

GRADE guidelines 11 Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes

Gordon Guyatt, Andrew D. Oxman, Shahnaz Sultan, Jan Brozek, Paul Glasziou, Pablo Alonso-Coello, David Atkins, Regina Kunz, Victor Montori, Roman Jaeschke, David Rind, Philipp Dahm, Elie A. Akl, Joerg Meerpohl, Gunn Vist, Elise Berliner, Susan Norris, Yngve Falck-Ytter, Holger J. Schünemann

【日本語訳】

GRADE ガイドライン: 11. 単一のアウトカムならびに全てのアウトカムに関する効果推定値への確信性の総合評価

抄録

GRADEはガイドライン作成者に対し、各重要または重大なアウトカムに関する効果推定値の確信性(エビデンスの質: 高、中、低、非常に低)を全体的に評価することを求めている。GRADEは、各アウトカムについて、効果推定値への確信性の評価を下げる5つの領域の要因を個別に初期検討することを提唱しており、これにより、システマティック・レビューの著者やガイドライン作成者はアウトカム特有の確信性評価を行うことが可能となる。この評価システムは順序尺度における離散的な値として表されるが、推定値への確信性を連続体として見ることに役立ち、最終的な確信性の評価は各領域の個別検討により示唆される評価とは異なるかもしれない。

効果推定値への確信性の全体的な評価は、推奨が行われる状況において関係するのみである。一般的に、確信性が最も低い重大なアウトカムに基づいて全体的な評価が行われる*。

*訳者解説: “効果推定値への確信性”とは従来の“エビデンスの質”のことです。

システマティック・レビューにおいては“効果推定値の確信性”を一つ一つのアウトカムについて評価するだけでアウトカム全体としては評価しません。一方、診療ガイドラインにおいては全てのアウトカムにわたる“効果推定値の確信性”を一つだけ提示しますが、その際のGRADEの原則は「重大なアウトカムに関する“効果推定値の確信性”を採用し、もしアウトカムの方向性が異なる場合は重大なアウトカムの中で“効果推定値の確信性”が最も低いものを採用する」です。

本JCE論文のTable 2(次頁)は、BMJシリーズやいくつかの他の論文でも利用されています。

例;書籍「診療ガイドラインのためのGRADEシステム」に記載しているエビデンス・プロファイル 2A-1 (膵臓癌患者における幽門輪温存膵頭十二指腸切除術 p77)

アウトカム全般に関する効果推定値への確信性 (JCE論文, p155より)

Determining the confidence in effect estimates across outcomes

～この例では、各アウトカムに関するエビデンスの質は、「中」から「非常に低」までさまざまである。ガイドライン作成者は治療による望ましくない帰結が最適な管理戦略に関する意思決定において重要だが重大ではないのか、もしくは重大なのかを考慮しなくてはならない。エビデンスの質が低いアウトカムが意思決定において重大なアウトカムである場合は、この質の低いエビデンスが全体的なエビデンスの質の評価に反映されていなければならない。確信性の低いアウトカムが重要ではあるが重大ではないアウトカムの場合、重大なアウトカムから得られた確信性の高い推定値が全体的な評価に反映される。したがってこの例では、推奨作成者が胃内容排出遅延を重大なアウトカムであると考えた場合、効果推定値の確信性についての全体的評価は「非常に低」となる。もし胃内容排出が重要ではあるが重大ではないアウトカムならば、5年生存率に関する確信性が「中」であっても、全体的な確信性は(明らかに重大なアウトカムである「周術期死亡率」に関する結果に基づき)「低」となる。

Table 2. GRADE Evidence Profile: different resection strategies for pancreatic carcinoma associated with different evidence quality of different outcomes (question: pylorus-preserving pancreaticoduodenectomy vs. standard Whipple pancreaticoduodenectomy in pancreatic or periampullary cancer?; settings: inpatients)

Quality assessment	Summary of findings										
	No. of patients					Absolute effect					
No of studies (design)	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	SWPD	PPPD	RR ^a (95% CI)	Control rate	Risk difference (95% CI)	Quality
Mortality at 5 years											
Three (RCT)	Serious limitations ^b	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	94/114	93/115	RR 0.98 (0.87–1.11)	825/1,000	20 fewer per 1,000 (-120 to +80)	⊕⊕⊕○ Moderate
In-hospital mortality											
Six (RCT)	Serious limitations ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	Undetected	12/244	4/246	RR 0.40 (0.14–1.13)	49/1,000	20 fewer per 1,000 (-50 to +10)	⊕⊕○○ Low
Biliary leaks											
Three (RCT)	Serious limitations ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	Undetected	0/133	2/135	RR 4.77 (0.23–97.96)	0/1,000	20 more per 1,000 (-20 to +50)	⊕⊕○○ Low
Delayed gastric emptying											
Five (RCT)	Serious limitations ^b	Serious inconsistency ^d	No serious indirectness	Serious imprecision ^c	Undetected	56/220	66/222	RR 1.52 (0.74–3.14)	255/1,000	110 more per 1,000 (-80 to +290)	⊕○○○ Very low
Blood transfusions (units)^e											
Five (RCT)	Serious limitations ^b	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	320		— ^e	Best estimate SWPD group 2.45	WMD -0.66 (-1.16 to -0.25)	⊕⊕⊕○ Moderate
Hospital stay (days)^e											
Five(RCT)	Serious limitations ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	Undetected	446		— ^e	19.17	WMD -1.45 (-3.28 to +0.38)	⊕⊕○○ Low

Abbreviations: SWPD, standard Whipple pancreaticoduodenectomy; PPPD, Pylorus-preserving pancreaticoduodenectomy; WMD, weighted mean difference.

^a All data based on random effect models.

^b Unclear allocation concealment in all studies, patients blinded in only one study, outcome assessors not blinded in any study, > 20% loss to follow-up in three studies, not analyzed using intention to treat in one study.

^c CI includes possible benefit from both surgical approaches.

^d Unexplained heterogeneity; $I^2 = 72.6\%$, $P = 0.006$.

^e Continuous outcome, therefore no relative effect is given.