote is the explanation why the recommendation is conditional in stead of strong.
84	Umesh N Jindal	156	4838	The safety concerns regarding dydrogestrone have been pointed out. There is only one pharmacovigilace study and not studies	This was adjusted.
111	T. Ramani Devi	156	4838	Agreed	Thank you.
115	Hassan Sallam	156	4838	• No RCTs are included to support the recommendations which will be very alarming to the physicians and public, despite lack of evidence-based support • On the contrary, findings of RCTs did not report any increased incidence of anomalies • Pharmacovigilant studies are not included in the hierarchy of the evidence pyramid • The methodologies used in the Li et al study (2024) and the Henry et al study (2025) do not support the generalization of the findings	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic systematic review, and where the EBM-pyramid does not apply.

NR	Reviewer	Page	Line	Comment	Action / Reply
				• This footnote combined with singling out dydrogesterone, omitting reports on other luteal support agents, inclusion of outdated studies (e.g., Kupferminc 1999) and omission of studies on long-term safety of progestogens (e.g. Carmichael et al, 2005) can send the wrong message to the physicians and to the public and may diminish from the value of this otherwise beautiful document • I suggest rephrasing the findings in a more evidence-based non-alarming manner to properly inform the physician who do not have the time to go into the details of the methodology and will jump to the wrong conclusions, while in the same time preserving the paramount safety interests of the patients • We do not want to repeat what happened with the WHI study	
124	Ayman Oraif	156	4838	LI ET AL. EXAMINES 18 DRUGS USED IN PREGNANCY, INCLUDING PROGESTERONE. IF IT IS CITED AS SAFETY EVIDENCE FOR DYDROGESTERONE, IT SHOULD BE EQUALLY APPLIED TO PROGESTERONE A PHARMACOVIGILIANCE STUDY CAN NOT BE RANKED ON THE EBM-PYRAMID AS IT NOT CONSIDERED TO BE CLINICAL EVIDENCE (REPORTING ODDS RATIO IS NOT A CLINICAL OUTCOME).	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
126	Emre Goksan Pabuccu	156	4838	For dydrogesterone: There are pharmacovigilance reports of association between dydrogesterone exposure and increased risk of congenital malformations, although the observed relations cannot necessarily be translated into a conclusion on causalityàWe do not support the inclusion of a safety warning based solely on a single pharmacovigilance report with methodological limitations. Relying on such data does not justify a cautionary statement—especially when level 1 evidence from a recent	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analyisis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-

NR	Reviewer	Page	Line	Comment	Action / Reply
				meta-analysis has demonstrated no increase in fetal risks associated with dydrogesterone use. 10	based studies can pick up an increase in congenital malformation rate.
151	Alexandra Kohl Schwartz	156	4838	This statement is alarmistic, too vague to be informative and borders on being trivial. Pharmacovigilance reports are an expected outcome of spontaneous reporting systems for any widely used drug, and their presence alone does not constitute a meaningful or actionable safety signal. The phrase “cannot necessarily be translated into a conclusion on causality” is awkward and imprecise; it is trivial that no causal inference can be drawn from such data. As currently written, the sentence acknowledges uncertainty about relevance but still implies a concern—without providing any context regarding the nature, frequency, or consistency of the reported malformations or how this reflects with other explicit safety evidence for the drug and within the field of ART. This creates a misleading impression of risk and may overstate what is, at most, a very speculative association. Selective inclusion of such remarks prominently in the summary of recommendations risks distorting the comparative safety profile of drugs used in ART and undermines the guideline’s neutrality and credibility. Suggested revision: Remove this safety alert from the guideline summary.	The GDG does not agree with the reviewer. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.
151	Alexandra Kohl Schwartz	156	4838	The current statement incorrectly and in an alarmistic way combines the terms ‘increased risk’ and ‘association’ in the context of a pharmacovigilance disproportionality analysis. This implies a risk-based association, which is misleading. Disproportionality analyses assess patterns in adverse event reporting—not incidence, risk, or causality. Apparent disproportionality can result from numerous artefacts, including notoriety bias, stimulated reporting, underreporting of other adverse events, or changes in reporting practices over time. Moreover,	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				practice patterns such as national licensing status, indication-specific use, and differences in surveillance intensity across countries can all distort PV signal strength. Disproportionality analyses on PV data are a hypothesis-generating tool for early signal detection. The statement is thus inappropriate. Suggested revision: Remove this safety alert from the guideline summary.	
151	Alexandra Kohl Schwartz	156	4838	The use of the plural reports implies to the reader that there are at least two, or possibly multiple, consistent pharmacovigilance (PV) signals, which is not the case. PV is a multifaceted discipline encompassing a wide range of data sources and methods beyond the spontaneous report-based disproportionality analysis of Henry et al. that the statement refers to. These include observational studies (e.g., cohort and case–control designs), clinical trial data, registries, in vitro and in vivo (animal) studies, mechanistic toxicology, and structured causality assessments. The disproportionality analysis relations cannot necessarily be translated into a conclusion on causality.” mentioned in the guideline (Henry et al.) represents only a minor and exploratory element within this broader framework. It is a hypothesis-generating tool for early signal detection and, on its own, cannot establish risk. Presenting it without specifying the nature of the study (i.e. early signal detection) and outcome (i.e. ROR) and proper context is therefore misleading. Furthermore, dydrogesterone has been in widespread clinical use for decades, and cumulative PV data—together with robust clinical evidence—have not demonstrated a consistent association with congenital anomalies. Isolating a single disproportionality analysis, without referencing the totality of evidence, risks overstating the signal and may compromise the neutrality and	The sentence in the justification was corrected. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				scientific impartiality of the guideline. It is also important to distinguish formal pharmacovigilance from observational pharmacoepidemiology. The DEBC pharmacoepidemiology study (Li et al., 2024), cited later in the document, is a prospective pregnancy cohort with several methodological limitations. It does not involve spontaneous reporting or signal detection, and describing it as PV data conflates distinct approaches and exaggerates the availability of PV evidence. Suggested revision: Remove this safety alert from the guideline summary.	
30	Dubrovina Svetlana	157	4838-4840	This study examined safety signals from the international adverse event reporting system Vigibase – this type of study by definition does not allow conclusions to be made about the cause-and-effect relationship between the use of the drug and the occurrence of adverse events. This is stated in the publication itself by A. Henry, as well as in the rules for using the studied Vigibase database.	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
90	Zürcher Kreis working group	156	4838-4840	We conclude that the hypothesis of an increased malformation risk in users of dydrogesterone during pregnancy is based essentially on just one disproportionality analysis from one pharmacovigilance database. This evidence is weak and does not allow any statement in this sense. In order not to unsettle the patients using dydrogesterone during their pregnancy, we suggest to reframe the paragraph where it is suggested that dydrogesterone might harm the offspring. Nevertheless, in pregnancy, the natural hormone progesterone should be preferred if possible, respecting the basic principal to avoid synthetic molecules in pregnant women see word document	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline. There is no suggestion in the justification that dydrogesterone might harm the offspring.

NR	Reviewer	Page	Line	Comment	Action / Reply
121	Srilatha Gorthi	156	4838-4840	I am a practicing Reproductive Medicine Consultant with over 20 years of experience, out of which over a Decade was spent in the UK. Duphaston has been safely used for over 60 years with the use in India for at least 2 decades where it's the most prescribed progesterone for pregnancy support. It is generally well tolerated and is considered safe in pregnancy through the experience of my peers and seniors as well as evidenced through LOTUS I and II trials. In our decade long experience Fertility Centre at Revive Clinics, Hyderabad, India, we retrospectively analyzed the fetal outcomes for over 3000 cycles in the last decade and found that 4 fetuses had hypospadias and 5 had cardiac anomalies. With only one woman linked to Dydrogesterone (Duphaston, Abbott pharma) during luteal phase and early pregnancy while the rest were on Vaginal progesterone gel 8% (Crinone gel, Merck pharma). This is a negligible incidence with no substantial increase in fetal anomalies. I urge you to consider removing the conditional clause and just mention the pharmacovigilance as a mere observational study. This would help not only the women but also the doctors, especially in the Indian subcontinent as ESHRE guidelines will impact our work in a major way.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
111	T. Ramani Devi	156	4839	Agreed	Thank you.
111	T. Ramani Devi	156	4840	Agreed	Thank you.
100	Reassure group	157	4842	The current statement that "oral dydrogesterone has similar live birth/ongoing pregnancy rates compared to progesterone" is inconsistent with the evidence presented earlier in the chapter (lines 4821–4828), with the methodology outlined in the ESHRE guideline (line 5436), and with the actual comparisons cited. According to the cited individual participant data (IPD) meta-analysis (Griesinger et al., 2020), oral dydrogesterone was associated with a statistically significantly	The IPD analysis of 2 RCTs shows a higher live birth rate, the combination of all 9 RCTs shows no significant difference in live birth rate (OR 1.14, 95% CI 0.99-1.32, 5 RCTs, 4470). Therefore, the text in the justification is correct.

NR	Reviewer	Page	Line	Comment	Action / Reply
				higher live birth rate (OR 1.28, 95% CI 1.04–1.57) and ongoing pregnancy rate (OR 1.32, 95% CI 1.08–1.61) compared to vaginal micronised progesterone. In addition, the same publication reports a combined IPD and aggregate data meta-analysis, which also shows a statistically significant increase in live birth rate and ongoing pregnancy in favour of dydrogesterone. As the guideline consistently relies on individual patient data (IPD) meta-analyses where available, this implicit methodological preference should be applied consistently across all comparisons—including those involving dydrogesterone. Furthermore, the focus should remain on live birth rate, as per the guideline's own stated outcome hierarchy (line 5436). It must also be noted that the comparison is not against "progesterone" in general, but specifically against vaginal micronised progesterone. Selectively downplaying statistically significant findings in favour of dydrogesterone, while misrepresenting the comparator group, risks undermining the objectivity and methodological consistency of the guideline. Suggested revision: Revise the statement to accurately reflect the statistically significant findings from both the IPD and the combined IPD/aggregate data meta-analyses. Refer specifically to the live birth rate as per ESHRE rules. Clarify that the comparison pertains only to oral dydrogesterone versus micronised vaginal progesterone	
140	German Society of Reproductive Medicine	157	4842	The current statement that "oral dydrogesterone has similar live birth/ongoing pregnancy rates compared to progesterone" is inconsistent with the evidence presented earlier in the chapter (lines 4821–4828), with the methodology outlined in the ESHRE guideline (line 5436), and with the actual comparisons cited. According to the cited individual participant data (IPD) meta-analysis (Griesinger et al., 2020), oral dydrogesterone was associated with a statistically significantly	The IPD analysis of 2 RCTs shows a higher live birth rate, the combination of all 9 RCTs shows no significant difference in live birth rate (OR 1.14, 95% CI 0.99-1.32, 5 RCTs, 4470). Therefore, the text in the justification is correct.

NR	Reviewer	Page	Line	Comment	Action / Reply
				higher live birth rate (OR 1.28, 95% CI 1.04–1.57) and ongoing pregnancy rate (OR 1.32, 95% CI 1.08–1.61) compared to vaginal micronised progesterone. In addition, the same publication reports a combined IPD and aggregate data meta-analysis, which also shows a statistically significant increase in live birth rate and ongoing pregnancy in favour of dydrogesterone. As the guideline consistently relies on individual patient data (IPD) meta-analyses where available, this implicit methodological preference should be applied consistently across all comparisons—including those involving dydrogesterone. Furthermore, the focus should remain on live birth rate, as per the guideline's own stated outcome hierarchy (line 5436). It must also be noted that the comparison is not against "progesterone" in general, but specifically against vaginal micronised progesterone. Selectively downplaying statistically significant findings in favour of dydrogesterone, while misrepresenting the comparator group, risks undermining the objectivity and methodological consistency of the guideline. Suggested revision: Revise the statement to accurately reflect the statistically significant findings from both the IPD and the combined IPD/aggregate data meta-analyses. Refer specifically to the live birth rate as per ESHRE rules. Clarify that the comparison pertains only to oral dydrogesterone versus micronised vaginal progesterone.	
151	Alexandra Kohl Schwartz	157	4842	The current statement that "oral dydrogesterone has similar live birth/ongoing pregnancy rates compared to progesterone" is inconsistent with the evidence presented earlier in the chapter (lines 4821–4828), with the methodology outlined in the ESHRE guideline (line 5436), and with the actual comparisons cited. According to the cited individual participant data (IPD) meta-analysis (Griesinger et al., 2020),	The IPD analysis of 2 RCTs shows a higher live birth rate, the combination of all 9 RCTs shows no significant difference in live birth rate (OR 1.14, 95% CI 0.99-1.32, 5 RCTs, 4470). Therefore, the text in the justification is correct.

NR	Reviewer	Page	Line	Comment	Action / Reply
				oral dydrogesterone was associated with a statistically significantly higher live birth rate (OR 1.28, 95% CI 1.04–1.57) and ongoing pregnancy rate (OR 1.32, 95% CI 1.08–1.61) compared to vaginal micronised progesterone. In addition, the same publication reports a combined IPD and aggregate data meta-analysis, which also shows a statistically significant increase in live birth rate and ongoing pregnancy in favour of dydrogesterone. As the guideline consistently relies on individual patient data (IPD) meta-analyses where available, this implicit methodological preference should be applied consistently across all comparisons—including those involving dydrogesterone. Furthermore, the focus should remain on live birth rate, as per the guideline's own stated outcome hierarchy (line 5436). It must also be noted that the comparison is not against "progesterone" in general, but specifically against vaginal micronised progesterone. Selectively downplaying statistically significant findings in favour of dydrogesterone, while misrepresenting the comparator group, risks undermining the objectivity and methodological consistency of the guideline. Suggested revision: Revise the statement to accurately reflect the statistically significant findings from both the IPD and the combined IPD/aggregate data meta-analyses. Refer specifically to the live birth rate as per ESHRE rules. Clarify that the comparison pertains only to oral dydrogesterone versus micronised vaginal progesterone.	
43	Jyothi G S	157	4843	Through extensive clinical application, I have observed remarkable results in my patients with recurrent pregnancy loss and threatened miscarriage. I have seen patients who have experienced multiple losses go on to have successful pregnancies, and those whom I have prescribed dydrogesterone for luteal phase support achieve positive outcomes due	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
				to the targeted action of the molecule. Its oral bioavailability and tolerability have made it my preferred choice. Therefore, I urge the ESHRE federation to change the recommendation for dydrogesterone from “conditional” to “strong” so that many patients can benefit from this treatment.	
43	Jyothi G S	157	4843	Maternal safety beyond observation of enhanced patient satisfaction, a body of evidence indicates that dydrogesterone demonstrates superior tolerability, which translates to an improved maternal safety profile during gestation. This finding consistent with supplementary data from Henry et al., suggests dydrogesterone is associated with diminished risks of adverse events such as cervical shortening, bacterial vaginosis, bloody discharge, abdominal pain and spontaneous abortion when compared to other progestogenic agents.	The GDG does not doubt the efficacy of dydrogesterone for LPS. However, the GDG has concerns about potential congenital malformations, therefore the recommendation is conditional.
41	Tian-Min Ye	157	4843-4863	A more comprehensive and balance inclusion of safety data will help clinics make proper decision for patients.	The GDG has included both the data from the systematic review by Katalinic 2024, as well as the 2 studies reporting a negative safety signal.
41	Tian-Min Ye	157	4843-4863	The citation of pharmacovigilance study (Henry et al., 2024) is not consistent with the methodology description of the guideline (page 179, line 5419). Reporting odds ratio is not even a clinical outcome actually. Low ranked studies in efficacy (Atarieh et al., 2024) and safety (Henry et al., 2024) are given more weight than high-quality clinical trials or meta-analysis, such as LOTUS I, LOTUS II, and meta-analysis by Katalinic et al., 2024. Reference: 1. Henry et al., Birth defects reporting and the use of dydrogesterone: a disproportionality analysis from the World Health Organization pharmacovigilance database (VigiBase). Hum Reprod Open 2025;2025:	The RCT by Atarieh et al. 2024 was taken out of the guideline because it has been retracted. The Henry 2025 study is cited in the justification, not the evidence section. One of the components to determine a recommendation is the balance between benefit and harms. Since the GDG does not want to disregard the signals from the Henry 2025 and Li 2024 studies, the recommendation cannot be strong.

NR	Reviewer	Page	Line	Comment	Action / Reply
				hoae072. 5018 2. Atarieh et al., Comparison of the effect of dydrogesterone and natural micronized progesterone for luteal-phase support in assisted reproductive technology cycles: A single-blind randomized clinical trial study. Health Sci Rep 2024;7: e2296. 3. Katalinic et al., No additional risk of congenital anomalies after first-trimester dydrogesterone use: a systematic review and meta-analysis. Hum Reprod Open 2024;2024: hoae004.	
41	Tian-Min Ye	157	4843-4863	In our previous study focusing on hormone replacement therapy FET (HRT-FET) cycles, we compared the clinical outcomes between dydrogesterone and vaginal microsomal progesterone (MVP) gel. Each group achieved 47 newborns, with no birth defect observed (Ye et al., 2024). In another study, we found that among 226 HRT-FET cycles with 54 newborns, dydrogesterone resulted in no birth defects. In contrast, among 263 cycles with 81 newborns where treatment included both dydrogesterone and MVP, 3 cases of birth defects were observed (P>0.05) (Huang et al., 2024). Similarly, propensity-matched analysis, conducted with 337 fresh cycles per group, revealed no birth defects among 92 newborns with dydrogesterone, in contrast to one polydactyly case observed among 89 live births with MVP gel (Huang et al., 2023). Reference: 1. Tian-Min Ye et al., Comparison between oral dydrogesterone versus micronized vaginal progesterone gel in clinical outcome within the first HRT-FET cycle: a retrospective analysis. Arch Gynecol Obstet. 2024 May;309(5):2167-2173. 2. Huang YF, Luo LD, Ding SF, Lin SX, Ye TM. Effectiveness of Oral Dydrogesterone Tablets during Hormone Replacement Therapy-Frozen Embryo Transfer. J Med Res 2024; 53: 141-145.	The incidence of congenital malformations in IVF/ICSI is very low. The studies mentioned by the reviewer are not powered to detect small differences in the congenital malformation rate. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				3. Huang YF, Ding SF, Lin SX, Huang LF, He SB, Ye TM. Clinical efficiency of oral dydrogesterone tablets on luteal phase support in antagonist protocol fresh embryo transfer cycles. Chin J Wom Chil Healt Res 2023; 34: 69-74.	
117	Tatyana Pestova	157	4846	In our work, we pay great attention to the health of children after ART. I focus on both the highest level of evidence (systematic reviews and meta-analyses) and real clinical practice data. We also analyze the identified disorders, congenital malformations in children born after ART in relation to population data. Today, there is a lot of scientific data that discusses the health of children conceived after ART. The most common factor that could lead to an increase in congenital malformations was the presence of concomitant pathology in mothers: endometriosis (1.16 aOR [95% CI 1.10-1.22], P < 0.0001), PCOS (1.20 aOR [95% CI 1.08-1.34], P = 0.001) or POF (1.52 aOR [95% CI 1.23-1.88], P = 0.0001), also underlying maternal infertility can contribute to an increased risk of IVF-associated defects1. The results of a systematic review and meta-analysis (Katalinic A, 2024) did not reveal an additional risk of congenital anomalies after the use of dydrogesterone in the first trimester2. The results of a real-life clinical practice study (Nazarenko T. et al., 2025) confirmed the previously established safety profile of dydrogesterone for the fetus and newborn3. Congenital malformations and anomalies were reported in 13 (3.7%) fetuses/neonates: 2 fetuses and 11 neonates, and 3 reports of congenital pneumonia in 3 neonates. In 1-12 months of neonatal observation, all cardiac anomalies resolved spontaneously in 9 neonates; 1 event was resolved by surgical treatment. In total, during the first year, all cardiac anomalies resolved (in 10 children)3. In addition, in the study by Nazarenko T et al., 2025, we observed a correlation between the use of dydrogesterone and favorable obstetric and perinatal outcomes: the frequency of preeclampsia was 3 times lower compared to the MOH data on this complication3.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				I suggest to reconsider the recommendation of dydrogesterone for luteal phase support in IVF cycles, assigning the same strength of recommendation as progesterone.	
152	Liudmila Stavinskaia	157	4846	Since dydrogesterone is a retroprogesterone, it differs from progesterone in its higher bioavailability, which ensures a prolonged and stable systemic effect. The daily dose of oral dydrogesterone is 20 times lower than that of progesterone. Since dydrogesterone has high selectivity for progesterone receptors, the likelihood of side effects caused by effects on other receptors is reduced. (Kuhl H. Pharmacology of estrogens and progestogens: influence of different routes of administration. Climacteric. 2005; 8(Suppl. 1): 3-63. Schindler A.E., Campagnoli C., Druckmann R., Huber J., Pasqualini J.R., Schweppe K.W., Thijssen J.H. Classification and pharmacology of progestins. Maturitas. 2008; 61(1-2): 171-80.) In our clinical practice, we recommend 10 mg of dydrogesterone three times a day starting on the day of egg retrieval and continuing for 10 weeks (if pregnancy is confirmed).	There are several hypotheses on how dydrogesterone would cause congenital malformations, one is related to its increased binding potential to the progesterone receptor, one other relates to the transformation of dydrogesterone into additional non-biological compounds such as 20-alpha-dihydrodydrogesterone.
100	Reassure group	157	4850	This statement is factually misleading. Of the 6 RCTs included in the Katalinic et al. (2024) meta-analysis, 48.2% of patients (764/1,584) were from studies in the ART/luteal phase support setting, and 51.8% from studies addressing miscarriage. Describing the dataset as “mainly in couples with recurrent miscarriage” therefore misrepresents the evidence base, which is roughly balanced between the two indications. Furthermore, when observational data are included by Katalinic (total N = 3,780), 78.3% of patients (2,960) were treated for ART-related indications, while only 21.7% were in miscarriage cohorts. These proportions should be clearly stated, as they are critical for interpreting the applicability of the safety signal to ART populations. The current wording underplays this, again questioning the neutrality of the document, and should be corrected for transparency and accuracy. There is also a lack of contextualization with broader safety data in the field. As	The text in the justification was adapted to reflect the different populations feeding the numbers in the Katalinic meta-analysis.

NR	Reviewer	Page	Line	Comment	Action / Reply
				highlighted in a comprehensive review by van Gelder et al. (2014), the evidence base for teratogenic risk is absent or severely limited for the majority of drugs used during early pregnancy — including many fertility treatments. In contrast, dydrogesterone is one of the comparatively better-studied compounds in this context, with multiple robust randomized controlled trials and observational datasets reporting on congenital outcomes. Therefore, the existing evidence on dydrogesterone—while not definitive—offers a stronger empirical foundation than that available for many other agents and other progestin formulations routinely used in ART. This context should be reflected in the guideline text to avoid disproportionate scrutiny of a compound with relatively robust safety data. Suggested revision, rephrase: 'A recent systematic review and meta-analysis including six randomized controlled trials (n = 1512) reported a relative risk of 0.92 (95% CI 0.55–1.55) for congenital malformations in pregnancies exposed to dydrogesterone compared to placebo, no treatment, or other interventions (Katalinic et al., 2024). While the confidence interval is wide and includes the possibility of both reduced and increased risk, the point estimate does not suggest an elevated risk. Taken together with observational data, this represents one of the more substantial clinical evidence bases in the field and clinical data do not indicate a signal of concern regarding offspring safety.'	
140	German Society of Reproductive Medicine	157	4850	This statement is factually misleading. Of the 6 RCTs included in the Katalinic et al. (2024) meta-analysis, 48.2% of patients (764/1,584) were from studies in the ART/luteal phase support setting, and 51.8% from studies addressing miscarriage. Describing the dataset as "mainly in couples with recurrent miscarriage" therefore misrepresents the evidence base, which is roughly balanced between the two indications. Furthermore, when observational data are included by Katalinic (total N = 3,780), 78.3% of patients (2,960) were treated for ART-related indications, while only 21.7% were in miscarriage cohorts. These proportions should be clearly stated, as they are critical for interpreting the applicability of the safety signal to ART populations. The current wording underplays this, again questioning the neutrality of the document, and should be corrected for transparency and	The text in the justification was adapted to reflect the different populations feeding the numbers in the Katalinic meta-analysis.

NR	Reviewer	Page	Line	Comment	Action / Reply
				accuracy. There is also a lack of contextualization with broader safety data in the field. As highlighted in a comprehensive review by van Gelder et al. (2014), the evidence base for teratogenic risk is absent or severely limited for the majority of drugs used during early pregnancy — including many fertility treatments. In contrast, dydrogesterone is one of the comparatively better-studied compounds in this context, with multiple robust randomized controlled trials and observational datasets reporting on congenital outcomes. Therefore, the existing evidence on dydrogesterone—while not definitive—offers a stronger empirical foundation than that available for many other agents and other progestin formulations routinely used in ART. This context should be reflected in the guideline text to avoid disproportionate scrutiny of a compound with relatively robust safety data. Suggested revision, rephrase: 'A recent systematic review and meta-analysis including six randomized controlled trials (n = 1512) reported a relative risk of 0.92 (95% CI 0.55–1.55) for congenital malformations in pregnancies exposed to dydrogesterone compared to placebo, no treatment, or other interventions (Katalinic et al., 2024). While the confidence interval is wide and includes the possibility of both reduced and increased risk, the point estimate does not suggest an elevated risk. Taken together with observational data, this represents one of the more substantial clinical evidence bases in the field and clinical data do not indicate a signal of concern regarding offspring safety.'	
131	Teraporn Vutyavanich	157	4855	The current pharmacovigilance safety database reports by Henry et al., 2025 suggests a potential safety concern. However, this should be interpreted with caution and weighed against evidence from other studies, especially meta-analysis The REASSURE Study II (Katalinic et al., 2024), a recent meta-analysis of 8 studies (including 6 RCTs), found no significant association between dydrogesterone use in the first trimester and congenital anomalies (RR 0.92; 95% CI 0.55–1.55).	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect an increase in congenital malformations. The meta-analyisis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up an increase in congenital malformation rate.

NR	Reviewer	Page	Line	Comment	Action / Reply
				In my IVF practice and experiences, I have not observed any such safety signals. A more balanced interpretation that incorporates both clinical evidence and real-world experience would strengthen the guideline's relevance to patient care.	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality".
135	Sadiah Ahsan	157	4855	There are pharmacovigilance reports of association between dydrogesterone exposure and increased risk of congenital malformations, although the observed relations cannot necessarily be translated into a conclusion on causality. We have been using Dydrogestrone for LPS (30mg /d) at Concept Fertility Center, Karachi, Pakistan (an Australian chain branch) since 2018. Our initial audit and comparison of two sets of patients of 4 different Consultants, (2 Consultant giving only vaginal Progesterone and 2 Consultants giving both Vaginal Progesterone and Dydrogesterone in 30 mg in 8 hrly divided doses per day in addition to Vaginal progesterone , confirmed the safety for both mother and newborn baby. There continues to a better pregnancy rate of around 5% in the Dydrogestrone group. This is particularly important as in a LMIC country like Pakistan IVF is funded privately and even 1% better results are important for the couple. Dydrogesterone LPS is a lot cheaper and easier to manage, therefore we must re-evaluate evidence against it for reasons of doubtful increased congenital anomaly risk, being attributed to Dydrogesterone alone in the quoted studies. I hope the guideline will remove "conditional" with this recommendation. We are compiling our data, hoping to publish soon.	Publishing your data would be very helpful. However, at this moment, a strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
22	Qinjie Tian	157	4855-4858	This guideline referred to a global pharmacovigilance study published in 2025. Based on case-non-case studies conducted in the pharmacovigilance database, the authors claimed that dydrogesterone should be cautiously used in luteal support of ART. The conclusion of this	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of

NR	Reviewer	Page	Line	Comment	Action / Reply
				study may be misleading and is not so scientific. This kind of case-non-case study is only used for the detection of early signals and provides direction for hypothesis, but not for the confirmation of causality. Therefore, it is not scientific to conclude that there is a risk from the results of this study. Further clinical studies are needed to confirm the signal. The authors did not follow the guidelines for the use of data from the VigiBase, a database managed by the Uppsala Monitoring Center (UMC, WHO Drug Monitoring Center), did not conduct a full scientific assessment for the diversity of data sources, and improperly concluded the causality. Reference of this study is not recommended.	ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
100	Reassure group	157	4855-4858	Lack of Contextualization of Pharmacovigilance Data The current presentation of the Henry et al. pharmacovigilance (PV) study in the guideline lacks essential scientific context and is methodologically unbalanced. No comparable PV data are cited for other progestins or ART-related medications, resulting in a selective focus on dydrogesterone. This lack of systematic comparison undermines the guideline's neutrality and conveys a misleading impression of uniquely elevated risk, which is not supported by the broader evidence base. Most critically, the guideline fails to acknowledge that the longstanding allegation linking therapeutic progestins to non-genital congenital malformations has been conclusively refuted already more than 20 years ago. As reviewed in Brent et al., 2005 (DOI: 10.1542/peds.2004-1951), over two decades of rigorous epidemiological, animal, and mechanistic studies have demonstrated no causal relationship between progestins and malformations such as cardiac or limb defects. In fact, the U.S. FDA rescinded its warning on progestin use in 1999 after concluding that "clinically utilized progestational drugs do not cause nongenital malformations." Citing a disproportionality signal from a PV database in isolation, without appropriate context or triangulation with clinical, mechanistic, and epidemiologic evidence, is not only scientifically flawed—it risks generating	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				undue alarm among clinicians and patients. This is particularly problematic in the context of fertility treatment, where reassurance, evidence-based decision-making, and trust in treatment protocols are essential. There is also an ethical responsibility to avoid presenting data in a way that could inadvertently generate unwarranted concern, psychological burden, or treatment hesitancy among patients and clinicians. Furthermore, the current presentation fails to address genital malformations, which represent the primary theoretical concern regarding structural malformations associated with hormonally active substances during pregnancy. The omission of this more biologically plausible outcome underscores a lack of conceptual coherence in how fetal safety is currently framed within the guideline. Finally, the guideline does not adequately consider potential long-term effects of prenatal hormone exposure on growth, metabolic function, or endocrine health in adolescence and adulthood—an evidence gap that applies not only to dydrogesterone but to all progestins used in ART (doi.org/10.1038/s41598-020-78976-x; doi.org/10.1111/1471-0528.16608.). Suggested revision: The single pharmacovigilance study (Henry et al.) should be appropriately contextualized. Most importantly, the guideline should acknowledge that concerns about progestational agents (including progestins used in contraceptive pills) and non-genital congenital malformations have been scientifically refuted (DOI: 10.1542/peds.2004-1951), and that regulatory authorities such as the U.S. FDA have withdrawn prior warnings. Patients should thus be reassured and not alarmed regarding non-genital malformations and progestin exposure in pregnancy. Without this context, the presentation is unbalanced and may lead to misinterpretation of risk.	
100	Reassure group	157	4855-4858	The statement citing a significant disproportionality in birth defect reporting associated with dydrogesterone (Henry et al., 2025) introduces a reporting odds ratio (ROR)—a metric derived from spontaneous reporting systems. However, ROR is not a clinical outcome, nor is it among the predefined outcome measures specified in the ESHRE	A dedicated search strategy for pharmacovigilance studies and re-calculation into explicit qualifiers is both not in line with the ESHRE methodology of guideline development. In 2019, the recommendation for

NR	Reviewer	Page	Line	Comment	Action / Reply
				guideline methodology. This presents several problems for interpretation, consistency and neutrality: 1. ROR reflects reporting patterns, not incidence, prevalence, or comparative risk. It is influenced by reporting behaviour, media attention, and regulatory actions. 2. Most clinical readers are not familiar with pharmacovigilance metrics like ROR or disproportionality analysis, and may mistakenly interpret a statistically significant ROR as evidence of an actual increase in risk. 3. The inclusion of ROR as a stand-alone result, without the necessary appropriate contextualization or methodological explanation, risks misleading readers and undermines the scientific clarity and clinical usability of the guideline. Suggested revision: If pharmacovigilance data using disproportionality analysis (such as RORs) are included in the guideline, this should be clearly justified in the methods section and supported by a dedicated subsection on search strategy and pharmacovigilance methodology, including a plain-language explanation of how RORs differ from clinical effect estimates and what limitations apply. Otherwise, the Henry et al. citation should be either removed or rewritten with explicit qualifiers and context.	dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.
107	Emad Darwish	157	4855-4858	The pharmacovigilance study is a case by case report descriptive study. The total number of cases with congenital anomalies is not large enough to give solid conclusion or a causality relation between the anomaly and the dyderogesterone . Also Pharmacovigilance studies, which report associations such as reporting odds ratios, do not constitute clinical evidence and are not represented on the evidence-based medicine (EBM) pyramid. These studies are inherently hypothesis-generating and cannot establish	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				causality. The guideline emphasizes fetal safety concerns due to dydrogesterone's synthetic nature, despite the fact that it is derived from natural sources and synthesized in a manner similar to natural progesterone. This distinction appears inconsistent and may contribute to biased interpretation	
140	German Society of Reproductive Medicine	157	4855-4858	Lack of Contextualization of Pharmacovigilance Data The current presentation of the Henry et al. pharmacovigilance (PV) study in the guideline lacks essential scientific context and is methodologically unbalanced. No comparable PV data are cited for other progestins or ART-related medications, resulting in a selective focus on dydrogesterone. This lack of systematic comparison undermines the guideline's neutrality and conveys a misleading impression of uniquely elevated risk, which is not supported by the broader evidence base. Most critically, the guideline fails to acknowledge that the longstanding allegation linking therapeutic progestins to non-genital congenital malformations has been conclusively refuted already more than 20 years ago. As reviewed in Brent et al., 2005 (DOI: 10.1542/peds.2004-1951), over two decades of rigorous epidemiological, animal, and mechanistic studies have demonstrated no causal relationship between progestins and malformations such as cardiac or limb defects. In fact, the U.S. FDA rescinded its warning on progestin use in 1999 after concluding that "clinically utilized progestational drugs do not cause nongenital malformations." Citing a disproportionality signal from a PV database in isolation, without appropriate context or triangulation with clinical, mechanistic, and epidemiologic evidence, is not only scientifically flawed—it risks generating undue alarm among clinicians and patients. This is particularly problematic in the context of fertility treatment, where reassurance, evidence-based decision-making, and trust in treatment protocols are essential. There is also an ethical responsibility to avoid presenting data in a way that could inadvertently	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				generate unwarranted concern, psychological burden, or treatment hesitancy among patients and clinicians. Furthermore, the current presentation fails to address genital malformations, which represent the primary theoretical concern regarding structural malformations associated with hormonally active substances during pregnancy. The omission of this more biologically plausible outcome underscores a lack of conceptual coherence in how fetal safety is currently framed within the guideline. Finally, the guideline does not adequately consider potential long-term effects of prenatal hormone exposure on growth, metabolic function, or endocrine health in adolescence and adulthood—an evidence gap that applies not only to dydrogesterone but to all progestins used in ART (doi.org/10.1038/s41598-020-78976-x; doi.org/10.1111/1471-0528.16608.) Suggested revision: The single pharmacovigilance study (Henry et al.) should be appropriately contextualized. Most importantly, the guideline should acknowledge that concerns about progestational agents (including progestins used in contraceptive pills) and non-genital congenital malformations have been scientifically refuted (DOI: 10.1542/peds.2004-1951), and that regulatory authorities such as the U.S. FDA have withdrawn prior warnings. Patients should thus be reassured and not alarmed regarding non-genital malformations and progestin exposure in pregnancy. Without this context, the presentation is unbalanced and may lead to misinterpretation of risk.	
140	German Society of Reproductive Medicine	157	4855-4858	The statement citing a significant disproportionality in birth defect reporting associated with dydrogesterone (Henry et al., 2025) introduces a reporting odds ratio (ROR)—a metric derived from spontaneous reporting systems. However, ROR is not a clinical outcome, nor is it among the predefined outcome measures specified in the ESHRE guideline methodology. This presents several problems for interpretation, consistency and neutrality: 1. ROR reflects reporting patterns, not incidence, prevalence, or comparative risk. It is influenced by reporting behaviour, media	The justification was expanded to provide some explanation on disproportionality analysis and the interpretation of the results.

NR	Reviewer	Page	Line	Comment	Action / Reply
				attention, and regulatory actions. 2. Most clinical readers are not familiar with pharmacovigilance metrics like ROR or disproportionality analysis, and may mistakenly interpret a statistically significant ROR as evidence of an actual increase in risk. 3. The inclusion of ROR as a stand-alone result, without the necessary appropriate contextualization or methodological explanation, risks misleading readers and undermines the scientific clarity and clinical usability of the guideline. Suggested revision: If pharmacovigilance data using disproportionality analysis (such as RORs) are included in the guideline, this should be clearly justified in the methods section and supported by a dedicated subsection on search strategy and pharmacovigilance methodology, including a plain-language explanation of how RORs differ from clinical effect estimates and what limitations apply. Otherwise, the Henry et al. citation should be either removed or rewritten with explicit qualifiers and context.	
151	Alexandra Kohl Schwartz	157	4855-4858	Lack of Contextualization of Pharmacovigilance Data The current presentation of the Henry et al. pharmacovigilance (PV) study in the guideline lacks essential scientific context and is methodologically unbalanced. No comparable PV data are cited for other progestins or ART-related medications, resulting in a selective focus on dydrogesterone. This lack of systematic comparison undermines the guideline's neutrality and conveys a misleading impression of uniquely elevated risk, which is not supported by the broader evidence base. Most critically, the guideline fails to acknowledge that the longstanding allegation linking therapeutic progestins to non-genital congenital malformations has been conclusively refuted already more than 20 years ago. As reviewed in Brent et al., 2005 (DOI: 10.1542/peds.2004-1951), over two decades of rigorous epidemiological, animal, and mechanistic studies have demonstrated	The Henry et al. pharmacovigilance study was included in the results of the literature search for dydrogesterone. Had a similar study been included in the results for any of the other compounds in the guideline that is administered in early pregnancy, this would also have been included in the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				no causal relationship between progestins and malformations such as cardiac or limb defects. In fact, the U.S. FDA rescinded its warning on progestin use in 1999 after concluding that “clinically utilized progestational drugs do not cause nongenital malformations.” Citing a disproportionality signal from a PV database in isolation, without appropriate context or triangulation with clinical, mechanistic, and epidemiologic evidence, is not only scientifically flawed—it risks generating undue alarm among clinicians and patients. This is particularly problematic in the context of fertility treatment, where reassurance, evidence-based decision-making, and trust in treatment protocols are essential. There is also an ethical responsibility to avoid presenting data in a way that could inadvertently generate unwarranted concern, psychological burden, or treatment hesitancy among patients and clinicians. Furthermore, the current presentation fails to address genital malformations, which represent the primary theoretical concern regarding structural malformations associated with hormonally active substances during pregnancy. The omission of this more biologically plausible outcome underscores a lack of conceptual coherence in how fetal safety is currently framed within the guideline. Finally, the guideline does not adequately consider potential long-term effects of prenatal hormone exposure on growth, metabolic function, or endocrine health in adolescence and adulthood—an evidence gap that applies not only to dydrogesterone but to all progestins used in ART (doi.org/10.1038/s41598-020-78976-x; doi.org/10.1111/1471-0528.16608.) Suggested revision: The single pharmacovigilance study (Henry et al.) should be appropriately contextualized. Most importantly, the guideline should acknowledge that concerns about progestational agents (including progestins used in contraceptive pills) and non-genital congenital malformations have been scientifically refuted (DOI: 10.1542/peds.2004-1951), and that regulatory authorities such as the U.S. FDA have withdrawn prior warnings. Patients should thus be reassured and not alarmed regarding non-genital malformations and	

NR	Reviewer	Page	Line	Comment	Action / Reply
				progestin exposure in pregnancy. Without this context, the presentation is unbalanced and may lead to misinterpretation of risk.	
151	Alexandra Kohl Schwartz	157	4855-4858	The statement citing a significant disproportionality in birth defect reporting associated with dydrogesterone (Henry et al., 2025) introduces a reporting odds ratio (ROR)—a metric derived from spontaneous reporting systems. However, ROR is not a clinical outcome, nor is it among the predefined outcome measures specified in the ESHRE guideline methodology. This presents several problems for interpretation, consistency and neutrality: 1. ROR reflects reporting patterns, not incidence, prevalence, or comparative risk. It is influenced by reporting behaviour, media attention, and regulatory actions. 2. Most clinical readers are not familiar with pharmacovigilance metrics like ROR or disproportionality analysis, and may mistakenly interpret a statistically significant ROR as evidence of an actual increase in risk. 3. The inclusion of ROR as a stand-alone result, without the necessary appropriate contextualization or methodological explanation, risks misleading readers and undermines the scientific clarity and clinical usability of the guideline. Suggested revision: If pharmacovigilance data using disproportionality analysis (such as RORs) are included in the guideline, this should be clearly justified in the methods section and supported by a dedicated subsection on search strategy and pharmacovigilance methodology, including a plain-language explanation of how RORs differ from clinical effect estimates and what limitations apply. Otherwise, the Henry et al. citation should be either removed or rewritten with explicit qualifiers and context.	The justification was expanded to provide some explanation on disproportionality analysis and the interpretation of the results.

NR	Reviewer	Page	Line	Comment	Action / Reply
151	Alexandra Kohl Schwartz	157	4855-4858	This statement is factually misleading. Of the 6 RCTs included in the Katalinic et al. (2024) meta-analysis, 48.2% of patients (764/1,584) were from studies in the ART/luteal phase support setting, and 51.8% from studies addressing miscarriage. Describing the dataset as “mainly in couples with recurrent miscarriage” therefore misrepresents the evidence base, which is roughly balanced between the two indications. Furthermore, when observational data are included by Katalinic (total N = 3,780), 78.3% of patients (2,960) were treated for ART-related indications, while only 21.7% were in miscarriage cohorts. These proportions should be clearly stated, as they are critical for interpreting the applicability of the safety signal to ART populations. The current wording underplays this, again questioning the neutrality of the document, and should be corrected for transparency and accuracy. There is also a lack of contextualization with broader safety data in the field. As highlighted in a comprehensive review by van Gelder et al. (2014), the evidence base for teratogenic risk is absent or severely limited for the majority of drugs used during early pregnancy — including many fertility treatments. In contrast, dydrogesterone is one of the comparatively better-studied compounds in this context, with multiple robust randomized controlled trials and observational datasets reporting on congenital outcomes. Therefore, the existing evidence on dydrogesterone—while not definitive—offers a stronger empirical foundation than that available for many other agents and other progestin formulations routinely used in ART. This context should be reflected in the guideline text to avoid disproportionate scrutiny of a compound with relatively robust safety data. Suggested revision, rephrase: ‘A recent systematic review and meta-analysis including six randomized controlled trials (n = 1512) reported a relative risk of 0.92 (95% CI 0.55–1.55) for congenital malformations in pregnancies exposed to dydrogesterone compared to placebo, no treatment, or other interventions (Katalinic et al., 2024). While the confidence interval is wide and includes the possibility of both reduced and increased risk, the point estimate does not	The text in the justification was adapted to reflect the different populations feeding the numbers in the Katalinic meta-analysis.

NR	Reviewer	Page	Line	Comment	Action / Reply
				suggest an elevated risk. Taken together with observational data, this represents one of the more substantial clinical evidence bases in the field and clinical data do not indicate a signal of concern regarding offspring safety.'	
6	Anagani Manjula	157	4855-4863	On the Safety Section Citing Henry et al. (2025) (Page 157, Lines 4855-4863): The conclusions drawn from the pharmacovigilance study by Henry et al. (2025), as presented in the safety section, stand in stark contrast to the extensive and reassuring real-world safety profile of dydrogesterone. This existing evidence encompasses clinical experience from its use in over 147 million women and during more than 20 million pregnancies, suggesting a significant discrepancy that warrants careful consideration.	The GDG can only rely on published data, in the setting of luteal phase support. This data is presented in the guideline. The real world data on large numbers of pregnancies have not been identified, as such it would serve the purpose well if the stakeholder would provide the data source.
9	Mita Aggarwal	157	4855-4863	I don't comply with the findings as I have not seen any such adverse reactions in my practice.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
11	Namita Kotia	157	4855-4863	I HAVE BEEN PRACTISING SINCE 30 YEARS USING DYDROGESTERONE IN A DOSAGE OF 20-30 MG DAILY AS LUTEAL SUPPORT AND IN WOMEN WITH FIRST TRIMESTER RECURRENT PREGNANCY LOSSES TILL 9-10 WEEKS OF PREGNANCY BUT NOT ENCOUNTERED ANY CONGENITAL MALFORMATIONS.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
14	Tetiana Tutchenko	157	4855-4863	The conditional recommendation for dydrogestrerone use was formulated because of pharmacovigilance reports published in 2025. But they concerned the use of dydrogesterone during pregnancy for miscarriage prevention. Is it rational to extrapolate this data on luteal phase support as long as efficacy studies show good results? Micronized progesterone and dydrogesterone should have the same status of recommendation for luteal phase support after ovarian stimulation.	A strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.
15	Fei Gong	157	4855-4863	To ensure comprehensive and balance of the guide, when discussing fetal safety of dydrogesterone use in early pregnancy, both positive and negative reports should be presented. Except for pharmacovigilance study and DEBC cohort, the following	The results of the Lotus II trial on congenital malformations is discussed in the justification.

NR	Reviewer	Page	Line	Comment	Action / Reply
				publications should be included: • In RCT of LOTUS II (Yang et al 2020), the incidence of congenital, familial, and genetic disorders in the fetal/neonatal population was similar between the dydrogesterone and MVP gel groups (6.3% and 5%, respectively). In the Chinese mainland subpopulation, the incidence of these disorders was numerically lower in the oral dydrogesterone group, compared with the MVP gel group (4.6% and 12.7%, respectively). Therefore, dydrogesterone is not inferior to MVP in the perspective of fetal/neonatal safety. • Network meta-analysis by Zhao et al.2025 showed that there was no statistically significant difference in congenital abnormality between oral dydrogesterone, vaginal progesterone, oral progesterone and placebo or no treatment (p>0.05). The SUCRA ranking results showed that oral dydrogesterone may be the most effective invention in reducing congenital abnormality (SUCRA = 82%). Retrospective cohort study by Yang et al (2025) found the similar results (6.05‰[dyd-exposed] vs. 7.90‰[dyd-unexposed], p=0.02). Considering the amount and level of evidence, a more equitable recommendation should be presented. Reference: 1. Yang et al., A Phase III randomized controlled trial of oral dydrogesterone versus intravaginal progesterone gel for luteal phase support in in vitro fertilization (Lotus II): results from the Chinese mainland subpopulation. Gynecol Endocrinol. 2020 Feb;36(2):175-183. doi: 10.1080/09513590.2019.1645110. Epub 2019 Aug 9. 2. Zhao et al., Efficacy and safety of different progestogens in women with first threatened miscarriage: A network meta-analysis. Int J Gynaecol Obstet. 2025 Mar;168(3):944-957.doi: 10.1002/ijgo.15987. Epub 2024 Dec 19. 3. Yang et al., CONGENITAL ANOMALIES AFTER FIRST-TRIMESTER DYDROGESTERONE THERAPY DURING IN VITRO FERTILIZATION. Abstract onlyVolume 120, Issue 4, Supplement e72October 2023Open Archive	The network meta-analyisis was published after the date of the final literature update and abstract-only studies are not eligible for inclusion in the guideline.
16	Sonia Naik	157	4855-4863	I HAVE BEEN USING DYDROGESTERONE IN ALL CASES OF IVF FOR SUP[PORT AND IMMUNOMODULATOR , IN CASES OF RECURRENT MISCARRIAGES, IN THREATENED MISCARRIAGES AND SO FAR I HJAVE NOT SEEN ANY BIRTH DEFECTS IN THE CHILDREN	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
22	Qinjie Tian	157	4855-4863	Given the limitation of Disease Surveillance systems, a higher evidence-based level evidence should be considered for the recommendations regarding dydrogesterone	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. As such the use of Dydrogesterone cannot be strongly supported.
29	Ritu Joshi	157	4855-4863	I have been using this molecule since 1992 and have not come across this any issues	Individual experience and practice, while valuable is not a substitute for published clinical studies.
29	Ritu Joshi	157	4855-4863	Critique of Dydrogesterone Safety Subheading (Recommendation 109; Full Text Line 4838): The subheading pertinent to dydrogesterone safety (Recommendation 109) seems to portray an incomplete and selective evidentiary landscape. Its reliance on a single study positioned outside the established Evidence-Based Medicine (EBM) pyramid is unlikely to be beneficial for clinical decision-making and may instead engender confusion among practitioners.	When formulating recommendations, one of the key elements, in addition to the evidence cited in the evidence section, is benefits vs harms. These considerations are explained in the justification, where the Li et al and the Henry et al studies are cited, as well as the Katalinic systematic review, and where the EBM-pyramid does not apply.
31	Debankur Barman	157	4855-4863	Dydrogesterone delivers good results. Highly endorsed for its effectiveness.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
32	Monu Pattanayak	157	4855-4863	In so many years of my experience, I have not found any co-relation of any adverse effects of dydrogesterone on foetus or mother and hence I consider it safe for administration.	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
33	Arnab Bhowmik	157	4855-4863	Dydrogesterone works efficiently. Strongly recommended for its efficacy.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
34	Bharat S.	157	4855-4863	dydrogesterone is very effective in 1st and 2nd trimester of Pregnancy. I highly recommend.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
35	Puja Kumari	157	4855-4863	Strong therapeutic impact and highly recommended. Offers beneficial results and generates effectiveness.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
36	Meeta Meeta	157	4855-4863	I agree with the one member regarding the strength of the recommendation on the historical use of dydrogesterone for early miscarriage prevention. My reasoning is as follows 1.Determination of casuality in pharmacovigilance studies is challenging and data from these form the low rung in the evidence for the strength of recommendation. Moreover, the genetic factors ,co-morbidities and the use of other drugs and their impact on the casuality cannot be ruled out. 2.In a guideline making ,the strength of recommendation is based on the available RCT results and additional safety data from recent Systematic review and Meta analysis, 2022 and 2024 respectively (katalinic et al) , predominantly focusing on safety aspect of dydrogestrone in first trimester on the risk of congenital anomalies proved that Dydrogestrone has favourable safety profile for recurrent pregnancy / threatened miscarriage and ART. 3.The word conditional is probably used based on a weak evidence of Henry et al. 2025 A significant methodological consideration for the Henry et al. (2025). (A) VigiBase data to not fully mirror the real-world patient population. The reported congenital anomaly incidence of 0.7% is substantially lower than the recognized baseline incidence in natural pregnancies (approximately 2%), a discrepancy strongly suggestive of considerable underreporting of adverse events within the database. (B)As highlighted by Parazzini et al. (2024), is that a disproportionately high frequency of non-malformative adverse events reported	The suggested meta-analysis was published after the final literature search for the guideline. The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				in the progesterone cohort may have the effect of diminishing the statistical signal for congenital birth defects, potentially obscuring a true association.(C) Exposure Bias--80% of the reported adverse event cases originate from Europe and North America, whereas dydrogesterone's approval for use in Assisted Reproductive Technology is currently restricted to only six European Union member states, indicating a potential mismatch between reporting regions and actual drug utilization patterns in ART. The data needs to be global for a guideline intended to be used globally. (D) Additional safety data from recent studies The comprehensive network meta-analysis by Zhao et al. (2024/2025) reached the conclusion that dydrogesterone exhibits a comparatively superior safety profile for both the mother and the developing fetus when contrasted with other progestogenic compounds, most notably when compared against vaginally administered progesterone. (E) Additional safety data from recent studies Corroborating these findings, several more recent clinical investigations, predominantly focusing on Frozen Embryo Transfer (FET) cycles, have not discerned any emergent safety signals or concerns within patient groups receiving dydrogesterone for luteal phase support (exemplified by studies such as Mackens et al. 2023, Vidal et al. 2023, and Zhang et al. 2025). 4.This statement is a disservice to the women, similar to the effects of WHI on midlife women 5. A suggestion based on the study by Henry et al. 2025 can be made that large multicentric RCT regarding the safety of Dydrogesterone vs Progesterone in early pregnancy use is needed. 6. With the available evidence today, use of dydrogesterone for early miscarriage prevention should be a strong recommendation.	
40	Ulughbek Jabborov	157	4855-4863	I disagree with the pharmacovigilance reports of an association between dydrogesterone exposure and an increased risk of birth defects. For many years we have been using dydrogesterone in our country and in our center and we have not had cases like those described in the report.	Individual experience and practice, while valuable is not a substitute for published clinical studies.

NR	Reviewer	Page	Line	Comment	Action / Reply
96	Sharda Jain	157	4855-4863	Methodological Concerns Regarding Li et al. (2025) (Page 157, Lines 4855-4863, Safety Section):The investigation by Li and colleagues (2024/2025) is characterized by significant methodological deficiencies in its design, statistical analysis, and reporting. These flaws have the potential to undermine the reliability of its conclusions for all pharmaceutical compounds evaluated therein.	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stands with the 2025 update of the guideline. RCTs are not powered to detect an increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.
105	Shaily Agarwal	157	4855-4863	The safety data reported by Henry et al. (2025) in the safety section (Page 157, Lines 4855–4863) contrasts sharply with the extensive real-world evidence on dydrogesterone, which includes its use in over 147 million women and more than 20 million pregnancies. This discrepancy warrants thorough examination and interpretation within the broader clinical context.	Individual experience and practice, while valuable is not a substitute for published clinical studies.
84	Umesh N Jindal	157	4857	Instead of writing reporting OR it should be ROR since OR stands out and gives a wrong impression of incidence	The justification states "reporting OR" and later in the text "ROR".
84	Umesh N Jindal	157	4858	The study by Henry et al 2025 on the basis of which this alert has been created has major concerns. 1. Reporting bias there is a geographical disproportionate reporting from world areas, " Birth defect cases were mostly reported from Europe (73%) and Asia (22%) for dydrogesterone, and from Europe (53%) and North America (33%) for progesterone (Table 1)" 2. Cluster Reporting: "Regarding dydrogesterone, these 48 cases contained 56 major anomalies, consisting mainly in genital defects such	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification

NR	Reviewer	Page	Line	Comment	Action / Reply
				as hypospadias (n = 18, 32%, including a cluster of 10 cases of hypospadias reported from United Arab Emirates in 2021) and CHD (n = 15, 27%) (Table 2)" Out of 48 cases 10 cases form a cluster and likely to lead to biased statistical analysis and flawed interpretation (type 1 error i.e. describing an association when there is no association.). Cluster of anomalies from a restricted geographical area can have different causative factors which need verification before ascribing it to dydrogesterone.	for the 'conditional' recommendation has been sufficiently nuanced.
22	Qinjie Tian	157	4858-4859	In DEBC study, the authors mentioned that this analysis did not rule out the influence of other maternal environmental and behavioral factors during pregnancy, except for maternal age and diseases, the causal association remains to be further verified. Due to the limitations of this cohort study, causative conclusions cannot be drawn. Based on the authors' description, this study should be cited with caution.	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
22	Qinjie Tian	157	4858-4859	Based on the clinical practice and latest evidence of dydrogesterone in luteal phase support, no causality was found between dydrogesterone and congenital malformation. In different meta-analysis [1-3] and RCTs[4-5], no severe adverse reaction was observed. [1]Katalinic, Alexander, et al. "No additional risk of congenital anomalies after first-trimester dydrogesterone use: a systematic review and meta-analysis." Human reproduction open 2024.1 (2024): hoae004. [2]Griesinger, Georg, et al. "Dydrogesterone as an oral alternative to vaginal progesterone for IVF luteal phase support: A systematic review and individual participant data meta-analysis." PloS one 15.11 (2020): e0241044.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate. As such the use of Dydrogesterone cannot be strongly supported.

NR	Reviewer	Page	Line	Comment	Action / Reply
				[3] Zhao, Hongqiong, et al. "Efficacy and safety of different progestogens in women with first threatened miscarriage: A network meta-analysis." International Journal of Gynecology & Obstetrics 168.3 (2025): 944-957. [4] Tournaye, Herman, et al. "A Phase III randomized controlled trial comparing the efficacy, safety and tolerability of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization." Human Reproduction 32.5 (2017): 1019-1027. [5] Griesinger, Georg, et al. "Oral dydrogesterone versus intravaginal micronized progesterone gel for luteal phase support in IVF: a randomized clinical trial." Human Reproduction 33.12 (2018): 2212-2221.	
24	Surinder Pal Singh Kochar	157	4858-4859	Specific Flaw in Li et al. (2025) (Page 157, Lines 4855-4863, Safety Section): A notable deficiency in the Li et al. study is the imbalanced and inequitable enumeration of compounds. The study fails to adequately differentiate agents prescribed for underlying conditions that could independently influence the observed safety events, rendering the juxtaposition and amalgamation of data statistically unsound and medically unjustifiable.	In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns by the synthetic nature of dydrogesterone. The safety concerns still stands with the 2025 update of the guideline. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only large registry-based studies can pick up a potential increase in congenital malformation rate.
27	Mukesh Gupta	157	4858-4859	Safety CONCERN: The systematic review and meta-analysis conducted by Katalinic and colleagues (2024) constitutes the most rigorous and highest echelon of available evidence concerning the safety profile of dydrogesterone. It is important to highlight positively safe data strongly.	The incidence of congenital malformations in IVF/ICSI is very low. RCTs are not powered to detect a small but relevant increase in congenital malformations. The meta-analysis of Katalinic et al. is mentioned in the justification, however, also only includes 1512 women. Only

NR	Reviewer	Page	Line	Comment	Action / Reply
					large registry-based studies can pick up a potential increase in congenital malformation rate. As such the use of Dydrogesterone cannot be strongly supported.
100	Reassure group	157	4858-4859	The DEBC cohort, while prospective in nature, is observational and based on administrative data, which limits control over key confounders such as maternal comorbidities, ART status, BMI, previous miscarriage, and, most importantly, drug indication. Importantly, the study was conducted in a general obstetric population, not in an ART setting, and therefore its findings should not be extrapolated to ART without caveat. The modest relative difference reported for dydrogesterone versus natural progesterone (aRR 1.13 vs. 1.05) comes with widely overlapping confidence intervals, suggesting that any difference between the two is statistically non-significant. Additionally, serious residual confounding, confounding by indication, and methodological flaws—including lack of adjustment for duplicate records or polypharmacy, absence of the outcomes in the unexposed reference group, uncorrected multiple testing, and imbalance in comparator group sizes—collectively undermine the validity of the findings. Any conclusion that dydrogesterone could pose a unique risk based on this study is therefore highly speculative. Suggested revision: Either remove the DEBC study from the guideline due to its limited relevance and methodological concerns, or—if retained—ensure that all available pregnancy cohort studies on dydrogesterone and progesterone are systematically and consistently reported and appraised for their relevance, quality, and applicability to clinical recommendations.	All studies reporting on safety of dydrogesterone, included in the literature search, were reviewed and included in the justification section of the guideline. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns. The safety concerns still stand with the 2025 update of the guideline.
100	Reassure group	157	4858-4859	The DEBC study by Li et al. (2024) does not provide a methodological basis for drawing comparative safety conclusions between	It was clarified in the justification that the comparison was to the unexposed population,

NR	Reviewer	Page	Line	Comment	Action / Reply
				dydrogesterone and progesterone. This sentence is thus misleading. The adjusted relative risks (aRRs) for each drug were calculated independently against an unexposed reference population, and without a direct head-to-head comparison between the two. Importantly, the confidence intervals for dydrogesterone (aRR 1.13, 95% CI 1.06–1.21) and for progesterone (aRR 1.05, 95% CI 0.97–1.13) overlap, indicating no statistically significant difference. Moreover, the study does not account for confounding by indication—particularly relevant as dydrogesterone is more frequently prescribed in higher-risk pregnancies, such as those involving recurrent miscarriage in the DEBC cohort. No stratification, matching, or adjustment was applied to compare the exposed groups directly. Therefore, the data do not support a strong or reliable risk signal for dydrogesterone relative to progesterone, and any inference of differential teratogenicity based on these results is methodologically unsound und inferentially unjustified. Suggested revision, rephrase: Data from the China maternal drug exposure birth cohort (DEBC) (Li et al., 2024) reported an association between first-trimester dydrogesterone exposure and birth defects relative to an unexposed reference group (aRR 1.13, 95% CI 1.06–1.21). A separate estimate for natural progesterone (aRR 1.05, 95% CI 0.97–1.13) was also reported; however, no direct comparison was made between the two drugs. The confidence intervals overlap, indicating no statistically significant difference between the two formulations, and substantial differences in sample size and insufficient control for confounding by indication limit the interpretability of these findings. Next, to ensure balance and methodological consistency, findings from other cohort studies on dydrogesterone must be presented, ideally on ART population (which are ample in the literature). The DEBC cohort	and that dydrogesterone was not directly compared to natural progesterone. Furthermore, the adjusted factors were added.

NR	Reviewer	Page	Line	Comment	Action / Reply
				should be contextualized as one source of observational data among others, not presented in isolation let alone as a comparative study of dydrogesterone versus progesterone.	
140	German Society of Reproductive Medicine	157	4858	The DEBC cohort, while prospective in nature, is observational and based on administrative data, which limits control over key confounders such as maternal comorbidities, ART status, BMI, previous miscarriage, and, most importantly, drug indication. Importantly, the study was conducted in a general obstetric population, not in an ART setting, and therefore its findings should not be extrapolated to ART without caveat. The modest relative difference reported for dydrogesterone versus natural progesterone (aRR 1.13 vs. 1.05) comes with widely overlapping confidence intervals, suggesting that any difference between the two is statistically non-significant. Additionally, serious residual confounding, confounding by indication, and methodological flaws—including lack of adjustment for duplicate records or polypharmacy, absence of the outcomes in the unexposed reference group, uncorrected multiple testing, and imbalance in comparator group sizes—collectively undermine the validity of the findings. Any conclusion that dydrogesterone could pose a unique risk based on this study is therefore highly speculative. Suggested revision: Either remove the DEBC study from the guideline due to its limited relevance and methodological concerns, or—if retained—ensure that all available pregnancy cohort studies on dydrogesterone and progesterone are systematically and consistently reported and appraised for their relevance, quality, and applicability to clinical recommendations.	It was clarified in the justification that the comparison was to the unexposed population, and that dydrogesterone was not directly compared to natural progesterone. Furthermore, the adjusted factors were added.
140	German Society of Reproductive Medicine	157	4858	The DEBC study by Li et al. (2024) does not provide a methodological basis for drawing comparative safety conclusions between dydrogesterone and progesterone. This sentence is thus misleading. The adjusted relative risks (aRRs)	It was clarified in the justification that the comparison was to the unexposed population, and that dydrogesterone was not directly

NR	Reviewer	Page	Line	Comment	Action / Reply
				for each drug were calculated independently against an unexposed reference population, and without a direct head-to-head comparison between the two. Importantly, the confidence intervals for dydrogesterone (aRR 1.13, 95% CI 1.06–1.21) and for progesterone (aRR 1.05, 95% CI 0.97–1.13) overlap, indicating no statistically significant difference. Moreover, the study does not account for confounding by indication—particularly relevant as dydrogesterone is more frequently prescribed in higher-risk pregnancies, such as those involving recurrent miscarriage in the DEBC cohort. No stratification, matching, or adjustment was applied to compare the exposed groups directly. Therefore, the data do not support a strong or reliable risk signal for dydrogesterone relative to progesterone, and any inference of differential teratogenicity based on these results is methodologically unsound und inferentially unjustified. Suggested revision, rephrase: Data from the China maternal drug exposure birth cohort (DEBC) (Li et al., 2024) reported an association between first-trimester dydrogesterone exposure and birth defects relative to an unexposed reference group (aRR 1.13, 95% CI 1.06–1.21). A separate estimate for natural progesterone (aRR 1.05, 95% CI 0.97–1.13) was also reported; however, no direct comparison was made between the two drugs. The confidence intervals overlap, indicating no statistically significant difference between the two formulations, and substantial differences in sample size and insufficient control for confounding by indication limit the interpretability of these findings. Next, to ensure balance and methodological consistency, findings from other cohort studies on dydrogesterone must be presented, ideally on ART population (which are ample in the literature). The DEBC cohort should be contextualized as one source of observational data among others, not presented in isolation let alone as a comparative study of dydrogesterone versus progesterone.	compared to natural progesterone. Furthermore, the adjusted factors were added.
151	Alexandra Kohl Schwartz	157	4858	The DEBC cohort, while prospective in nature, is observational and based on administrative data, which limits control over key confounders such as maternal comorbidities, ART status, BMI, previous miscarriage, and, most importantly, drug indication. Importantly, the study was conducted in a	It was clarified in the justification that the comparison was to the unexposed population, and that dydrogesterone was not directly

NR	Reviewer	Page	Line	Comment	Action / Reply
				general obstetric population, not in an ART setting, and therefore its findings should not be extrapolated to ART without caveat. The modest relative difference reported for dydrogesterone versus natural progesterone (aRR 1.13 vs. 1.05) comes with widely overlapping confidence intervals, suggesting that any difference between the two is statistically non-significant. Additionally, serious residual confounding, confounding by indication, and methodological flaws—including lack of adjustment for duplicate records or polypharmacy, absence of the outcomes in the unexposed reference group, uncorrected multiple testing, and imbalance in comparator group sizes—collectively undermine the validity of the findings. Any conclusion that dydrogesterone could pose a unique risk based on this study is therefore highly speculative. Suggested revision: Either remove the DEBC study from the guideline due to its limited relevance and methodological concerns, or—if retained—ensure that all available pregnancy cohort studies on dydrogesterone and progesterone are systematically and consistently reported and appraised for their relevance, quality, and applicability to clinical recommendations.	compared to natural progesterone. Furthermore, the adjusted factors were added.
151	Alexandra Kohl Schwartz	157	4858	A major concern with the current guideline draft is the selective inclusion of evidence regarding fetal safety. The DEBC cohort, conducted in a general obstetric population, is the only observational dataset cited in support of concerns about dydrogesterone. However, other relevant evidence from non-ART populations is systematically omitted, particularly data suggesting a favorable safety profile. For example, the guideline does not reference a recent network meta-analysis (DOI: 10.1002/ijgo.15987), which found that oral dydrogesterone had the highest probability of being the safest intervention in terms of congenital anomaly risk. The SUCRA ranking indicated oral dydrogesterone (82%) outperformed oral progesterone (67%), placebo or no treatment (47%), and vaginal progesterone (4%). These findings should be	The suggested meta-analysis was published after the final literature search for the guideline. All studies reporting on safety of dydrogesterone, included in the literature search, were reviewed and included in the justification section of the guideline. In 2019, the recommendation for dydrogestrerone was already conditional because of safety concerns. The safety concerns still stand with the 2025 update of the guideline.

NR	Reviewer	Page	Line	Comment	Action / Reply
				presented and balanced against other evidence. Omitting positive safety data from non-ART studies while emphasizing a single cohort which is likewise not an ART study (i.e. the DEBC study in which dosages predominantly used are not LPS dosages) with methodological limitations introduces bias and undermines the evidence balance expected in guideline development. Suggested revision: For the guideline to remain balanced, neutral, and methodologically consistent, safety data from all available sources and for all relevant interventions must be systematically searched, transparently evaluated, and consistently integrated. This includes data from pharmacovigilance databases, observational cohorts, randomized trials, and meta-analyses—both favorable and unfavorable. Given the complexity and sensitivity of interpreting reproductive safety data, the guideline group may wish to consider seeking advice from an independent expert panel with recognized expertise in pharmacovigilance, teratology, and perinatal epidemiology to support evidence appraisal and ensure that recommendations are based on a comprehensive and unbiased assessment of the totality of evidence.	
151	Alexandra Kohl Schwartz	157	4858	The DEBC study by Li et al. (2024) does not provide a methodological basis for drawing comparative safety conclusions between dydrogesterone and progesterone. This sentence is thus misleading. The adjusted relative risks (aRRs) for each drug were calculated independently against an unexposed reference population, and without a direct head-to-head comparison between the two. Importantly, the confidence intervals for dydrogesterone (aRR 1.13, 95% CI 1.06–1.21) and for progesterone (aRR 1.05, 95% CI 0.97–1.13) overlap, indicating no statistically significant difference. Moreover, the study does not account for confounding by indication—particularly relevant as dydrogesterone is more frequently prescribed in higher-risk pregnancies, such as those involving recurrent miscarriage in the DEBC cohort. No stratification, matching, or adjustment was applied to compare the exposed groups directly. Therefore, the data do not support a strong or reliable risk signal for dydrogesterone relative to progesterone, and any inference of differential teratogenicity based on these	It was clarified in the justification that the comparison was to the unexposed population, and that dydrogesterone was not directly compared to natural progesterone. Furthermore, the adjusted factors were added.

NR	Reviewer	Page	Line	Comment	Action / Reply
				results is methodologically unsound und inferentially unjustified. Suggested revision, rephrase: Data from the China maternal drug exposure birth cohort (DEBC) (Li et al., 2024) reported an association between first-trimester dydrogesterone exposure and birth defects relative to an unexposed reference group (aRR 1.13, 95% CI 1.06–1.21). A separate estimate for natural progesterone (aRR 1.05, 95% CI 0.97–1.13) was also reported; however, no direct comparison was made between the two drugs. The confidence intervals overlap, indicating no statistically significant difference between the two formulations, and substantial differences in sample size and insufficient control for confounding by indication limit the interpretability of these findings. Next, to ensure balance and methodological consistency, findings from other cohort studies on dydrogesterone must be presented, ideally on ART population (which are ample in the literature). The DEBC cohort should be contextualized as one source of observational data among others, not presented in isolation let alone as a comparative study of dydrogesterone versus progesterone.	
42	Yan Gong	157	4858-4863	Due to the limitations of the research, DEBC needs to carefully consider its weight in the guideline recommendations	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
42	Yan Gong	157	4858-4863	In DEBC, the authors explicitly pointed out that due to the limitations of the cohort study, it is impossible to draw a causal relationship between drug exposure and birth defects. This article has certain limitations: 1. The drug exposure in this study was evaluated through self-reported questionnaires of pregnant women, which may lead to recall bias; 2. The	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC

NR	Reviewer	Page	Line	Comment	Action / Reply
				current preliminary analysis has not excluded factors such as maternal age and drug treatment that may affect the safety of the offspring; 3. The majority of the population exposed to dydrogesterone in the article were patients with threatened abortion, who are themselves a high-risk group for birth defects. If the association in the results is wrongly interpreted as a causal relationship, it may cause panic among clinicians. Therefore, the results of this article should be viewed with caution. Reference: Li, L., Wang, K., Wang, M. et al. The maternal drug exposure birth cohort (DEBC) in China. Nat Commun 15, 5312 (2024). https://doi.org/10.1038/s41467-024-49623-0	studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.
81	Dongzi Yang	157	4861-4868	Although the current guideline notes a lack of consensus, it retains a safety-related footnote (lines 4838-4840) that could raise unwarranted concerns without unified expert agreement. Specifically, the cited study by Henry et al. (2025) employs the Reporting Odds Ratio (ROR), a pharmacovigilance signal detection metric. It is crucial to recognize that ROR is methodologically distinct from the Odds Ratio (OR) derived from controlled clinical trials. The ROR denominator is based on the count of adverse event reports associated with dydrogesterone use, not the number of women exposed to the drug. This methodological approach, particularly for rare non-malformation events, inherently risks overestimating the reporting association for malformation events. RECOMMENDATION 2: The guideline should remove the footnote on lines 4838-4840. It should instead provide physicians with clear, evidence-based guidance for daily practice, ensuring continued access to dydrogesterone as a trusted and effective option for luteal phase support.	It is stated in the justification: "It needs to be pointed out here that the observed relations from these two studies cannot be translated into a conclusion on causality". The option of ignoring the pharmacovigilance and the DEBC studies, based on potential biases in the work, was deemed inappropriate, and the justification for the 'conditional' recommendation has been sufficiently nuanced.

NR	Reviewer	Page	Line	Comment	Action / Reply
83	Shamugiya Nato	157	4861-4868	Scientific evidence from RCTs and non-RCTs has shown a favourable safety profile of progestogens, particularly dydrogesterone, which is also observed in real clinical practice. In the RCT (Tournaye H, 2017), infant safety data collected at delivery were similar between the two treatment groups (dydrogesterone and micronized progesterone), with most infants born with normal physical examination findings (93.4% in the dydrogesterone group and 92.4% in the MVP group). The rate of infants experiencing at least one serious adverse event was also similar between groups: 4.2% in the dydrogesterone group and 5.7% in the MVP group1. In the non-RCT (Jiang X, 2025, Nazarenko T, 2025), one of the study outcomes was the assessment of neonatal outcomes of all live births. Cardiovascular defects (atrial septal defects, ventricular septal defects, congenital heart defects, etc.) were the most common, with the incidence of detected disorders not exceeding population thresholds and not having a causal relationship with the drug intake. These studies did not reveal any new safety concerns associated with oral DYD2,3. According to the data of a prospective randomized study comparing the effectiveness of progestogens (dydrogesterone and vaginal micronized progesterone) for the support of LF, comparable pregnancy outcomes were obtained. The use of dydrogesterone is associated with better tolerability and satisfaction among patients4. In view of the scientific evidence on comparable efficacy and safety of dydrogesterone in LF support in IVF cycles compared with other progestogens, I propose to assign them the same strength of recommendation.	The GDG has no doubts with regards to the efficacy of dydrogesterone for LPS. However, a strong recommendation would indicate that the GDG is confident that most patients would benefit from the intervention. Since the GDG has concerns about potential congenital malformations, the recommendation is conditional.

NR	Reviewer	Page	Line	Comment	Action / Reply
19. Prevention of OHSS					
132	The Chinese Expert Review Panel for ESHRE OS Guideline	22	R101-103, 117	Do not totally agree with the statement that the dual trigger for final oocyte maturation is probably not recommended for normal/low/high responders. Because According to the recommendation, there will be no suitable population for dual trigger.	The GDG sees your concern. However, it seems to be the case according to the available evidence.
106	Rishma Dhillon Pai	23	R115	A GnRH agonist trigger is recommended for final oocyte maturation in women at risk of OHSS combined with a freeze-all strategy to minimise the risk of severe OHSS. ((ADD IN A GnRH ANTAGONIST PROTOCOL	A GnRH agonist trigger can be used in any protocol other than a GnRH agonist protocol, i.e., a GnRH antagonist protocol or PPOS, preference depending on the intention for a fresh embryo transfer at the start of the stimulation. Since this is regarded as basic knowledge and mentioned earlier we did not feel it is necessary to reiterate that GnRH agonists should not be used in anticipated hyper responders at risk of OHSS.
106	Rishma Dhillon Pai	23	R119	GnRH agonist trigger for final oocyte maturation with or without a freeze-all strategy is preferred over a coasting strategy in patients at risk of OHSS. (ADD IN A GnRH ANTAGONIST PROTOCOL)	A GnRH agonist trigger can be used in any protocol other than a GnRH agonist protocol, i.e., a GnRH antagonist protocol or PPOS, preference depending on the intention for a fresh embryo transfer at the start of the stimulation. Since this is regarded as basic knowledge and mentioned earlier we did not feel it is necessary to reiterate that GnRH

NR	Reviewer	Page	Line	Comment	Action / Reply
					agonists should not be used in anticipated hyper responders at risk of OHSS.
1	Raj Mathur	164	5104	A 'reduced' dose is not evidence based in Antagonist cycles – see above	Evidence comes from 2 dose-finding studies in GnRH antagonist protocol. However, the GDG realises that the evidence is not very strong, therefore a conditional recommendation was formulated.
39	Ahmed Elsayed Hassan Hamed Elbohoty	169	5253	The guideline currently omits the use of early luteal GnRH antagonist administration as a preventive strategy for OHSS in non-transfer cycles, particularly in patients who receive an hCG trigger. I recommend including GnRH antagonist (e.g., cetrorelix 0.25 mg daily for 2–3 days post-OPU) as a potential adjunct in high responders undergoing freeze-all cycles. Although freeze-all significantly reduces OHSS risk, patients with extreme ovarian response (e.g., >25 follicles, E2 >6,000 pg/mL) may still develop moderate to severe OHSS. Early luteal GnRH antagonist administration has been shown to suppress VEGF expression, reduce luteal corpora lutea activity, and promote faster ovarian involution, thereby mitigating OHSS progression. This approach is supported by Aboulghar et al. (2011), Seyhan et al. (2013), and Fatemi et al. (2008). Including it would enhance the practical safety guidance for high-risk patients in freeze-all cycles.	GnRH antagonist administration in the luteal phase happens after ovarian stimulation, so this is not covered in the present guideline.
7	Sujoy Dasgupta	170	5260-5262	In addition, young age is a risk factor for OHSS. Regarding body mass index (BMI) , studies show conflicting results, with some studies connecting low BMI with risk of OHSS while other studies could not confirm the same. Practice Committee of the American Society for Reproductive Medicine. Electronic address: ASRM@asrm.org; Practice Committee of the American Society for Reproductive Medicine. Prevention and treatment	The GDG does not agree with the reviewer. Neither age or BMI are risk factors independent of ovarian reserve.

NR	Reviewer	Page	Line	Comment	Action / Reply
				of moderate and severe ovarian hyperstimulation syndrome: a guideline. Fertil Steril. 2016 Dec;106(7):1634-1647.	
Annexes					
79	Mitranovici Melinda Ildiko	13	25 in table	Regarding Annex 6 table 13 a, b, c, d there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	20	37 in table	Regarding Annex 6 table 20 a, b, c, there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
38	Padmaja veeramachaneni	25	/	23 Metformin compared to placebo/no intervention as adjunct during ovarian stimulation for women with PCOS Instead of metformin : Risk with testosterone By mistake	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	25	47 in table	Regarding Annex 6 table 23 there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
38	Padmaja veeramachaneni	26	/	24a Growth hormone compared to placebo/no intervention as adjunct during ovarian stimulation for normal responders The risk is with GH , testosterone typed in my mistake	A mistake in the table was corrected.
38	Padmaja veeramachaneni	26	/	24b Growth hormone compared to placebo/no intervention as adjunct during ovarian stimulation for low responders The risk is with GH , testosterone typed in my mistake	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	28	54 in table	Regarding Annex 6 table 27 there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	28	56 in table	Regarding Annex 6 table 28 a, b there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
38	Padmaja veeramachaneni	31	/	32a Dual trigger compared to hCG for final oocyte maturation Instead of dual trigger : ghrha is typed	A mistake in the table was corrected.

NR	Reviewer	Page	Line	Comment	Action / Reply
38	Padmaja veeramachaneni	32	/	32b Dual trigger compared to GnRH agonist for final oocyte maturation Instead of dual trigger : ghrha is typed	A mistake in the table was corrected.
38	Padmaja veeramachaneni	32	/	32c Dual trigger compared to hCG for final oocyte maturation for low responders Instead of dual trigger : ghrha is typed	A mistake in the table was corrected.
38	Padmaja veeramachaneni	33	/	Double trigger compared to hCG for final oocyte maturation Instead of double trigger : ghrha is typed	A mistake in the table was corrected.
38	Padmaja veeramachaneni	44	/	45 Dopamine agonists compared to placebo/no treatment for prevention of OHSS Risk with albumin Risk with Freeze-all : latter should be dopamine agonist	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	44	120 in table	Regarding Annex 6 table 45 there are inconsistencies between the populations compared in the text and those in the tables, please explain.	A mistake in the table was corrected.
79	Mitranovici Melinda Ildiko	178	5365	Recommendations for research: would you consider endometrium thickness evaluation in women>40 years old. Also do you take into account different methods to improve oocyte quality during stimulation?	Thank you for the suggestion, the GDG will consider it.
General comments					
1	Raj Mathur	/	/	This is an excellent and comprehensive piece of work, well organized. By and large, it is well-written and clear. I am sure it will be helpful to clinicians and the GDG should be congratulated.	Thank you.
7	Sujoy Dasgupta	/	/	Overall, the guideline provides comprehensive view of ovarian stimulation including its various aspects starting from prediction of response to luteal phase support. This version is more extensive than its previous version published in 2019.	Thank you.

NR	Reviewer	Page	Line	Comment	Action / Reply
37	Ahmed Samy Abdelazim Saad	/	/	Overall, congratulations for such an updated elaborate discussion of every detail in ovarian stimulation	Thank you.
39	Ahmed Elsayed Hassan Hamed Elbohoty	/	/	The current update presents a well-structured and literature-informed approach to ovarian stimulation. The inclusion of stakeholder perspectives and recent evidence strengthens its overall value.	Thank you.
39	Ahmed Elsayed Hassan Hamed Elbohoty	/	/	The guideline would benefit from more explicit recommendations on safety considerations related to women's age and BMI. Advanced age (>43 years) is associated with low success rates and increased obstetric risks, while BMI extremes (>30 or <18.5) may complicate anesthesia, affect oocyte retrieval efficiency, and alter drug pharmacodynamics. Recognizing these parameters as safety indicators may improve patient selection and clinical planning. Broekmans FJ, et al. Female reproductive ageing: current knowledge and future trends. Hum Reprod Update. 2009;15(1):23–37. (Maheshwari A, et al. Hum Reprod Update. 2007; Jungheim ES, et al. Obstet Gynecol Clin North Am. 2015).	Good point but it relates more generally to the moment that the couple is offered IVF/ICSI treatment, and judgement of obstetrical risks of a pregnancy need to be considered and where age is an important factor. Such PICO needs to be discussed in a guideline on Infertility work up and treatment choice. The risks associated with high BMI are at the same way a bit out of the scope of this guideline, with the possible exception of drug dynamics. This PICO could be introduced in a next update.
39	Ahmed Elsayed Hassan Hamed Elbohoty	/	/	Cost-related aspects are underexplored. Incorporating data on the economic impact of different stimulation protocols—such as PPOS versus GnRH antagonist and rFSH/LH versus hMG—would enhance the guideline's relevance in diverse healthcare settings, especially those with limited resources. (Shi Y, et al. Reprod Biol Endocrinol. 2023; Stimpfel M, et al. Cost-effectiveness of PPOS vs antagonist protocols in donor cycles. J Assist Reprod Genet. 2024; Leijdekkers JA, et al. Reprod Biomed Online. 2022).	As stated in the ESHRE manual for guideline development: ESHRE Guidelines will not include a formal analysis of cost effectiveness of recommended as compared to established practice, as this is not the main aim, and is sometimes impossible because of the obvious differences in current European economic and healthcare systems. The clinical and organizational impact of costs on recommendations will be considered in GDG meetings, and if relevant, described in the justification section.

NR	Reviewer	Page	Line	Comment	Action / Reply
					The economic feasibility of recommendations will not be covered.
39	Ahmed Elsayed Hassan Hamed Elbohoty	/	/	The current guideline does not address individualized stimulation for poor/low responders using the POSEIDON classification, which has gained broad international acceptance. Incorporating subgroup-specific recommendations for POSEIDON Groups I–IV (e.g., tailored gonadotrophin dosing, dual stimulation, minimum oocyte yield goals) would enhance clinical utility and alignment with contemporary practice. (Alviggi C, et al., Hum Reprod. 2016; Conforti A, et al. Reprod Biol Endocrinol. 2020) Additionally, the guideline lacks recommendations for specific populations, such as those with endometriosis, estrogen-sensitive tumors (e.g., hormone receptor-positive breast cancer), or individuals seeking fertility preservation for medical indications. Addressing these scenarios would broaden the guideline's applicability and inclusiveness. (Dolmans MM, et al., Clin Obstet Gynecol. 2021; Kim JH, et al., Fertil Steril. 2022)	The GDG is very aware that there are more nuanced definitions of a low ovarian response but decided to try and keep it simple: in the guideline there is only low-normal-high responder. Moreover, the Poseidon classification has not delivered comparative studies to show, within each of the classes, the benefit of one treatment approach over the other. In contrast, such studies do exist in low and high responder groups as defined in the guideline.
73	Juan-Enrique Schwarze Shiv Gupta Susana Montenegro*	/	/	The term r-hFSH (recombinant human Follicle Stimulating Hormone) & r-hLH (recombinant human luteinising hormone) is the correct and internationally accepted abbreviation, since we should always use the INN conventional name. Guidelines and Sources Supporting this Convention: • International Nonproprietary Names (INN) by WHO: The WHO INN system recommends r-hFSH or r-hLH for recombinant human follicle-stimulating hormone or recombinant human luteinising hormone. The "h" explicitly denotes the human origin of the protein cDNA (as opposed to other species). • European Medicines Agency (EMA) and FDA documents: These	Thank you for pointing this out, this was corrected in the guideline text.

NR	Reviewer	Page	Line	Comment	Action / Reply
				regulatory authorities use r-hFSH or r-hLH in their SmPCs (Summary of Product Characteristics) and labeling for GONAL-f, Bemfola, Ovaleap, Luveris, Pergoveris etc. • Scientific Publications & Textbooks: High-quality journals and IVF literature consistently use r-hFSH or r-hLH for recombinant products. In contrast, just “FSH” is often used more generically when not distinguishing between urinary and recombinant forms. Reference: See Smpc of the product for correct name GONAL-f: https://ec.europa.eu/health/documents/community-register/2016/20161213136695/anx_136695_en.pdf Luveris: https://www.ema.europa.eu/en/documents/product-information/luveris-epar-product-information_en.pdf Pergoveris: https://ec.europa.eu/health/documents/community-register/2015/20150203130959/anx_130959_en.pdf	
89	Johannes Ott	/	/	We appreciate the effort of the guideline team and thank you for your valuable work!	Thank you.
112	Pedro Augusto Araujo Monteleone	/	/	A methodological review of the coherence between the strength of the recommendation and the level of evidence is recommended, especially when the evidence is ⨁□□□.	The decision on a strong or a weak recommendation depends on 5 key factors: benefits vs harms, quality of the evidence, patient preferences, acceptability to stakeholders and resource use. The factors that played into the decision on the strength of the recommendations are detailed in the justification of the recommendations.
112	Pedro Augusto Araujo Monteleone	/	/	The terminology 'normal/high/low responders' could be replaced by more neutral and inclusive language.	As explained in the introduction, there is a need to classify patients into ovarian response categories and the terms 'normal/high/low' are

NR	Reviewer	Page	Line	Comment	Action / Reply
					already more neutral than the previously used 'poor/excessive'.
112	Pedro Augusto Araujo Monteleone	/	/	Greater openness to compassionate use and/or research of adjuvants in subgroups with poor reproductive prognosis is suggested.	All potential adjuvants that could alter the outcome of the stimulation, mostly in low responder/prognosis patients have been addressed in the current update of the guideline.
114	Monica Varma	/	/	Heartiest congratulations to the Guideline Development Group for a detailed comprehensive document, a great help for clinical practice.	Thank you.
115	Hassan Sallam	/	/	• A deeply researched and well written document (as expected) and thanks to all involved, but nobody is perfect • My specific comments are included (vide infra) regarding my main interest (luteal phase support) in addition to embryo transfer	Thank you. Your comments are much appreciated.
115	Hassan Sallam	/	/	• I suggest the use of the term “regimen” when referring to medication as the word “regime” is a more political term (e.g. Pituitary suppression regimens rather than pituitary suppression regimes (page 92)	The reviewer has a point, the text was adapted.
122	Surveen Ghumman	/	/	The guideline is very well done dealing with all aspects of ovarian stimulation. Congratulations to the GDG for this detailed effort. The data and studies are many and very detailed It requires hard work and experience to go through each one and have achieved this and come out with guidelines on ovarian stimulation. It was a much-needed guideline.	Thank you.
127	ESHRE SIG RE	/	/	We would like to express our sincere appreciation to the ESHRE Guideline Development Group for the immense effort, expertise, and perseverance that go into crafting evidence-based clinical guidelines. The observations and suggestions that follow are offered in this collaborative spirit. Our aim is to help strengthen the clarity, consistency, and practical	Thank you.

NR	Reviewer	Page	Line	Comment	Action / Reply
				applicability of the guideline so it can serve the reproductive-medicine community even more effectively.	
127	ESHRE SIG RE	/	/	The guideline proposes replacing “poor responder” with “low responder,” yet “poor responders” still appears multiple times (e.g., pregnancy-prediction section). Consider using “low responder” consistently across the entire document.	In the evidence sections, the terminology from the published research studies was used, in the other sections of the guideline, the terminology recommended by the GDG was used.
129	Roberto de Azevedo Antunes	/	/	There should be a guideline recommendation points specifically for the luteal phase support (LPS) regimen in frozen embryo transfer (FET) cycles. Specific topics comparing artificial, natural and modified natural cycles would be greatly appreciated. There is important good quality evidence to recommend that artificial FET performed under vaginal LPS present with poorer clinical outcomes when low progesterone threshold are detected on the day of the embryo transfer1 Also, the supplementation of either intramuscular or subcutaneous progesterone, as well as, of oral dydrogesterone seem to rescue the clinical outcomes 2, 3 Another import point to be addressed is the comparison between artificial versus natural/modified natural (N/mN) FET cycles. There is evidence suggesting better obstetrical outcomes in favor of the N/mN FET cycles4,5,6,7,8,9 Also, there is good data showing that progesterone supplementation for luteal phase support in natural/modified natural FET cycles render better clinical outcomes10	In the scope of the guideline is stated: "the following issues were outside the scope of the current document: [...], frozen embryo transfer, [...]. ESHRE has recently approved the development of an evidence-based guideline on frozen embryo transfer.
133	Suresh Nair	/	/	The guideline is well-structured and logically organized, allowing easy navigation across different patient subgroups. The subdivision into specific populations such as fertility preservation, elective oocyte cryopreservation, and oocyte donation greatly enhances clinical	Thank you.

NR	Reviewer	Page	Line	Comment	Action / Reply
				applicability. The guideline demonstrates rigorous evidence appraisal and provides clear justification for the strength and direction of recommendations. The distinction between 'updated' versus 'reworded' guidance improves transparency	
135	Sadiah Ahsan	/	/	Thank you for preparing an excellent guideline	Thank you.
137	Jayesh Amin	/	/	Heading is missing, it is directly written as PART -D after PART- B	This was corrected.
144	Karolina Palinska-Rudzka	/	/	This is a very valuable and thoughtfully developed guideline. My comments are offered in support of ESHRE's goal to promote safe, effective, and evidence-based reproductive care. I hope these points are helpful and contribute positively to the final document.	Thank you.
153	Stefan Matik	/	/	Excellent guideline, very thorough, concise and up to date, covers all aspects of OS	Thank you.
156	Adrija Kumar Datta Stuart Campbell Geetta Nargund	/	/	Definition of 'conventional' IVF The guideline development group (GDG) has referred "conventional IVF" to a protocol where a stimulation dose of 150-225 IU/ day is used. Our understanding is that a stimulation dose which is personalized, adjusting for age, BMI, ovarian reserve etc with the intention of generating maximum number of oocytes could be regarded as "conventional IVF". However, we appreciate that the GDG recommends a stimulation dose that is more close to that defined as Mild stimulation IVF in the recent publication from International Society for Mild Approaches Assisted Reproduction (ISMAAR) (Nargund, et al., 2022).	The term 'Conventional' in the context of FSH dosing, was applied to describe the normal FSH dosage range (150-225 IU per day) by which, instead of one, several to many follicles will grow into dominance, and as such will guarantee the efficiency of the ART process (getting oocytes, making embryo's and putting the high quality embryos back, one by one). In many different ways and based on various criteria, this normal range could be changed into a higher or lower FSH daily dosage than the conventional range (150-225), for the purposes of safety, or efficacy.
156	Adrija Kumar Datta Stuart	/	/	Safety and efficacy We appreciate that the GDG has put due stress on the safety aspects which may be compromised while chasing for a high oocyte yield.	ESHRE Guidelines will not include a formal analysis of cost effectiveness of recommended as compared to established practice, as this is

NR	Reviewer	Page	Line	Comment	Action / Reply
	Campbell Geetta Nargund			Higher gonadotropin dose also has other implications including increased treatment burden and treatment cost (Datta, et al., 2021). The latter is linked with affordability and access to private treatment or public-funded treatment. It will be appreciated it if GDG would consider addressing this aspect in relation to the recommendations when appropriate.	not the main aim, and is sometimes impossible because of the obvious differences in current European economic and healthcare systems. The clinical and organizational impact of costs on recommendations will be considered in GDG meetings, and if relevant, described in the justification section. The economic feasibility of recommendations will not be covered. The treatment burden effects of high response have not been very well researched, still there is acknowledgement of this item in the guideline.
1	Raj Mathur	172	5335	The 'definition' of mild ovarian stimulation in the glossary is not a definition of treatment but of the intention of treatment ('of limiting the number of oocytes'). This leads to much confusion among clinicians and patients alike. Mild stimulation becomes a moveable feast, based on what the clinician intends rather than what the clinician does. Would the GDG consider putting a dose of FSH in addition to the intention as part of the definition of mild ovarian stimulation? I note that the guideline already states that a dose of 150 to 225 iu FSH is 'conventional'. Surely this implies that a dose less than 150 iu is 'mild'?	It is this confusion on Definition that has lead to the decision to not include mild stimulation in the Guideline. The GDG has added to the recommendation on reduced dose, a dose between 100 and <150 IU gonadotropins.