APPENDIX

Appendix 1. Needs assessment survey.

CUA Guideline on Azoospermia: Needs assessment/semi-structured interview

CUA Guideline Committee Member:_____

Interviewee Demographics:

	Interviewee #1	Interviewee #2
Practice Location		
Practice type (acad/comm)		
Years in practice		
Do you currently see outpatient consults for Infertility?		
Do you perform sperm retrievals?		
Do you perform vasectomy reversals?		
Do you refer your patients with azoospermia to another urologist who specializes in infertility		
Do you perform investigations of azoospermic men? If so which do you order?		

How would you describe your current level of comfort in treating men with azoospermia? (Rate 1-5, with 1=very uncomfortable, 5=very comfortable)

Interview Questions:

1. Have you made use of existing clinical practice guidelines on infertility (CUA, AUA, EAU, ASRM etc)? If so, how helpful have these been in guiding your clinical care? Any comments?

Interviewee 1:

Interviewee 2:

2. Within your current scope of practice as it relates to management of azoospermia, what are some topic areas where you would like to have more information and guidance (ie gaps in your knowledge or skills)?

Interviewee 1:

Interviewee 2:

3. What topics would you like to see addressed in the new CUA guideline on azoospermia? (can provide prompts to interviewee with examples such as: role of the urologist in doing the initial lab evaluation etc, how to optimize hormone status, surgical sperm retrieval, surgical reconstruction techniques for vasal or epididymal obstruction, sperm retrieval for men lacking ejaculation and emission)

Interviewee 1:

Interviewee 2:

4. If the new CUA guideline provided evidence-based recommendations on aspects of azoospermia management that you do NOT currently perform (eg. hormone therapy to optimize non-obstructive azoospermia, surgical sperm retrieval, genetic investigations etc), do you think you would potentially expand the scope of your practice?

Interviewee 1:

Interviewee 2:

5. Any other comments/suggestions?

		JUDGEMENT											
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large		Varies	Don't know						
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial		Varies	Don't know						
CERTAINTY OF EVIDENCE	Very low	Low	Moderate	High			No include studies						
VALUES	Important uncertainty or variability	Possibly important uncertainty or variability	Probably no important uncertainty or variability	No important uncertainty or variability									
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	Don't know						
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings	Varies	Don't knov						
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High			No include studies						
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	No include studies						
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased	Varies	Don't knov						
ACCEPTABILITY	No	Probably no	Probably yes	Yes		Varies	Don't know						
FEASIBILITY	No	Probably no	Probably yes	Yes		Varies	Don't knov						

Appendix 2. Summary of judgements for cryopreservation of surgically retrieved sperm.

*Note: The intervention in this table is the use of fresh sperm, while the comparison is use of cryopreserved sperm.

Appendix 3. Summary of findings for cryopreservation of surgically retrieved sperm.

Summary of findings:

Sperm cryopre	eservation co	mpared to fres	sh sperm in n	nen with non	-obstructiv	ve azoospermia
Patient or populat Setting: Intervention: sperr Comparison: fresh	n cryopreservatior	-obstructive azoosper	rmia			
		bsolute effects[*] 5% Cl)		Nº of	Certainty of	
Outcomes	Risk with fresh sperm	Risk with sperm cryopreservation	Relative effect (95% Cl)	participants (studies)	the evidence (GRADE)	Comments
Clinical pregnancy	383 per 1,000	345 per 1,000 (307 to 387)	RR 0.90 (0.80 to 1.01)	2084 (22 observational studies)	⊕OOO Very low ^{a,b,c}	The evidence is very uncertain about the effect of sperm cryopreservation on clinical pregnancy. In every 1000 NOA couples who use cryopreserved sperm compared to fresh sperm, 38 fewer (95% Cl from 76 fewer to 4 more) couples have clinical pregnancies.
Live birth	339 per 1,000	261 per 1,000 (227 to 302)	RR 0.77 (0.67 to 0.89)	1973 (21 observational studies)	⊕OOO Very low ^{a,b}	The evidence is very uncertain about the effect of sperm cryopreservation on live birth. In every 1000 NOA couples who use cryopreserved sperm compared to fresh sperm, 78 fewer (95% CI from 112 fewer to 37 fewer) couples have live birth.

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

Cl: confidence interval; RR: risk ratio

GRADE Working Group grades of evidence High certainty: we are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect. Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Studies are at high risk of bias, mainly for a lack of confounder measurement and adjustment. b. As the individual study results vary considerably with very wide confidence intervals, we decided to rate down by one level for inconsistency and imprecision. c. Based on the review of the funnel plot, the chances of publication bias are suspected.

••	1		5		•	1			
	Cryopreserved	Sperm	Fresh S	perm		Risk Ratio		Risk Ratio	Risk of Bias
itudy or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year	M–H, Random, 95% CI	ABCDEFGH
riedler 1997	1	11	5	23	0.5%	0.42 [0.06, 3.16]	1997	· · ·	
en-Yosef 1999	7	24	3	15	1.4%	1.46 [0.44, 4.79]	1999		- ••••••••
labermann 2000	3	9	1	3	0.6%	1.00 [0.16, 6.35]	2000		— ••••••
riedler 2002	15	63	16	65	5.4%	0.97 [0.52, 1.79]	2002		•••••
lauser 2005	2	13	2	13	0.6%	1.00 [0.16, 6.07]	2005		— ••••••
/u 2005	10	24	2	6	1.3%	1.25 [0.37, 4.26]	2005		••••••
onc 2008	14	93	10	64	3.6%	0.96 [0.46, 2.03]	2008		••••
karsu 2009	0	2	3	4	0.3%	0.24 [0.02, 3.19]	2009	← • • • • • • • • • • • • • • • • • • •	
alsi 2010	4	7	13	41	3.3%	1.80 [0.82, 3.94]	2010		•••••
aeacan 2013	25	110	27	99	9.1%	0.83 [0.52, 1.34]	2013		
aheem 2013	9	64	6	31	2.3%	0.73 [0.28, 1.86]	2013		••••
avukcuoglu 2013	12	39		43	5.4%	0.83 [0.45, 1.52]	2013		₽₽₽₽₽₽ ?
Madureira 2014	4	17		20	2.2%	0.47 [0.18, 1.23]	2014		
ark 2015	2	49	5	61	0.8%	0.50 [0.10, 2.46]	2015	· · · · · · · · · · · · · · · · · · ·	••••
chachter-safrai 2017	9	48		22	0.5%	4.13 [0.56, 30.58]			→ ������?
)kuyama 2018 (KF)	3	18		19	1.1%	0.79 [0.21, 3.06]	2018		
0 NOA) 002 (NOA)	12	78		71	5.3%	0.47 [0.26, 0.88]	2018		
alah 2019	1	36		32	0.3%			• •	─→ ? � ₽₽₽₽₽
Cavoussi 2020	20	38		29	11.7%	0.85 [0.56, 1.28]			?
arros 2021	4	21		23	2.2%			·	
hang 2021 b.	12	30		40	9.0%	0.53 [0.33, 0.86]			
hang 2021.	40	110		234	25.1%	0.79 [0.59, 1.05]			
/ang 2022	14	43	32	68	8.1%	0.69 [0.42, 1.14]	2022		
otal (95% CI)		947		1026	100.0%	0.77 [0.67, 0.89]		•	
otal events	223		348						
leterogeneity: Tau ² = 0	0.00; Chi ² = 19.96	, df = 22	P = 0.5	9); I ² = (О%			0.2 0.5 1 2	<u>+</u>
est for overall effect: Z	f = 3.62 (P = 0.00)	03)						Favours Fresh Favours Cryopr	eserved
tisk of bias legend									
A) Selection bias									
B) Exposure measurem	ent bias								
C) Outcome ascertion b									
D) Confounder adjustn									
E) Confounder assessm									

Appendix 4A. Forest plot of adjuvant cryopreservation on live birth rates.

Appendix 4B. Forest plot demonstrating a sensitivity analysis of adjuvant cryopreservation on live birth rates among studies using an intention-to-treat-like methodology.

	Cryopreserved	Sperm	Fresh S	perm		Risk Ratio	Risk Ratio	Risk of Bias
tudy or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% CI	ABCDEFGH
karsu 2009	0	2	3	4	0.5%	0.24 [0.02, 3.19]	←	
en-Yosef 1999	7	22	3	14	2.6%	1.48 [0.46, 4.81]		- •••••••••
riedler 1997	1	14	5	18	0.9%	0.26 [0.03, 1.96]	←	
riedler 2002	15	50	16	50	10.4%	0.94 [0.52, 1.68]		
avoussi 2020	21	38	18	29	21.8%	0.89 [0.59, 1.33]		? • • • • • • • •
Madureira 2014	4	13	10	19	4.2%	0.58 [0.23, 1.47]		
hang 2021	40	116	108	222	43.8%	0.71 [0.53, 0.94]		
hang 2021 b	12	30	30	40	15.9%	0.53 [0.33, 0.86]		
otal (95% CI)		285		396	100.0%	0.73 [0.60, 0.88]	•	
otal events	100		193				-	
leterogeneity: $Tau^2 = 0$	0.00; Chi ² = 6.7	5, df = 7	(P = 0.46)	5); $I^2 = 0$)%			<u>+</u>
est for overall effect: Z	Z = 3.26 (P = 0.)	001)					0.2 0.5 1 2 Favours Fresh Favours Cryopi	5 reconved
							ravours riesir ravours cryopi	eserveu

(**B**) Exposure measurement bias

(F) Outcome measurement bias
 (G) Sufficient Follow-up
 (H) Co-intervention bias

(C) Outcome ascertion before exposure

(D) Confounder adjustment bias

(E) Confounder assessment bias

(F) Outcome measurement bias

(G) Sufficient Follow-up

(H) Co-intervention bias

	Cryopreserved	Sperm	Fresh S	perm		Risk Ratio		Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year	M-H, Random, 95% CI	ABCDEFGI
riedler 1997	3	14	6	25	0.9%	0.89 [0.26, 3.03]	1997		
Ben-Yosef 1999	9	42	4	15	1.3%	0.80 [0.29, 2.23]	1999		
labermann 2000	6	9	1	3	0.5%	2.00 [0.38, 10.58]	2000		
riedler 2002	19	63	19	65	4.8%	1.03 [0.61, 1.76]	2002		
ousa 2002	9	37	17	50	2.9%	0.72 [0.36, 1.42]	2002		
/erheyen 2004	6	42	7	44	1.4%	0.90 [0.33, 2.45]	2004		
Vu 2005	15	24	2	6	1.0%	1.88 [0.58, 6.06]	2005		→ ����●��
lauser 2005	2	13	2	13	0.4%	1.00 [0.16, 6.07]	2005	•	\rightarrow 0 0 0 0 0
onc 2008	22	93	20	64	5.1%	0.76 [0.45, 1.27]	2008		
karsu 2009	0	2	3	4	0.2%	0.24 [0.02, 3.19]	2009	·	
alsi 2010	4	7	15	41	2.4%	1.56 [0.73, 3.33]	2010		
Kaeacan 2013	28	110	29	99	6.9%	0.87 [0.56, 1.35]	2013		
Raheem 2013	16	46	9	31	3.0%	1.20 [0.61, 2.36]	2013		
avukcuoglu 2013	17	39	19	43	5.6%	0.99 [0.60, 1.61]	2013		
Aadureira 2014	4	17	12	20	1.6%	0.39 [0.15, 0.99]	2014	< <u> </u>	
'ark 2015	16	49	9	61	2.6%	2.21 [1.07, 4.57]	2015		
chachter-safrai 2017	9	48	4	22	1.2%	1.03 [0.36, 2.99]	2017		
Okuyama 2018 (NOA)	33	78	36	71	11.2%	0.83 [0.59, 1.18]	2018		
Okuyama 2018 (KF)	7	18	5	19	1.5%	1.48 [0.57, 3.82]	2018		
alah 2019	5	36	6	32	1.2%	0.74 [0.25, 2.20]	2019		? + + + + + +
hang 2021 b	21	30	30	40	15.2%	0.93 [0.70, 1.25]	2021		
Barros 2021	4	21	14	23	1.5%	0.31 [0.12, 0.80]	2021	·	
hang 2021	47	110	116	234	20.6%	0.86 [0.67, 1.11]	2021		
Vang 2022	17	43	34	68	7.0%	0.79 [0.51, 1.23]	2022		
otal (95% CI)		991		1093	100.0%	0.90 [0.80, 1.01]		•	
otal events	319		419						
Heterogeneity: Tau ² = 0	.00; Chi ² = 23.23	df = 23	(P = 0.4)	5); $I^2 = 3$	1%			0.2 0.5 1 2	<u>-</u> +
est for overall effect: Z	= 1.72 (P = 0.09)))						0.2 0.5 1 2 Favours Fresh Favours Cryopr	
tisk of bias legend									
A) Selection bias									
B) Exposure measurem									
C) Outcome ascertion b									
D) Confounder adjustm									
E) Confounder assessm									
F) Outcome measureme	ent bias								

Appendix 5A. Forest plot of adjuvant cryopreservation on clinical pregnancy rates.

Appendix 5B. Forest Plot demonstrating a sensitivity analysis of adjuvant cryopreservation on clinical pregnancy rates among studies using an intention-to-treat-like methodology.

	Cryopreserved	Sperm	Fresh S	perm		Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl	ABCDEFGH
Akarsu 2009	0	2	3	4	0.4%	0.24 [0.02, 3.19]	←	
Ben-Yosef 1999	9	22	4	14	3.1%	1.43 [0.54, 3.77]		· ●●●●●●●
Friedler 1997	3	14	6	18	2.0%	0.64 [0.19, 2.13]		
Friedler 2002	19	50	19	50	11.7%	1.00 [0.61, 1.65]		
Madureira 2014	4	13	12	19	3.7%	0.49 [0.20, 1.18]		
Zhang 2021	47	116	116	222	45.4%	0.78 [0.60, 1.00]		
Zhang 2021 b	21	30	30	40	33.6%	0.93 [0.70, 1.25]		€€€€€€€
Total (95% CI)		247		367	100.0%	0.84 [0.71, 1.00]	•	
Total events	103		190					
Heterogeneity: Tau ² :	= 0.00; Chi ² $= 5.1$	6, $df = 6$	(P = 0.52)	2); $ ^2 = ($)%			<u></u>
Test for overall effect	t: $Z = 1.94 (P = 0.1)$	05)					0.2 0.5 1 2 Favours Fresh Favours Cryop	5 reserved
<u>Risk of bias legend</u>								
(A) Selection bias								
(P) Exposure mossure	amont bias							

(B) Exposure measurement bias

(C) Outcome ascertion before exposure

(D) Confounder adjustment bias

(E) Confounder assessment bias

(F) Outcome measurement bias (G) Sufficient Follow-up

(H) Co-intervention bias

(G) Sufficient Follow-up (H) Co-intervention bias

		JUDGEMENT									
		T	JU	DGEMENT		Γ	T				
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large		Varies	Don't know				
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial		Varies	Don't know				
CERTAINTY OF EVIDENCE	Very low	Low	Moderate	High			No includeo studies				
VALUES	Important uncertainty or variability	Possibly important uncertainty or variability	Probably no important uncertainty or variability	No important uncertainty or variability							
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	Don't know				
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings	Varies	Don't know				
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High			No included studies				
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	No includeo studies				
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased	Varies	Don't know				
ACCEPTABILITY	No	Probably no	Probably yes	Yes		Varies	Don't know				
FEASIBILITY	No	Probably no	Probably yes	Yes		Varies	Don't knov				

Appendix 6. Summary of judgements for neoadjuvant varicocele repair in NOA.

Note: Intervention in this table represents performing a neo-adjuvant varicocele repair in NOA, while the comparison represents observation of the varicocele.

Appendix 7. Summary of findings for neoadjuvant varicocele repair in NOA.

Surgical varicoco and varicocele	ele repair cor	npared to no	varicocele re	epair in men v	with non-ol	bstructive azoospermia
Patient or population Setting: Intervention: surgical Comparison: no varice	varicocele repair	ostructive azoosper	mia and varicoce	le		
		solute effects [*] % CI)				
Outcomes	Risk with no varicocele repair	Risk with surgical varicocele repair	Relative effect (95% Cl)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
Clinical pregnancy	95 per 1,000	167 per 1,000 (40 to 692)	RR 1.75 (0.42 to 7.27)	87 (1 observational study)	⊕OOO Very low ^{a,b}	The evidence is very uncertain about the effect of surgical varicocele repair on clinical pregnancy. In ever 1000 NOA men who undergo varicocele repair, 72 more men (95% Cl from 55 fewer to 597 more) have clinical pregnancy.
Live birth	95 per 1,000	136 per 1,000 (32 to 582)	RR 1.43 (0.34 to 6.11)	87 (1 observational study)	⊕OOO Very low ^{a,b}	The evidence is very uncertain about the effect of surgical varicocele repair on live birth rate. In every 1000 NOA men who undergo varicocele repair, 41 more men (95° Cl from 63 fewer to 487 more) have Live birth.
Sperm retrieval	514 per 1,000	612 per 1,000 (421 to 889)	RR 1.19 (0.82 to 1.73)	260 (3 observational studies)	⊕OOO Very low ^{c,d}	The evidence is very uncertain about the effect of surgical varicocele repair on sperm retrieval. In every 1000 NOA men who undergo varicocele repair, 98 more men (95° Cl from 93 fewer to 375 more) have sperm retreival.

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

Cl: confidence interval; RR: risk ratio

GRADE Working Group grades of evidence

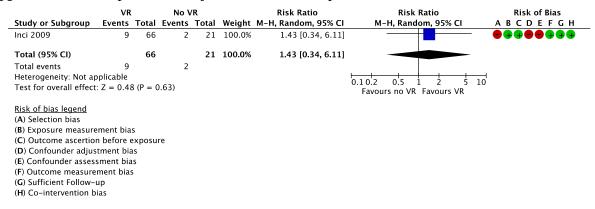
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect. Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

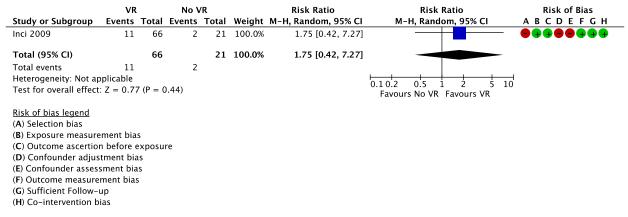
Explanations

a. The only included study is at a high risk of bias.
b. Extremely few events and wide confidence interval.
c. None of the included studies is at low risk of bias.
d. Wde confidence interval with different boundary interpretations.

Appendix 8. Forest plot of neoadjuvant varicocele repair on live birth rate.



Appendix 9. Forest plot of neoadjuvant varicocele repair on clinical pregnancy rate.



Appendix 10. Forest plot of neoadjuvant varicocele repair on sperm retrieval rate.

I I		1				1	1	
	VR		No V	′R		Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% CI	ABCDEFGH
Inci 2009	35	66	9	21	29.7%	1.24 [0.72, 2.13]		
Schlegel 2004	41	68	42	70	57.0%	1.00 [0.77, 1.32]	-	
Zampieri 2013	11	19	4	16	13.3%	2.32 [0.91, 5.88]		$\bullet \bullet \bullet ? ? \bullet \bullet \bullet$
Total (95% CI)		153		107	100.0%	1.19 [0.82, 1.73]	•	
Total events	87		55					
Heterogeneity: Tau ² =	= 0.04; Cł	1i ² = 3.	21, df =	2 (P =	0.20); I ² =	= 38%	0.05 0.2 1 5	20
Test for overall effect	z = 0.94	4 (P = 0)).35)				Favours No VR Favours VR	20
<u>Risk of bias legend</u>								

- (A) Selection bias
- (B) Exposure measurement bias (C) Outcome ascertion before exposure
- (D) Confounder adjustment bias
- (E) Confounder assessment bias
- (F) Outcome measurement bias
- (G) Sufficient Follow-up
- (H) Co-intervention bias

	JUDGEMENT												
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large		Varies	Don't know						
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial		Varies	Don't know						
CERTAINTY OF EVIDENCE	Very low	Low	Moderate	High			No included studies						
VALUES	lmportant uncertainty or variability	Possibly important uncertainty or variability	Probably no important uncertainty or variability	No important uncertainty or variability									
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	Don't know						
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings	Varies	Don't know						
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High			No included studies						
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	No included studies						
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased	Varies	Don't know						
ACCEPTABILITY	No	Probably no	Probably yes	Yes		Varies	Don't know						
FEASIBILITY	No	Probably no	Probably yes	Yes		Varies	Don't know						

Appendix 11. Summary of judgements for neoadjuvant hormone therapy in NOA.

Note: Intervention in this table represents use of neo-adjuvant hormone therapy in NOA, where the comparison represents no treatment with respect to hormone therapy for the purpose of improving semen analysis parameters alone.

Appendix 12. Summary of findings for neoadjuvant hormone therapy in NOA.

Summary of findings:

Hormonal treatment compared to no hormonal treatment in men with non-obstructive azoospermia

Patient or population: men with non-obstructive azoospermia Setting: Intervention: hormonal treatment Comparison: no hormonal treatment

		osolute effects [*] % Cl)				
Outcomes	Risk with no hormonal treatment	Risk with hormonal treatment	Relative effect (95% Cl)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
Clinical pregnancy	224 per 1,000	211 per 1,000 (128 to 343)	RR 0.94 (0.57 to 1.53)	582 (3 observational studies)	⊕OOO Very low ^{a,b,c}	The evidence is very uncertain about the effect of hormonal treatment on clinical pregnancy.In the two observational studies that used hCG FSH, a mixed approach to improve male infertility through hormonal treatments, 13 fewer couples (95% CI from 96 fewer to 119 more) had a clinical pregnancy in every 1000 men who received the treatment compared to no treatment.
Live birth rate	257 per 1,000	193 per 1,000 (134 to 278)	RR 0.75 (0.52 to 1.08)	615 (3 observational studies)	⊕⊖⊖⊖ Very low ^{a,c}	The evidence is very uncertain about the effect of hormonal treatment on the live birth rate. In every 1000 NOA men who received hormonal treatment compared to no treatment, 64 fewer couples (95% Cl from 123 fewer to 21 more) had a live birth.
Sperm retrieval	362 per 1,000	506 per 1,000 (365 to 698)	RR 1.40 (1.01 to 1.93)	1540 (11 observational studies)	000 Very low ^{a,c,d}	The evidence is very uncertain about the effect of hormonal treatment on sperm retrieval rate. In every 1000 NOA men who received hormonal treatment compared to no treatment, 144 more men (95% CI from 3 more to 336 more) had sperm retreival.

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

CI: confidence interval; RR: risk ratio

GRADE Working Group grades of evidence High certainty: we are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. The studies included in the analysis are at high risk of bias especially because of lack of confounder measurement and adjustment.

- b. Results of the study using FSH was considerably different from the other subgroups
 c. The confidence interval of the absolute effect crosses the line of no impact and the boundaries fall into different interpretation zones.
- d. Visual inspection of the confidence intervals demonstrates heterogeneity.

	Hormone Th		No Trea			Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight M	I-H, Random, 95% CI	M-H, Random, 95% Cl	ABCDEFGH
1.1.1 AI								
Subtotal (95% CI)		0		0		Not estimable		
Total events	0		0					
Heterogeneity: Not applicable								
Test for overall effect: Not appl	licable							
1.1.2 hCG								
Gul 2016	10	34	17	49	32.2%	0.85 [0.44, 1.62]		
Guo 2020 (non-mosaic KF)	15	134	7	50	19.3%	0.80 [0.35, 1.85]		
Subtotal (95% CI)		168		99	51.5%	0.83 [0.50, 1.38]		
Total events	25		24				-	
Heterogeneity: $Tau^2 = 0.00$; Ch	$hi^2 = 0.01, df$	= 1 (P =	0.91); I ² =	= 0%				
Test for overall effect: $Z = 0.72$	P = 0.47							
1.1.3 SERMs								
Subtotal (95% CI)		0		0		Not estimable		
Total events	0		0					
Heterogeneity: Not applicable								
Test for overall effect: Not appl	licable							
1.1.4 Mix								
Reifsnyder 2012 (low T level)	60	307	12	41	48.5%	0.67 [0.39, 1.13]		
Subtotal (95% CI)		307		41	48.5%	0.67 [0.39, 1.13]		
Total events	60		12					
Heterogeneity: Not applicable								
Test for overall effect: $Z = 1.50$	(P = 0.13)							
Total (95% CI)		475		140	100.0%	0.75 [0.52, 1.08]		
Total events	85		36				-	
Heterogeneity: Tau ² = 0.00; Ch	$hi^2 = 0.35, df$	= 2 (P =	0.84); I ² =	= 0%			0.1 0.2 0.5 1 2 5	
Test for overall effect: Z = 1.56							0.1 0.2 0.5 1 2 5 Favours no tratment Favours hormone	10 troatmont
Test for subgroup differences:	$Chi^2 = 0.33, c$	df = 1 (P	= 0.56), I	$^{2} = 0\%$			ravours no traument ravours normone	treatment
Risk of bias legend								
(A) Selection bias								
(B) Exposure measurement bias	s							
(C) Outcome ascertion before e								
(D) Confounder adjustment bia	15							
(E) Confounder assessment bia	s							
(F) Outcome measurement bias								
(G) Sufficient Follow-up								
(H) Co-intervention bias								

Appendix 13. Forest plot of neoadjuvant hormone therapy on live birth rates.

(H) Co-intervention bias

Churcher and Carls and an	Hormone Th		No Treat		M/-:	Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup 1.3.1 Al	Events	Total	Events	Total	weight	M-H, Random, 95% CI	M-H, Random, 95% CI	ABCDEFGH
Subtotal (95% CI)		0		0		Not estimable		
Total events Heterogeneity: Not applicable	0	Ū	0	Ū		Hot estimatic		
Test for overall effect: Not app	licable							
1.3.2 hCG								
Guo 2020 (non-mosaic KF) Subtotal (95% CI)	22	134 134	9	50 50	33.0% 33.0%	0.91 [0.45, 1.84] 0.91 [0.45, 1.84]		
Fotal events Heterogeneity: Not applicable Fest for overall effect: Z = 0.20	22 6 (P = 0.80)		9					
	0 (1 = 0.00)							
1.3.3 FSH								
Cocci 2018 Subtotal (95% CI)	7	25 25	3	25 25	13.8% 1 3.8%	2.33 [0.68, 8.01] 2.33 [0.68, 8.01]		? 🗣 🗣 🗣 🗣 🗣
Total events Heterogeneity: Not applicable Test for overall effect: Z = 1.3!	7 5 (P = 0.18)		3					
1.3.4 SERMs								
Subtotal (95% CI)		0		0		Not estimable		
Total events Heterogeneity: Not applicable	0		0					
Test for overall effect: Not app	licable							
1.3.5 Mix								
Reifsnyder 2012 (low T level) Subtotal (95% CI)	79	307 307	14	41 41	53.2% 53.2%	0.75 [0.47, 1.20] 0.75 [0.47, 1.20]		
Total events Heterogeneity: Not applicable	79		14				-	
Test for overall effect: $Z = 1.19$	9 (P = 0.23)							
Total (95% CI)		466		116	100.0%	0.94 [0.57, 1.53]		
Total events	108		26				T	
Heterogeneity: Tau ² = 0.06; Cl		= 2 (P =	0.24); I ² =	31%			0.1 0.2 0.5 1 2	5 10
Test for overall effect: $Z = 0.20$							Favours no tratment Favours horm	
Test for subgroup differences:	$Chi^{2} = 2.84, c$	at = 2 (P)	= 0.24), I	= 29.5	%			
<u>Risk of bias legend</u> (A) Selection bias								
 B) Exposure measurement bia 	IC .							
(C) Outcome ascertion before								
(D) Confounder adjustment bia								
(E) Confounder assessment bia								

Appendix 14. Forest plot of neoadjuvant hormone therapy on clinical pregnancy rates.

(E) Confounder assessment bias
 (F) Outcome measurement bias
 (G) Sufficient Follow-up
 (H) Co-intervention bias

	Hormone T		No Treat			Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup 1.2.1 Al	Events	Total	Events	Total	weight M	I-H, Random, 95% CI	M-H, Random, 95% CI	ABCDEFGI
		0		0		Not estimable		
Subtotal (95% CI)		0		0		Notestimable		
Fotal events	0		0					
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
1.2.2 hCG								
Gul 2016	17	34	28	49	14.3%	0.88 [0.58, 1.32]		
Guo 2020 (non-mosaic KF)	58	134	22	50	15.0%	0.98 [0.68, 1.42]	_	
Sen 2020	9	12	4	12	8.0%	2.25 [0.95, 5.34]		? • ? • • ? •
Shiraishi 2013 (redo mTESE)	2	13	Ó	10	1.1%	3.93 [0.21, 73.71]		
Shiraishi 2013 (redo mTESE) [+ recFSH]	4	15	0	10	1.2%	6.19 [0.37, 103.71]		
Subtotal (95% CI)		208	•	131	39.6%	1.15 [0.76, 1.73]		•••••
Fotal events	90		54				-	
Heterogeneity: Tau ² = 0.07; Chi ² = 6.36	, df = 4 (P = 0)	.17); I ² =	37%					
Test for overall effect: $Z = 0.67$ ($P = 0.50$	D)							
1.2.3 SERMs								
Subtotal (95% CI)		0		0		Not estimable		
Total events	0		0					
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
1.2.4 FSH								
Aydos 2003	40	63	15	45	13.6%	1.90 [1.21, 3.00]		
Cocci 2018	10	25	7	25	8.8%	1.43 [0.65, 3.15]		
Subtotal (95% CI)	10	88		70		1.77 [1.20, 2.63]		
Total events	50		22		2211/0	107 [1120, 2105]		
Heterogeneity: Tau ² = 0.00; Chi ² = 0.38 Test for overall effect: Z = 2.85 (P = 0.00	, df = 1 (P = 0)	.54); I ² =						
1.2.5 Mix								
Amer 2020	2	20	0	20	1.1%	5.00 [0.26, 98.00]		─→ ? ₽₽ ₽ ₽₽₽
Majzoub 2016	6	16	0	4	1.3%	3.82 [0.26, 56.78]		
Reifsnyder 2012 (low T level)	157	307	25	41	16.6%	0.84 [0.64, 1.10]		
Subtotal (95% CI)		343		65	19.0%	1.32 [0.41, 4.26]		
Total events	165		25					
Heterogeneity: Tau ² = 0.47; Chi ² = 2.90 Test for overall effect: Z = 0.47 (P = 0.64		.23); I ² =	31%					
1.2.6 Hormonal Treatment Strategy								
Hussain 2012	306	496	39	116	16.6%	1.83 [1.41, 2.39]		
Sujenthiran 2019	6	15	1	8	2.4%	3.20 [0.46, 22.16]		
Subtotal (95% CI)	-	511	-	124	19.0%	1.85 [1.43, 2.41]		
Total events	312		40					
Heterogeneity: Tau ² = 0.00; Chi ² = 0.31 Test for overall effect: Z = 4.61 (P < 0.00		.58); I ² =	0%					
Fotal (95% CI)		1150		300	100.0%	1.40 [1.01, 1.93]		
Fotal events	617	1150	141	550	-00.070	1.10 [1.01, 1.55]	-	
Heterogeneity: $Tau^2 = 0.14$; $Chi^2 = 32.9$		= 0.0005						
Test for overall effect: $Z = 2.04$ (P = 0.04							0.1 0.2 0.5 1 2 5	10
Test for subgroup differences: $Chi^2 = 3.9$		0.26) 12	= 24.7%				Favours no tratment Favours hormone to	eatment
Risk of bias legend								
A) Selection bias								
B) Exposure measurement bias								
(b) Exposule measurement bias								

Appendix 15. Forest plot of neoadjuvant hormone therapy on sperm retrieval rates.

(B) Exposure measurement bias
(C) Outcome ascertion before exposure
(D) Confounder adjustment bias
(E) Confounder assessment bias
(F) Outcome measurement bias
(G) Sufficient Follow-up
(H) Co-intervention bias