SUMMARY OF FINDINGS TABLES IN CGF REVIEWS

How to avoid common problems

Jane Marjoribanks, Helen Nagels, Toby Lasserson and Newton Opiyo

Cochrane Gynaecology and Fertility 1996-2016

Advancing women's health through evidence
Oxford 1-3 April 2016

AN SoF WITH PROBLEMS

Ensure headings make sense – edit as required.

Fill in setting.

Should surgery versus LNG-IUS for heavy menstrual bleeding?

Patient or population: Heavy menstrual bleeding
Setting:
Intervention: Surgery
Comparison: LNG-IUS

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with LNG-IUS	Risk with Surgery				
Objective control of bleeding at one year: menstrual loss < 80 ml per cycle – Hysterectomy versus LNG-IUS	**Study population** 465 per 1000	**516 per 1000** (488 to 553)	**RR 1.11** (1.05 to 1.19)	73 (1 RCT)	⊕⊕⊕⊕ HIGH	
	Moderate 465 per 1000	516 per 1000 (488 to 553)				
Final PGWBI score: thermal balloon ablation versus LNG-IUS	The mean final PGWBI score: thermal balloon ablation versus LNG-IUS was 0	The mean final PGWBI score: thermal balloon ablation versus LNG-IUS in the intervention group was 10.3 fewer (26.54 fewer to 5.94 more)	–	161 (2 RCTs)	⊕⊕◯◯[1,3,4] LOW	
Quality of life: any surgery versus LNG-IUS		The mean Quality of Life score: surgery versus LNG-IUS in the intervention group was SMD 0.39 fewer (0.75 fewer to 0.03 fewer)		389 (4 RCTs)	⊕⊕◯◯[2] LOW	
Additional surgery for HMB received by two years - Thermal balloon versus LNG-IUS	**Study population** 205 per 1000	**132 per 1000** (62 to 279)	**RR 0.64** (0.30 to 1.36)	142 (2 RCTs)	⊕◯◯◯[2,3,4,5] VERY LOW	
	Moderate 209 per 1000	**134 per 1000** (63 to 284)				

*__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; **MD:** Mean difference.

Adverse events are a priority outcome but no studies reported it.

Add a line to the SoF to highlight this lack of evidence.

Level of evidence is implausible due to small sample size, which indicates imprecision.

Delete three superfluous lines – choose median (Moderate line) or mean (Study pop line) and explain choice in footnote.

Only include multiple lines if there is a clinical rationale, which must be explained in footnotes.

Specify whether mean or median used.

Continuous data always require interpretation.

Specify what measure is used, what the scale is, what the direction of effect is and what the effect estimate means.

GRADE – Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: 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 quality: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

1. Downgraded one level for serious risk of bias: no blinding.
2. Downgraded two levels for serious risk of bias: 25% attrition rate, no blinding.
3. Wide confidence interval.
4. Crosses line of no effect.
5. Small sample size

Use GRADE terminology and explain downgrading decisions.

In the example, footnotes 3-5 are all indicators of imprecision and can share a footnote.

State the follow-up time for each outcome.

THE SAME SoF AFTER EDITING IN REVMAN

Surgery versus levonorgestrel intrauterine device (LNG-IUS) in women with heavy menstrual bleeding

Population: Women with heavy menstrual bleeding
Setting: Inpatient or outpatient
Intervention: Surgery
Comparison: LNG-IUS

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with LNG-IUS	Risk with Surgery				
Objective control of bleeding at one year: menstrual loss < 80 ml per cycle - Hysterectomy versus LNG-IUS	465 per 1000	**516 per 1000** (488 to 553)	**RR 1.11** (1.05 to 1.19)	73 (1 RCT)	⊕⊕◯◯[1] LOW	
Adverse events	No usable data on adverse events reported in any of the included studies, therefore no conclusions can be drawn					
Final Psychological General Wellbeing Index (PGWBI) score at one to two years: thermal balloon ablation versus LNG-IUS	There was no evidence of a difference between the groups. The mean final PGWBI score was 10.3 points lower in the thermal balloon ablation group than in the LNG-IUS group (95% CI 26.54 lower to 5.94 more).		—-	161 (2 RCTs)	⊕⊕◯◯[2,3] LOW	PGWBI: possible range 0-110. High score is positive.
Quality of life: any surgery versus LNG-IUS at two to three years. Measured with SMD as studies used four different questionnaires.	Quality of life was better in the surgical group. Findings were imprecise and compatible with an effect that ranged from clinically meaningless to moderately large (SMD 0.39, 95% CI 0.03 to 0.75)			389 (4 RCTs)	⊕⊕◯◯[4] LOW	In all questionnaires a high score was positive. Effect sizes can be interpreted as small (SMD = 0.2), moderate (SMD = 0.5), and large (SMD ≥ 0.8).
Additional surgery for HMB received by two years – Thermal balloon versus LNG-IUS	209 per 1000	**134 per 1000** (63 to 284)	**RR 0.64** (0.30 to 1.36)	142 (2 RCTs)	⊕◯◯◯[3,4] VERY LOW	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the median risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk ratio; **SMD:** Standardised mean difference.

GRADE – Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: 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 quality: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

1. Downgraded two levels for very serious imprecision: single small study with few events, findings compatible with benefit in surgical arm or with no clinically meaningful effect.
2. Downgraded one level for serious risk of bias: no blinding.
3. Downgraded one level for serious imprecision: small sample, wide confidence interval which crosses line of no effect.
4. Downgraded two levels for very serious risk of bias: 25% attrition rate, no blinding.

HANDY HINT

Use these numbers to express absolute effect:

e.g. "If 47% of women have objective control of bleeding with LNG-IUS, between 49% and 55% will do so after surgery".

KEY MESSAGES

- **Include all priority outcomes in the SoF and report the same set in the abstract, plain language summary and main conclusions.**
  Choose priority outcomes early on, preferably at protocol stage or before starting an update.
- **Explain the clinical meaning of effect estimates from continuous data.**
- **Explain in footnote which GRADE consideration the outcome has been downgraded for, and whether by one or two levels.**

Cochrane Gynaecology and Fertility 1996-2016