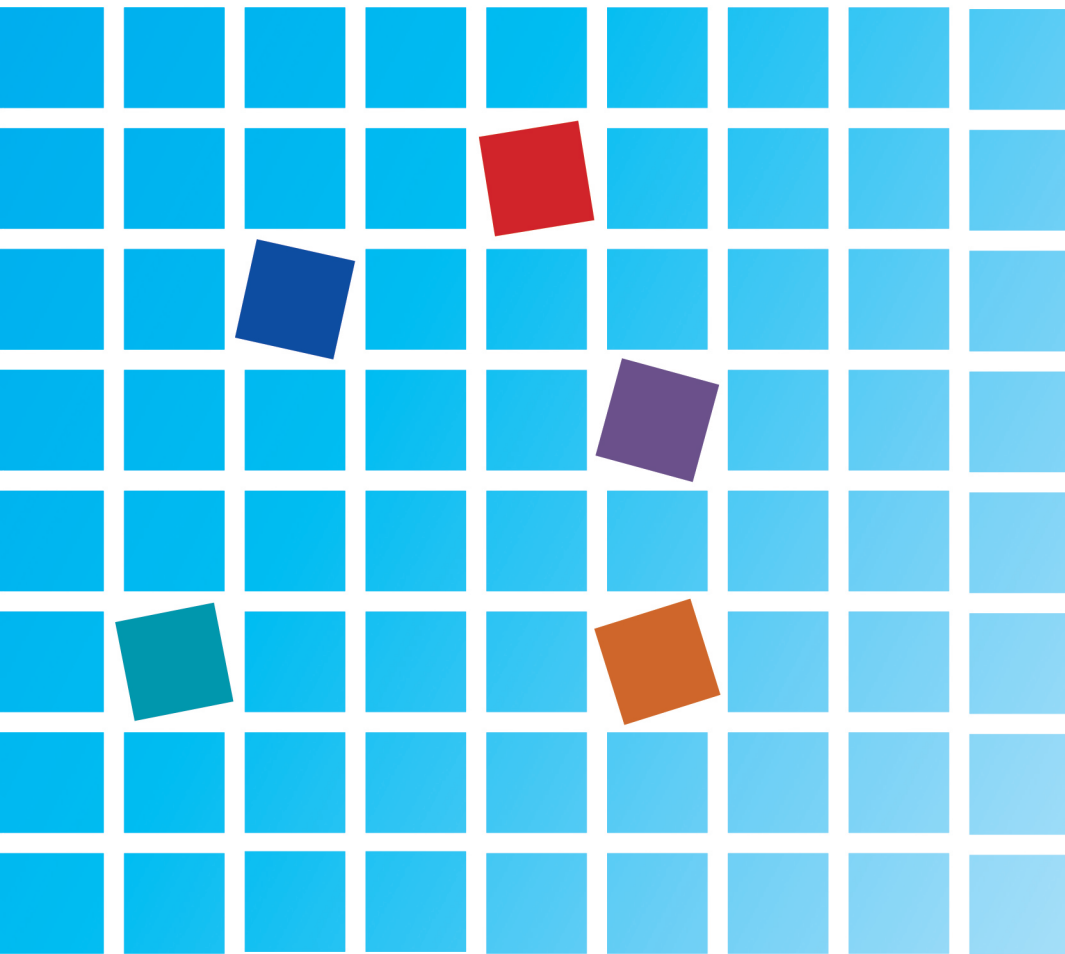




COHM-MOPE

GRADE

наръчник
за експерти

Аганмация по

Holger Schünemann

Jan Brožek

Gordon Guyatt

Andrew Oxman

artTracer
Publishing

COHM-MOPE

GRADE

наръчник
за експерти

Агантация по

Holger Schünemann

Jan Brožek

Gordon Guyatt

Andrew Oxman

Варна, 2025

GRADE НАРЪЧНИК ЗА ЕКСПЕРТИ

Аганмация no: Holger Schünemann, Jan Brožek, Gordon Guyatt, Andrew Oxman
<https://gdt.gradeapro.org/app/handbook/handbook.html#h.wlsfq2lmj0gb>

© **СОНМ** – *принципал*, 2025

© **Димитър КАЛЕВ** – *съставител*, 2025

© **Арт Трейсър ООД** – *издател*, 2025

ISBN 978-619-7094-76-3

ИЗПОЛЗВАНИ СЪКРАЩЕНИЯ

CI	confidence interval, доверителен интервал
EB-CPG	evidence-based clinical practice guideline ръководство за клинична практика, основано на доказателства
EP	evidence profile профил на доказателствата
GRADE	Grading of recommendations assessment, development and evaluations Степенуване на оценки за препоръки, разработване и оценявания
HR	hazard ratio коефициент на вероятност
HTA	health technology assessment оценка на здравна технология
ICEMAN	Instrument to assess the credibility of effect modification analyses инструмент за оценка на достоверността на анализите за промяна на ефекта
OIS	optimal information size оптимален размер на информация
OR	odds ratio коефициент на шанс
OS	observational study наблюдателно проучване
PICO	patient / intervention / comparison / outcome пациент / интервенция / компаратор / резултат

RCT randomized controlled trial
рандомизирано контролирано проучване

RR risk ratio
коефициент на риск

SoF summary of findings
обобщение на констатациите

SR systematic review
систематичен преглед

ВЪВЕДЕНИЕ

Подходът GRADE (**G**rating of **R**ecommendations **A**ssessment, **D**evelopment and **E**valuations) е система за **оценка на качеството на доказателства** (в систематични прегледи и други синтези на доказателства) и за **степенуване на препоръки в здравеопазването** (в клинични ръководства и оценки на здравни технологии). Методът предлага прозрачен и структуриран процес за разработване и представяне на обобщения на доказателства, както и за последователно извършване на стъпките, включени в разработване на препоръки. Може да се използва за съставяне на ръководства за клинична практика (CPG) и други препоръки за здравеопазването (напр. в общественото здраве, здравните политики, здравните системи и здравните фондове).

Подходът GRADE е одобрен и използван в световен мащаб от повече от 100 организации в 19 държави. Въпреки че е разработен главно за препоръки в клинични ръководства, той също е одобрен и използван (със или без модификации или адаптации) в практиката на систематични прегледи (systematic reviews/SRs) и като една от системите за оценка на доказателства при оценка на здравни технологии (health technology assessment/HTA). Cochrane Collaboration е приела системата GRADE от 2009 г. и използването ѝ вече е препоръчителна стъпка в рамките на Cochrane SRs за улесняване на прозрачна оценка на качеството на доказателствата.



Методологичен акцент

Подходът GRADE е епистема върху четири стълба: (1) PICO структуриран въпрос, (2) систематичен преглед/метаанализ, (3) обобщение и оценка на качеството на доказателствата и (4) категоризирана по сила препоръка

1. РАМКИРАНЕ НА КЛИНИЧЕН ВЪПРОС

Експертният панел за клинични ръководства трябва да определи обхвата на клиничното ръководство, основано на доказателства (EB-CPG), и планираните препоръки. Всяка препоръка трябва да отговаря на целенасочен и разумен въпрос, който води до клинично поведение. По същия начин и авторите на SRs трябва да формулират фокусирани въпроси за здравни грижи, на които прегледът да отговори.

Рамката PICO представлява добре структурирана методология за формулиране на въпроси, свързани със здравни грижи, и налага внимателно определяне на четири компонента:

- Пациент (P – patient): пациенти или популация, за които са предназначени препоръките;
- Интервенция (I – intervention): терапевтична, диагностична или друга интервенция, която се изследва;
- Сравнение (C – comparison): алтернативна интервенция; интервенция в контролна група;
- Резултат (O – outcome): резултат(и), който представлява интерес.

Когато се задава въпросът, проучванията и техният дизайн, които предоставят отговор, често не са известни. Това означава, че когато рандомизирани контролирани проучвания (RCTs) не са налични или не са свързани с висока степен на доверие в оценките, ще трябва да се използва информация от наблюдателни проучвания (OSs). Поради това обикновено не е разумно предварително да се дефинира дизайн на проучването.

Въпросите в клиничните ръководства често включват и друга специфика: условията, в която ръководството ще бъде прилагано. Например, ръководства, предназначени за богати на ресурси здравни системи, често биват неприложими към бедни системи. Така социалната среда може да се определи като част от дефинирането на популацията (напр. пациенти в страни с ниски доходи или в първична здравна помощ, или в селски здравни заведения).

Грешките, които често се допускат при формулиране на PICO въпрос, са невключване на всички важни за пациента резултати (напр. неблагоприятни ефекти или токсичност) или непълно разглеждане на всички съответни алтернативи (особено проблематично, когато ръководствата са адресирани към глобална аудитория).

Разработчиците на ръководства трябва да вземат предвид всички съответни резултати едновременно и да изберат **два различни формата** на въпроси за клинично поведение:

- При *[клиничен проблем и/или популация]* трябва ли *[интервенция]* спрямо *[сравнение]* да се използва за *диагностициране на [клиничен проблем]*?
- При *[клиничен проблем и/или популация]* трябва ли *[интервенция]* срещу *[сравнение]* да се използва за *лечение относно [целово състояние]*?

2. СИСТЕМАТИЧЕН ПРЕГЛЕД И МЕТААНАЛИЗ (SYSTEMATIC REVIEW & META-ANALYSIS)

Това е аналитично разглеждане на ясно формулиран въпрос, който използва систематични и прозрачни методи за идентифициране, избор и критична оценка на съответни изследвания и за събиране и анализ на данни от включените проучвания. Всеки резултат в SRs се разглежда отделно. За анализиране и обобщаване на резултатите по правило се използват статистически методи (метаанализ).

Метаанализът е комбиниране на резултатите от отделните проучвания в SR, за да се получи обща статистика и да се генерира среден резултат. Целта е да се предостави по-точна оценка на ефектите от интервенцията и да се намали несигурността. Ако дизайните на включените проучвания са твърде различни, ако измерените резултати не са достатъчно сходни или има опасения за качеството на проучванията, определящи значимост на осреднения резултат, извършване на метаанализ е неуместно, т.е. не всеки SR съдържа метаанализ.

3. ПРЕДСТАВЯНЕ НА ДОКАЗАТЕЛСТВА И ПРЕПОРЪКИ

Повечето експертни панели за клинични ръководства използват букви и цифри, за да обобщят своите препоръки. Поради твърде променливото им използване такова представяне може да е объркващо. Подходът GRADE също предлага предпочитаните символни представяния, но потребителите на GRADE ръководства трябва да виждат главно **изразяване на качеството на доказателствата и силата на препоръката (Табл. 1).**

Таблица 1

Представяне на качество на доказателствата и сила на препоръките			
Дизайн на проучване	Качество на доказателства	Символ	Буква (варианти)
Рандомизирано	Високо	⊕⊕⊕⊕	A
	Умерено	⊕⊕⊕○	B
Наблюдателно	Ниско	⊕⊕○○	C
	Много ниско	⊕○○○	D
Сила на препоръка		Символ	Цифра
Силна за интервенция		↑↑	1
Слаба за интервенция		↑?	2
Слаба против интервенция		↓?	2
Силна против интервенция		↓↓	1

4. ОБОБЩЕНИЕ НА ДОКАЗАТЕЛСТВАТА

Експертният панел, изготвящ клинични ръководства, трябва да основава своите препоръки на най-добрите налични доказателства. GRADE препоръчва SRs да формират основата за изготвяне на препоръките. Може да се използват вече съществуващи висококачествени SRs или да се провеждат собствени SRs според специфичните обстоятелства.

Крайна цел на SRs е обобщение на доказателствата, оценка на качеството за всеки резултат (outcomes) и оценка на ефекта. За разработчиците на ръководства и HTA, които предоставят съвети на политици, обобщението на доказателствата представлява крайъгълен камък по пътя към препоръка. Доказателствата, събрани от SRs, се използват за създаване на GRADE профил на доказателства (evidence profile/EP) и таблица за обобщение на констатациите (summary of findings/SoF).

4.1. Таблицы с доказателства

Те са ключов инструмент при представяне на доказателства и съответни резултати. Таблиците с доказателства са метод за представяне на качеството на наличните доказателства, преценките, които имат отношение към степента на качеството и ефектите на алтернативните стратегии за управление върху представляващите интерес резултати. Клиницистите, пациентите, обществеността, разработчиците на клинични ръководства и политиците изискват кратки и прозрачни обобщения на доказателства в подкрепа на своите решения. Следователно форматът на всяка таблица с доказателства може да бъде различен според нуждите на потребителя.

Налични са **два подхода** за таблици с доказателства, служещи за различни цели и предназначени за различни аудитории:

(i) Профил на доказателствата (evidence profile/EP);

(ii) Таблица за обобщение на констатациите (summary of findings/SoF).

4.1.2. Профил на доказателствата (EP)

Съдържа подробна информация за качеството на доказателствата за всеки от включените резултати. Предназначен е за автори на ръководства, за тези, които изготвят таблици за SoF, и за всеки, който поставя под съмнение оценката на качеството. На практика EP помага на тези, които изготвят SoF, за да гарантират, че преценките, които се правят, са систематични и прозрачни, и позволява на други да ги проверяват. **Експертният панел трябва да използва EP,** за да гарантира, че е съгласен с преценките, които са в основата на оценките на качеството.

4.1.3. Обобщение на констатациите (SoF)

Представлява таблица за всеки от включените резултати и оценка на качеството на доказателствата за всеки резултат в бърз и достъпен формат. Предназначена е за по-широка аудитория, включително крайни потребители на SRs и ръководства. Предоставя кратко резюме на ключовата информация, необходима на този, който взема решение, или в контекста на клинично ръководство предоставя резюме на ключовата информация, която е в основата на препоръка.

Стандартният формат на таблица за **SoF** включва (**Табл. 2**):

- Списък на **резултатите (outcomes)** (**Табл. 3**);
- **Предполагам риск (assumed risk)**: типични рискове от възникване на резултат без намеса, основаващи се или на рисковете от контролна група, докладвани във включените проучвания, или на епидемиологични данни от другаде. Когато е предоставен само един риск за контролна група, обикновено това е средният риск за контролната група в проучванията, които предоставят данни за този резултат (брой събития на всеки 1000 души);
- **Съответстващ риск (corresponding risk)**: риск от възникване на резултат в групата, получаваща интервенцията (брой събития на всеки 1000 души);
- **Относителен ефект (relative effect)**: за дихотомни резултати с предоставяне на коефициент на риск (risk ratio/RR), коефициент на шанс (odds ratio/OR) или коефициент на вероятност (hazard ratio/HR);
- **Брой на участници (number of participants)** и брой на проучвания (**number of studies**) и техните **дизайни (designs)**;
- Оценка на **общо качество на доказателствата (overall quality of evidence)** за всеки резултат (което може да варира според резултата);
- **Коментари (comments)** (ако е необходимо);
- **Бележки под линия или обяснения (footnotes or explanations)**, ако е необходимо, за предоставяне на обяснение относно информацията в таблицата.

Систематични прегледи за повече от едно основно сравнение (напр. ефекти от редица интервенции) изискват отделни таблици за SoF. За всяко сравнение на алтернативни стратегии всички резултати трябва да бъдат представени заедно в един PE или таблица за SoF. Тъй като повечето съществуващи SRs не адресират адекватно всички съответни резултати, GRADE може да изисква разчитане на повече от един SR.

Таблица 2

Систематичен преглед/метаанализ DOI:						
Пациенти/популация:						
Интервенция:						
Компаратор:						
Резултат	Илюстративен сравнителен риск (95% CI)		Относителен ефект (95% CI)	Брой участници (проучвания)	Качество на доказателствата	Коментари
	Предполагам риск	Съответстващ риск				

Таблица 2 (пример)

Систематичен преглед/метаанализ DOI: https://doi.org/10.1038/s41598-021-04201-y						
Пациенти/популация: Здрави безсимптомни жени						
Интервенция: Начален hrHPV тест (PCR)						
Компаратор: Начална самостоятелна цервикална цитология (ASCUS)						
Резултат	Илюстративен сравнителен риск (95% CI)		Относителен ефект (95% CI)	Брой участници (проучвания)	Качество на доказателствата	Коментари
	Предполагам риск	Съответстващ риск				
Чувствителност	95% CI: 0.216–0.497	95% CI: 0.994–1.0	RR 0.73 (95% CI: 0.59–0.92)	185 269 (27 проспективни проучвания)	⊕⊕⊕○ Умерено	
Специфичност	95% CI: 0.990–0.996	95% CI: 0.753–0.907	RR 0.98 (95% CI: 0.96–1.00)	185 269 (27 проспективни проучвания)	⊕⊕⊕○ Умерено	

Таблица 3

Термини, използвани в обобщението на констатациите (SoF)	
Термин	Обяснение
Резултати	Това са всички резултати (потенциални ползи или вреди), считани за важни при експерименталната популация и за формулиране на препоръки или вземане на решения. Може да се направи литературен преглед, за да се определи изборът на важните резултати. Трябва да се вземе също предвид важността (или стойността) на всеки резултат спрямо другите резултати, т.е. относителната важност на резултата.
95% доверителен интервал (95% CI)	Това е диапазонът около оценката, който показва колко прецизна е тя, т.е. CI е ръководство за това колко сигурни могат да бъдат експертите относно количеството, което ги интересува: колкото по-тесен е диапазонът между двете числа, толкова по-уверени могат да бъдат и обратно – колкото по-широк е диапазонът, толкова по-малко сигурни могат да бъдат. Ширината на 95% CI отразява степента, до която шансът може да е отговорен за наблюдаваната оценка (с по-широк интервал, отразяващ по-голям шанс). 95% CI означава, че експертите могат да бъдат 95% уверени, че истинският размер на ефекта е между долната и горната граница на доверителност, и обратно – има 5% шанс истинският ефект да е извън този диапазон.
Относителен ефект или RR (Risk Ratio, коефициент на риск)	Изразява се като съотношение (коефициент) на риск. Рискът е вероятността за възникване на резултат, а RR е съотношение между риска в експерименталната и в контролната група. Ако RR е точно 1.0, означава, че няма разлика между появата на резултат в експерименталната и контролната групи; ако $RR > 1.0$, интервенцията увеличава риска от резултата; ако $RR < 1.0$, интервенцията намалява риска от резултата.
Сигурност на доказателствата (certainty of the evidence) (GRADE)	Това е оценка на индикацията, която проучването дава за вероятния ефект, т.е. вероятността ефектът да бъде достоверно различен от сравнението в проучването. Под „достоверно различен“ се има предвид достатъчно голяма разлика, която може да повлияе на клиничното решение. Тази оценка се основава на цялостна оценка на причините за по-голяма или по-малка сигурност, използвайки подхода GRADE. Могат да се използват и други синонимни термини: „качество на доказателствата“, „увереност в оценката“ и „сила на доказателствата“ (виж Табл. 4).

Таблица 4

Дефиниции за степени на сигурност на доказателствата (ratings of the certainty of the evidence)	
Оценка	Дефиниции
⊕⊕⊕⊕ Висока	Проучването предоставя много добра индикация за вероятния ефект; вероятността ефектът да бъде съществено различен е ниска.
⊕⊕⊕○ Умерена	Проучването предоставя добра индикация за вероятния ефект; вероятността ефектът да бъде съществено различен е умерена.
⊕⊕○○ Ниска	Проучването предоставя известна индикация за вероятния ефект; вероятността обаче да е съществено различна (достатъчно голяма разлика, че да има ефект върху решение) е висока.
⊕○○○ Много ниска	Проучването не дава надеждна индикация за вероятния ефект; вероятността ефектът да бъде съществено различен (достатъчно голяма разлика, че да има ефект върху решение) е много висока.



Методологичен акцент

Подходът GRADE не е количествена система за изрично степенуване на качество на доказателства и на препоръки, а е интерпретация на континуум от параметри с подвижни прагове на решения

5. КАЧЕСТВО НА ДОКАЗАТЕЛСТВАТА

GRADE предоставя конкретно определение за качество на доказателствата, което е различно в контекста на изготвяне на препоръки и в контекста на обобщаване на констатации от SR. Насоките са специфични за всяка група – клинични ръководства или HTA.

За експертни панели/клинични ръководства качеството на доказателствата отразява степенята, до която **увереността на експерта в оценката на ефекта** е достатъчна, за **да подкрепи конкретна препоръка**. Експертните панели трябва да правят преценки за качеството на доказателствата в конкретния контекст, за който ги използват. Подходът GRADE включва отделно оценяване на качеството на доказателствата за всеки важен за пациента резултат, последвано от определяне на общото качество на доказателствата за всички резултати.

Въпреки че качеството на доказателствата представлява континуум, подходът GRADE изисква оценка на качеството на набора от доказателства в една от четири степени с цел простота, прозрачност и яркост (**Табл. 5**).

Таблица 5

Степени на качество на доказателства (Quality of Evidence Grades)	
Степен	Определение
Високо	Панелът е силно уверен, че истинският ефект е близък до този на оценката на ефекта.
Умерено	Панелът е умерено уверен в оценката на ефекта: истинският ефект вероятно ще бъде близък до оценката на ефекта, но има възможност да е чувствително различен.
Ниско	Увереността на панела в оценката на ефекта е ограничена: истинският ефект може да е съществено различен от оценката на ефекта.
Много ниско	Панелът има много малко доверие в оценката на ефекта: истинският ефект вероятно ще бъде значително различен от оценката на ефекта.

5.1. Фактори, определящи качеството на доказателствата

Подходът GRADE започва оценяването на качеството на доказателствата с **формата на проучванията (рандомизирано или наблюдателно)** и след това разглежда 5 причини за евентуално понижаване (**Табл. 6**) и 3 за евентуално повишаване (**Табл. 7**) на качеството.

Таблица 6

Фактори, които могат да снижат качеството на доказателствата		
№	Фактор	Следствие
1	Rиск от отклонение (risk of bias) – ограничения в дизайна или в изпълнението на проучването	↓ 1 или 2 нива
2	Хетерогенност (inconsistency) на резултатите	↓ 1 или 2 нива
3	Косвеност (indirectness) на доказателствата	↓ 1 или 2 нива
4	Неточност (imprecision) на резултатите	↓ 1 или 2 нива
5	Пристрастност на публикацията (publication bias)	↓ 1 или 2 нива

Таблица 7

Фактори, които могат да повишат качеството на доказателствата		
№	Фактор	Следствие
1	Високостепенен ефект (large magnitude)	↑ 1 или 2 нива
2	Степен доза-отговор (dose-response)	↑ 1 ниво
3	Объркващи правдоподобни фактори (plausible confounding) (снижават ефекта или предлагат фалшив ефект, ако резултатите показват неефективност)	↑ 1 ниво

Докато факторите, влияещи върху качеството на доказателствата, са добавъчни (намаляването или увеличаването на всеки отделен фактор се добавя заедно с другите фактори, за да се намали или повиши качеството за даден резултат), степенуването на качеството на доказателствата включва преценки, които не са неизменни. Всеки фактор за понижаване или повишаване отразява не отделни категории, а непрекъснатост (континуум) във всяка категория и между категориите. Когато съвкупността от доказателства е междинна по отношение на конкретен фактор, решението дали дадено изследване попада над или под прага за повишаване или понижаване на качеството зависи от преценката.

Например, ако има известна несигурност за трите фактора: отклонение, хетерогенност и неточност, но не достатъчно сериозни, за да се понижи всеки от тях, може разумно да се направи или да не се направи аргумент за понижаване. Така че във всяка категория даден експерт може да отчете полза от недоверието и да тълкува доказателствата като висококачествени, а друг експерт, който решава да намали качеството с едно ниво, би го оценил като умерено. Експертите трябва да оценяват качеството на доказателствата, вземайки предвид отделните фактори в контекста на други преценки, направени за същия резултат. В такъв случай трябва да се избере една или две категории фактори, които да се предложат като причини за понижаване, и да се обясни изборът в бележка под линия. По подобен начин трябва да се постави бележка под линия за даден фактор, отбелязвайки, че е решено да не се понижава нивото поради наличие на несигурност, но вече е понижено нивото за другия фактор, така че по-нататъшното снижение на качеството за този резултат би изглеждало неподходящо. Тоест GRADE силно насърчава авторите на ръководства да бъдат ясни и прозрачни, когато попаднат в тези ситуации, вземайки гранични решения. Въпреки ограниченията при разделянето на континуумите на категории, третирането на всеки критерий за оценка на качеството нагоре или

надолу като отделни категории подобрява прозрачността. Накратко, голямата заслуга на GRADE не е, че осигурява възпроизводими преценки, а че **изисква изрична преценка, която е прозрачна за потребителите**.



Методологичен акцент

Разглеждането на фактори, снижаващи нивото на качеството на доказателствата, трябва да предхожда анализа на причините за повишаването им и задължително интерпретира възможността за 6 ограничения: (1) форма на проучванията, включени в систематичния преглед/метаанализ, (2) риск от отклонения в дизайна и изпълнението на проучванията, (3) хетерогенност на резултатите, (4) косвеност на доказателствата, (5) неточност на резултатите и (6) пристрастност на публикациите

5.1.1. Форми на проучванията

За препоръки относно клинично поведение (за разлика от прогноза или точност на диагностични тестове) RCTs предоставят като цяло много по-силни доказателства от OSs, а строгите OSs предоставят по-силни доказателства от неконтролираните серии от случаи. В подхода GRADE началната оценка на качеството на доказателствата използва следната рамка:

- **рандомизираните проучвания** без съществени отклонения предоставят **високо качество**;
- **наблюдателните проучвания** без специални силни страни или важни отклонения предоставят доказателства с **ниско качество**
- **нерандомизираните експериментални проучвания** (квази-RCT) без съществени отклонения също предоставят висококачествени доказателства, но автоматично трябва да бъдат понижени поради ограничения в дизайна (риск от отклонения), като липса на прикриване на разпределението и връзка с доставчика (напр. номер на диаграма);
- **сериите от случаи и докладите за случаи** (case reports) са наблюдателни проучвания, които изследват само пациенти, изложени на интервенция; източникът на резултати от контролна група е подразбиращ се или неясен, поради което те обикновено изискват понижаване от ниско до много ниско качество;
- **експертното мнение** не е категория за качество на доказателствата, а тълкуване на доказателства в контекста на личен опит и знания; експертите могат да имат мнение за доказателства, които се основават на тълкуване на проучвания, вариращи от неконтролирани серии от случаи (напр. наблюдения в собст-

вената практика) до RCT и SRs – важно е да се опише какъв тип доказателства (независимо публикувани или непубликувани) се използват като основа за тълкуване.

6. ФАКТОРИ, КОИТО МОГАТ ДА СНИЖАТ КАЧЕСТВОТО НА ДОКАЗАТЕЛСТВАТА

6.1. Ограничения на проучването (risk от отклонение)

Ограниченията в дизайна и изпълнението на проучването могат да променят оценките на ефекта от лечение. Увереността на експертите в оценката на ефекта намалява, ако проучванията страдат от големи ограничения, които могат да доведат до пристрастна оценка на ефекта от интервенцията. Отклоненията, които вероятно биха довели до пристрастен резултат при рандомизирани и нерандомизирани проучвания, са представени съответно на **Табл. 8** и **Табл. 9**.

Таблица 8

Отклонения в рандомизирани контролирани проучвания	
Отклонение	Обяснение
Липса на прикриване на разпределението (allocation concealment)	Включваните пациенти са наясно с групата (или периода в кръстосано изпитване), към която ще бъде включен следващият пациент (основен проблем при „псевдо“ или „квази“ рандомизирани проучвания с разпределение по ген от седмица, дата на раждане, номер на диаграма и т.н.).
Липса на заслепяване (blinding)	Пациентът, лекарят, техническите лица или анализаторите на данни са наясно с рамото, в което са включени пациентите (или с лекарството, което в момента се проучва в кръстосано изпитване).
Непълно отчитане на пациенти и крайни цели (събития) (incomplete accounting)	Загуба на проследяване и неспазване на принципа за намерение за лечение при проучвания за превъзходство; или загуба на проследяване и неуспех при провеждане на двата анализа при проучвания за нонинфериорити, като се имат предвид само тези пациенти, които са се придържали към лечението, и всички, за които има налични данни за резултати. Колкото по-висока е пропорцията, загубена за проследяване, по отношение на честота на събития на експерименталната и контролната група, и разликата между тях, толкова по-голяма е заплахата от отклонение.

Отклонения в рандомизирани контролирани проучвания	
Отклонение	Обяснение
Отчитане на селективни резултати (selective outcome reportin)	Непълно или липсващо отчитане на някои ефекти, а не на други (поради резултатите).
Други ограничения (other limitations)	- По-ранно прекратяване на проучване за полза; значително надценяване е вероятно при проучвания с по-малко от 500 събития и голямо надценяване е вероятно при проучвания с по-малко от 200 събития. Емпиричните данни показват, че официалните правила за прекратяване не намаляват това отклонение. - Използване на невалидирани измервания на резултатите (напр. резултати, докладвани от пациенти). - Ефекти на пренасяне при кръстосано проучване. - Отклонения при набиране в клъстерни рандомизирани проучвания.

Таблица 9

Отклонения в наблюдателни проучвания	
Отклонение	Обяснение
Липса на разработване и прилагане на подходящи включващи критерии (включване на контролна популация) (eligibility criteria)	- В проучвания случай-контрола – недостатъчно или прекомерно съвпадение. - В кохортни проучвания от различни популации – избор на третирани и нетретирани.
Погрешно измерване на експозиция и на резултат (exposure and outcome)	- Разлики в измерване на експозицията (напр. повтарящо се отклонение в проучвания случай-контрола). - Диференцирано наблюдение за изход при третирани и нетретирани в кохортни проучвания.
Неспособност за адекватен контрол на объркващи променливи (control confounding)	- Неуспешно точно измерване на всички известни прогностични фактори. - Несъвпадение на прогностични фактори и/или корекция в статистическия анализ.
Непълно или неадекватно кратко проследяване (short follow-up)	И двете групи трябва да бъдат проследявани за еднакъв период от време, особено в рамките на проспективни кохортни проучвания.

Преминаването от критерии за риск от отклонение за всяко отделно проучване **към преценка за риск от отклонение в група проучвания**, насочени към конкретен резултат, представлява предизвикателство и подходът GRADE предлага 5 принципа.

1. При вземане на решение за цялостно качество на доказателства не се прави усредняване на проучванията (например, ако някои изследвания нямат сериозни ограничения, докато други имат сериозни ограничения, а трети – много сериозни ограничения, не се оценява автоматично качеството с едно ниво надолу поради средна оценка на сериозни ограничения). По-скоро е оправдано разумно разглеждане на приноса на всяко изследване с обща нагласа за фокусиране върху висококачествените изследвания.
2. Разумното разглеждане изисква оценка на степента, в която всяко проучване допринася за оценка на степента на ефекта и обикновено отразява размера на извадката и броя на крайните цели – по-големи проучвания с много събития допринасят повече, а много по-големи проучвания с много повече събития допринесат много повече.
3. Експертите трябва да бъдат консервативни в преценката за понижаване на нивото, т.е. трябва да бъдат уверени, че е налице значим риск от отклонение в по-голямата част от наличните доказателства.
4. Рискът от отклонение (risk of bias) трябва да се разглежда в контекста на други ограничения; например, ако експертите се окажат в сходна ситуация за два проблема с качеството (да речем, риск от отклонение и неточност), трябва да снижат оценката за поне един от двата.
5. В оценките си експертите могат да се сблъскват лице в лице и трябва честно да признават, че наистина са в такава ситуация, да изясняват защо смятат, че това е така, и да направят очевидни мотивите за крайната си преценка.

Експерти, работещи в контекста на **систематични прегледи на Cochrane (Cochrane Reviews)**, могат да използват техните насоки за оценяване на ограниченията на изследването (risk of bias) и да ги включват директно в своето SoF: (i) „нисък риск“ означава „без отклонение“; (ii) „неясен риск“ означава или „без отклонение“, или „сериозно отклонение“; (iii) „висок риск“ означава или „сериозно отклонение“, или „много сериозно отклонение“ (**Табл. 10**).

Таблица 10

Оценка на отклонения (risk of bias) в Cochrane систематични прегледи и съответна GRADE оценка на качеството на доказателствата				
Risk of bias	От проучванията	Интерпретация	Съображения	GRADE оценка на отклонения
Нисък (low)	Повечето информация е от проучвания с нисък риск от отклонение	Възможното отклонение е малко вероятно да промени сериозно резултатите	Без видими ограничения	Няма сериозни отклонения, без понижаване на нивото
Неясен (unclear)	Повечето информация е от проучвания с нисък или неясен риск от отклонение	Възможно отклонение, което поражда известно съмнение относно резултатите	Малко вероятно е потенциалните отклонения да намалят доверието в оценката на ефекта	Няма сериозни отклонения, без понижаване на нивото
			Потенциалните отклонения вероятно ще намалят доверието в оценката на ефекта	Сериозни отклонения, понижаване с едно ниво
Висок (high)	Делът на информацията от проучвания с висок риск от отклонение е достатъчен, за да повлияе на тълкуването на резултатите	Възможно отклонение, което сериозно отслабва доверието в резултатите	Решаващо отклонение за един критерий или някои ограничения за множество критерии, достатъчни да намалят доверието в оценката на ефекта	Сериозни отклонения, понижаване с едно ниво
			Решаващо отклонение за един или повече критерии, достатъчно да намали значително доверието в оценката на ефекта	Много сериозни отклонения, понижаване с две нива

6.2. Хетерогенност на резултатите

Експертите трябва да проучват обясненията за хетерогенност и ако не могат да идентифицират правдоподобно обяснение, качеството на доказателствата трябва да бъде понижено с едно или две нива (според величината на несъответствие в резултатите). Когато се говори за хетерогенност в измерването на ефекта, **имат се предвид относителни параметри**: risk ratios (RR), hazard ratios (HR) или odds ratios (OR).

Когато характеристиките на пациента позволяват уверено класифициране в субпопулации със значително различен риск, абсолютните разлики в резултатите между експерименталните и контролните групи биха се различавали достоверно и **оправдават различия в препоръките за отделни субпопулации**, вместо да се понижават нивата на качеството на доказателствата за размера на ефекта.

Критериите за понижаване на нивата на качеството поради хетерогенност могат да се прилагат, когато резултатите са от повече от едно проучване, и включват 6 критерии.

1. Голяма вариация на точковите оценки в проучванията (посоката на ефекта не е критерий за несъответствие).
2. Минимално или никакво припокриване на 95% CI, което предполага, че вариацията е повече от очакваната.
3. Статистическите критерии, включително тестове за хетерогенност, които тестват нулева хипотеза (т.е., че всички изследвания имат една и съща изходна степен на ефект), имат **ниска p -стойност ($p < 0.05$)**, което показва отхвърляне на нулевата хипотеза.
4. Висока стойност от I^2 **статистиката**, която количествено определя дела на вариациите в точковите оценки, дължащи се на разликите между изследванията. Определянето на „висока стойност“ на I^2 е субективно, но могат да се използват следните практически правила:
 - < 40% може да е ниска;
 - 30-60% може да е умерена;
 - 50-90% може да е значителна;
 - 75-100% може да е много висока.

Припокриването в тези диапазони и използването на изречение „може да е“ илюстрират несигурността, свързана

с такива преценки. Важно е да се отбележат имплицитните ограничения в тази статистика: когато размерът на отделни извадки е малък, точковите оценки може да варират значително, но тъй като вариацията може да се обясни със случайност, I^2 може да е нисък; обратно, когато извадката е голяма, относително малка разлика в точковите оценки може да доведе до висок I^2 .

5. Статистика τ^2 (**may kwagram**) е мярка за вариабилност с предимство пред другите, тъй като не зависи от размера на извадката.
6. Инструмент за оценка на достоверността на анализа за промяна на ефекта (Instrument to assess the credibility of effect modification/**ICEMAN**), който бе създаден допълнително през 2023 г. и отговаря на въпроса дали съществува истински подгрупов ефект, а не дали неговата величина е важна, т.е. разглежда дали съществува промяна на ефекта, а не нейната значимост.

Всички статистически подходи имат ограничения и техните резултати трябва да се разглеждат в контекста на субективно изследване на вариабилността в точковите оценки и припокриването им в 95% CI.

При **интерпретиране на метаанализи** се препоръчва да се включват формални тестове за това дали априорните хипотези обясняват несъответствието между важни подгрупи, дори ако съществуващата вариабилност изглежда случайна (напр. високи p -стойности в тестове за хетерогенност и ниски стойности на I^2). Ако размерът на ефекта се различава в различните проучвания, обясненията за хетерогенността може да се дължат на 4 разлики.

- популации (напр. лекарства могат да имат по-големи относителни ефекти при по-тежко болни);
- интервенции (напр. по-големи ефекти с по-високи дози лекарства);
- резултати (напр. продължителност на проследяване);
- методи на изследване (напр. RCTs с по-висок и по-нисък риск от отклонения).

Ако методите на изследване предоставят убедително обяснение за разликите в резултатите между проучванията, авторите трябва да обмислят фокусиране върху оценки на ефекта от проучвания с по-нисък риск от отклонения. Ако голямата вариабилност в степента на ефекта остане необяснена и авторите не успеят да я свържат с разлики в една от тези четири променливи, тогава качеството на доказателствата намалява.

При **използване на оценки от подгрупов анализ** за предпочитане е намиране на обяснение за хетерогенността, която може да се основава на разлики в популацията, интервенцията или резултатите, т.е. може да се наложат две или повече оценки на ефекта, евентуално с отделни препоръки. Повечето предполагаеми подгрупови ефекти в

крайна сметка се оказват фалшиви, така че е оправдана предупредителна бележка под линията относно анализите на подгрупите и тяхното представяне. Тоест разработчиците на клинични ръководства трябва да упражняват **висока степен на скептицизъм** по отношение на потенциални подгрупови ефекти, като обръщат специално внимание на 7 критерия.

1. Дали променливата на подгрупата е характеристика, определена в началото или след рандомизация (хипотезите за подгрупи трябва да бъдат разработени *a priori*).
2. Дали разликата в подгрупите се предполага от сравнения в рамките на проучването, или между тях.
3. Дали статистическият анализ предполага, че случайността е малко вероятно обяснение за разликата в подгрупите.
4. Дали хипотезата предхожда, а не следва анализа и дали включва хипотетична насока, която впоследствие е потвърдена.
5. Дали хипотезата е тествана за подгрупа с по-малък брой.
6. Дали разликата в подгрупите е консистентна в проучванията и във важните резултати.
7. Дали външни доказателства (биологична или социологическа обосновка) подкрепят хипотетичната подгрупова разлика.

Достоверността на подгруповите ефекти не е въпрос на „да“ или „не“, а на непрекъснатост (континуум), т.е. необходима е преценка, за да се определи колко убедителен е подгруповият анализ въз основа на горните критерии.

■ 6.3. Индиректност (косвеност) на доказателствата

Преките доказателства се състоят от изследвания, които директно сравняват параметри от интерес за експерта: интервенции в популациите и измервания на важните за пациентите резултати. Авторите на клинични ръководства трябва да обмислят степента на индиректност, в която не са сигурни за приложимостта на доказателствата към техния PICO въпрос, и да понижат оценката на качеството с едно или дори две нива. Експертите трябва да отговорят на PICO въпроса, който са задали, и така биха оценили директността на откритите от тях доказателства. Съображенията, направени от авторите на SRs, може да са различни от тези на панелите за ръководства,

които използват SRs. Колкото по-ясно и изрично е формулиран PICO въпросът, толкова по-лесно е за потребители-те да разберат преценките на авторите на SRs.

Съществуват 4 източника на индиректност.

- 1. Разлики в популацията (приложимост)** е често срещан проблем за авторите на SRs и за авторите на ръководства и тогава доказателствата са косвени. Тоест ефектът върху общото качество на доказателствата би могъл да варира според това колко различни са изследваните популации, тъй като качеството на резултата може да не намалее, да намалее с едно ниво или в крайни случаи – с две нива.
- 2. Разлики в интервенциите (приложимост).** Авторите на SRs трябва да са гарантирали, че са включили само проучвания с пряко релевантни интервенции. Когато са включени интервенции, косвено свързани с изследването, качеството на доказателствата трябва да бъде намалено. В някои случаи използваната интервенция може да бъде същата, но да се извършва по различен начин според обстановката.
- 3. Разлики в измерванията на резултати (сурогатни резултати).** Наличните проучвания може да измерват въздействието на интересуваша ни интервенция върху резултати, свързани със, но различни от тези с първостепенно значение за пациентите. Разликата между желани и измерени резултати може да се отнася и до времевата рамка (напр. резултат, измерен на 3-ти месец срещу 12-ти месец). Друг източник на индиректност е използване на сурогатни крайни цели вместо важни за пациента резултати.
- 4. Косвени сравнения** възникват, когато не е налично сравнение на интервенция А срещу В, а А е сравнена с С, а В е сравнена със С, т.е. това доказателство е с по-ниско качество, отколкото биха предоставили непосредствените сравнения на А и В. Валидността на непрякото сравнение се основава на предположението, че факторите в дизайна на проучването и качеството на методологията не са достатъчно различни, за да доведат до различни ефекти, т.е. всички очевидни разлики се обясняват с истинските разлики в ефекта. Някои автори наричат това „предположение за подобие“, но тъй като то винаги е под съмнение, косвените сравнения винаги изискват оценка на качеството с едно ниво надолу. Дали да се снижи с две нива надолу зависи от правдоподобността, че алтернативните фактори (популация, интервенции, съвместни интервенции, резултати и методи на изследване) обясняват или прикриват разликите в ефекта.

6.4. Неточност на резултатите

Като цяло, резултатите са неточни, когато проучванията включват относително малко пациенти и малко събития и по този начин имат широк 95% CI около оценката на ефекта. В допълнение към използването на 95% CI подходът GRADE въвежда **оптимален размер на информация (OIS)** като втори необходим критерий за определяне на адекватна точност. Тъй като GRADE определя по различен начин качеството на доказателствата за SRs и за ръководства, критериите за понижаване на оценката за неточност се различават по това, че панелите за ръководства трябва да вземат предвид и други резултати в контекста на препоръката, докато преценките за конкретни резултати в SR са без този контекст.

В панелите за клинични ръководства всички резултати се разглеждат заедно, като се обръща внимание дали са критични, или са важни, но не критични. Решението да се намали качеството на доказателствата за неточност зависи от прага, който представлява **основа за решение за клинично поведение** и за разглеждане на компромиса между желани и нежелани последици. Определянето на приемлив праг неизбежно включва преценка, която трябва да бъде изрично изразена.

За дихотомни резултати разработчиците на ръководства трябва да вземат предвид контекста на конкретната препоръка, за да определят дали резултатите от дихотомния (двоичен) резултат са достатъчно точни, за да я подкрепят. Проучването на долна и горна граница на 95% CI във връзка с прага, зададен от панела, последвано от определяне дали критериите за оптимален размер на информацията са изпълнени, биха помогнали за решение дали да се намали оценката за неточност. С тази цел разработчиците на ръководства трябва да обмислят 4 стъпки.

1. Първо да се обсъди дали границите на 95% CI са от една и съща страна на техния праг за вземане на решения и дали 95% CI преминава прага на клинично решение между препоръчване и не препоръчване на лечение. Ако отговорът е „да“ (т.е. 95% CI преминава прага), оценката трябва да се намали поради неточност, независимо къде се намират точковата оценка и 95% CI.
2. Ако прагът не е преминал, трябва да се обсъди дали са изпълнени критериите за оптимален размер на информацията (виж OIS),

или

3. процентът на събитията е много нисък и размерът на извадката е много голям (поне 2000, а може би 4000 пациенти).
4. Ако нито един от критериите не е изпълнен, оценката трябва да се намали поради неточност.

Въпреки че 95% CI улавя най-вече степенята на неточност, той може да бъде подвеждащ при определени обстоятелства, т.е. може да изглежда стабилен, но малък брой събития могат да направят резултатите крехки. По правило 95% CI предполага, че всички пациенти са изложени на еднакъв риск (т.е. има прогностичен баланс) – предположение, което е невярно. Рандомизирането облекчава проблема чрез балансиране на прогностичните фактори между експерименталните и контролните групи, но увереност, че е постигнат прогностичен баланс, съществува само ако размерите на извадката са големи. Големите ефекти при малки размери на извадката, дори в RCTs, може да се дължат на прогностичен дисбаланс и изискват повишено внимание. Систематичният преглед на тези проучвания впоследствие също би генерирал надценен ефект от лечението. Съществуват примери за метаанализи, генериращи очевидни ползи или вредни ефекти, опровергавани от последващи по-големи проучвания. Подобно обстоятелство възниква, когато проучванията са прекратени рано поради полза (т.е. преди достигане на общ брой събития или необходим размер на извадка, както е изчислено за проучване с подходяща мощност). Когато ефектът от лечението е надценен, 95% CI може фалшиво да изглежда подходящ за покриване на критерия за праг на клинично решение, т.е. необходим е критерият за OIS.

По правило OIS се прилага, ако общият брой пациенти, включени в SR, е по-малък от броя на пациентите, генериран чрез конвенционално изчисление на размера на извадката за адекватно набирано проучване – тогава трябва да се обърне внимание на оценката за неточност. Налични са много онлайн калкулатори за изчисляване на размера на извадката: <http://www.stat.ubc.ca/rollin/stats/ssize/b2.html>. Изключение за необходимост от OIS са ниски проценти на събития с голям размер на извадката, т.е. преценката за точност може да се основава на 95% CI за абсолютния ефект и не трябва да се понижава качеството на доказателствата за неточност. Подходът GRADE предлага за изчисляване на OIS да се използва редукция на относителния риск (RRR) от 20% до 30%. Изборът на RRR е въпрос на преценка и може да има случаи, в които убедителна предварителна информация би предложила избор на по-малка или по-голяма стойност на RRR за изчисляване на OIS.

За непрекъснати (континуум) резултати съображенията за понижаване на качеството поради неточност следват същата логика като за дихотомни променливи. Процесът започва с понижаване на качеството за неточност, ако долната спрямо горната граница на 95% CI включва истинския основен ефект. Ако 95% CI не преминава този праг, но доказателствата не отговарят на критерия OIS, авторите на ръководства трябва да обмислят понижаване на качеството на доказателствата за неточност; в този случай преценката на OIS критерия ще изисква изчисление на размера на извадката за непрекъснатата променлива. В контекста на ръководството прагът за вземане на решение за приемлива оценка изисква разглеждане на пълния контекст на препоръката, включително други резултати – всички потенциални ползи и важни нежелани ефекти.

Когато са налице много малко събития и малко 95% CIs за относителни и абсолютни оценки на ефект, включващи както осезаема полза, така и забележима вреда, разработчиците на ръководства трябва да обмислят понижаване на качеството на доказателствата с две нива.

6.5. Пристрастност на публикацията (publication bias)

Дефинира се като систематично подценяване или надценяване на основния полезен или вреден ефект поради селективно публикуване на проучвания. Доверието в комбинираните оценки на ефектите от SR може да бъде намалено, когато се подозира пристрастност на публикацията, дори когато самите включени проучвания имат нисък риск от отклонения (risk of bias). За разлика от това, когато цялото проучване остане непубликувано (недокладвано), експертът може да прецени вероятността от пристрастност на публикацията само като разгледа група от проучвания. Понастоящем GRADE следва подхода на Cochrane Collaboration и разглежда селективното докладване като проблем с риск от отклонение в отделни проучвания (**Табл. 11**).

Емпиричните доказателства сочат, че проучванията, отчитащи статистически значими констатации, е по-вероятно да бъдат приети за публикуване, отколкото тези, отчитащи статистически незначими констатации („негативни проучвания“). Пристрастност на публикацията възниква, когато цели проучвания остават недоокладвани или непубликувани, или неясно публикувани (например в списания с ограничен тираж, които не са индексирани от основните бази данни, като резюмета на конференции или тезиси) и авторите на SRs може да не успеят да ги идентифицират, дори ако използват най-строги техники за търсене. Рискът от пристрастност на публикацията може да бъде по-висок при SRs на OSs, отколкото при прегледи на RCTs, особено ако OSs се провеждат автоматично от

регистри или медицински досиета. Трябва да се подозира пристрастност на публикацията, когато публикуваните доказателства са ограничени до малък брой малки проучвания, особено ако повечето от тях показват ползи от определена интервенция.

Таблица 11

Възможни източници на пристрастност на публикацията	
Фази на публикуване	Действия, допринасящи/водещи до пристрастност
Предварителни и пилотни проучвания	Малки проучвания, които е по-вероятно да бъдат „негативни“ (напр. тези с отхвърлени или неуспешни хипотези), остават непубликувани; някои ги класифицират като частна информация.
Подаване на статията	Авторите решават, че докладването на „негативно“ проучване е безинтересно, и не инвестират време и усилия, необходими за подаване.
Избор на списание	Авторите решават да изпратят „негативния“ доклад в неиндексирано, неанглоезично списание или списание с ограничен тираж.
Редакционно съображение	Редакторът решава, че „негативното“ проучване не изисква партньорска рецензия и отхвърля ръкописа.
Партньорски рецензии (peer reviewer)	Партньорските рецензенти заключават, че „негативното“ проучване не допринася за темата и препоръчват отхвърляне на ръкописа. Авторът се отказва или преминава към списание с по-нисък импакт фактор; публикуването се забавя.
Авторска редакция и повторно изпращане	Авторът на отхвърления ръкопис решава да се откаже от подаване на „негативното“ изследване или да го изпрати отново по-късно в друго списание (виж по-горе „Избор на списание“).
Публикация на проучването	Списанието забавя публикуването на „негативното“ проучване. Личните интереси водят до изпращане на ръкописи и приемане от няколко различни списания.

За откриване на пристрастност на публикацията в SRs се използват 3 метода.

- Визуална проверка и тестове за асиметрия на **фуниевидните графики (funnel plots)**. Практически трябва да се използват фуниевидна диаграма и статистически тестове за асиметрия за откриване на пристрастност на публикацията, ако има поне 10 проучвания, включени в метаанализа (някои изискват поне 5 проучвания).
- **Метод „погрязване и запълване“ (trim and fill)**. Тази техника е разширение на фуниевидните графики и започва с премахване на малки „позитивни“ изследвания, които нямат негативен аналог, оставяйки симетрична фуния. Следваща стъпка е добавяне на хипотетични проучвания, отразяващи резултати от позитивните проучвания, но все пак запазващи новата оценка на обединения ефект. За укрепване на заключенията се препоръчва използване на множество тестове.
- **Рекурсивен кумулативен метаанализ**. Използва се за откриване на отклонение във времето на забавяне (lag time bias), и се извършва в края на всяка година, като отбелязва промените в оценките на ефекта. Ако ефектите от интервенцията непрекъснато намаляват, има силна индикация за отклонение във времето на забавяне.

Независимо от използвания тест, разработчиците на ръководства трябва да са наясно, че те могат да бъдат склонни към грешки и резултатите им трябва да се тълкуват предпазливо. За експертите е изключително трудно да бъдат уверени, че липсва пристрастност на публикацията и почти е толкова трудно да се постави праг кога да се понижи качеството на доказателствата. Затова GRADE предлага за пристрастност на публикацията да се понижава качеството на доказателствата **с максимум едно ниво**.

7. ФАКТОРИ, КОИТО МОГАТ ДА ПОВИШАТ КАЧЕСТВОТО НА ДОКАЗАТЕЛСТВОТА

Решение за повишаване на нивото трябва да се взема рядко. Например при големи или много големи ефекти трябва да се вземат предвид не само точковата оценка, но и ширината на нейния 95% CI: рядко трябва да се повишава нивото за големи ефекти, ако 95% CI се припокрива значително с ефекти, по-малки от избрания праг. Чрез рамката GRADE съвкупността от доказателства от OSs първоначално се класифицира като ниско качество (т.е. позволяващо ниска степен на доверие в очаквания ефект). Обаче има ситуации, когато експертите имат голяма увереност в оценката на ефекта на OSs (включително кохортни, случай-контрола, проучвания преди-след, времеви серии и т.н.) и нерандомизирани експериментални проучвания (напр. квазирандомизирани и нерандомизирани контролирани проучвания). Въпреки че теоретично е възможно да се повишат резултатите от RCTs, все още не са намерени убедителни примери за такива случаи.



Методологичен акцент

Разглеждането на фактори, повишаващи нивото на качеството на доказателствата, се налага при систематичен преглед на наблюдателни проучвания (OSs) и задължително включва интерпретиране на възможност за 3 ситуации: (1) голяма величина на ефекта, (2) степен на доза-отговор и (3) остатъчно правдоподобно обръкване

7.1. Голяма величина на ефекта

Когато съвкупността от доказателства от OSs, които не са понижени за нито един от 5-те фактора, дават големи или много големи оценки на мащаба на ефекта от интервенцията, тогава експертите могат да са по-уверени в резултатите. В тези ситуации, въпреки че има вероятност OSs да осигурят надценяване на истинския ефект, дизайнът на проучването, който е по-склонен към отклонения, е малко вероятно да обясни цялата очевидна полза (или вреда). Решението за повишаване на качеството трябва да отчита не само точковата оценка, но и прецизността (ширина на 95% CI) за този ефект: рядко и много предпазливо трябва да се повишава качеството на доказателствата, ако 95% CI значително се припокрива с ефекти, по-малки от избрания праг на клинично значение (**Табл. 12**).

Таблица 12

Дефиниции за голям и много голям ефект		
Степен на ефекта	Дефиниция	Качество на доказателствата
Голям	RR* > 2 или < 0.5 (въз основа на преки доказателства, без обръквачи променливи)	може да се увеличи с 1 ниво
Много голям	RR* > 5 или < 0.2 (въз основа на преки доказателства без сериозни проблеми за риск от отклонение или неточност, т.е. с достатъчно тесен 95% CI)	може да се увеличи с 2 нива
* <i>Забележка:</i> тези правила се прилагат, когато мярката за ефект се изразява като коефициент на риск (RR) или коефициент на вероятност (HR). Те не винаги могат да бъдат приложени, когато мярката за ефекта се изразява като коефициент на шанс (OR). Предлага се OR да се преобразува в RR и едва след това да се оценява степенята на ефекта.		

По-вероятно е експертът да повиши качеството на доказателствата поради голяма или много голяма степен на ефект, когато: (i) ефектът е бърз; (ii) ефектът е консистентен при различни субекти; (iii) обърната е предишната траектория на болестта; (iv) голямата величина на ефекта се подкрепя от косвени доказателства. Ако резултатите са субективни, експертите трябва да бъдат предпазливи, когато обмислят надграждане заради големи ефекти. Това е особено вярно, когато екипите, оценяващи резултатите, са наясно към коя група принадлежат изследваните субекти (т.е. не са били заслепени).

■ 7.2. Степен гоза-отговор

Наличието на тази категория е признато като важен критерий за увереност в предполагаема причинно-следствена връзка. Затова тя може да повиши увереността на експертите в констатации от OSs и по този начин да повиши качеството на доказателствата.

■ 7.3. Ефект на остатъчно правдоподобно объркване (plausible confounding)

Понякога вероятните остатъчни объркващи фактори от OSs може да работят за намаляване на демонстрирания ефект или за увеличаване на ефекта, ако такъв не е наблюдаван. Причината, поради която в повечето случаи се счита, че OSs предоставят само доказателства с ниско качество, е, че неизмерените или неизвестни детерминанти на резултата, които не са отчетени в коригирания анализ, вероятно се разпределят неравномерно между експерименталните и контролните групи, произвеждайки т.нар. „остатъчно объркване“ или „остатъчни отклонения“, и биха довели до подценяване на очевидния лечебен ефект. Например, ако само по-тежко болни пациенти получават експериментална интервенция или лекарствена експозиция, но въпреки това се справят по-добре, вероятно е действителният ефект да е по-голям, отколкото предполагат данните. Съществува паралелна ситуация, когато OSs не успяват да демонстрират връзка.

■ 7.4. Общо качество на доказателствата

Това е комбинирана оценка на качеството на доказателствата за всички резултати, считани за критични за отговор на PICO въпроса (т.е. вземане на решение или препоръка). GRADE предупреждава срещу механистичен подход

при прилагане на критериите за оценка на качеството на доказателствата „нагоре“ или „надолу“. Въпреки че GRADE предлага **първоначално отделно разглеждане на 5 категории причини за понижаване и 3 категории за повишаване на оценката с решение „да/не“**, окончателната оценка на общото качество на доказателствата се получава в непрекъсната степен на доверие в оценките на ефектите. Тъй като SRs не правят (или поне не трябва да правят) препоръки, авторите им не трябва да оценяват общото качество на доказателствата според резултатите, а да оценяват качеството на доказателствата само за всеки резултат поотделно. Експертните панели за клинични ръководства трябва да определят общото качество на доказателствата за всички критични резултати, които са от съществено значение за препоръката, която правят, т.е. за всяка препоръка трябва да предоставят единична степен на качество на доказателствата, но силата на препоръката обикновено зависи от доказателствата не само за един, а за редица важни резултати и от качеството на доказателствата за всеки от тях. Затова при определяне на общото качество трябва да се спазват 3 правила.

1. Да се вземат под внимание само тези резултати, които се считат за критични.
2. Ако качеството на доказателствата е еднакво за всички критични резултати, тогава то става общо качество на доказателствата, подкрепящи отговора на PICO въпроса.
3. Ако качеството на доказателствата се различава при различните критични резултати, логично е общата достоверност на оценките да не е по-висока от най-ниската за всеки резултат, който е критичен за решение. Следователно най-ниското качество на доказателствата за който и да е от критичните резултати определя общото качество на доказателствата.

8. ПРЕМИНАВАНЕ ОТ ДОКАЗАТЕЛСТВА КЪМ ПРЕПОРЪКИ

8.1. Препоръки и категоризиране по сила

Силата на дадена препоръка отразява степента, в която експертният панел е уверен, че желаните ефекти от интервенцията надвишават нежеланите, или обратното, в диапазона от пациенти, за които е предназначена препоръката.

GRADE определя **две категории за сила** на дадена препоръка и предлага използване на термините „**силна**“ и „**слаба**“, но експертите могат да изберат различна формулировка за тяхното характеризиране. В специални случаи може да се препоръча дадена интервенция да се използва само в клинични проучвания (докато се генерират повече данни, позволяващи по-изчерпателна препоръка) или изобщо да не се прави препоръка. Повечето организации, изготвящи ръководства, приемат, че предимствата на изричната степен на препоръка надвишават недостатъците, защото предоставя ясна насока на пациенти, клиницисти и политици.

За да могат разработчиците на ръководства да предложат силна препоръка, трябва да са сигурни относно различните фактори, които ѝ влияят, и да разполагат със съответна информация, която поддържа ясен баланс между желаните ефекти от интервенцията (за да я препоръча) или нежеланите ефекти (за да препоръча отказ от нея). Когато панелът не е сигурен дали балансът е ясен, или когато информацията за различните фактори, влияещи върху силата на дадена препоръка, не е налична, в повечето случаи експертите би трябвало да направят слаба препоръка (**Табл. 13**).

Прецизирането на вземане на клинични решения при слаби препоръки остава предизвикателство и клиницистите винаги трябва да вземат предвид предпочитанията и ценностите на пациентите. Дори силни препоръки, базирани на висококачествени доказателства, не могат да се прилагат за всички обстоятелства и всички пациенти. Затова клиницисти, пациенти, платци, институционални комисии, други заинтересовани страни или съдилища никога не трябва да гледат на препоръките като на диктат. В този смисъл никоя оценка на действията на лекаря не трябва да цитира препоръките наизуст или въобще.

Таблица 13

Фактори, определящи силата и посоката на препоръките	
Домейни	Коментар
Баланс между желани и нежелани резултати (компромиси), като се вземат предвид: - най-добри оценки на степента на ефектите върху желаните и нежеланите резултати - важност на резултатите (приблизителни типични стойности и предпочитания)	Колкото по-големи са разликите между желани и нежелани ефекти, толкова по-вероятно е да бъде оправдана силна препоръка. Колкото по-малка е нетната полза и колкото по-ниска е сигурността за нея, толкова по-вероятно е да бъде оправдана слаба препоръка.

Фактори, определящи силата и посоката на препоръките	
Домейни	Коментар
Доверие в мащаба на оценките на ефекта от интервенциите върху важните резултати (общо качество на доказателствата за резултатите)	Колкото по-високо е качеството на доказателствата, толкова по-вероятно е да бъде оправдана силна препоръка.
Увереност в ценностите и предпочитанията и в тяхната променливост	Колкото по-голяма е променливостта в ценностите и предпочитанията или в несигурността на типичните ценности и предпочитания, толкова по-вероятно е да бъде оправдана слаба препоръка.
Използване на ресурси	Колкото по-високи са разходите за интервенция (повече изразходвани ресурси), толкова по-малко е вероятно да бъде оправдана силна препоръка.



Методологичен акцент

Категоризирането на препоръките по сила трябва да отразява непрекъснатостта (континуума) на качеството на доказателства, затова експертите трябва предпазливо да преценяват баланса между категориите „силна“ и „слаба“ единствено в контекста на ценностите за пациента и яснотата за потребителите на ръководства

8.2. Силна препоръка

Това е препоръка, за която експертният панел за клинично ръководство е уверен, че желаните ефекти от дадена интервенция надвишават нейните нежелани ефекти (силна препоръка за интервенция) или че нежеланите ефекти от дадена интервенция надвишават желаните ефекти (силна препоръка срещу интервенция). Силните препоръки не са непременно препоръки с висок приоритет, а предполагат, че повечето или всички индивиди ще бъдат най-добре обслужени от препоръчаното клинично поведение.

■ 8.3. Слаба препоръка

Това е препоръка, при която желаните ефекти вероятно надвишават нежеланите ефекти (слаба препоръка за интервенция) или нежеланите ефекти вероятно надвишават желаните ефекти (слаба препоръка срещу интервенция), но съществува значителна несигурност. Слабата препоръка предполага, че не всички индивиди ще бъдат най-добре обслужени от препоръчаното клинично поведение, а е необходимо да се разглеждат по-внимателно от обикновено обстоятелствата, предпочитанията и ценностите на отделния пациент, т.е. доставчиците на здравни грижи трябва да отделят повече време за съвместно вземане на решения и да са уверени, че ясно и изчерпателно обясняват потенциалните ползи и вреди за пациента.

Някои експертни панели са загрижени, че терминът „слаба препоръка“ поражда нежелана негативна конотация с думата „слаб“ и често се идентифицира със „слаби“ доказателства. За да се избегне объркване, слабите препоръки могат да бъдат описвани със следните алтернативни термини:

- условна (conditional) (според стойностите на пациента, наличните ресурси или настройката);
- по преценка (discretionary) (въз основа на мнение на пациент или практикуващ лекар);
- квалифицирана (qualified) (чрез обяснение по въпросите, които биха довели до различни решения).

Ако се използва алтернативен термин, от съществено значение е авторите да упражняват последователност в употребата му във всички препоръки на ръководството.

■ 8.4. Препоръка за използване на интервенция само в клинично проучване

Обещаващи интервенции (обикновено нови) с недостатъчни доказателства за полза могат да бъдат свързани със значителни вреди или разходи. Експертите могат да се притесняват от преждевременни препоръки, насърчаващи бързо разпространение на потенциално неефективни или вредни интервенции и предотвратяващи набиране на вече започнали проучвания. Давайки препоръки за използване на интервенция само в контекста на клинично проучване, те осигуряват стимул за отговор на важни изследователски въпроси и така разрешават несигурността за оптимално поведение. Формулировката може да бъде изразена като изрична силна препоръка – експерименталната интервенция да не се използва извън изследователски контекст.

Препоръки за използване на интервенции само в клинични проучвания са подходящи, когато са изпълнени 3 условия.

1. Засега липсват достатъчно доказателства в подкрепа на решение „за“ или „против“ използване на интервенцията.
2. По-нататъшните изследвания имат голям потенциал за намаляване на несигурността относно ефектите от интервенцията.
3. По-нататъшните изследвания имат добра стойност за очакваните разходи.

Препоръките трябва да бъдат придружени от подробни предложения за специфичните изследователски въпроси, които трябва да бъдат разглеждани, особено за измерваните резултати от важност за пациентите.

8.5. Липса на препоръка

Има три причини, поради които разработчиците на ръководства може да не са склонни да направят препоръка „за“ или „против“ определена стратегия за поведение или да заключат, че препоръката за използване на интервенцията само в клинични проучвания не е подходяща.

1. Доверието в оценките на ефекта е толкова ниско, че препоръката е твърде спекулативна.
2. Независимо от увереността в оценките на ефекта, компромисите са много тясно балансирани, а стойностите, предпочитанията и ресурсните вложвания са неизвестни или твърде променливи.
3. Сравняваните интервенции имат много различни нежелани последици и реакциите на отделните пациенти към тях вероятно ще бъдат много различни. Експертите могат да преценят, че в такива ситуации единствената разумна препоръка е дискусия между пациент и лекар, за да се установят предпочитанията на пациента.

От друга страна, потребителите на ръководства могат да бъдат разочаровани от липсата на препоръка и да твърдят, че вземащите решения нямат лукса да чакат определени доказателства и въпреки че са недостатъчни, клиникът трябва да даде съвет, пациентите трябва да направят избор, а политиките трябва да установят политики. Затова подходът GRADE насърчава панелите да се справят с този дискомфорт и **да дават препоръки дори когато увереността в оценката на ефекта е ниска и/или желаните и нежеланите последици са много тясно балансирани.** Такива препоръки неизбежно ще бъдат слаби и могат да бъдат придружени от квалификации. При необичайни обстоятелства, при които експертите избират да не направят препоръка, те трябва да посочат причината за това решение.

9. СОНМ-МОРЕ МОДИФИКАЦИИ НА ПОДХОДА GRADE

Работната група на GRADE препоръчва да не се правят модификации на оригиналния подход, тъй като елементите на процеса са взаимосвързани и модификациите могат да объркат някои потребители на EP и SoF. Въпреки това броят на съобщенията за различни подходи при прилагане на GRADE нараства и са полезни от прагматична гледна точка.

9.1. Мотиви за модификации

Експертните бордове на СОНМ-МОРЕ изработват клинични ръководства за различни ракови локализации, изискващи включване на сложна и обемна мултимодалност. Необходимостта от препоръки за много голям брой диагностични и терапевтични интервенции налага формулиране на изключително много PICO въпроси и изработване на съответни EPs, SoFs и SRs, за което липсват ресурси в националната клинична онкология. От друга страна, достъпните биомедицински дигитални база данни и директории със статии, референции и резюмета предоставят готови SRs, в които сравнително се интерпретират ефективността и качеството на доказателствата за повечето интервенции от интерес за пациенти и експерти (MEDLINE/PubMed, EMBASE, Cochrane Library, включително CENTRAL, и др.). Систематичните прегледи на Cochrane са структурирани изцяло върху принципите на подхода GRADE и са препоръчителни за използване от експертния борд на СОНМ-МОРЕ.



Методологичен акцент

Модификацията на подхода GRADE, използвана от експертните бордове на СОНМ-МОРЕ, предпоставя намиране на публикация на качествен систематичен преглед/метаанализ с последващо формулиране на PICO въпрос според неговия дизайн

9.2. Модифицирани процеси в подхода GRADE

Модификацията на СОНМ-МОРЕ предоставя рамка със 7 последователни стъпки и процеси.

1. Експертът дефинира задачата си чрез **ключови думи за вид рак и интервенция** (диагностична или терапевтична). Могат да се добавят ключови думи за специфични пациентски популации.

2. Експертът извършва **търсене на SRs** по ключовите думи в дигиталните база данни.
3. Експертът **формулира PICO въпрос** според дизайна на намерения SR – популация (P), компаратор (C) и резултат (O). При дизайни с косвени сравнения и при наличие на повече от един SR трябва да се обмисли формулиране на няколко съответни PICO въпроси.
4. Експертът **съставя SoF** за всеки SR и за всеки резултат в него, както е показано на Табл. 2.
5. Експертът извършва **оценка на качеството на доказателствата** за всеки SR по таблиците с фактори за повишаване или намаляване на нивото, маркирайки всеки един от тях с „да/не“ (виж Табл. 6 и Табл. 7).
6. Експертът **формулира препоръка** в една от двете категории за сила (виж Табл. 13).
7. Експертният борд дискутира качеството на доказателствата за формулираните препоръки, **дефинирайки степени на сигурност и качество на доказателствата**, както е показано на Табл. 4 и Табл. 5.

9.3. Недостатъци и предимства на модификацията

Основен източник на отклонения в модифицирания процес са готовите SRs, върху чиито дизайн и качество експертът няма влияние. Затова модифицираният подход на COHМ-MOPE насърчава отделните експерти, автори на специални раздели в клиничното ръководство, внимателно да подбират SRs, селектирайки ги според време на публикуване, включени проучвания, автори и школи. Препоръчително е да се използват SRs, изготвени след 2011 г., когато подходът GRADE получава по-глобално утвърждаване. Най-подходящи са Cochrane SRs, в чиято методология са включени всички GRADE критерии за оценка на качеството на доказателствата.

Макар и рядко се налага експертът да включи в ръководството раздел (препоръка) за интервенция, за която не се открива публикуван SR. В тези случаи се препоръчва търсене на RCTs или в краен случай – на OSs. Независимо от използвания източник на резултати, задължително се провежда оценка на качеството на доказателствата според оригиналните критерии на подхода GRADE.

Ключова стъпка в модифицирания подход на COHМ-MOPE са **работните дискусии на експертните бордове**. Екипните и мултидисциплинарни дебати са най-висша форма и последна инстанция за контрол върху стриктното

приложение на всички изисквания на оригиналния подход GRADE, свързани преди всичко с оценката на качеството на доказателствата. От разработчиците на COHM-MOPE клинични ръководства се изисква детайлно познаване на всички теоретични понятия и статистически параметри, залежали в оригиналните таблици за оценка на качеството (*Табл. 6 и Табл. 7*). Също така е необходимо свободно боравене с оригиналните степени на сигурност на доказателствата (*Табл. 4*).

Главно предимство на модифицирания подход на COHM-MOPE е спестяването на ресурси за изготвяне на голям брой собствени SRs – дейност, непосилна засега в организацията на българската клинична онкология. От друга страна, модификацията контекстуализира оригиналния подход GRADE в българската здравна система, насърчавайки използването ѝ за стратегически решения за здравни политики, здравни технологии, здравни фондове и др.



GRADE НАРЪЧНИК ЗА ЕКСПЕРТИ

Агантация no: Holger Schünemann, Jan Brožek, Gordon Guyatt, Andrew Oxman
<https://gdt.gradepro.org/app/handbook/handbook.html#h.wlsfq2lmj0gb>

Първо издание

Димитър КАЛЕВ – *агантация, съставител и редактор*

Петър ЖЕЛЕВ – *графичен дизайн и предпечат*

Печат: „Хеликс прес“

Варна, 2025

ISBN 978-619-7094-76-3



