SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

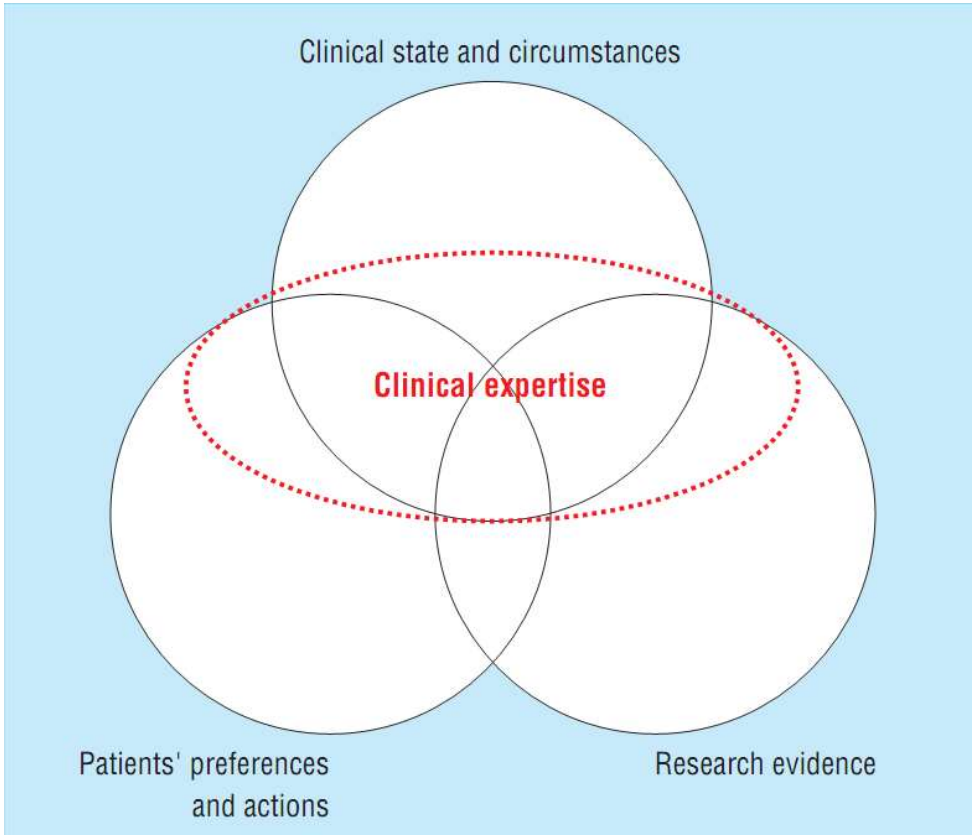
Una Premessa…

Physicians' and patients' choices in evidence based practice

Evidence does not make decisions, people do

R Brian Haynes P J Devereaux Gordon H Guyatt

BMJ 2002;324:1350

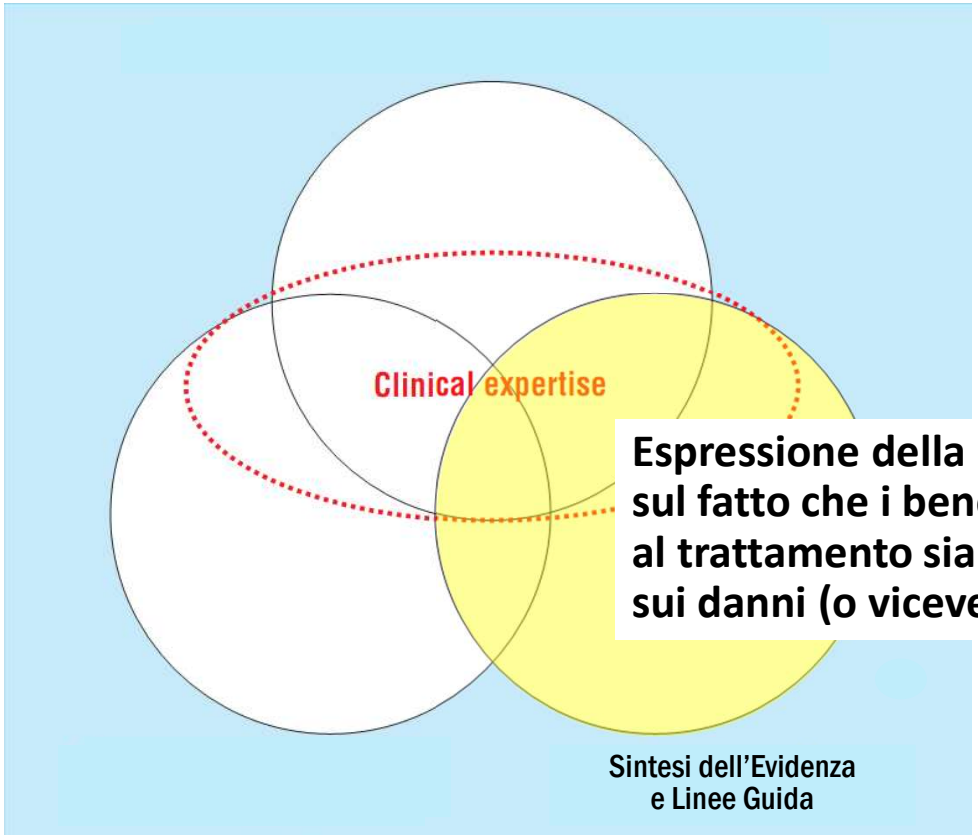


Physicians' and patients' choices in evidence based practice

Evidence does not make decisions, people do

R Brian Haynes P J Devereaux Gordon H Guyatt

BMJ 2002;324:1350



Espressione della "confidenza" sul fatto che i benefici ascrivibili al trattamento siano prevalenti sui danni (o viceversa)

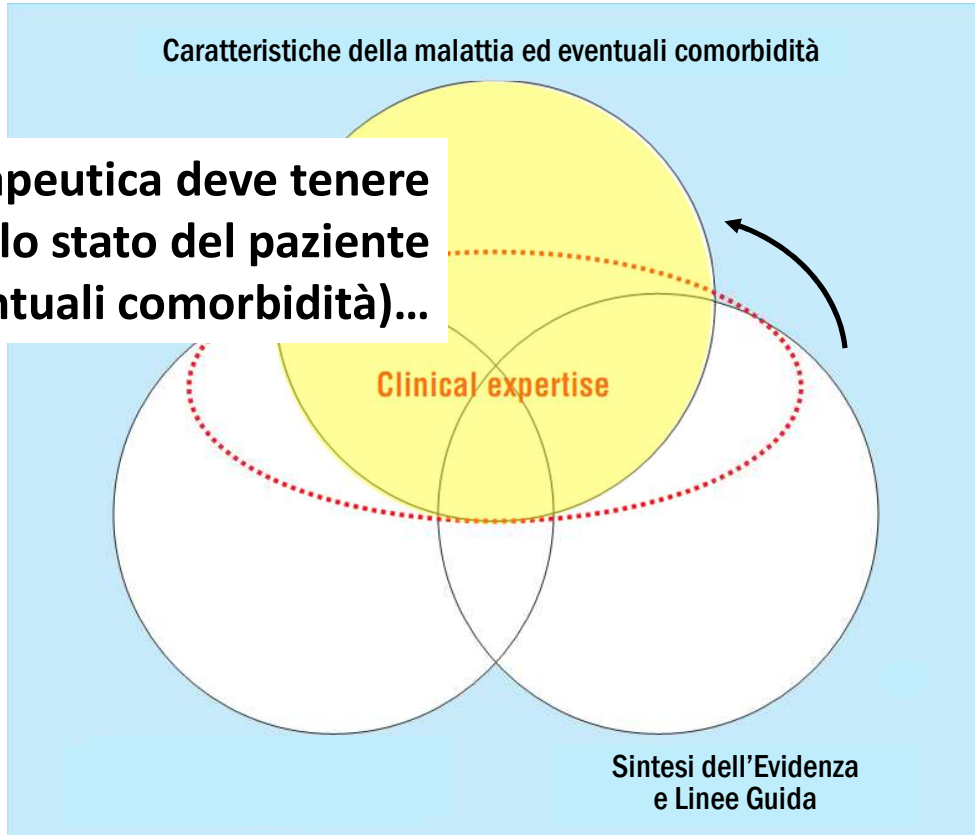
Physicians' and patients' choices in evidence based practice

Evidence does not make decisions, people do

R Brian Haynes P J Devereaux Gordon H Guyatt

BMJ 2002;324:1350

La proposta terapeutica deve tenere conto dello stato del paziente (malattia ed eventuali comorbidità)…

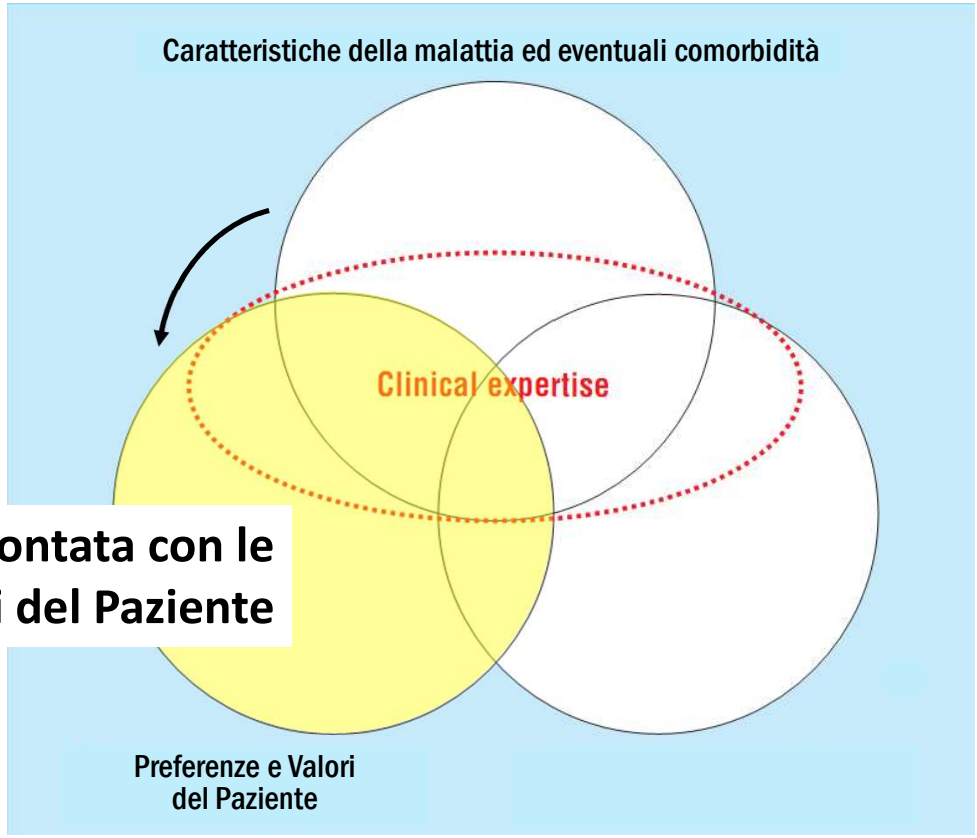


Physicians' and patients' choices in evidence based practice

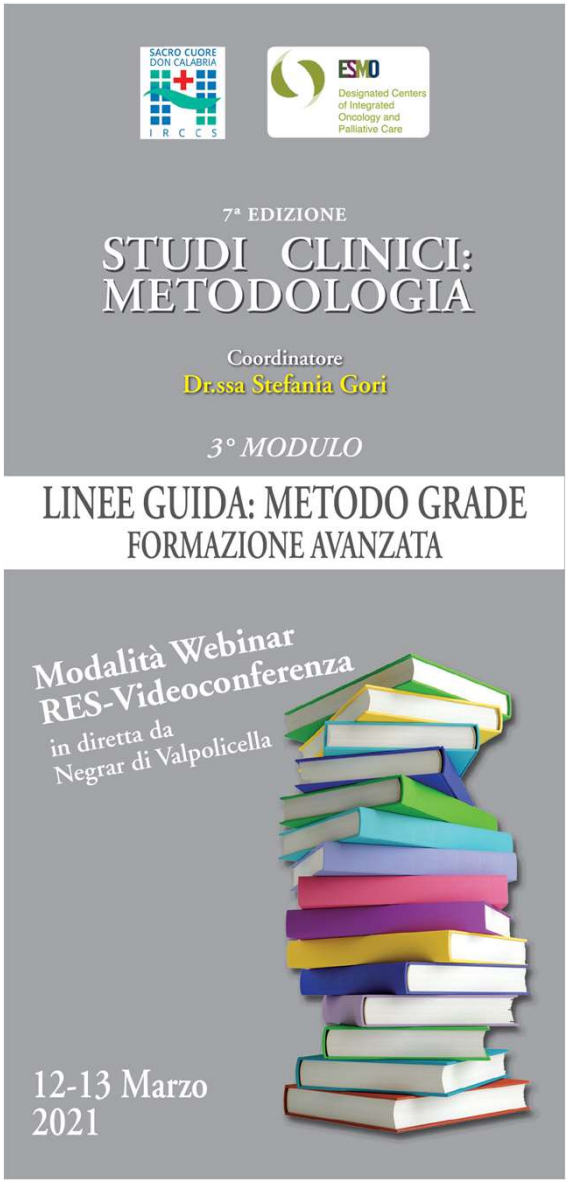
Evidence does not make decisions, people do

R Brian Haynes P J Devereaux Gordon H Guyatt

BMJ 2002;324:1350



… e deve essere confrontata con le Preferenza e i Valori del Paziente



***Obiettivo da perseguire*:**

determinare la "confidenza" sul fatto che i benefici ascrivibili al trattamento siano prevalenti sui danni (o viceversa)

La valutazione delle Evidenze prima del Sistema GRADE…

National Comprehensive Cancer Network

Your Best Resource in the Fight Against Cancer®

NCCN Categories of Evidence and Consensus

Category 1: The recommendation is based on high-level evidence (e.g. randomized controlled trials) and there is uniform NCCN consensus.

Category 2A: The recommendation is based on lower-level evidence and there is uniform NCCN consensus.

Category 2B: The recommendation is based on lower-level evidence and there is nonuniform NCCN consensus (but no major disagreement).

Category 3: The recommendation is based on any level of evidence but reflects major disagreement.

All recommendations are category 2A unless otherwise noted.

Oxford Centre for Evidence-based Medicine Levels of Evidence
(March 2009)

Level	Therapy/Prevention, Aetiology/Harm
1a	SR (with homogeneity*) of RCTs
1b	Individual RCT (with narrow Confidence Interval‡)
1c	All or none§
2a	SR (with homogeneity*) of cohort studies
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)
2c	"Outcomes" Research; Ecological studies
3a	SR (with homogeneity*) of case-control studies
3b	Individual Case-Control Study
4	Case-series (and poor quality cohort and case-control studies§§)
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998. Updated by Jeremy Howick March 2009.

Oxford Centre for Evidence-based Medicine Levels of Evidence
(March 2009)

Level	Therapy/Prevention, Aetiology/Harm
1a	SR (with homogeneity*) of RCTs
1b	Individual RCT (with narrow Confidence Interval‡)
1c	All or none§
2a	SR (with homogeneity*) of cohort studies
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)
2c	"Outcomes" Research; Ecological studies
3a	SR (with homogeneity*) of case-control studies
3b	Individual Case-Control Study
4	Case-series (and poor quality cohort and case-control studies§§)
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies or extrapolations from level 1 studies
C	level 4 studies or extrapolations from level 2 or 3 studies
D	level 5 evidence or troublingly inconsistent or inconclusive studies of any level

Levels of evidence

I	Evidence from at least one large randomised, controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity
II	Small randomised trials or large randomised trials with a suspicion of bias (lower methodological quality) or meta-analyses of such trials or of trials with demonstrated heterogeneity
III	Prospective cohort studies
IV	Retrospective cohort studies or case–control studies
V	Studies without control group, case reports, experts opinions

SIGN
Scottish Intercollegiate Guidelines Network

SIGN 50: A guideline developer's handbook *December 2011*

Disegno dello Studio

1^{++}	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1^{+}	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1^{-}	Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2^{++}	High quality systematic reviews of case control or cohort studies High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2^{+}	Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2^{-}	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

SIGN
Scottish Intercollegiate Guidelines Network

SIGN 50: A guideline developer's handbook *December 2011*

Qualità (affidabilità) dei risultati

Level of evidence	
1^{++}	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1^{+}	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1^{-}	Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2^{++}	High quality systematic reviews of case control or cohort studies High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2^{+}	Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2^{-}	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Note: The grade of recommendation relates to the strength of the evidence on which the recommendation is based. It does not reflect the clinical importance of the recommendation.

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Note: The grade of recommendation relates to the strength of the evidence on which the recommendation is based. It does not reflect the clinical importance of the recommendation.

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Healthcare Improvement Scotland | SIGN

RECOMMENDATIONS

Some recommendations can be made with more certainty than others. The wording used in the recommendations in this guideline denotes the certainty with which the recommendation is made (the 'strength' of the recommendation).

The 'strength' of a recommendation takes into account the quality (level) of the evidence. Although higher-quality evidence is more likely to be associated with strong recommendations than lower-quality evidence, a particular level of quality does not automatically lead to a particular strength of recommendation.

Other factors that are taken into account when forming recommendations include: relevance to the NHS in Scotland; applicability of published evidence to the target population; consistency of the body of evidence, and the balance of benefits and harms of the options.

R For '**strong**' recommendations on interventions that '**should**' be used, the guideline development group is confident that, for the vast **majority** of people, the intervention (or interventions) will do more good than harm. For **'strong'** recommendations on interventions that **'should not'** be used, the guideline development group is confident that, for the vast **majority** of people, the intervention (or interventions) will do more harm than good.

R For **'conditional'** recommendations on interventions that should be '**considered**', the guideline development group is confident that the intervention will do more good than harm for **most** patients. The choice of intervention is therefore more likely to vary depending on a person's values and preferences, and so the healthcare professional should spend more time discussing the options with the patient.

GUIDELINES INCLUDING GRADING OF NOVEL DRUGS WITH ESMO-MCBS

George Pentheroudakis

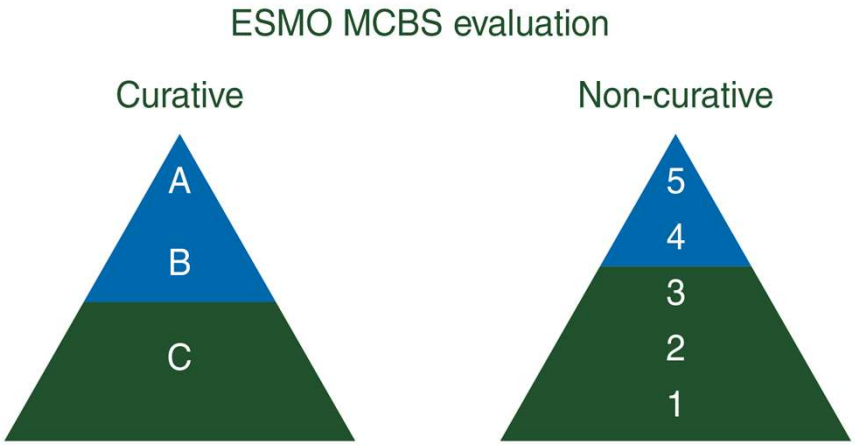
Chair, ESMO Guidelines Committee

A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS)

N. I. Cherny[1*], R. Sullivan[2], U. Dafni[3], J. M. Kerst[4], A. Sobrero[5], C. Zielinski[6], E. G. E. de Vries[7] & M. J. Piccart[8,9]

Annals of Oncology 26: 1547–1573, 2015

This tool uses a rational, structured and consistent approach to derive a relative ranking of the magnitude of clinically meaningful benefit that can be expected from a new anti-cancer treatment.

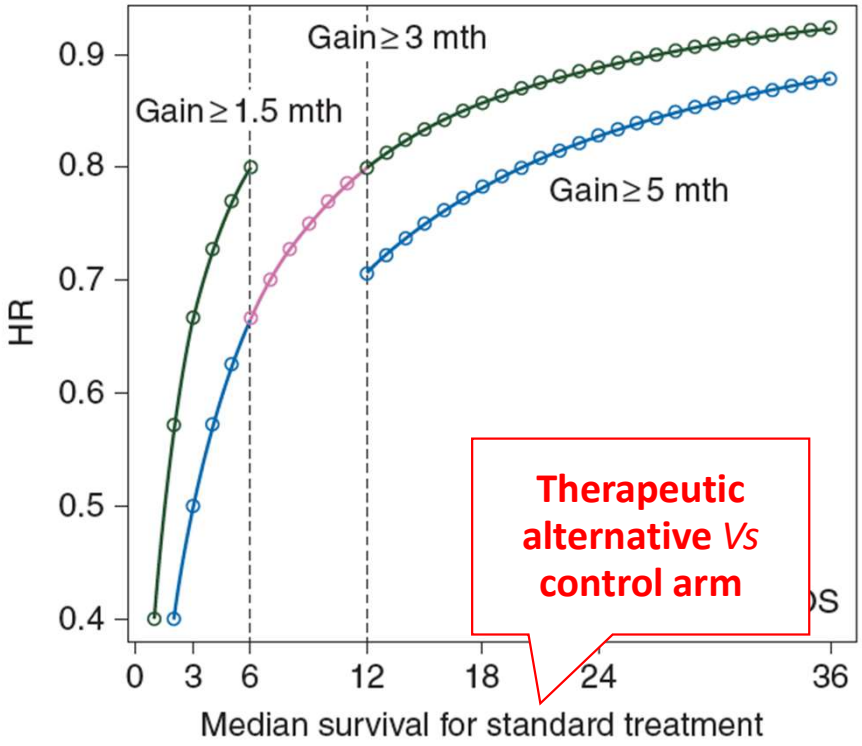


Visualisation of ESMO-MCB scores for curative and non-curative setting. A & B and 5 and 4 represent the grades with substantial improvement.

ESMO-MCBS

What if non proportional hazards?

(metastatic disease, no long-survivors; plateau of long-survivors; etc.)



Incomplete efficacy estimator; correlation with HR?

ESMO-MCBS

*fo*his version is used for therapies evaluated using a pr*imary* come of OS. The form is stratified by median OS of the control arm ≤12 and >12 months. Preliminary grading takes into consideration HR and median survival gain as well as late survival advantage and is reported on a 4-poin*t scale*. When there is differen*tial grading* between the median an*d* ival gain, the higher ails. Preliminary scores ca*n* ded by 1 point wh*en the* perimental arm demonstrates improved QoL or delayed deterioration in QoL using a validated scale or substantial reduction in grade 3 or 4 toxicity. A score of 5 can only be *a*hieved when optimal surv*i*l outcomes are further enhanced indicating reduc or improved QoL.

... as I thought

M.I.D. or P<0.05?

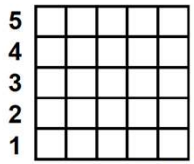
M.I.D. or P<0.05?

Clinical relevance or P<0.05?

What if grade 2 diarrhea?

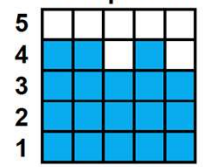
NCCN Evidence Blocks™

NCCN EVIDENCE BLOCKS CATEGORIES AND DEFINITIONS



E = Efficacy of Regimen/Agent
S = Safety of Regimen/Agent
Q = Quality of Evidence
C = Consistency of Evidence
A = Affordability of Regimen/Agent

Example Evidence Block



E = 4
S = 4
Q = 3
C = 4
A = 3

Efficacy of Regimen/Agent

5	Survival
4	
3	
2	
1	

Quality of Evidence

5	Study Design
4	
3	
2	
1	

Safety of Regimen/Agent

5	Interference with ADLs
4	
3	
2	
1	

Note: For significant chronic or long-term toxicities, score decreased by 1

Consistency of Evidence

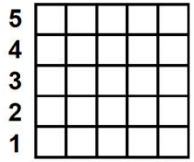
5	Consistency
4	
3	
2	
1	

Affordability of Regimen/Agent (includes drug cost, supportive care, infusions, toxicity monitoring, management of toxicity)

5	Costs
4	
3	
2	
1	

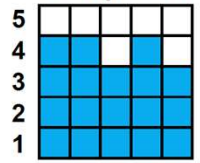
NCCN Evidence Blocks™

NCCN EVIDENCE BLOCKS CATEGORIES AND DEFINITIONS



E = Efficacy of Regimen/Agent
S = Safety of Regimen/Agent
Q = Quality of Evidence
C = Consistency of Evidence
A = Affordability of Regimen/Agent

Example Evidence Block



E = 4
S = 4
Q = 3
C = 4
A = 3

Efficacy of Regimen/Agent

5	**Highly effective:** Cure likely and often provides long-term survival advantage
4	**Very effective:** Cure unlikely but sometimes provides long-term survival advantage
3	**Moderately effective:** Modest impact on survival, but often provides control of disease
2	**Minimally effective:** No, or unknown impact on survival, but sometimes provides control of disease
1	**Palliative:** Provides symptomatic benefit only

Safety of Regimen/Agent

5	**Usually no meaningful toxicity:** Uncommon or minimal toxicities; no interference with activities of daily living (ADLs)
4	**Occasionally toxic:** Rare significant toxicities or low-grade toxicities only; little interference with ADLs
3	**Mildly toxic:** Mild toxicity that interferes with ADLs
2	**Moderately toxic:** Significant toxicities often occur but life threatening/fatal toxicity is uncommon; interference with ADLs is frequent
1	**Highly toxic:** Significant toxicities or life threatening/fatal toxicity occurs often; interference with ADLs is usual and severe

Note: For significant chronic or long-term toxicities, score decreased by 1

Quality of Evidence

5	**High quality:** Multiple well-designed randomized trials and/or meta-analyses
4	**Good quality:** One or more well-designed randomized trials
3	**Average quality:** Low quality randomized trial(s) or well-designed non-randomized trial(s)
2	**Low quality:** Case reports or extensive clinical experience
1	**Poor quality:** Little or no evidence

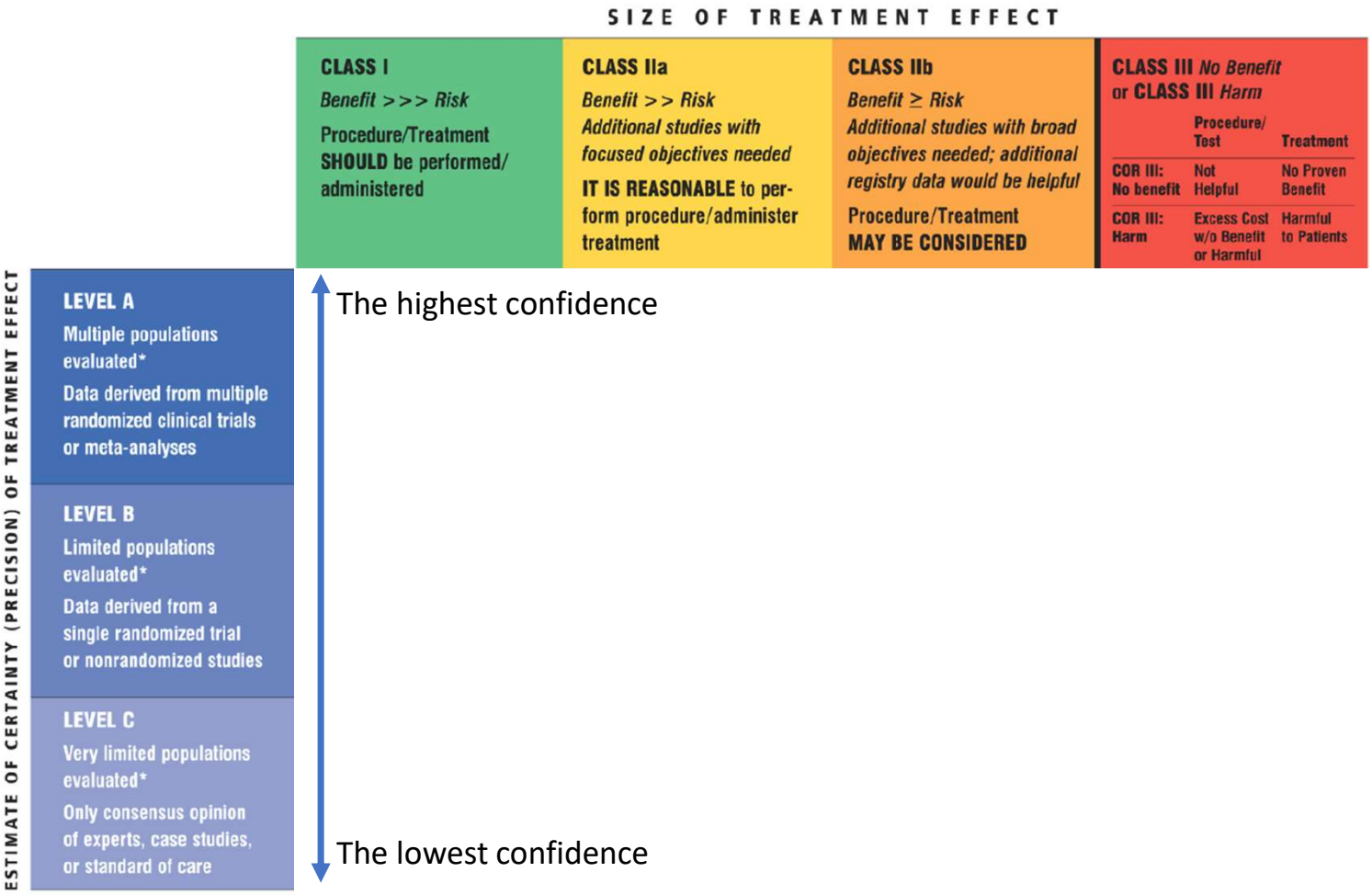
Consistency of Evidence

5	**Highly consistent:** Multiple trials with similar outcomes
4	**Mainly consistent:** Multiple trials with some variability in outcome
3	**May be consistent:** Few trials or only trials with few patients, whether randomized or not, with some variability in outcome
2	**Inconsistent:** Meaningful differences in direction of outcome between quality trials
1	**Anecdotal evidence only:** Evidence in humans based upon anecdotal experience

Affordability of Regimen/Agent (includes drug cost, supportive care, infusions, toxicity monitoring, management of toxicity)

5	**Very inexpensive**
4	**Inexpensive**
3	**Moderately expensive**
2	**Expensive**
1	**Very expensive**

2013 ACC/AHA Cardiovascular Risk Guideline



2013 ACC/AHA Cardiovascular Risk Guideline

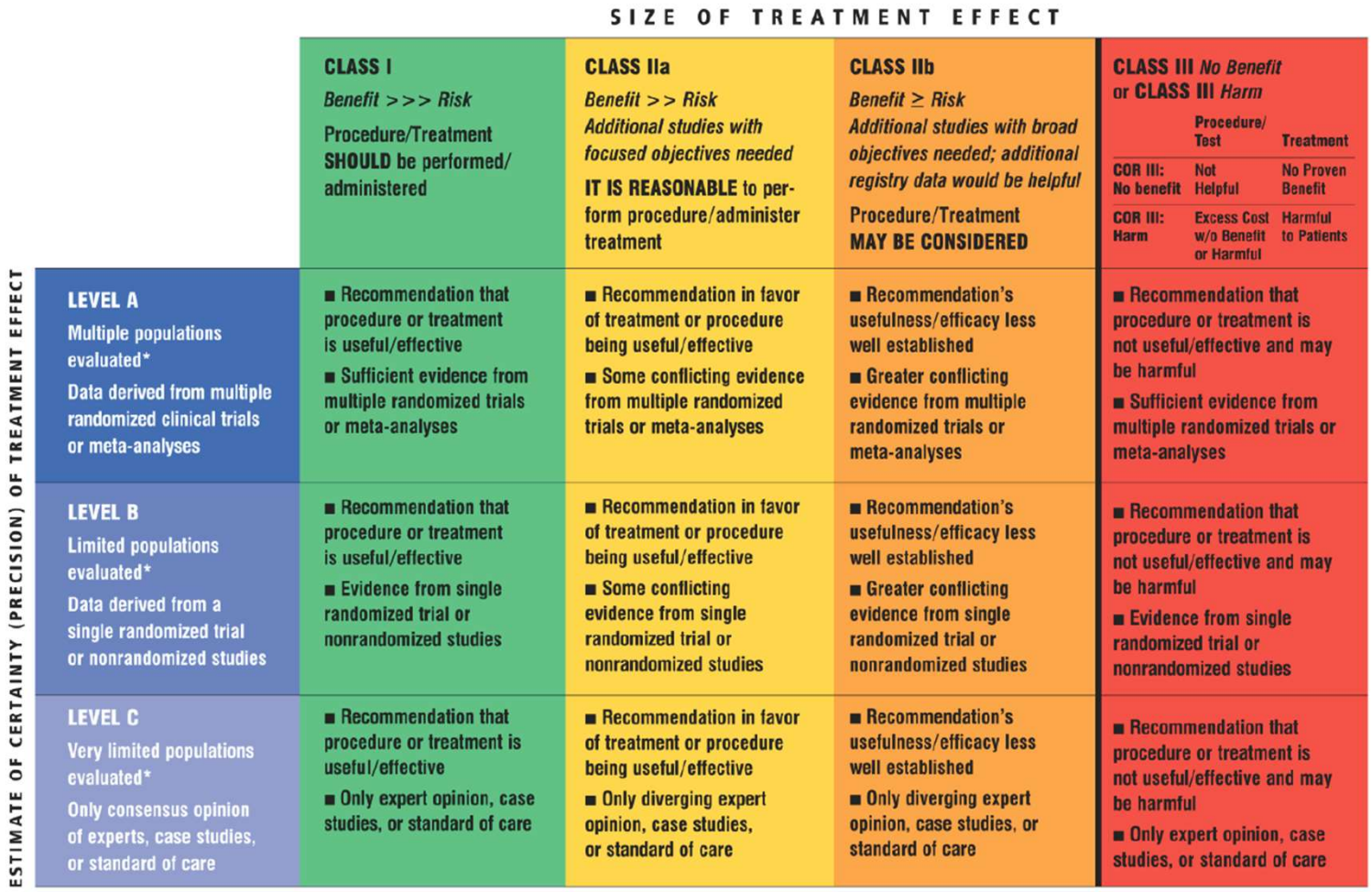
SIZE OF TREATMENT EFFECT

CLASS I	CLASS IIa	CLASS IIb	CLASS III *No Benefit* or CLASS III *Harm*		
				Procedure/ Test	Treatment
Benefit >>> Risk	*Benefit >> Risk*	*Benefit ≥ Risk*	COR III: No benefit	Not Helpful	No Proven Benefit
Procedure/Treatment **SHOULD** be performed/ administered	*Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	*Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	COR III: Harm	Excess Cost w/o Benefit or Harmful	Harmful to Patients

Benefit outweights harm

Harm outweights benefit

2013 ACC/AHA Cardiovascular Risk Guideline



SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Il Sistema GRADE…

World Health Organization

GRADE

Working Group

Grades of **R**ecommendation **A**ssessment, **D**evelopment and **E**valuation

McMaster University

- Aim: to develop a common, transparent and sensible system for grading the quality of evidence and the strength of recommendations (over 100 systems)
- International group of guideline developers, methodologists & clinicians from around the world (>200 contributors) – since 2000
- International group: WHO, ACCP, AHRQ, Australian NMRC, BMJ Clinical Evidence, CC, CDC, McMaster, NICE, Oxford CEBM, SIGN, UpToDate, USPSTF

CMAJ 2003, BMJ 2004, BMC 2004, BMC 2005, AJRCCM 2006, Chest 2006, BMJ 2008

Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group

David Atkins[1], Martin Eccles[2], Signe Flottorp[3], Gordon H Guyatt[4], David Henry[5], Suzanne Hill[5], Alessandro Liberati[6], Dianne O'Connell[7], Andrew D Oxman[3], Bob Phillips[8], Holger Schünemann[4,9], Tessa Tan-Torres Edejer[10], Gunn E Vist*[3], John W Williams Jr[11] and The GRADE Working Group[3]

BMC Health Services Research 2004, **4**:38

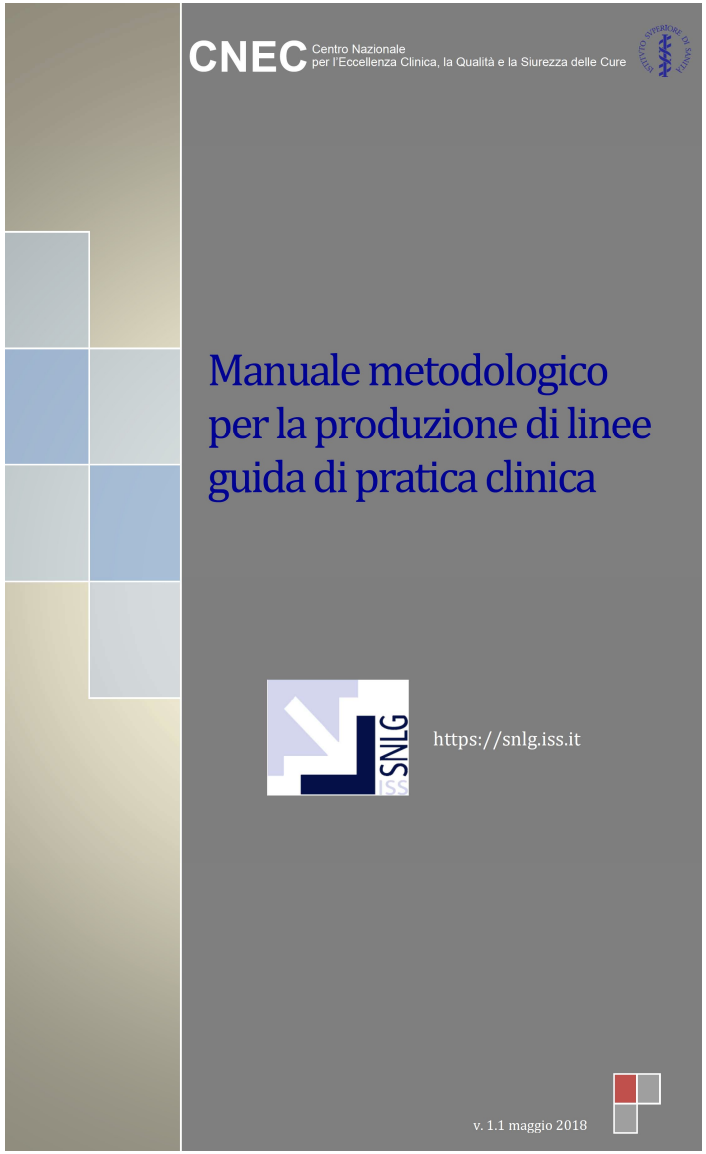
Table 1: Summary of assessments of the sensibility of six approaches to rating levels of evidence and strength of recommendation

Criteria[1]	ACCP			ANHMRC[2]			USTFCPS			OCEBM			SIGN			USPSTF[3]		
	No		Yes	No		Yes	No		Yes	No		Yes	No		Yes	No		Yes
1. Applicable to different questions:																		
Effectiveness			12		2	8		1	11			12	1		11		2	9
Harm		1	11		5	5	1	7	4		1	11	1	3	8	2	2	7
Diagnosis	7	3	2	4	4	2	9	3				12	5	2	5	2	2	7
Prognosis	6	3	3	2	5	3	9	2	1			11	4	3	5	3	3	5
2. Can be used by:																		
Professionals		1	11	1	5	3		7	4	1	6	5		5	7		3	8
Policy makers	1	5	6	1	5	3	1	2	9	3	7	2	2	6	4	1	4	6
Patients	4	5	3	5	5		6	3	3	9	3		7	5		4	6	1
3. Clear and simple	1	5	6	2	6	1	2	8	2	2	4	5	1	5	6	1	4	7
4. Information not available		8	4	1	5	3	1	6	5		4	8	1	7	4	1	9	2
5. Subjective decisions	2	10		5	2	2	5	5	2		7	5	5	7		2	9	
6. Inappropriate dimensions	1	3	8		1	6	2	4	6		1	10	1	2	8	1	4	6
7. Missing dimensions	1	6	5	2	2	4	5	4	3	9	3	1	5	4	3	2	5	4
Aggregation of dimensions:																		
8. Clear and simple	1	5	6	4	1	2	2	2	7	3	4	4		6	6	2	7	2
9. Appropriate		6	5	3	1	1	3	4	4	2	5	4	1	4	6	1	5	5
10. Sufficient categories	1	4	6	4	2	1		5	7	2	2	7	1	2	9		1	10
11. Likely to discriminate		7	5	2	5	1	1	9	2	2	4	6		5	7		4	7
12. Assessments reproducible	1	8	3	4	4		2	7	2		7	4	1	8	2		10	

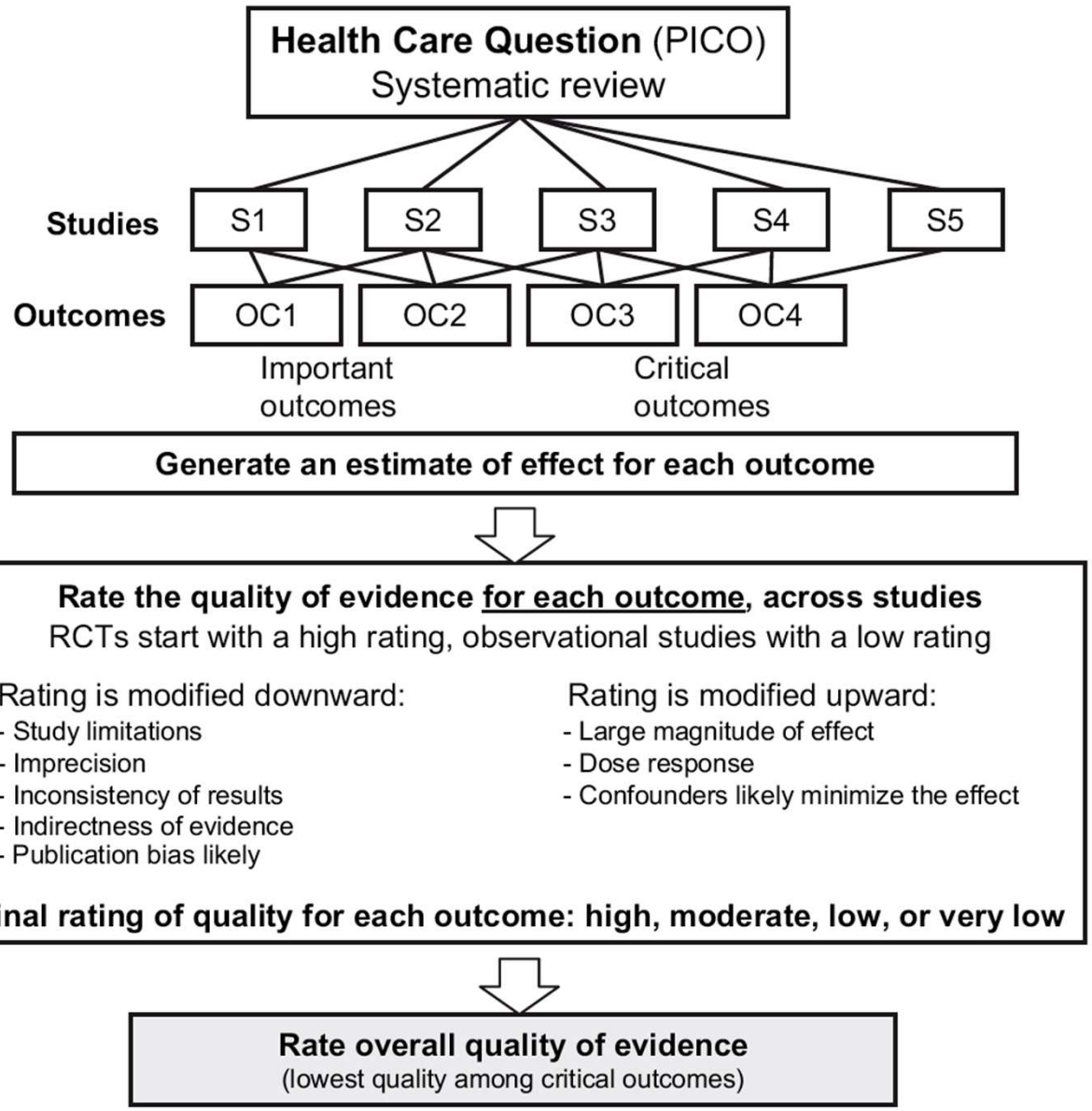
Based on discussions of the strengths and limitations of current approaches to grading levels of evidence and the strength of recommendations, we agreed to develop an approach that addresses the major limitations that we identified.

The GRADE approach

- Considers
  - all factors to determine how confident we are in the results – quality of evidence
  - the evidence for each outcome in the review separately
  - magnitude of the effect
- Ensures
  - systematic process
  - transparency



Il metodo GRADE propone una valutazione della qualità delle prove più ampia e articolata di quella proposta da tutti gli altri sistemi di grading e rappresenta *lo standard metodologico di riferimento per la produzione di linee guida adottato sempre più diffusamente a livello internazionale.*


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

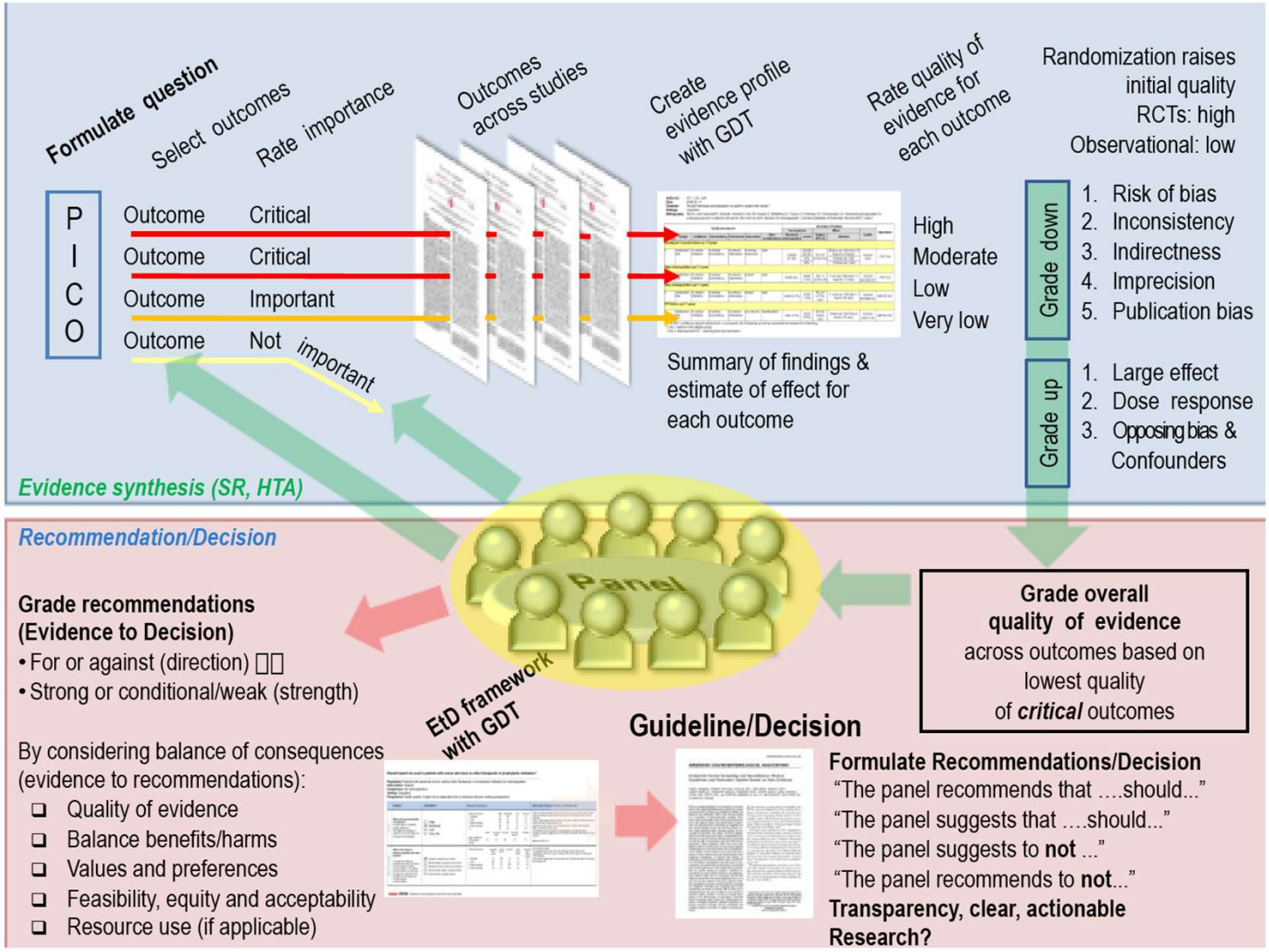
LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

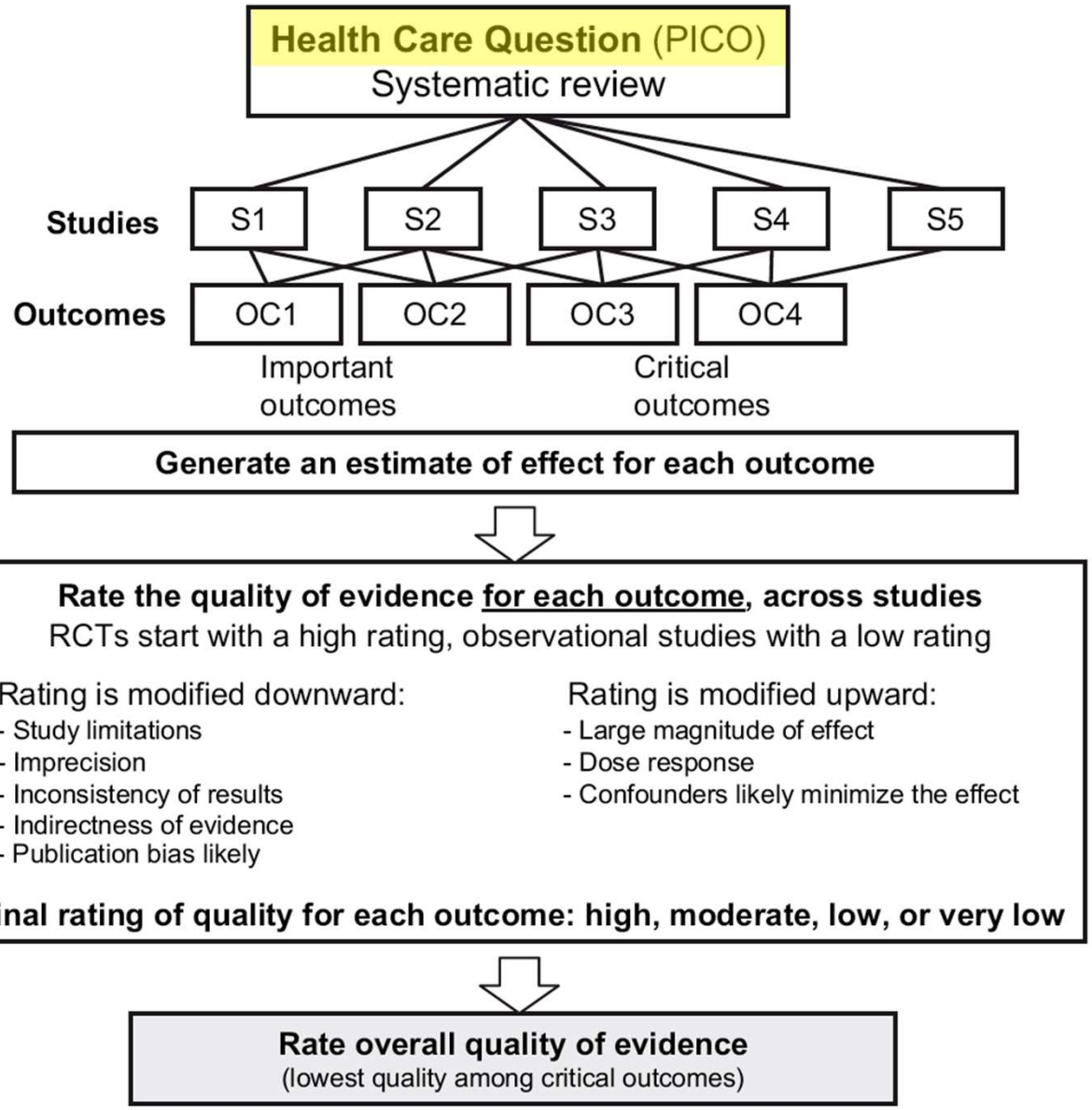
Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

La strutturazione del quesito clinico. Scelta e classificazione degli outcome.

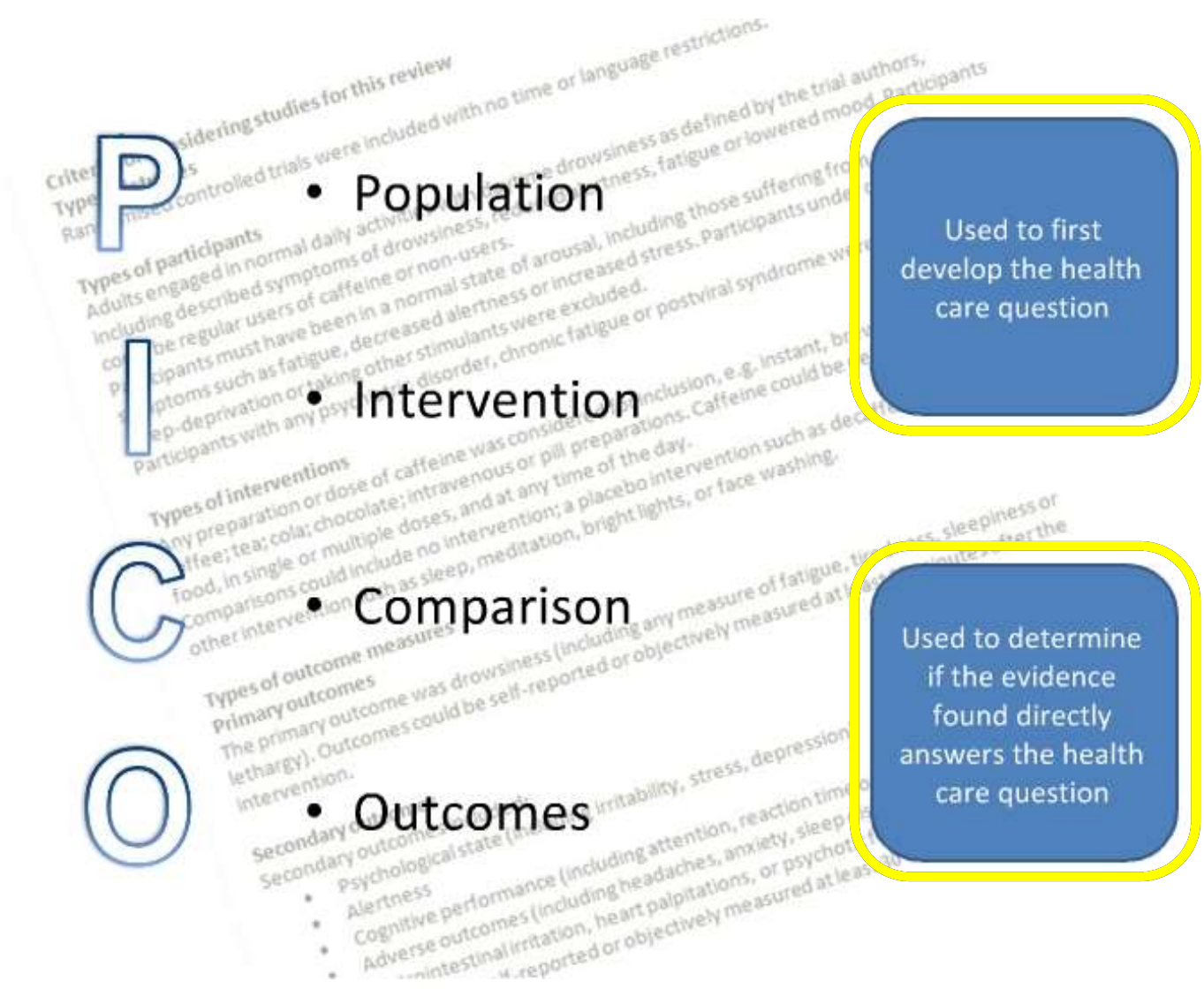

Formulate question
Select outcomes
Rate importance
Outcomes across studies
Create evidence profile with GDT
Rate quality of evidence for each outcome
Randomization raises initial quality
RCTs: high
Observational: low
P
I
C
O
Outcome Critical
Outcome Critical
Outcome Important
Outcome Not important
High
Moderate
Low
Very low
Summary of findings & estimate of effect for each outcome
Grade down
1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias
Grade up
1. Large effect
2. Dose response
3. Opposing bias & Confounders
Evidence synthesis (SR, HTA)
Recommendation/Decision
Panel
Grade overall quality of evidence across outcomes based on lowest quality of critical outcomes
Grade recommendations (Evidence to Decision)
• For or against (direction)
• Strong or conditional/weak (strength)
By considering balance of consequences (evidence to recommendations):
Quality of evidence
Balance benefits/harms
Values and preferences
Feasibility, equity and acceptability
Resource use (if applicable)
EtD framework with GDT
Guideline/Decision
Formulate Recommendations/Decision
"The panel recommends that ….should..."
"The panel suggests that ….should..."
"The panel suggests to not ..."
"The panel recommends to not..."
Transparency, clear, actionable
Research?


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)

Important Questions

Should be
from practice
NOT
evidence driven

GRADE



- Population
- Intervention
- Comparison
- Outcomes

Used to first develop the health care question

Used to determine if the evidence found directly answers the health care question

Strutturazione del Quesito Clinico sec. modello P.I.C.O.

P	Nei **P**azienti con…	Specifiche caratteristiche di malattia (stadio, classe di rischio, ecc.)
I	l'**I**ntervento…	Intervento terapeutico oggetto del quesito clinico
C	(è suscettibile di impiego) in **C**onfronto con…	Trattamento altrimenti considerabile in alternativa all'intervento in esame
O	riguardo agli **O**utcome di beneficio/danno…	Parametri clinico-laboratoristici ritenuti essenziali per la proposta terapeutica

Strutturazione del Quesito Clinico sec. modello P.I.C.O.

P	Nei **P**azienti con…	Specifiche caratteristiche di malattia (stadio, classe di rischio, ecc.)
I	l'**I**ntervento…	Intervento terapeutico oggetto del quesito clinico
C	(è suscettibile di impiego) in **C**onfronto con…	Trattamento altrimenti considerabile in alternativa all'intervento in esame
O	riguardo agli **O**utcome di beneficio/danno…	Parametri clinico-laboratoristici ritenuti essenziali per la proposta terapeutica

Strutturazione del Quesito Clinico sec. modello P.I.C.O.

P	Nei **P**azienti con…	Specifiche caratteristiche di malattia (stadio, classe di rischio, ecc.)
I	l'**I**ntervento…	Intervento terapeutico oggetto del quesito clinico
C	(è suscettibile di impiego) in **C**onfronto con…	Trattamento altrimenti considerabile in alternativa all'intervento in esame
O	riguardo agli **O**utcome di beneficio/danno…	Parametri clinico-laboratoristici ritenuti essenziali per la proposta terapeutica

P. pazienti adulti sottoposti a resezione chirurgica per carcinoma polmonare non a piccole cellule (NSCLC) e presenza di mutazioni attivanti il recettore per il fattore di crescita epidermico (EGFR).

I. Osimertinib

C. EGFR-tki? (gefitinib, erlotinib)

P. pazienti adulti sottoposti a resezione chirurgica per carcinoma polmonare non a piccole cellule (NSCLC) e presenza di mutazioni attivanti il recettore per il fattore di crescita epidermico (EGFR).

I. Osimertinib

C. non trattamento con inibitori di EGFR (in presenza o meno di trattamento CT adiuvante)

LG tumori del polmone 2020

Qualità globale delle evidenze	Raccomandazione clinica	Forza della raccomandazione
Moderata	Nei pazienti affetti da NSCLC radicalmente operato allo stadio II-IIIA e mutazione attivante di EGFR il trattamento con inibitori di EGFR non dovrebbe essere presa in considerazione come terapia adiuvante al di fuori degli studi clinici.	**Negativa debole**

Strutturazione del Quesito Clinico sec. modello P.I.C.O.

P	Nei **P**azienti con…	Specifiche caratteristiche di malattia (stadio, classe di rischio, ecc.)
I	l'**I**ntervento…	Intervento terapeutico oggetto del quesito clinico
C	(è suscettibile di impiego) in **C**onfronto con…	Trattamento altrimenti considerabile in alternativa all'intervento in esame
O	riguardo agli **O**utcome di beneficio/danno…	Parametri clinico-laboratoristici ritenuti essenziali per la proposta terapeutica

Outcomes

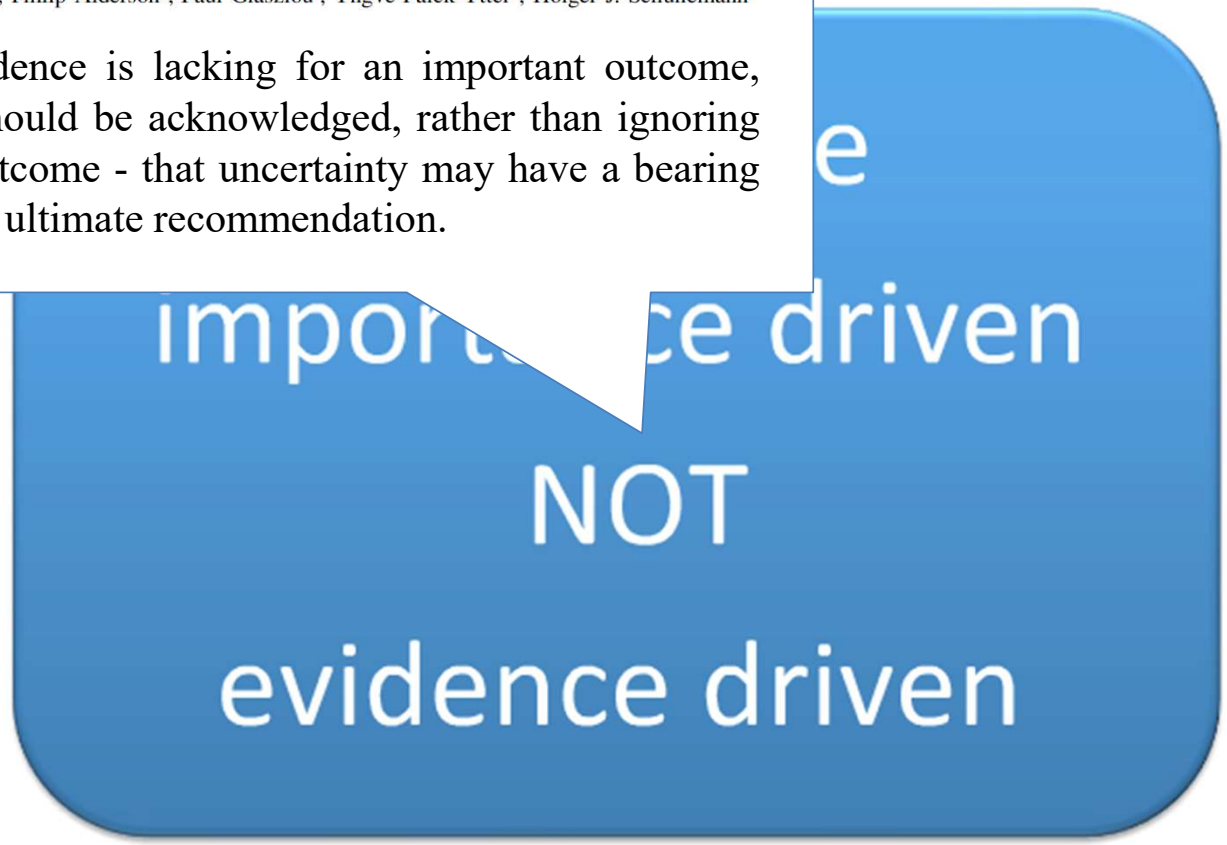
Should be
importance driven
NOT
evidence driven

Journal of Clinical Epidemiology 64 (2011) 395–400

GRADE guidelines: 2. Framing the question and deciding on important outcomes

Gordon H. Guyatt[a,*], Andrew D. Oxman[b], Regina Kunz[c], David Atkins[d], Jan Brozek[a], Gunn Vist[b], Philip Alderson[e], Paul Glasziou[f], Yngve Falck-Ytter[g], Holger J. Schünemann[a]

If evidence is lacking for an important outcome, this should be acknowledged, rather than ignoring the outcome - that uncertainty may have a bearing on the ultimate recommendation.

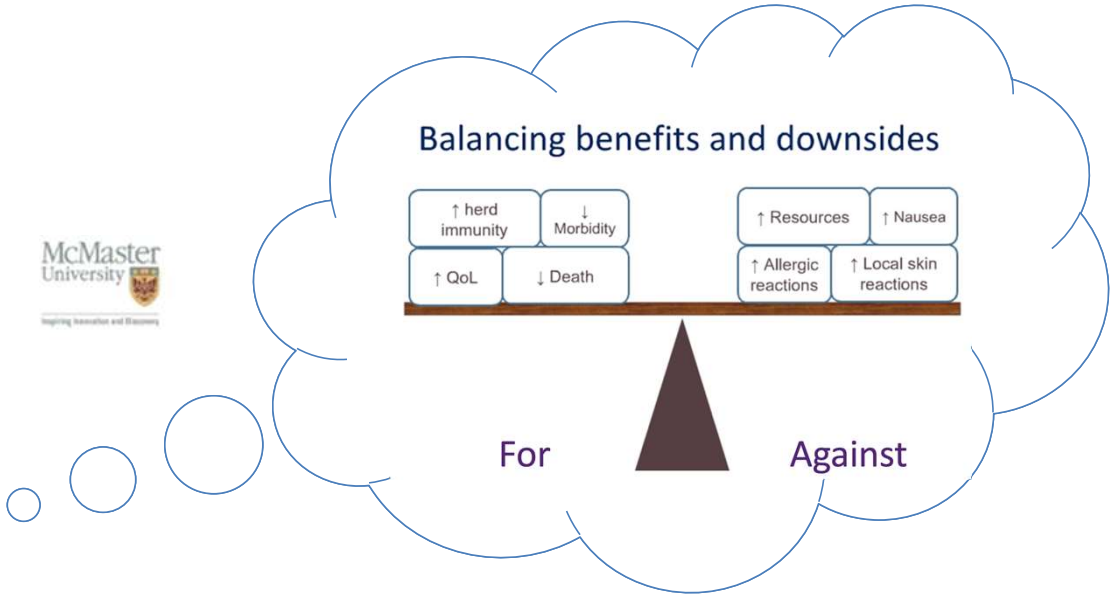


Choosing outcomes

Desirable outcomes
- lower mortality
- reduced hospital stay
- reduced duration of disease
- reduced resource expenditure

Undesirable outcomes
- adverse reactions
- the development of resistance
- costs of treatment

Recommendations must consider desirable and undesirable outcomes

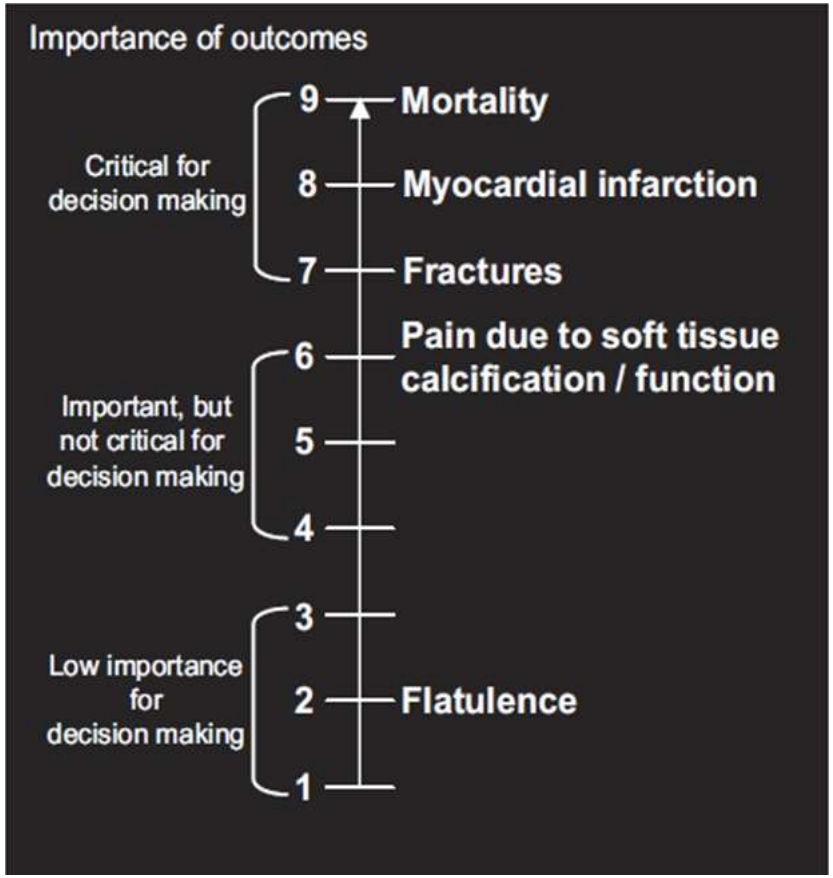


Valutazione dell'importanza degli *outcome* mediante votazione individuale dei panelisti, utilizzando una scala a 9 punti e assegnando l'outcome a una delle tre categorie sulla base del punteggio mediano ottenuto.

Classificazione degli *outcome* proposta dal metodo GRADE

Rating **(mediana del voto)**	**Importanza**	**Incluso in**
7 8 9	*outcome* importanti ed essenziali	tabelle sulla qualità delle prove: SÌ raccomandazione: SÌ
4 5 6	*outcome* importanti ma non essenziali	tabelle sulla qualità delle prove: SÌ raccomandazione: NO
1 2 3	*outcome* non importanti	tabelle sulla qualità delle prove: NO raccomandazione: NO

Importanza degli Outcomes



Hierarchy of outcomes according to their importance to assess the effect of phosphate-lowering drugs in patients with renal failure and hyperphosphatemia.

G.H. Guyatt et al. / Journal of Clinical Epidemiology 64 (2011) 395–400

P. pazienti adulti sottoposti a resezione chirurgica per carcinoma polmonare non a piccole cellule (NSCLC) e presenza di mutazioni attivanti il recettore per il fattore di crescita epidermico (EGFR).

I. Osimertinib

C. non trattamento con inibitori di EGFR (in presenza o meno di trattamento CT adiuvante)

O.

DFS in overall population	(critical)	Mean change from baseline in PCS T-score of SF-36	(important)	Drug-related AEs of CTCAE Grade ≥3	(critical)
DFS in stage IB patients	(important)	Time to deterioration in PCS T-score of SF-36	(important)	Drug-related AEs leading to discontinuation of treatment	(critical)
DFS in stage II-IIIA patients	(important)	Mean change from baseline in MCS T-score of SF-36	(important)	Drug-related Cardiac Failure / Cardiomyopathy	(important)
DFS in patients who received adjuvant CT	(important)	Time to deterioration in MCS T-score of SF-36	(important)	Drug-related Interstitial Lung Disease / Pneumonia	(important)
DFS in patients who did not receive adjuvant CT	(important)	OS in overall population	(critical)	Drug-related Diarrhea of CTCAE Grade ≥2	(important)
CNS recurrence-free survival in overall population	(critical)	OS in stage II-IIIA patients	(important)	Drug-related Paronychia of CTCAE Grade ≥2	(important)
CNS recurrence-free survival in stage II-IIIA patients	(important)				

P. pazienti con neoplasia a cellule renali in fase metastatica (mRCC), a rischio intermedio-alto

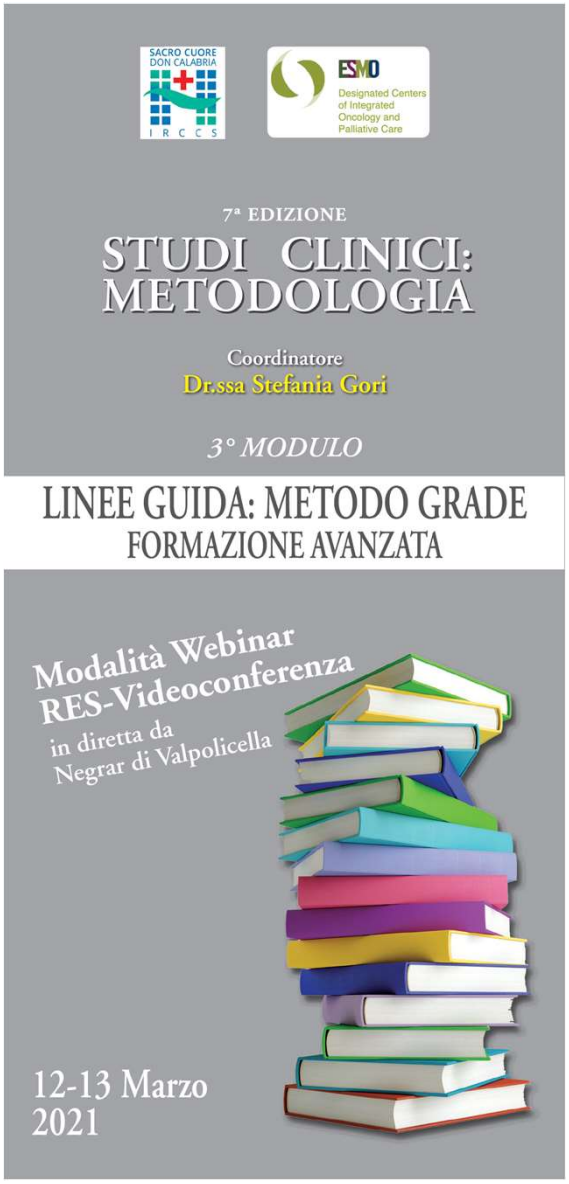
I. Cabozantinib 60mg/die

C. Sunitinib 50mg/die (4 settimane on / 2 settimane off)

O. outcome “essenziali” o “importanti” per la formulazione della proposta terapeutica:

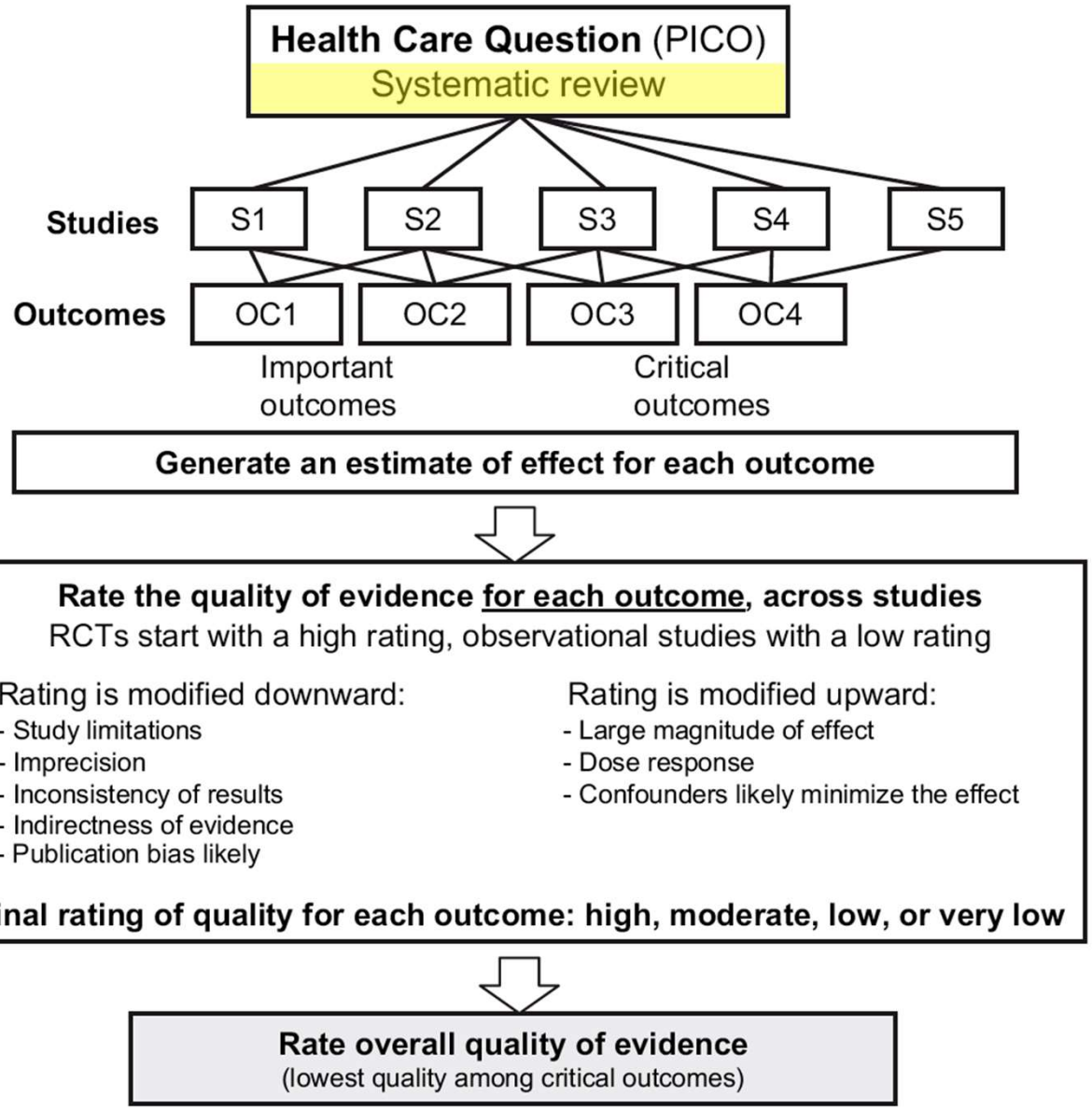
a) per quanto riguarda la definizione dei benefici sono stati presi in considerazione:
   - Progression Free Survival, Independent Central Review (PFS-IRC) (outcome essenziale);
   - Progression Free Survival, Investigator Assessment (PFS-INV) (outcome importante);
   - Overall Survival (OS) (outcome essenziale);
   - Objective Response Rate, Investigator Assessment (ORR-INV) (outcome importante);
   - Quality-Adjusted Time Without Symptoms (TWIST) (outcome essenziale);

b) per quanto riguarda la definizione dei danni, sono stati presi in considerazione:
   - qualsiasi evento avverso (TEAE) di grado 3-4 sec. CTC-AE v.4 (outcome essenziale);
   - qualsiasi TEAE che abbia condotto alla interruzione della terapia in atto (outcome essenziale);
   - evidenza di ipertensione di ogni grado sec. CTC-AE v.4 (outcome importante);
   - evidenza di eritrodisestesia palmo-plantare (PPE) di ogni grado sec. CTC-AE v.4 (outcome importante);
   - evidenza di diarrea di ogni grado sec. CTC-AE v.4 (outcome importante);

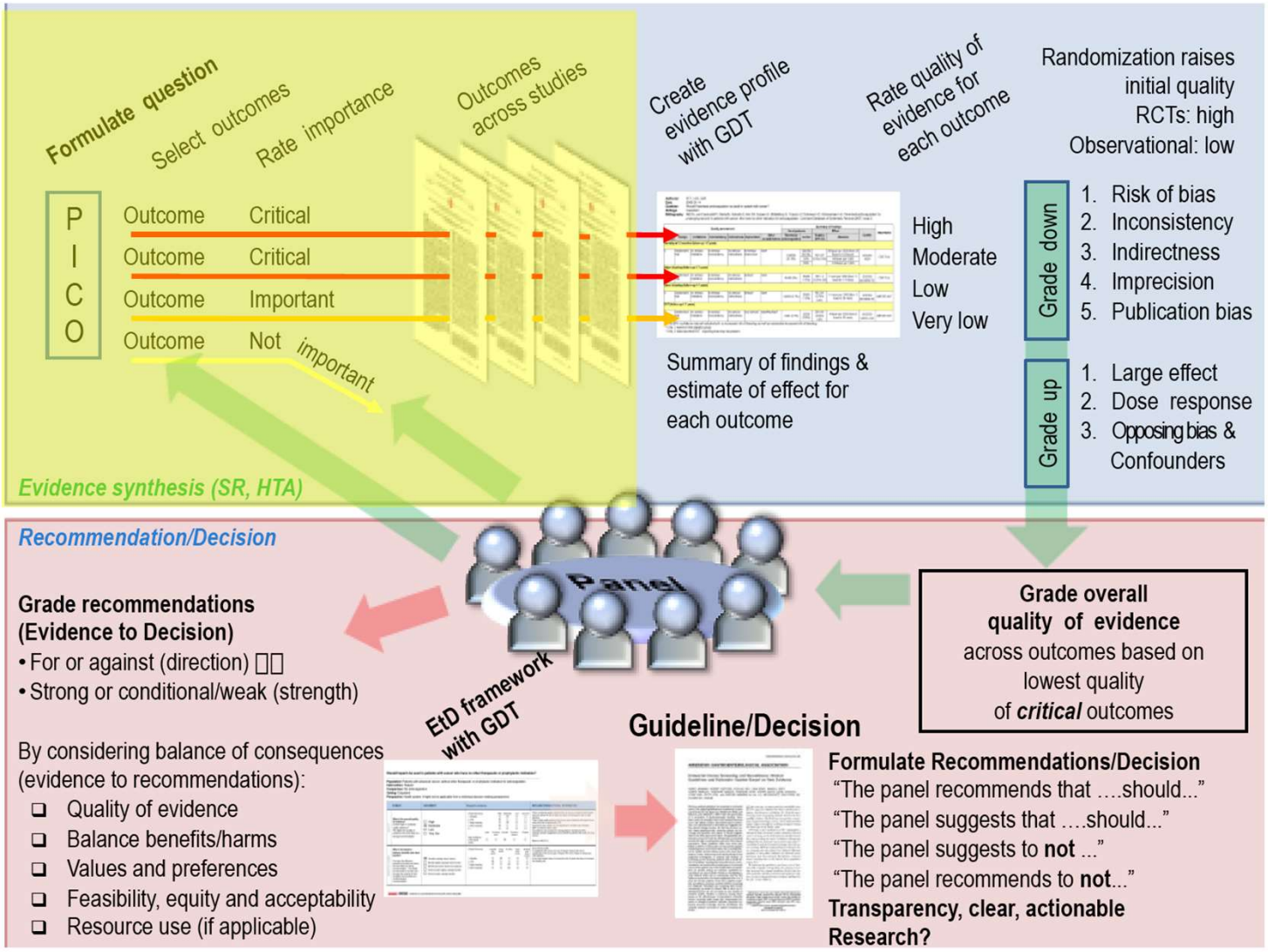


GRADE

Ricerca delle prove di beneficio e danno (sulla base del quesito clinico)


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)

The GRADE process in developing guidelines



Revisione Sistematica

Metodo esplicito e trasparente per identificare, valutare e riassumere i risultati di singoli studi sugli effetti di un intervento sanitario.

Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation

BMJ 2015;349:g7647

Larissa Shamseer[1], David Moher[1], Mike Clarke[2], Davina Ghersi[3], Alessandro Liberati (deceased)[4], Mark Petticrew[5], Paul Shekelle[6], Lesley A Stewart[7], the PRISMA-P Group

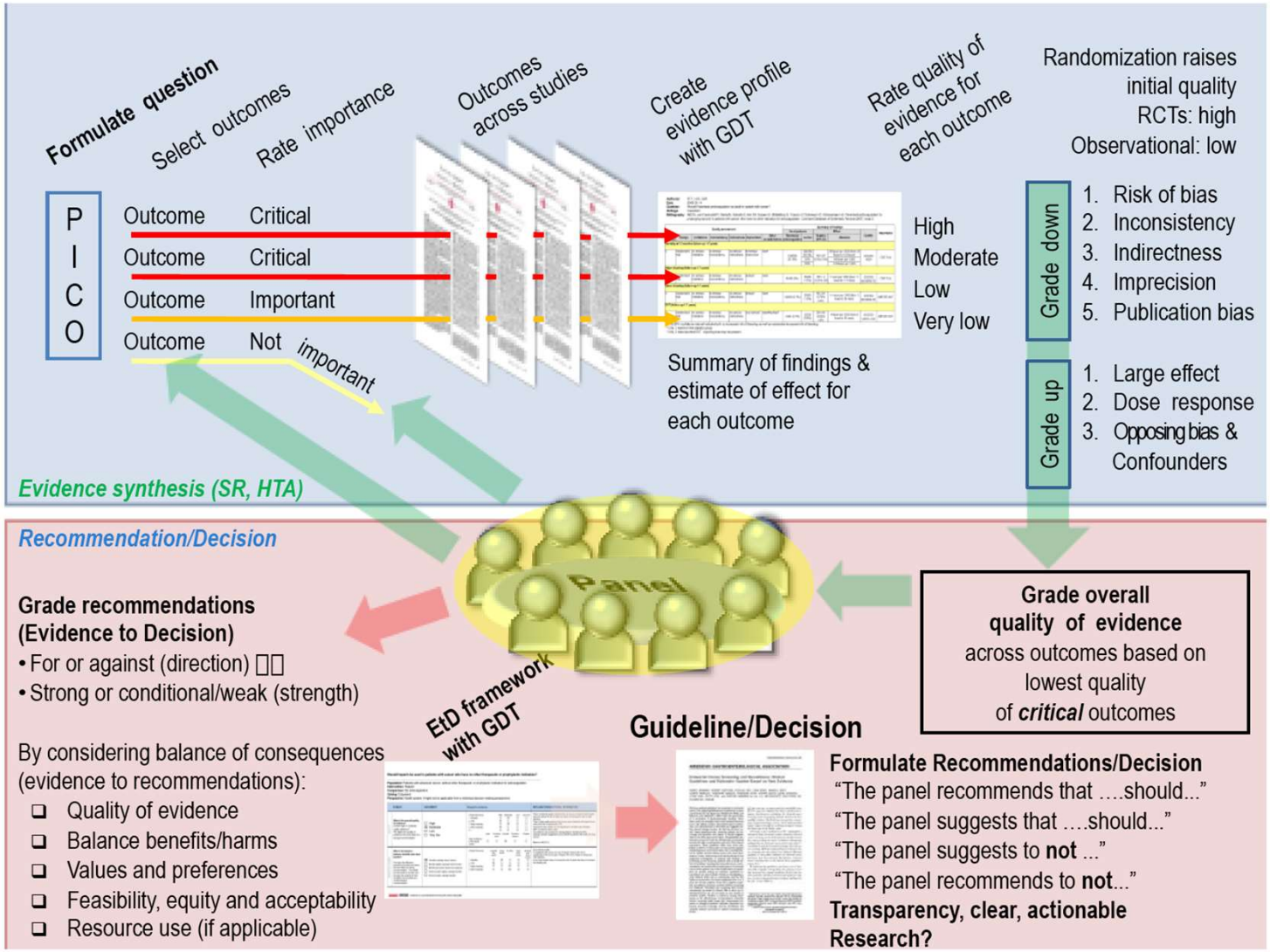
[1]Ottawa Hospital Research Institute and University of Ottawa, Canada; [2]Queen's University Belfast, Ireland; [3]National Health and Medical Research Council, Australia; [4]University of Modena, Italy; [5]London School of Hygiene and Tropical Medicine, UK; [6]Southern California Evidence-based Practice Center, USA; [7]Centre for Reviews and Dissemination, University of York, UK

An international group of experts has created a guideline to improve the transparency, accuracy, completeness, and frequency of documented systematic review and meta-analysis protocols—PRISMA-P (for protocols) 2015.

The PRISMA-P checklist contains 17 items considered to be essential and minimum components of a systematic review or meta-analysis protocol.

This PRISMA-P 2015 Explanation and Elaboration paper provides readers with a full understanding of and evidence about the necessity of each item as well as a model example from an existing published protocol.

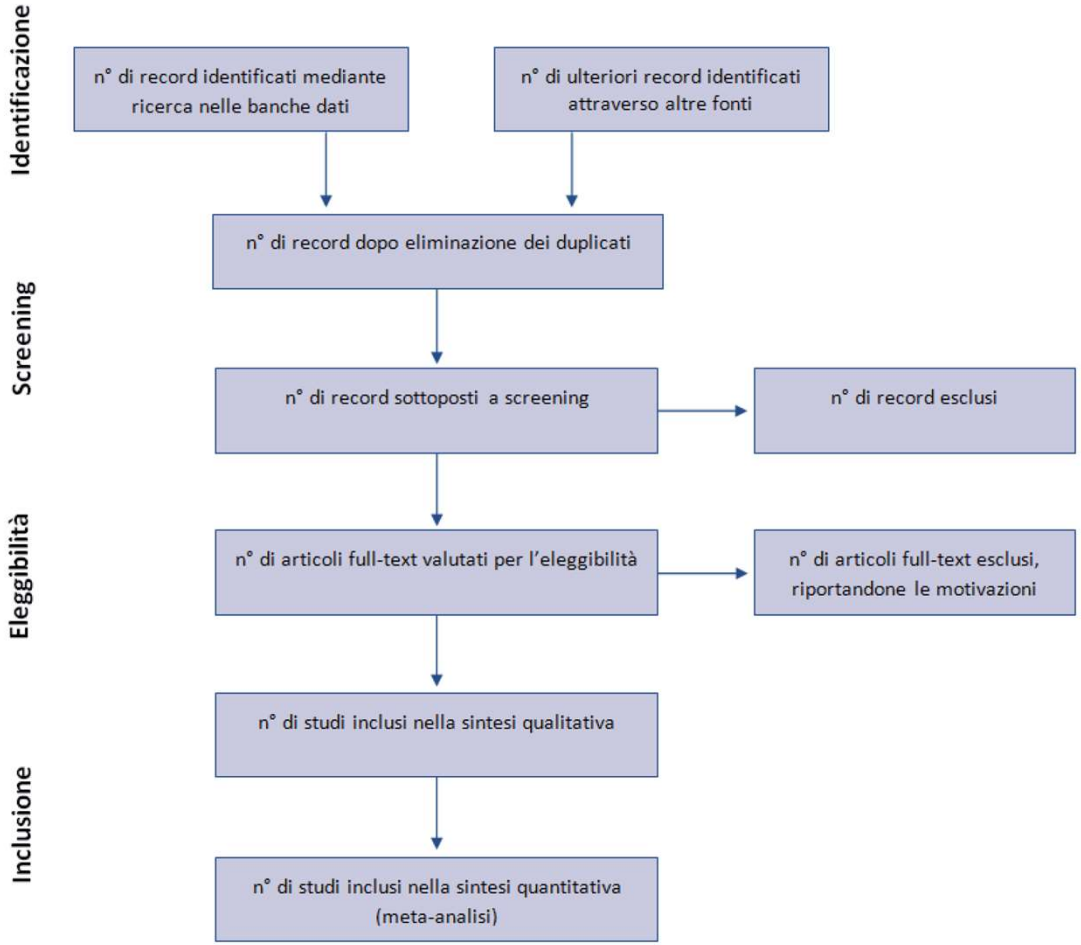
The GRADE process in developing guidelines



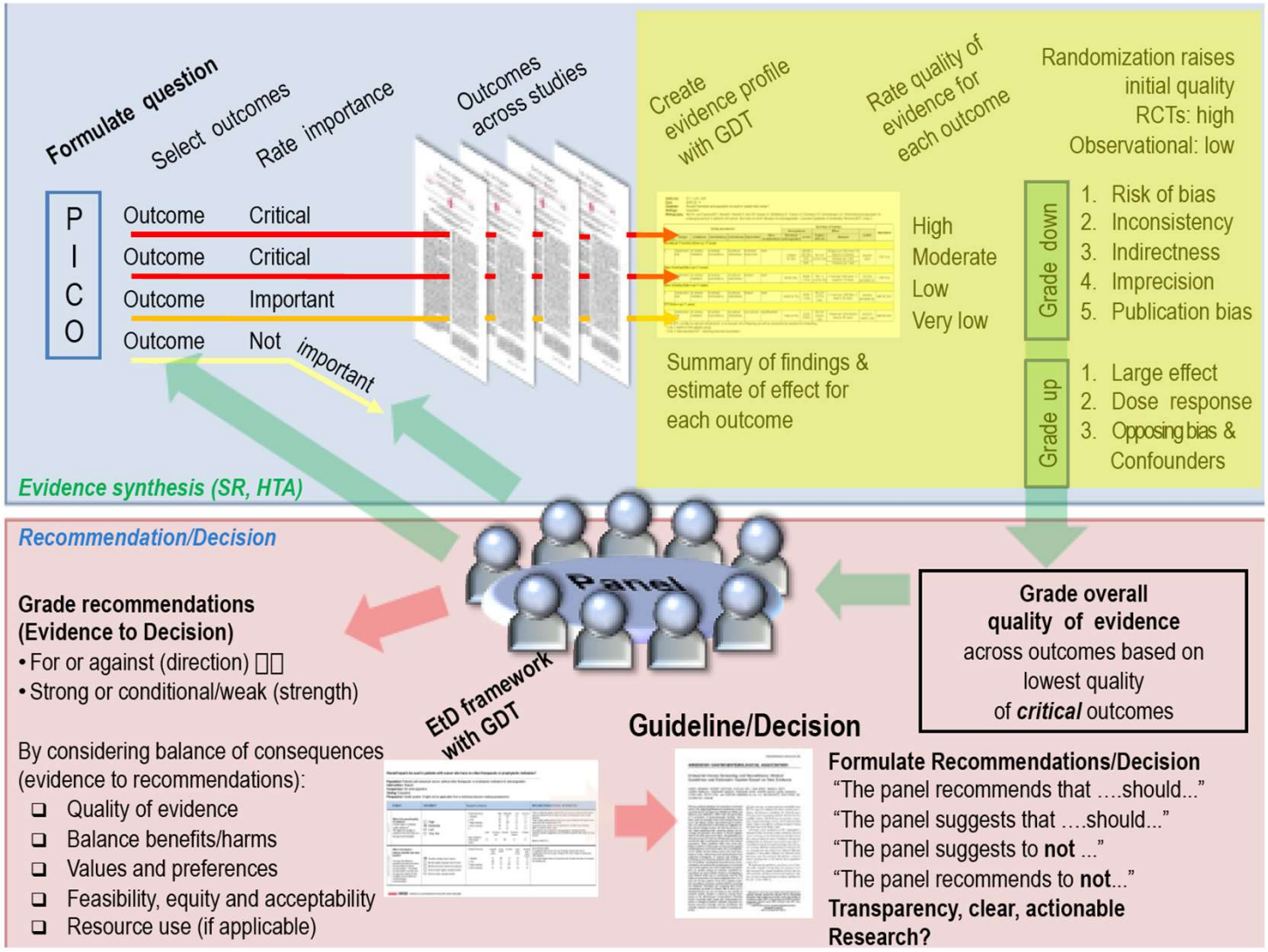
Linee guida per il reporting di revisioni sistematiche e meta-analisi: il PRISMA Statement

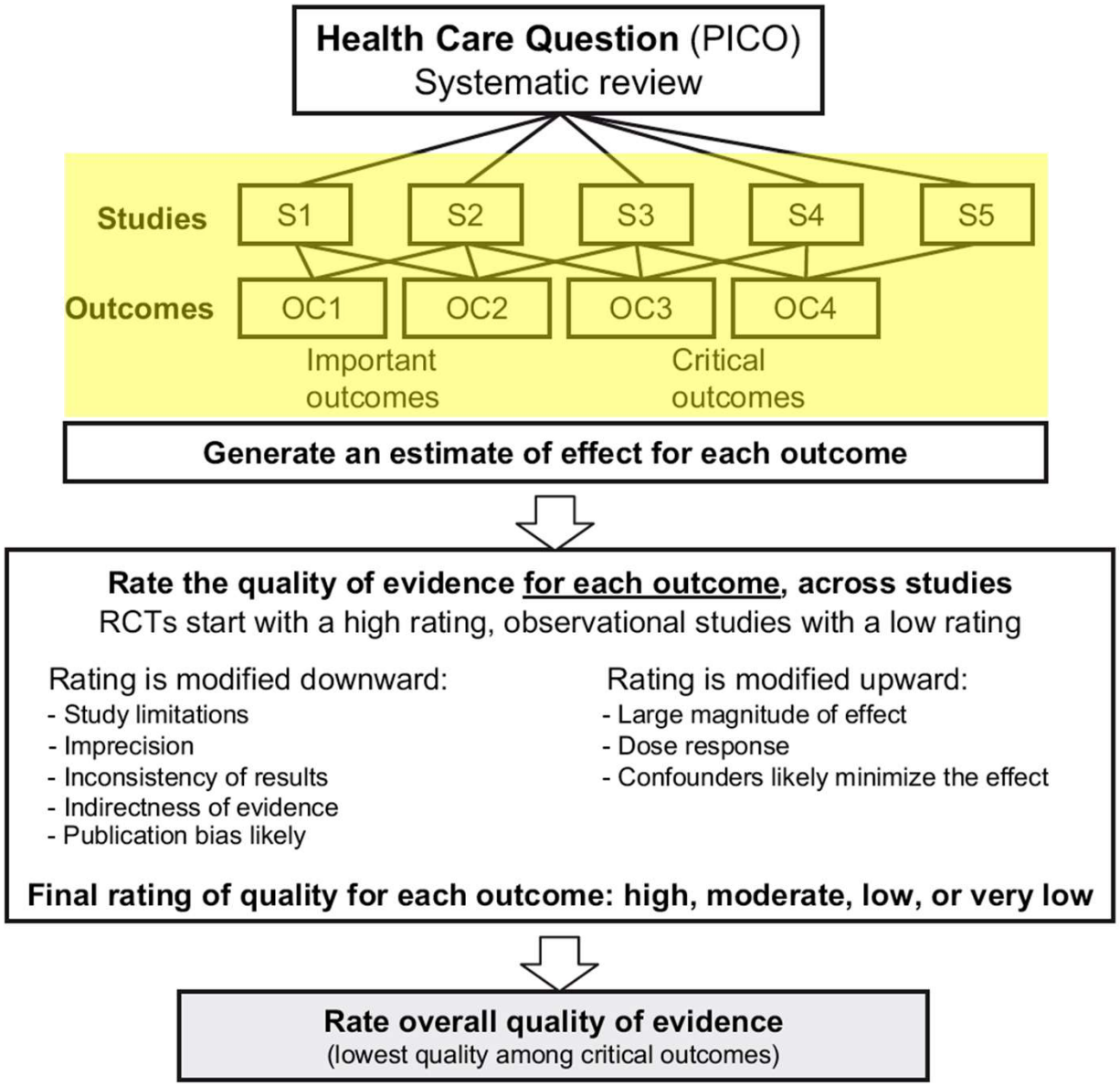
David Moher[1,2*], Alessandro Liberati[3,4], Jennifer Tetzlaff[1], Douglas G. Altman[5], The PRISMA Group[6]

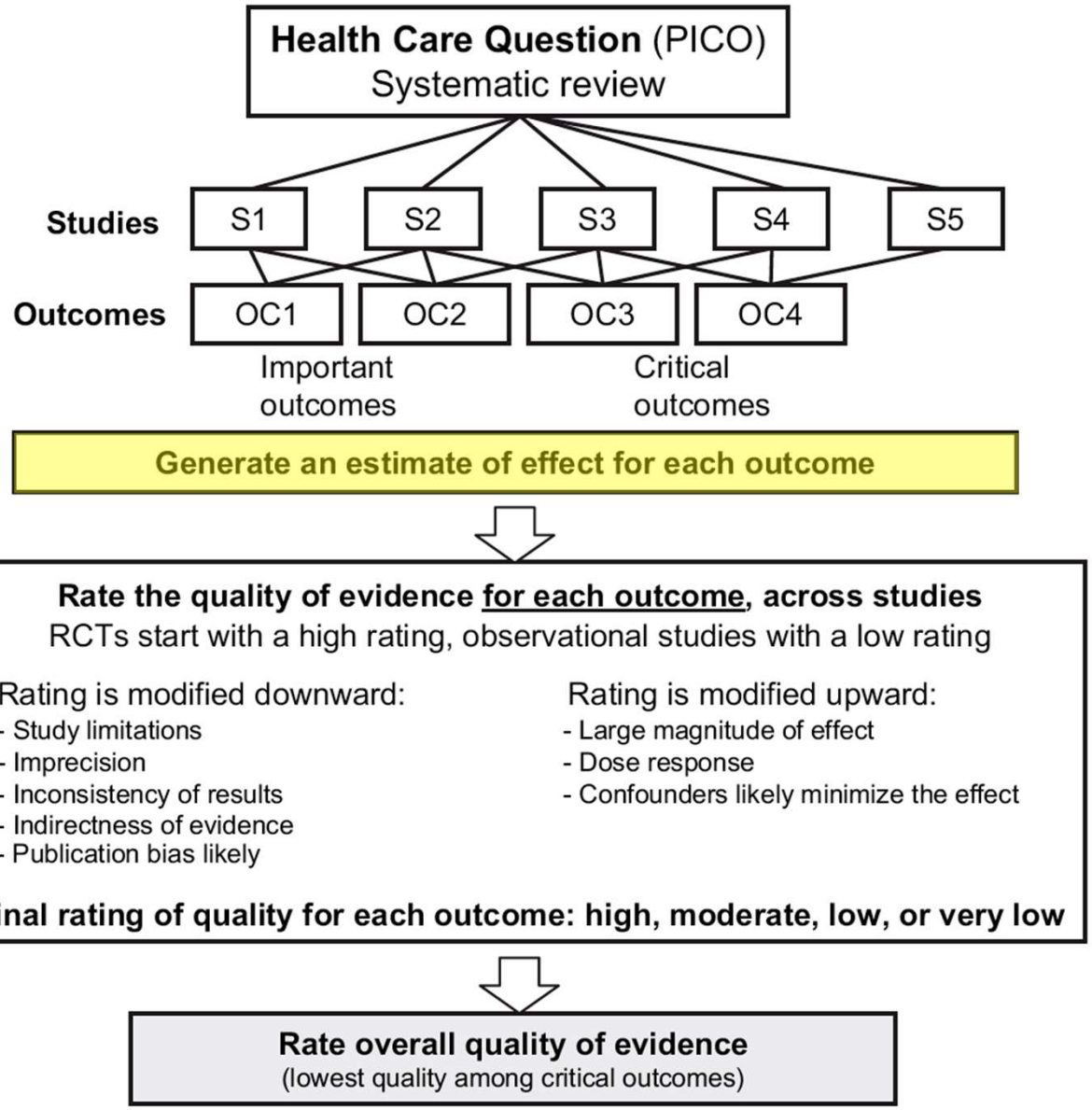
Evidence 2015;7(6): e1000114

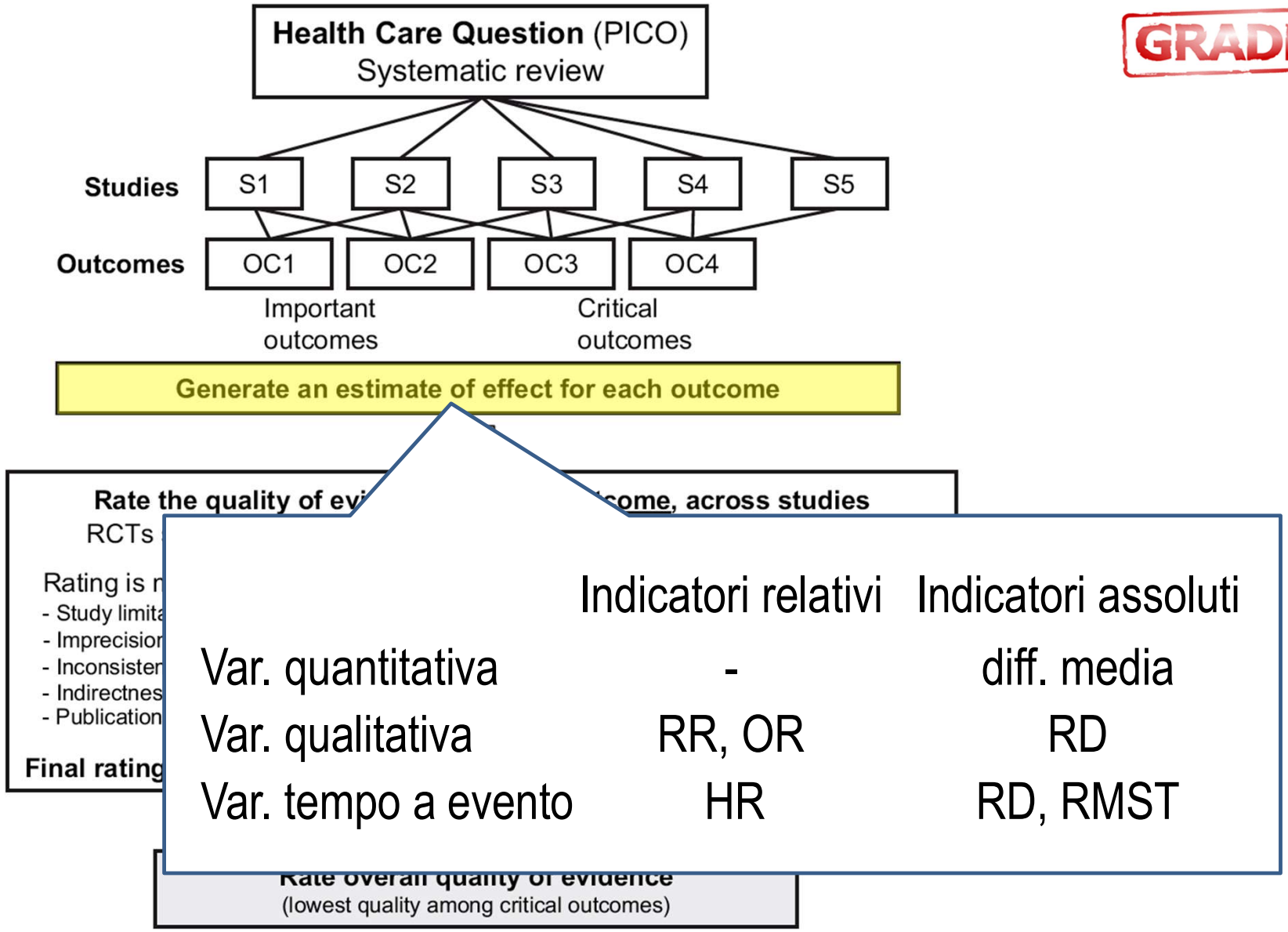


The GRADE process in developing guidelines




Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evi
come, across studies
RCTs
Rating is r
- Study limita
- Imprecisior
- Inconsister
- Indirectnes
- Publication
Final rating
Rate overall quality of evidence
(lowest quality among critical outcomes)
Indicatori relativi
Indicatori assoluti
Var. quantitativa
-
diff. media
Var. qualitativa
RR, OR
RD
Var. tempo a evento
HR
RD, RMST

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

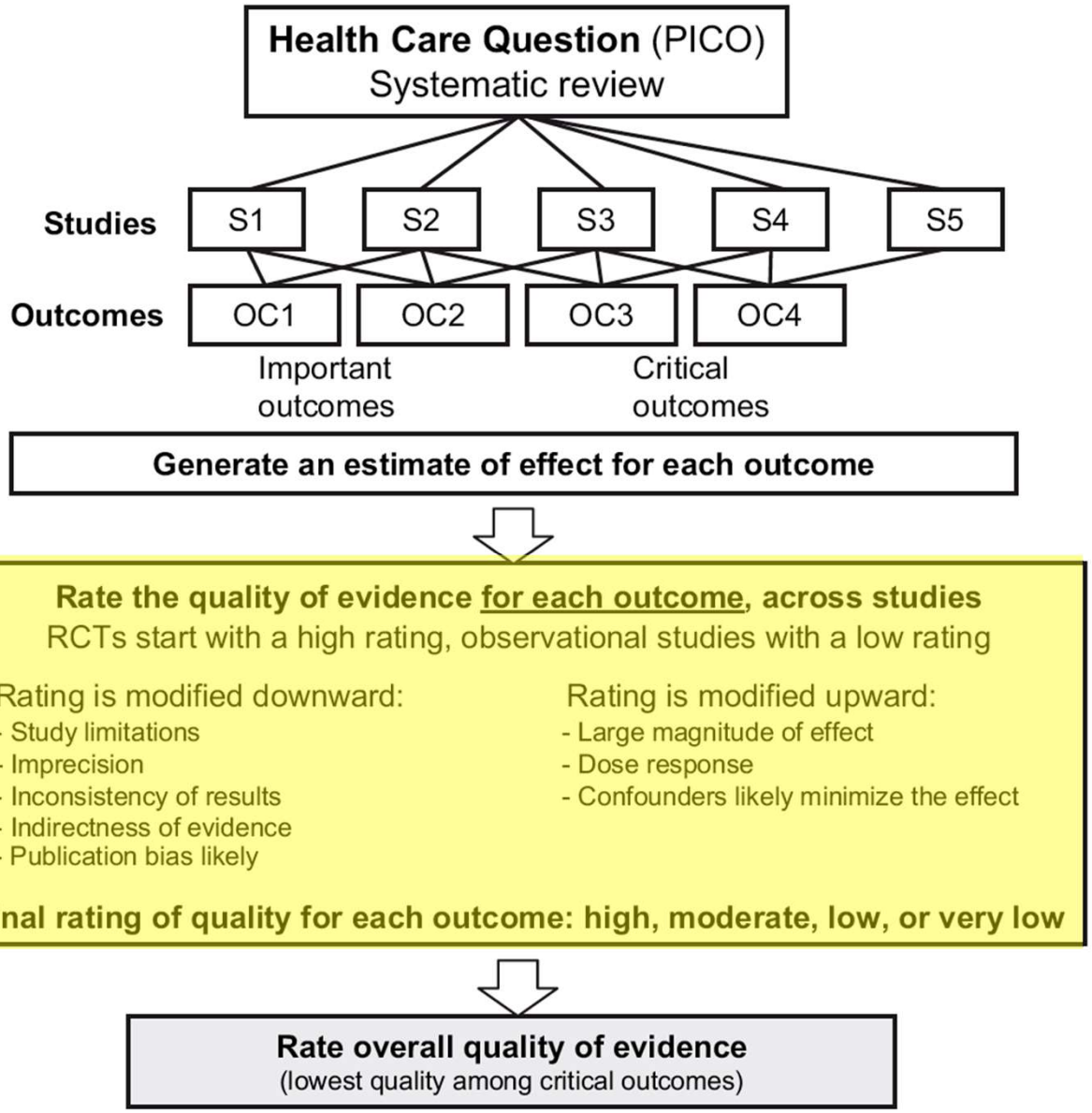
LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Valutazione della qualità delle prove: *generalità*


Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)

GRADE Quality of Evidence

In the context of making recommendations:

- The quality of evidence reflects the extent of our confidence that the estimates of an effect are adequate to support a particular decision or recommendation.

Quality of the body of evidence

Four levels

⊕⊕⊕⊕ High	We are very confident that the true effect lies close to that of the estimate of the effect
⊕⊕⊕◯ Moderate	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	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
⊕◯◯◯ Very low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Tabella 5- Criteri* per l'aumento (*upgrading*) o la diminuzione (*downgrading*) del giudizio di qualità (alta, moderata, bassa, molto bassa) delle prove

Tipo di prove	Studio controllato e randomizzato = alta Studio osservazionale = bassa Qualsiasi altro tipo di informazione = molto basso
A. Diminuzione della categoria di attribuzione (es. da "alta" a "moderata")	1. Limiti gravi (-1 livello) o molto gravi (-2 livelli) nella qualità di conduzione dello studio 2. Incoerenza nei risultati tra studi diversi sullo stesso quesito (-1 o -2 livelli) 3. Alcune (-1 livello) o importanti (-2 livelli) incertezze circa la diretta trasferibilità dei risultati (directness) 4. Imprecisione o dati insufficienti (*sparse data*) (-1 o -2 livelli) 5. Possibilità di pubblicazione selettiva dei dati (*publication* e *reporting bias*) (-1 o-2 livelli)
B. Aumento della categoria di attribuzione (es. da "bassa" a "moderata")	1. Associazione intervento-outcome forte, ovvero con rischio relativo >2 (<0,5), sulla base di prove concordanti provenienti da due o più studi osservazionali, senza alcun fattore di confondimento plausibile (+1 livello) 2. Associazione intervento-outcome molto forte, ovvero con rischio relativo >5 (<0,2) (+2 livelli) 3. Presenza di un gradiente dose-risposta (+1 livello) 4. Aver considerato tutti i possibili fattori di confondimento che avrebbero potuto alterare le stime di effetto (+1 livello)

* non sono da considerarsi un algoritmo

Tabella 5- Criteri* per l'aumento (*upgrading*) o la diminuzione (*downgrading*) del giudizio di qualità (alta, moderata, bassa, molto bassa) delle prove

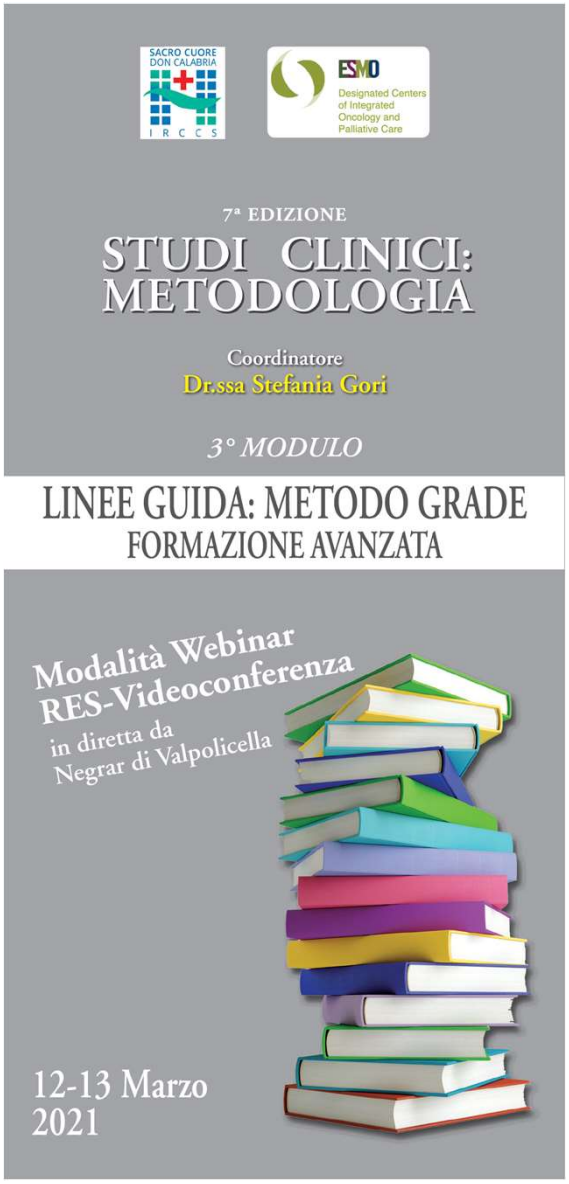
Tipo di prove	Studio controllato e randomizzato = alta Studio osservazionale = bassa Qualsiasi altro tipo di informazione = molto basso
A. Diminuzione della categoria di attribuzione (es. da "alta" a "moderata")	1. Limiti gravi (-1 livello) o molto gravi (-2 livelli) nella qualità di conduzione dello studio 2. Incoerenza nei risultati tra studi diversi sullo stesso quesito (-1 o -2 livelli) 3. Alcune (-1 livello) o importanti (-2 livelli) incertezze circa la diretta trasferibilità dei risultati (directness) 4. Imprecisione o dati insufficienti (*sparse data*) (-1 o -2 livelli) 5. Possibilità di pubblicazione selettiva dei dati (*publication* e *reporting bias*) (-1 o-2 livelli)
B. Aumento della categoria di attribuzione (es. da "bassa" a "moderata")	1. Associazione intervento-outcome forte, ovvero con rischio relativo >2 (<0,5), sulla base di prove concordanti provenienti da due o più studi osservazionali, senza alcun fattore di confondimento plausibile (+1 livello) 2. Associazione intervento-outcome molto forte, ovvero con rischio relativo >5 (<0,2) (+2 livelli) 3. Presenza di un gradiente dose-risposta (+1 livello) 4. Aver considerato tutti i possibili fattori di confondimento che avrebbero potuto alterare le stime di effetto (+1 livello)

* non sono da considerarsi un algoritmo

Tabella 5- Criteri* per l'aumento (*upgrading*) o la diminuzione (*downgrading*) del giudizio di qualità (alta, moderata, bassa, molto bassa) delle prove

Tipo di prove	**Studio controllato e randomizzato = alta** **Studio osservazionale = bassa** **Qualsiasi altro tipo di informazione = molto basso**
A. Diminuzione della categoria di attribuzione (es. da "alta" a "moderata")	1. Limiti gravi (-1 livello) o molto gravi (-2 livelli) nella qualità di conduzione dello studio 2. Incoerenza nei risultati tra studi diversi sullo stesso quesito (-1 o -2 livelli) 3. Alcune (-1 livello) o importanti (-2 livelli) incertezze circa la diretta trasferibilità dei risultati (directness) 4. Imprecisione o dati insufficienti (*sparse data*) (-1 o -2 livelli) 5. Possibilità di pubblicazione selettiva dei dati (*publication* e *reporting bias*) (-1 o-2 livelli)
B. Aumento della categoria di attribuzione (es. da "bassa" a "moderata")	1. Associazione intervento-outcome forte, ovvero con rischio relativo >2 (<0,5), sulla base di prove concordanti provenienti da due o più studi osservazionali, senza alcun fattore di confondimento plausibile (+1 livello) 2. Associazione intervento-outcome molto forte, ovvero con rischio relativo >5 (<0,2) (+2 livelli) 3. Presenza di un gradiente dose-risposta (+1 livello) 4. Aver considerato tutti i possibili fattori di confondimento che avrebbero potuto alterare le stime di effetto (+1 livello)

* non sono da considerarsi un algoritmo



GRADE

Valutazione della qualità delle prove: *risk of bias*

- **RCTs** ⊕⊕⊕⊕
- **observational studies** ⊕⊕◯◯
- **5 factors that can lower quality**
  1. limitations in study design, execution and reporting *(risk of bias criteria)*
  2. Inconsistency *(or heterogeneity)*
  3. Indirectness *(PICO and applicability)*
  4. Imprecision
  5. Publication bias
- **3 factors can increase quality**
  1. large magnitude of effect
  2. opposing plausible residual bias or confounding
  3. dose-response gradient

ELSEVIER

Journal of Clinical Epidemiology 64 (2011) 407–415

Journal of Clinical Epidemiology

GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias)

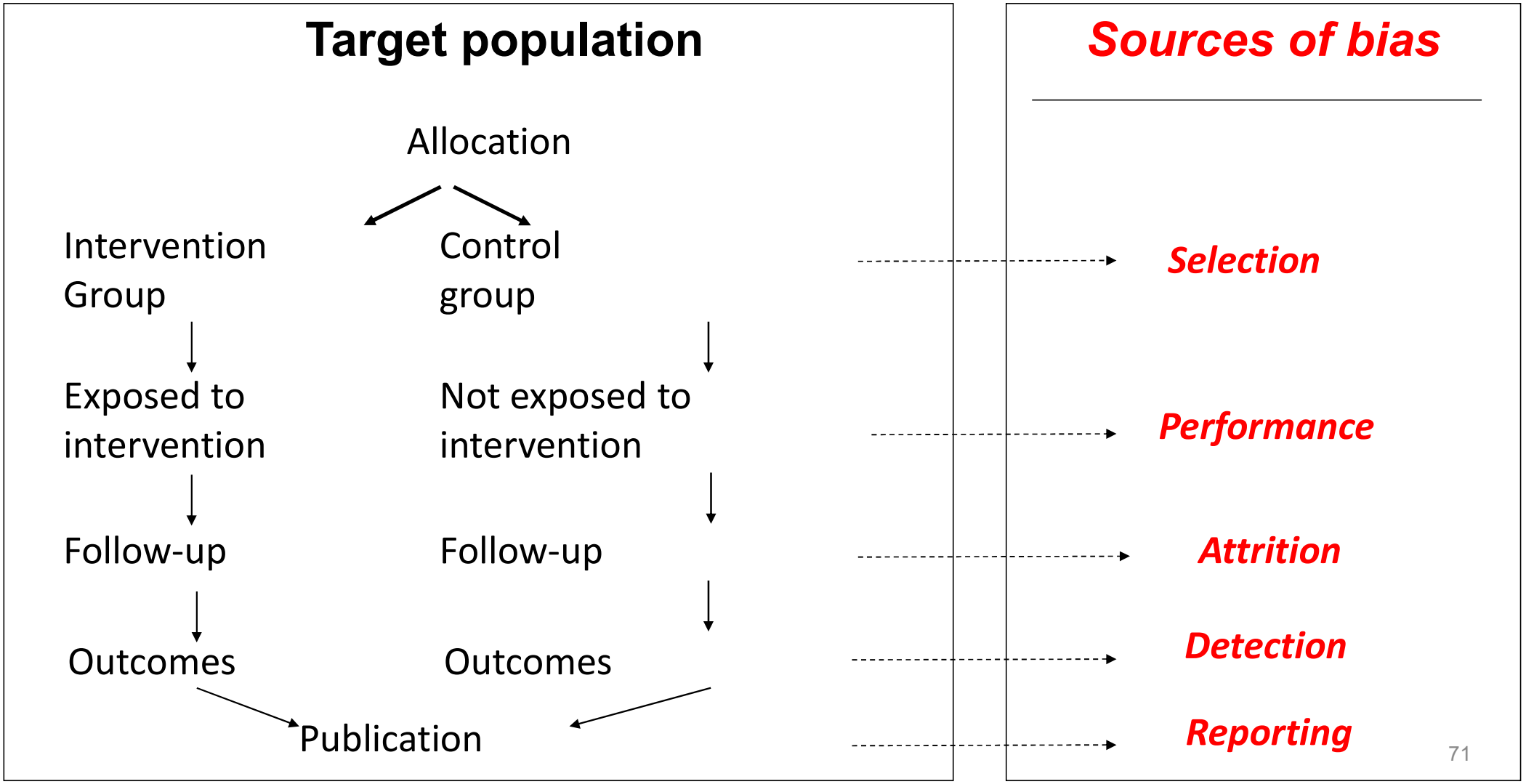
Gordon H. Guyatt[a,*], Andrew D. Oxman[b], Gunn Vist[b], Regina Kunz[c], Jan Brozek[a], Pablo Alonso-Coello[d], Victor Montori[e], Elie A. Akl[f], Ben Djulbegovic[g,h,i], Yngve Falck-Ytter[j], Susan L. Norris[k], John W. Williams Jr.[l], David Atkins[m], Joerg Meerpohl[n,o], Holger J. Schünemann[a]

Bias

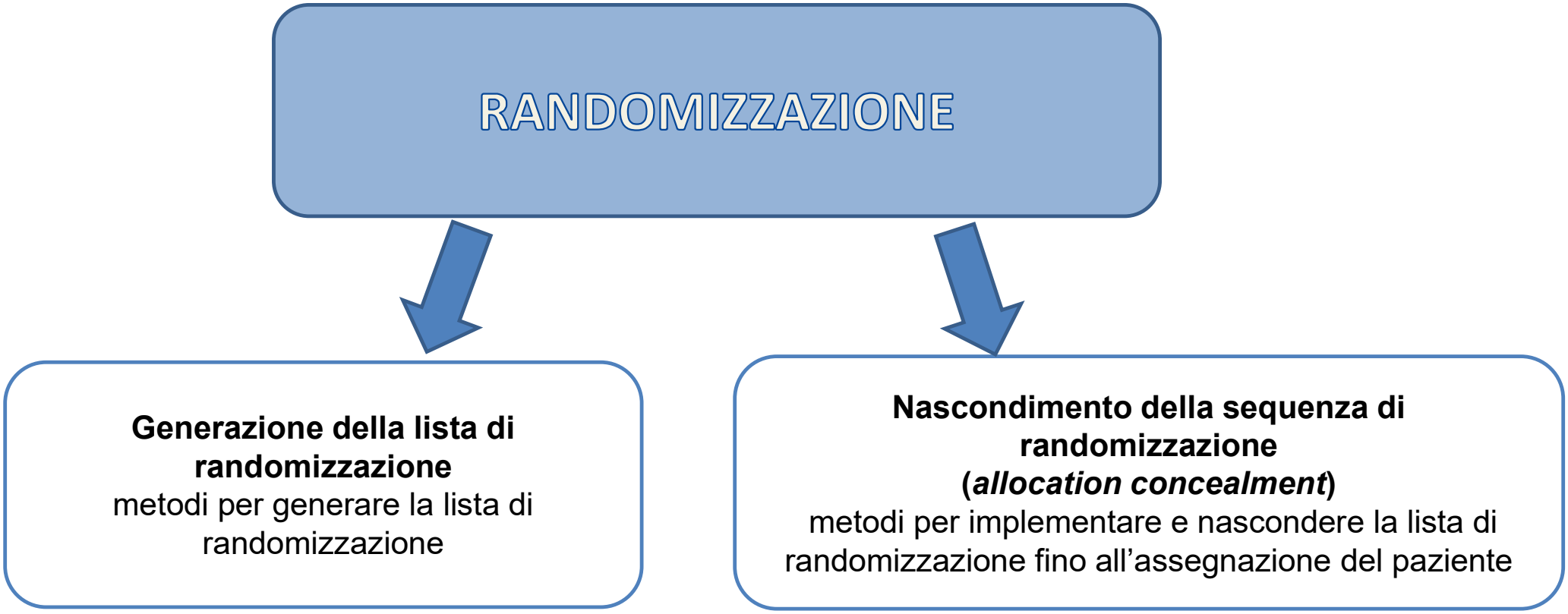
Systematic distortion of the estimated intervention effect away from the truth, caused by **inadequacies** in the **design**, **conduct**, or **analysis** of a trial , or in the **publication of its results**. In other words, in a biased trial, the results observed reflect other factors in addition to (or, in extreme cases, instead of) the effect of the tested therapeutic procedure alone.

Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;134:663–94

Trial as a flow



Selection bias: due componenti



Generazione lista di randomizzazione

- **Basso rischio di bias**. Uso di metodi realmente casuali come ad esempio: tavole di numeri random, sistemi computerizzati, lancio di una moneta o di un dado, sorteggio.
- **Alto rischio di bias.** Uso di metodi **NON** realmente casuali come ad esempio: giorno di nascita o di ammissione in ospedale, giudizio del medico, preferenze del paziente, risultati di test di laboratorio, disponibilità del trattamento, alternanza
- **Rischio incerto**. Quando non sono disponibili sufficienti informazioni per definire il rischio (bassa qualità del reporting)

Nascondimento della sequenza di randomizzazione*

- **Basso rischio di bias**. Sperimentatori che arruolano i pazienti non possono prevedere in quale gruppo verrà inserito il paziente perché si usa uno dei seguenti metodi: randomizzazione centralizzata (telefonica, via web, o gestita da personale esterno alla sperimentazione - farmacista, statistico); buste chiuse e opache.
- **Alto rischio di bias.** Sperimentatori che arruolano i pazienti possono prevedere in quale gruppo verrà inserito il paziente perché si usa uno dei seguenti metodi: liste di randomizzazione, buste aperte o non opache, alternanza, data di nascita, numero di cartella, ect.
- **Rischio incerto**. Non sono disponibili sufficienti informazioni per definire il rischio (bassa qualità del reporting)

*allocation concealment

Esempio 1

To prevent the introduction of bias, randomisation will be via an online system, accessed via http://www.ctu.co.uk, hosted by the Clinical Trials Unit (CTU), King's College London (KCL). Eligible participants will be randomised to either the nilvadipine or placebo treatment group. The

***Generazione della sequenza**: basso rischio*

***Allocation concealment:** basso rischio*

Esempio 2

Eligible patients with even-number hospital records were assigned to treatment (5000 U twice daily), those with odd-number records served as controls

***Generazione della sequenza**: alto rischio*

***Allocation concealment:** alto rischio*

Esempio 3

Subjects were assigned at random in a 2:1 ratio to naltrexone or control (Cornish 1977)

***Sequence generation**: rischio incerto*

***Allocation concealment:** rischio incerto*

Esempio 4

Patients were randomly assigned in a 1:1:1 ratio to one of three treatment groups: ledipasvir–sofosbuvir for 8 weeks, ledipasvir–sofosbuvir plus ribavirin for 8 weeks, or ledipasvir–sofosbuvir for 12 weeks. Randomization was stratified according to HCV genotype (1a or 1b).

***Generazione della sequenza**: rischio incerto*

***Allocation concealment:** rischio incerto*

Blinding

- Sperimentatori e partecipanti non conoscono gruppo di allocazione (*performance bias*)
- Valutatori degli esiti non conoscono gruppo di allocazione (*detection bias*)

Singolo cieco
i pazienti inclusi nello studio non conoscono il gruppo al quale sono stati assegnati

Doppio cieco
i pazienti e gli sperimentatori non conoscono il gruppo al quale (i pazienti) sono stati assegnati

Triplo cieco
i pazienti, gli sperimentatori e i valutatori degli esiti non conoscono il gruppo di allocazione

…

Non sempre il significato è questo … è sempre bene valutare chi è davvero in cieco!

Performance bias

Si verifica quando i partecipanti allo studio (sperimentatori o pazienti) modificano i loro comportamenti perché sanno a quale gruppo è assegnato un dato paziente

Esempi:

Lo sperimentatore controlla la presenza di effetti avversi più frequentemente nei pazienti assegnati al gruppo di trattamento.

Un paziente nel gruppo placebo assume altri farmaci, fa più (o meno) visite di controllo.

Detection bias

Si verifica quando la valutazione degli esiti dello studio viene influenzata dalla conoscenza del gruppo al quale è assegnato un dato paziente

Esempi:

Interpretazione di esiti radiologici, risoluzione dei sintomi, valutazione delle ricadute di malattia diversa nei pazienti assegnati al trattamento e al controllo

Performance bias

Cecità di pazienti e sperimentatori

- **Basso rischio di bias.** Pazienti e sperimentatori non conoscono l'assegnazione dei pazienti al gruppo di controllo o di trattamento oppure è poco probabile che la mancanza di cecità influenzi la performance di pazienti e sperimentatori
- **Alto rischio di bias.** Pazienti e sperimentatori conoscono l'assegnazione dei pazienti o, durante lo studio, diventa chiaro a quale gruppo di trattamento sono allocati (rottura del cieco). Studi definiti come "open label"
- **Rischio incerto**. Quando non sono disponibili sufficienti informazioni per definire il rischio (bassa qualità del reporting)

Detection bias

Cecità del valutatore degli esiti dello studio (outcome)

- **Basso rischio di bias.** L'esito dello studio è valutato senza conoscere l'assegnazione dei pazienti al gruppo di controllo o di intervento; oppure è poco probabile che la mancanza di cecità influenzi la valutazione
- **Alto rischio di bias.** L'esito dello studio è valutato conoscendo l'assegnazione dei pazienti al gruppo di controllo o di intervento ed è probabile che la mancanza di cecità influenzi la valutazione
- **Rischio incerto**. Quando non sono disponibili sufficienti informazioni per definire il rischio (bassa qualità del reporting)

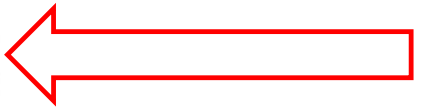
Performance and detection bias

- Impatto diverso su outcome **soggettivi** e **oggettivi** (quindi la valutazione va fatta separatamente)
- Se studio su **farmaco in doppio cieco** e dice che tutti gli operatori erano all'oscuro dell'assegnazione è probabile che sia in cieco anche l'outcome assessor, anche se non espressamente detto
- Se studio su **interventi che non possono essere in doppio cieco** (psicosociali, educativi, chirurgici, riabilitativi) importante che sia in cieco l'outcome assessor e deve essere specificato
  - Performance: high risk per outcomes soggettivi sempre
  - Detection: low risk se c'è blinding of outcome assessor anche per outcomes soggettivi

Esempio 1

Eligible participants will be randomised to either the nilvadipine or placebo treatment group. The nilvadipine capsules and placebo capsules will be packaged and labelled identically. Randomisation will be at the level of the individual patient, using block randomisation with randomly varying block sizes and stratified by country site. Once the patient has been randomised, the online system will automatically recognise which treatment packs are located in each study pharmacy at the recruiting study site and will randomly select a pack in the appropriate trial arm to be dispensed to the patient. All study staff at all sites will be blinded to treatment allocation and will remain blind until the end of the trial.

The primary outcome measure is the change from baseline to week 78 in cognitive function, as assessed by the ADAS-Cog 12.



Performance bias: *basso rischio*

Detection bias: *basso rischio*

Esempio 2

PROCEDURES

After providing written informed consent, all trial participants were randomly assigned to receive one of two doses of dabigatran, or to receive warfarin, by means of a central, interactive, automated telephone system. Dabigatran was administered, in a blinded fashion, in capsules containing either 110 mg or 150 mg of the drug, to be taken twice daily. Warfarin was administered, in an unblinded fashion, in tablets of 1, 3, or 5 mg and was adjusted locally to an international normalized ratio (INR) of 2.0 to 3.0, with the INR measured at least monthly. The time that the INR was with-

.....

group and were centrally blinded. Each primary and secondary outcome event was adjudicated by two independent investigators who were unaware of the treatment assignments.

Performance bias: *alto rischio*

Detection bias: *basso rischio*

Esempio 3

Study medications were prepared by a research pharmacist, who had no direct contact with participants. Buprenorphine mono tablets (containing only buprenorphine) and placebo tablets that appeared identical were provided by the manufacturer. Naltrexone was purchased for the study: tablets were crushed, and the study pharmacist placed naltrexone or placebo inside capsules that appeared identical. To mask slight taste differences between active and placebo buprenorphine tablets, participants gargled with a mentholated antiseptic mouthwash before taking the sublingual tablets (Schottenfeld 2008)

Performance bias: basso rischio

Detection bias: basso rischio

Attrition bias

- Quando non tutti i soggetti randomizzati completano lo studio
- i soggetti non escono a caso dallo studio: è possibile che quelli che escono siano sistematicamente diversi da quelli che non escono: i gruppi non sono più randomizzati
- **Validità esterna** : es: escono tutti i più giovani, o i meno gravi, o i maschi: posso trarre conclusioni solo su quelli che rimangono
- **Validità interna (Bias):** se la probabilità di uscire dallo studio è legata all'intervento o all'outcome, cioè se quelli che escono hanno sistematicamente probabilità più alte o più basse di avere l'outcome di quelli che restano

Attrition bias

Low risk of bias

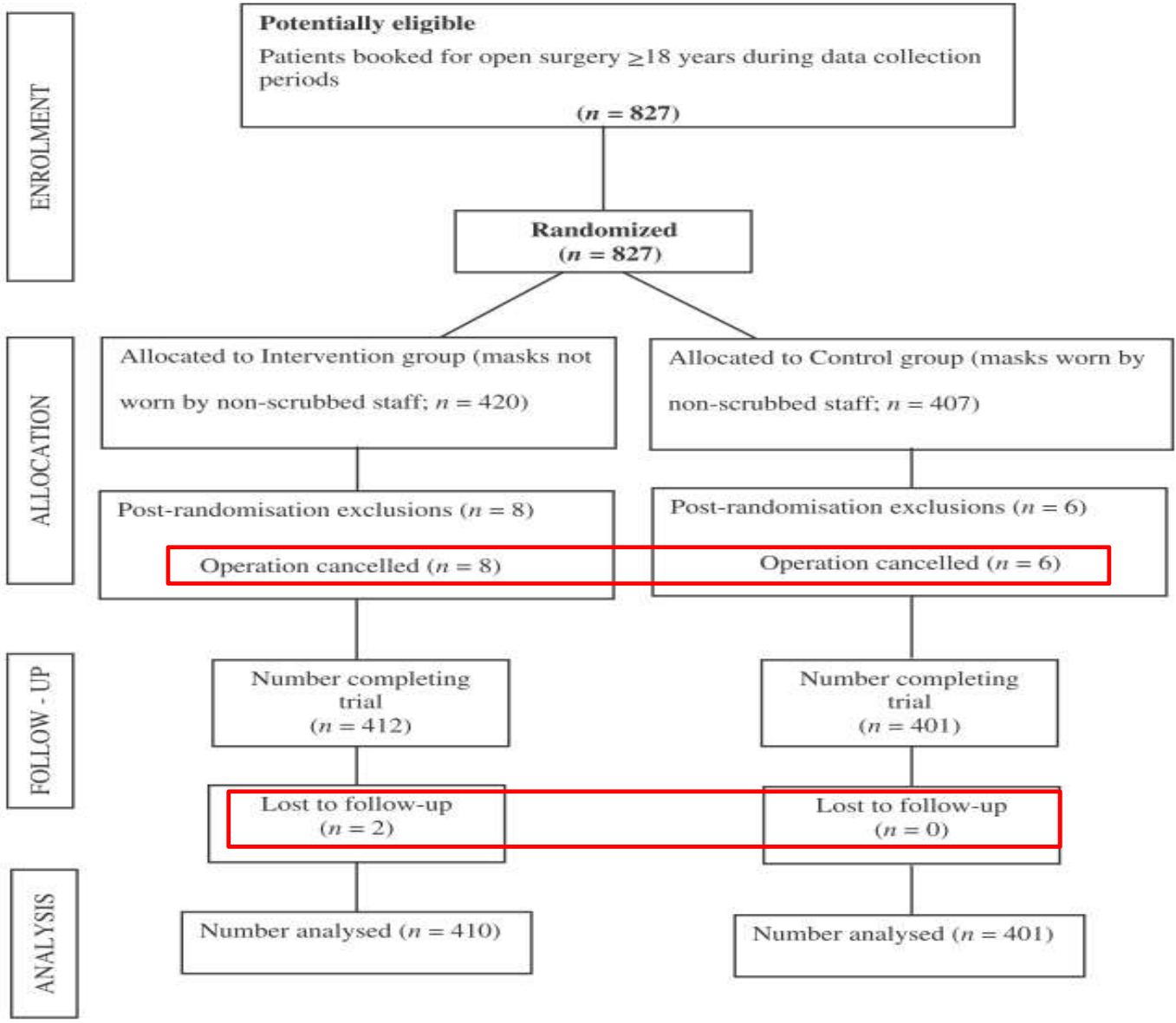
- No missing outcome data;
- the proportion of missing outcomes compared with observed event risk not enough to have a relevant impact on the intervention effect;
- Missing outcome data balanced in numbers across intervention groups, with similar reasons across groups;
- Missing data imputed using appropriate methods
- All patients analysed in the group they were allocated to by randomisation irrespective of non-compliance and co-interventions (intention to treat)

High risk of bias:

- the proportion of missing outcomes compared with observed event risk enough to induce relevant bias in intervention effect estimate
- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups;

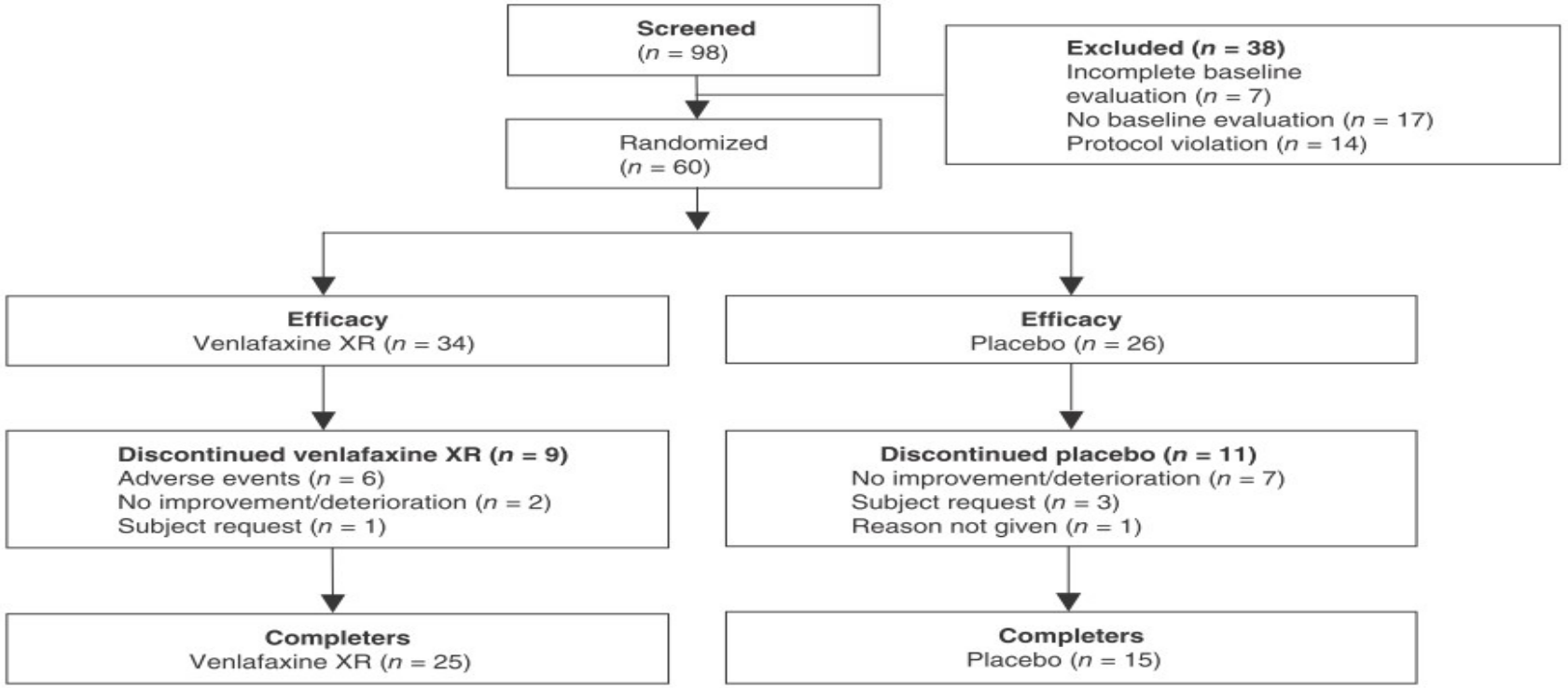
Esempio 1

A tota
enrolled and 811 (98.1%) patients complete
group and 410 No Mask group (Fig. 1).



Attrition bias:
basso rischio

Esempio 2



Totale:
33%
(20 su 60)

Intervento: 26%
(9 su 34)

Placebo: 42%
(11 su 26)

Attrition bias: *alto rischio*

Esempio 3

L’informazione nella sezione risultati

A total of 1395 (91.3%) randomized patients completed the doubleblind treatment phase. There were no patterns or trends suggesting any differences in discontinuation rates or the reasons for discontinuation between the treatment groups (Figure 1). (Bays 2004)

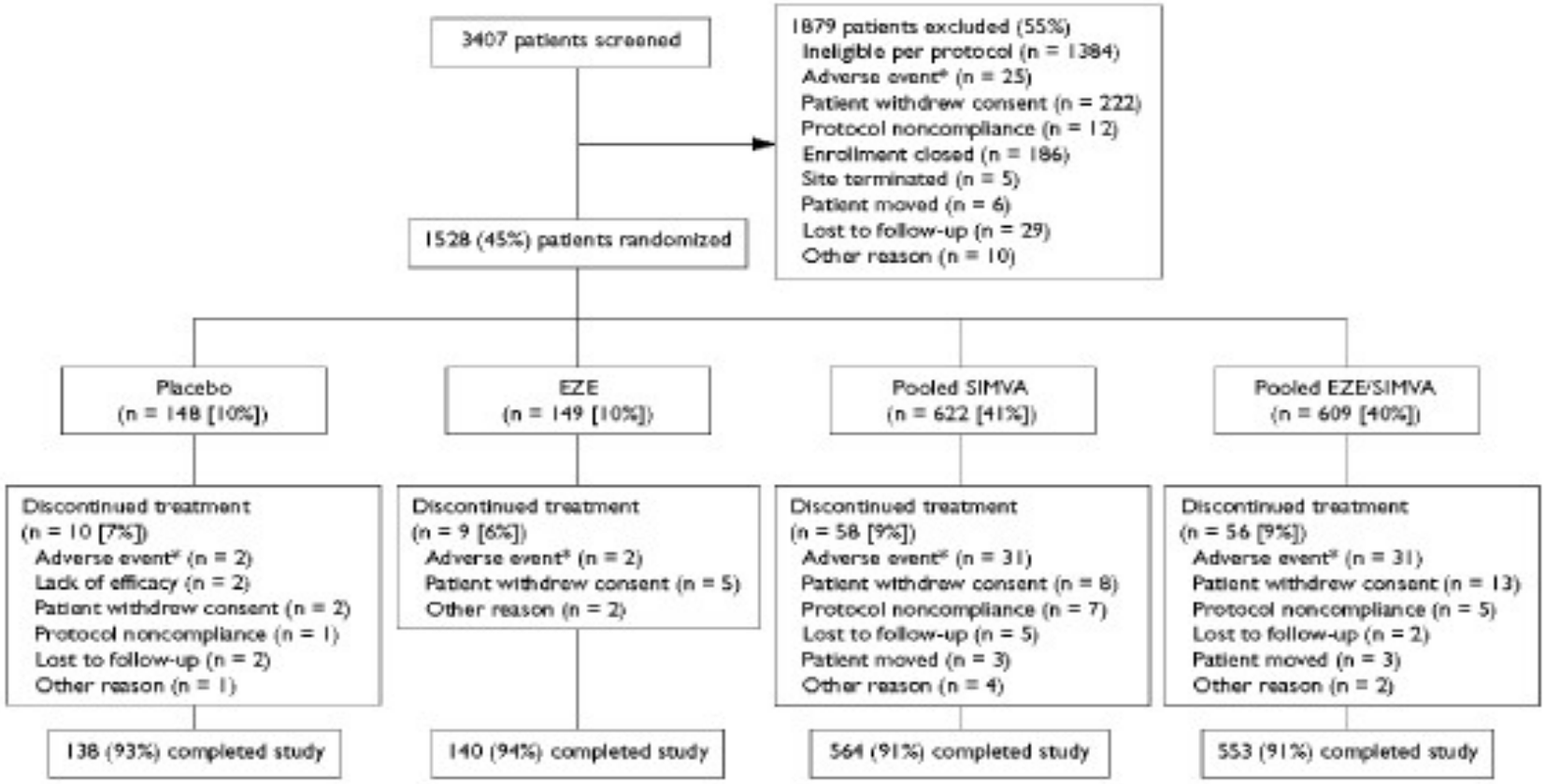
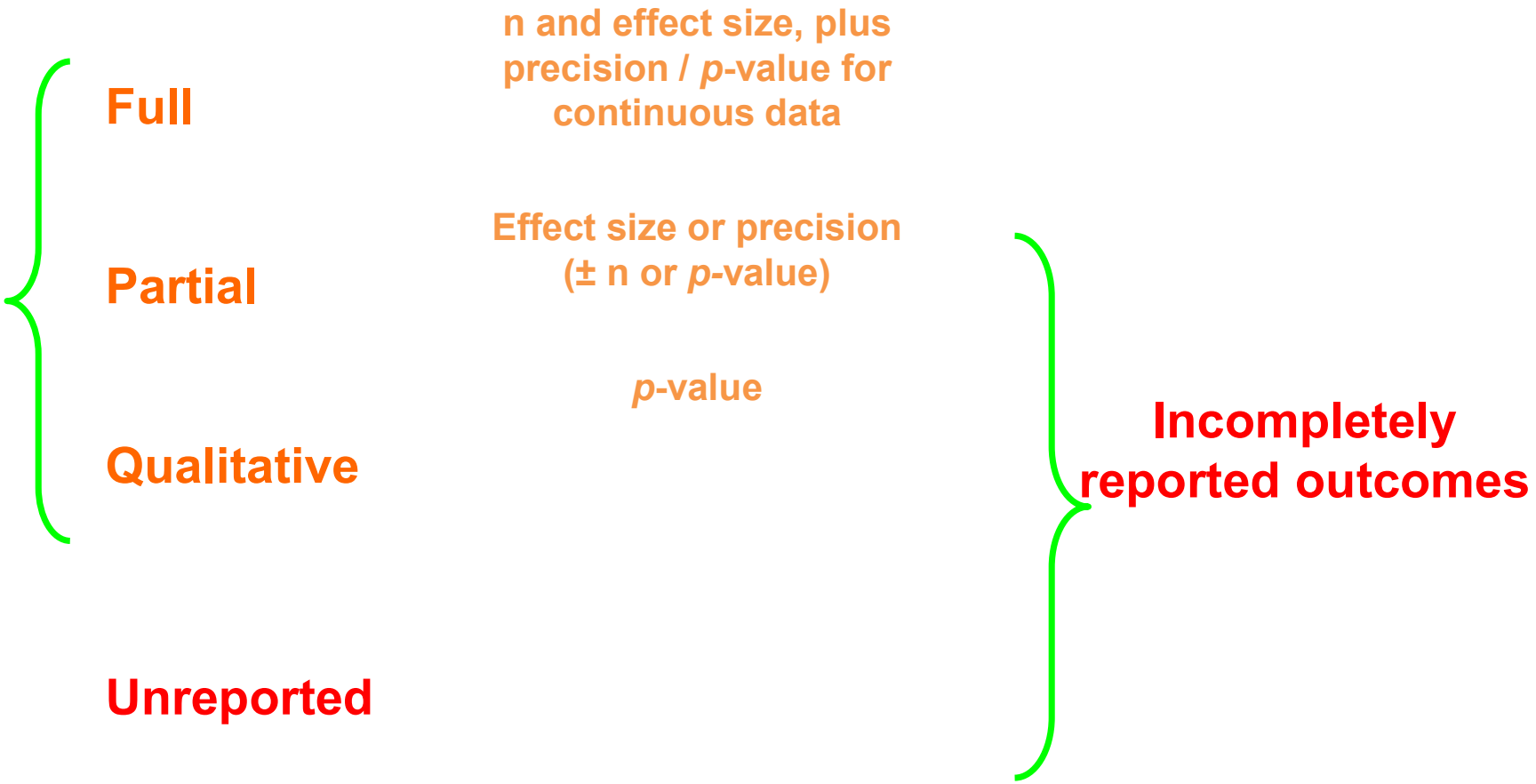


Figure 1. Disposition of patients in the study. Number of patients who were randomized, discontinued prematurely, and completed the study are shown for the placebo, ezetimibe monotherapy (EZE), pooled simvastatin monotherapy (pooled SIMVA), and pooled ezetimibe/simvastatin tablet (pooled EZE/SIMVA) groups. *Number of patients with clinical and laboratory adverse events.

Attrition bias: *basso rischio*

Reporting bias is selection bias

- Reporting bias is perhaps the greatest source of selection bias
- Originally defined as the publication or non-publication of studies depending on the direction and statistical significance of the results
- Is a complex phenomenon


Full
n and effect size, plus precision / p-value for continuous data
Partial
Effect size or precision (± n or p-value)
Qualitative
p-value
Unreported
Incompletely reported outcomes

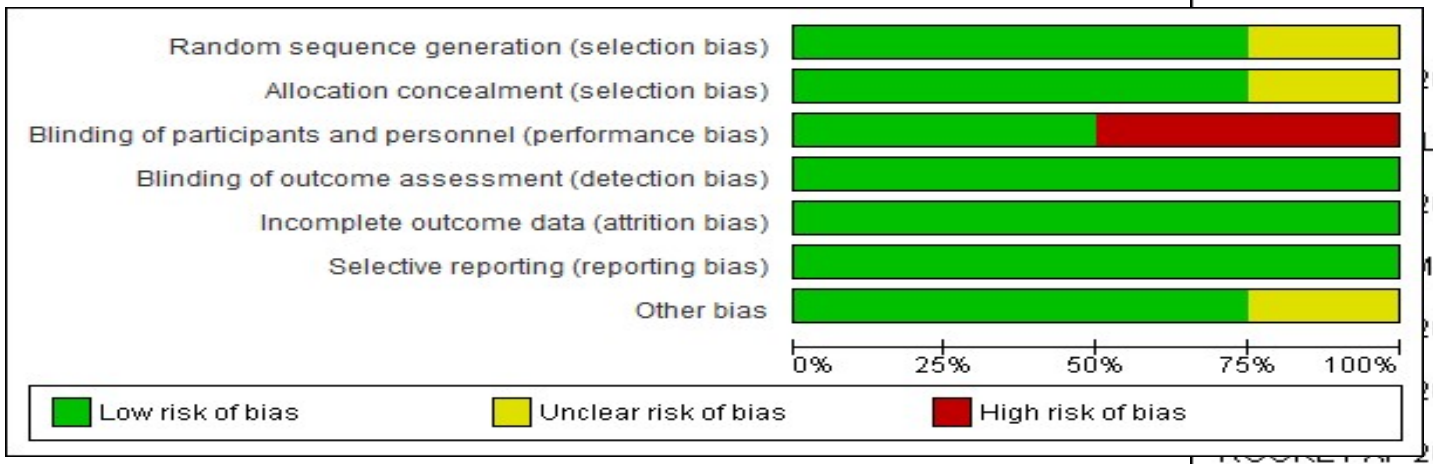
(Chan, 2004)

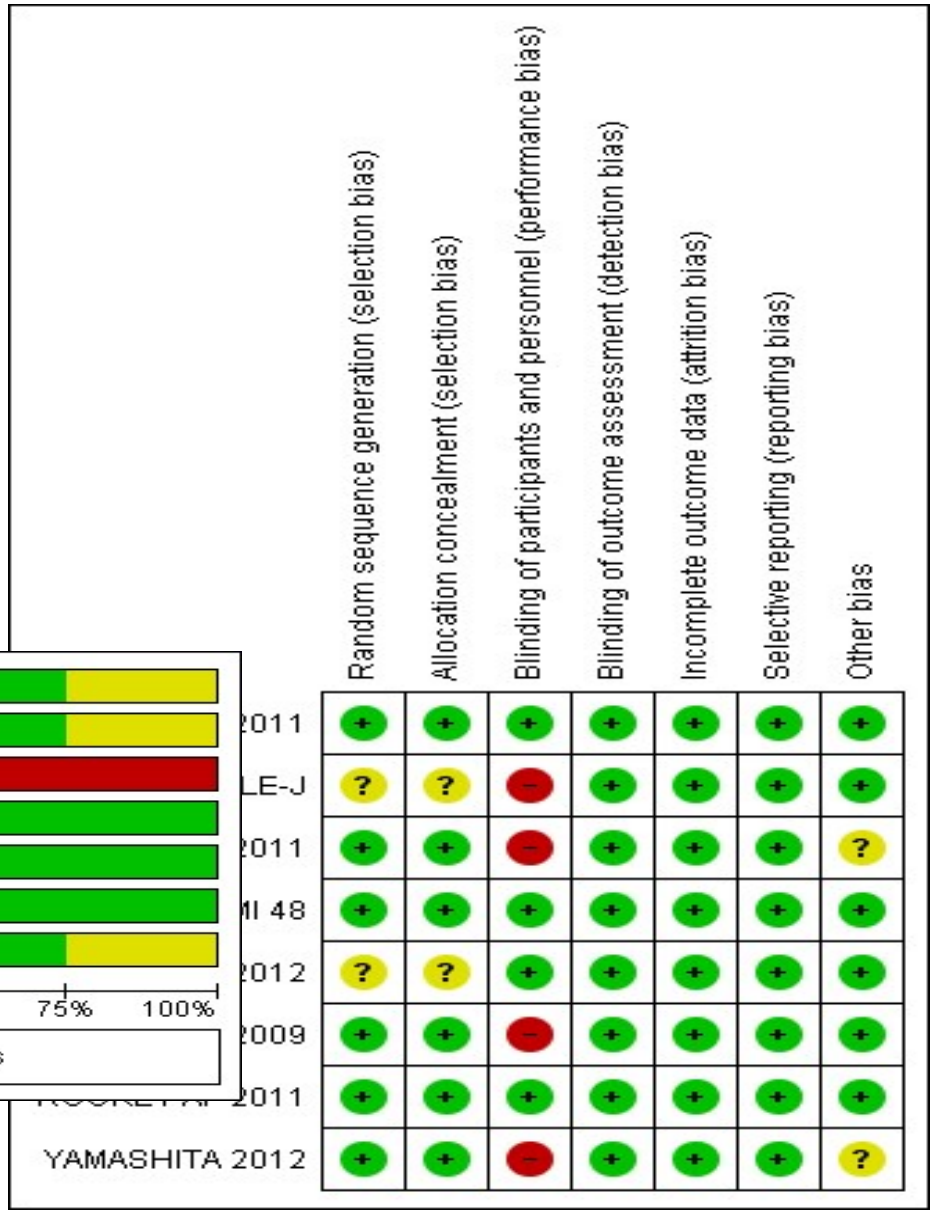
Risk of bias in one study

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	NR
Allocation concealment (selection bias)	Unclear risk	NR
Blinding of participants and personnel (performance bias)	High risk	open label
Blinding of outcome assessment (detection bias)	Low risk	An independent blinded endpoint committee adjudicated all reported bleeding and efficacy events
Incomplete outcome data (attrition bias)	Low risk	ITT. all patients followed up
Selective reporting (reporting bias)	Low risk	
Other bias	Low risk	

Risk of bias across studies/domains





Dal Risk of bias al GRADE

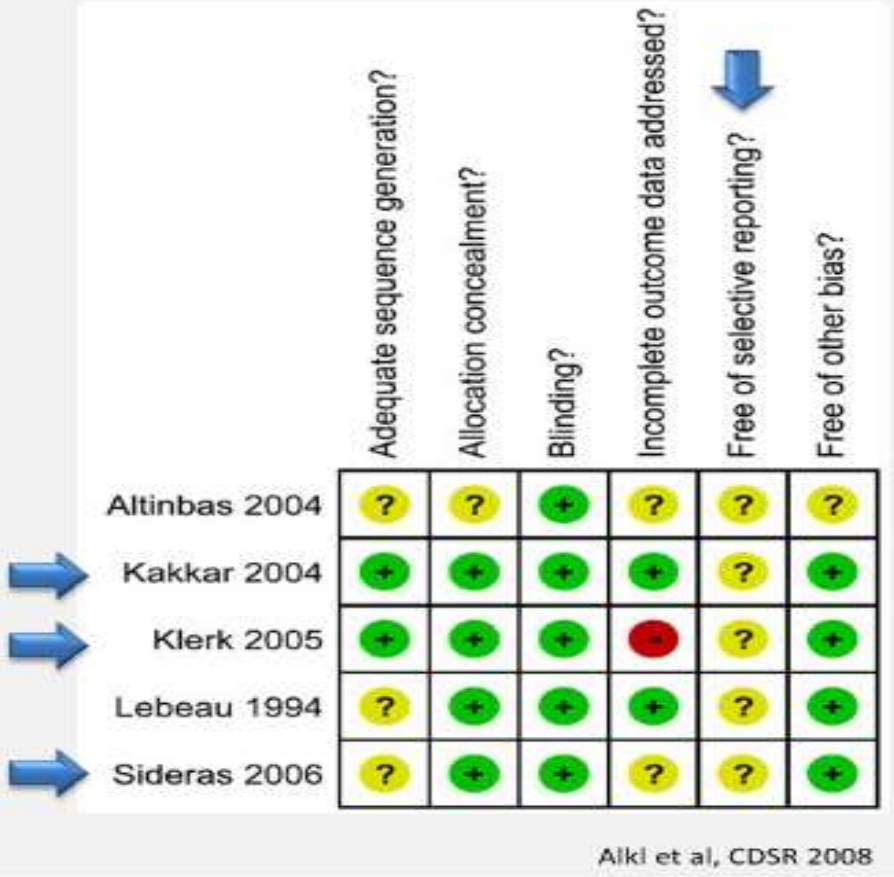
- La fiducia nelle prove complessiva, riferita ad un **singolo outcome**, può essere abbassata se si ritiene che esista un **importante rischio di bias** negli studi che contribuiscono a quelle evidenze

- Fiducia (certainty) minore nell'affidabilità dei risultati

- **APPROCCIO GRADE**
  - Se non si sospettano bias importanti, non si abbassa «not serious» (no change in quality)
  - Se si sospettano bias importanti, si abbassa di 1 livello «serious» (downgrade quality of evidence 1 level)
  - Se si sospettano bias molto importanti, si abbassa di 2 livelli «very serious» (downgrade quality of evidence 2 level)

Example: Major Bleeding with anticoagulation

Population: *people with cancer*
Intervention: *anti-coagulants, such as heparin*
Outcome: *major bleeding*



Alki et al, CDSR 2008

Concern maggiore su selective reporting of outcome
Solo tre studi riportano l'outcome major bleeding

!!! Do not count number of green, yellow, red spots!!!

Would you downgrade for risk of bias?

☐ No, there are no serious limitations

☒ Yes, there are serious limitations

Most people would agree that selective reporting of outcome is concern

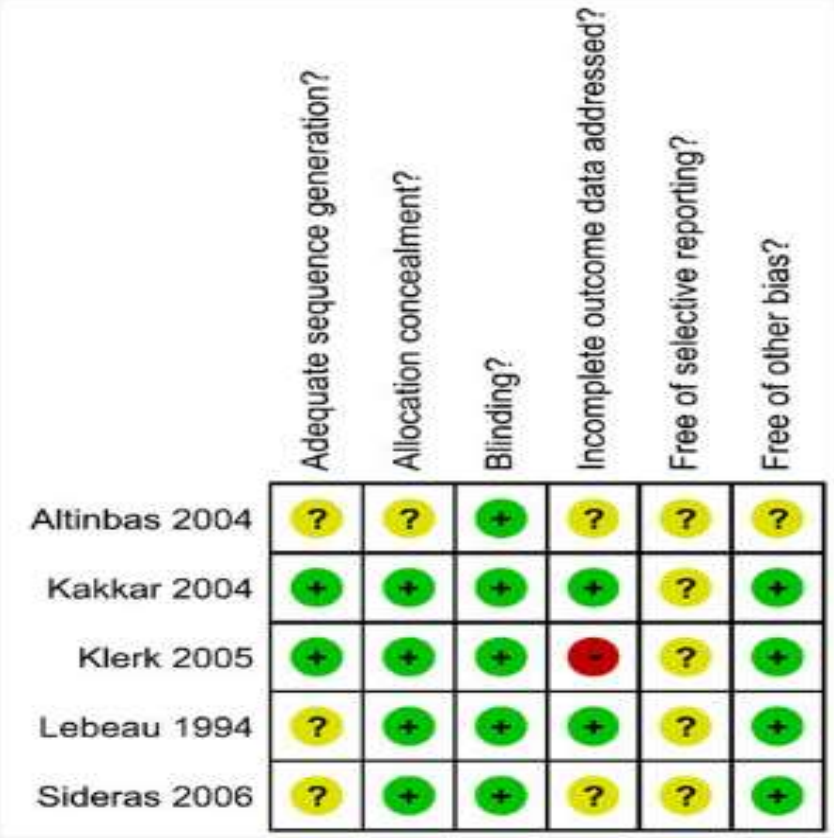
☐ Yes, there are very serious limitations

Example: Major Bleeding with anticoagulation

Population: *people with cancer*
Intervention: *anti-coagulants, such as heparin*
Outcome: *major bleeding*

Limitations likely to reduce confidence in effect
Downgrade 1 level

Footnotes
[1] Only 3 of 5 studies reported major bleeding, selective reporting likely



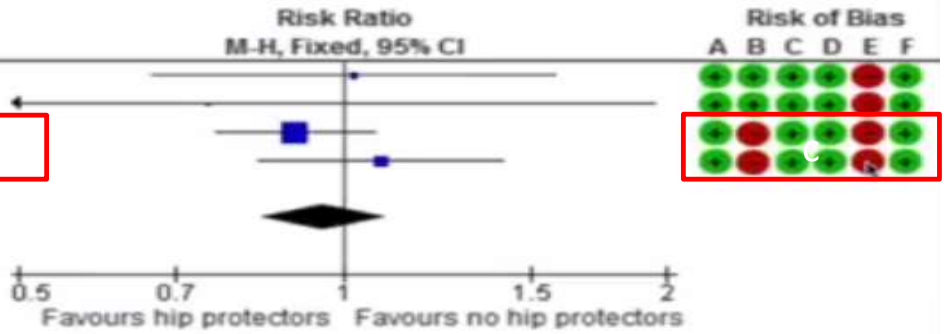
Alkl et al, CDSR 2008

Mortality (dichotomous outcome)

Mortality

Study or Subgroup	Hip protectors Events	Total	No hip protectors Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Cameron 2001	28	86	28	88	9.5%	1.02 [0.66, 1.58]
Jantti 1996	6	36	8	36	2.7%	0.75 [0.29, 1.94]
Meyer 2003	157	459	183	483	61.1%	0.90 [0.76, 1.07]
Van Schoor 2003	83	276	79	285	26.6%	1.08 [0.84, 1.41]
Total (95% CI)		**857**		**892**	**100.0%**	**0.96 [0.84, 1.10]**
Total events	274		298			

Heterogeneity: Chi² = 1.69, df = 3 (P = 0.64); I² = 0%
Test for overall effect: Z = 0.62 (P = 0.54)



A: random sequence generation
B: allocation concealment
C: Blinding of participants and personnel
D: Blinding of outcome assessor
E: Incomplete outcome data
F: selective outcome reporting

Would you downgrade for risk of bias?

☐ No, there are no serious limitations

☒ Yes, there are serious limitations

Some concerns with allocation concealment and high losses to follow up

☐ Yes, there are very serious limitations

Dal Risk of bias al GRADE

- Non fare medie, non contare i domini a rischio o quelli adeguati
- Tutto si basa su giudizi, spesso soggettivi
- Calarsi nel contesto clinico e metodologico
- Trasparenza nelle decisioni (soprattutto nelle situazioni intermedie)
- Focus sugli studi a basso rischio di bias
- Focus sugli studi che portano un maggiore contributo informativo (peso nella meta-analisi)
- Siate conservativi
- Rischio di bias nel contesto delle altre dimensioni della qualità

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Valutazione della qualità delle prove: *imprecision*

Imprecision of the overall estimate

- **Wide confidence intervals** (CIs inform the impact of random error on evidence quality; CI expresses the range in which the truth plausibly lies)
- **Small number of events** (total number of events is less than 300)
- **Small sample size**
- recommendation or **clinical course of action would differ** if the **upper versus the lower boundary** of the CI represented the truth

Imprecision for guideline developers

- Guideline developers should consider the clinical threshold between recommending and not recommending (consider also importance of the outcome, undesirable consequences – risks, AEs), not only the OIS

- GRADE's primary criterion for judging precision is to focus on the 95% confidence interval (CI) around the difference in effect between intervention and control for each outcome.

- In general, the CIs to consider are those around the absolute, rather than the relative effect.

- If a recommendation or clinical course of action would differ if the upper versus the lower boundary of the CI represented the truth, consider the rating down for imprecision.

- Even if CIs appear satisfactorily narrow, when effects are large and both sample size and number of events are modest, consider the rating down for imprecision.

Key Points

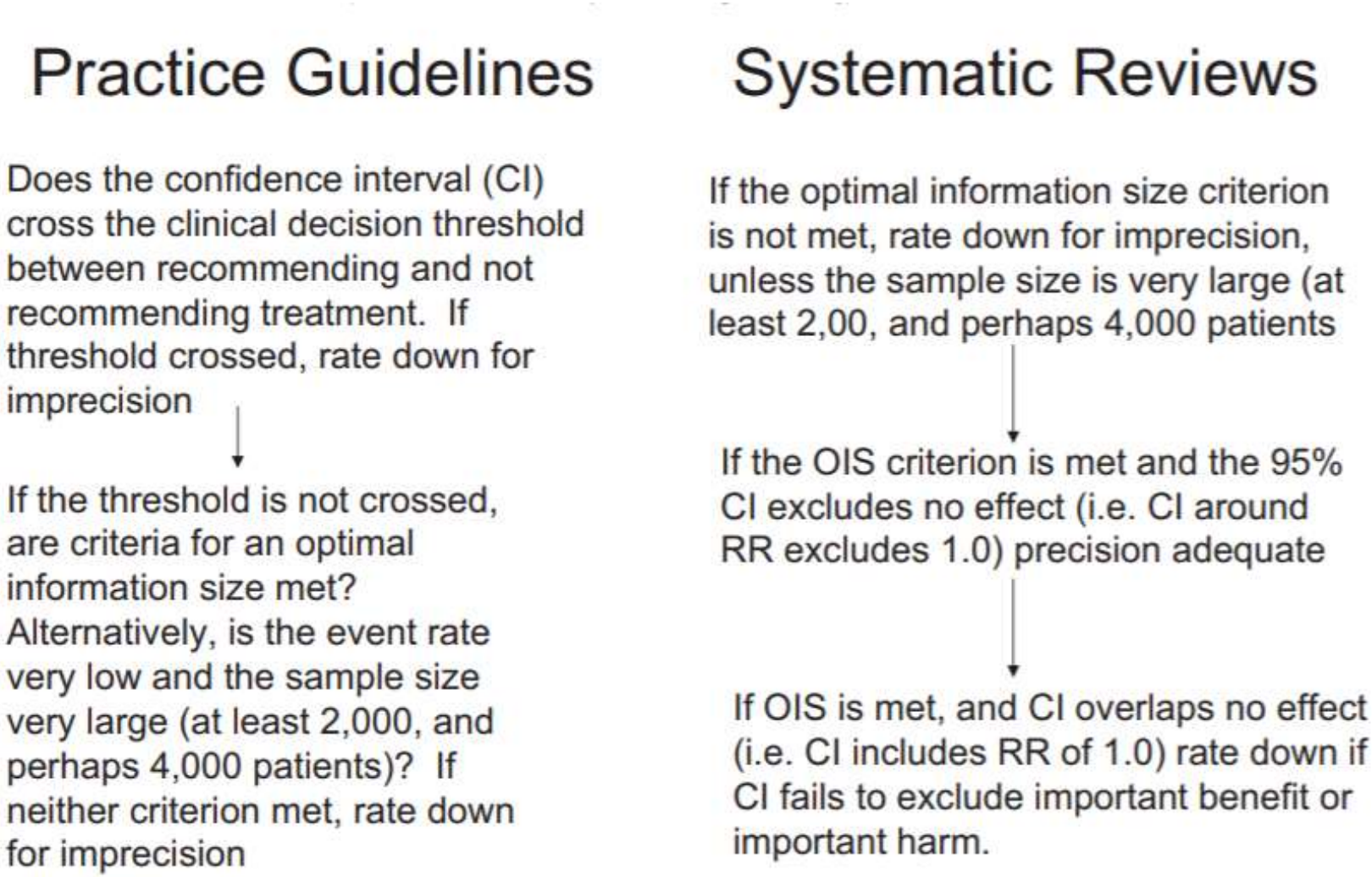


Fig. 3. Deciding whether to rate down for imprecision in guidelines and systematic reviews of binary variables.

OIS for continuous outcomes

- Authors can calculate the OIS for continuous variables in exactly the same way they can for binary variables by specifying the α and β errors (we have suggested 0.05 and 0.2) and the Δ (i.e. the difference one wishes to detect as clinically relevant), and choosing an appropriate standard deviation from one of the relevant studies.
- A particular challenge in calculating the OIS for continuous variables arises when studies have used different instruments to measure a construct, and the pooled estimate is calculated using a standardized mean difference.
- we suggest authors choose one of the available instruments (ideally, one in which an estimate of the minimally important difference is available) and calculate an OIS using that instrument

OIS for continuous outcomes

whenever there are **sample sizes** that are less than 400, review authors and guideline developers should certainly consider rating down for imprecision.

OIS for dicotomous outcomes

whenever there are **number of event** that are less than 300, review authors and guideline developers should certainly consider rating down for imprecision.

Rare events: an exception

- When event rates are very low, Cis around relative effects may be wide, but if sample sizes are sufficiently large, it is likely that prognostic balance has indeed been achieved, and rating down for imprecision becomes inappropriate
- Rare event: <5% nel gruppo di controllo
- Large sample size: circa 4000 soggetti

Imprecision

No effect

Control event rate: 50%

Total events: n:511

OIS: met

Lower CI: 25% reduction

Downgrade: yes

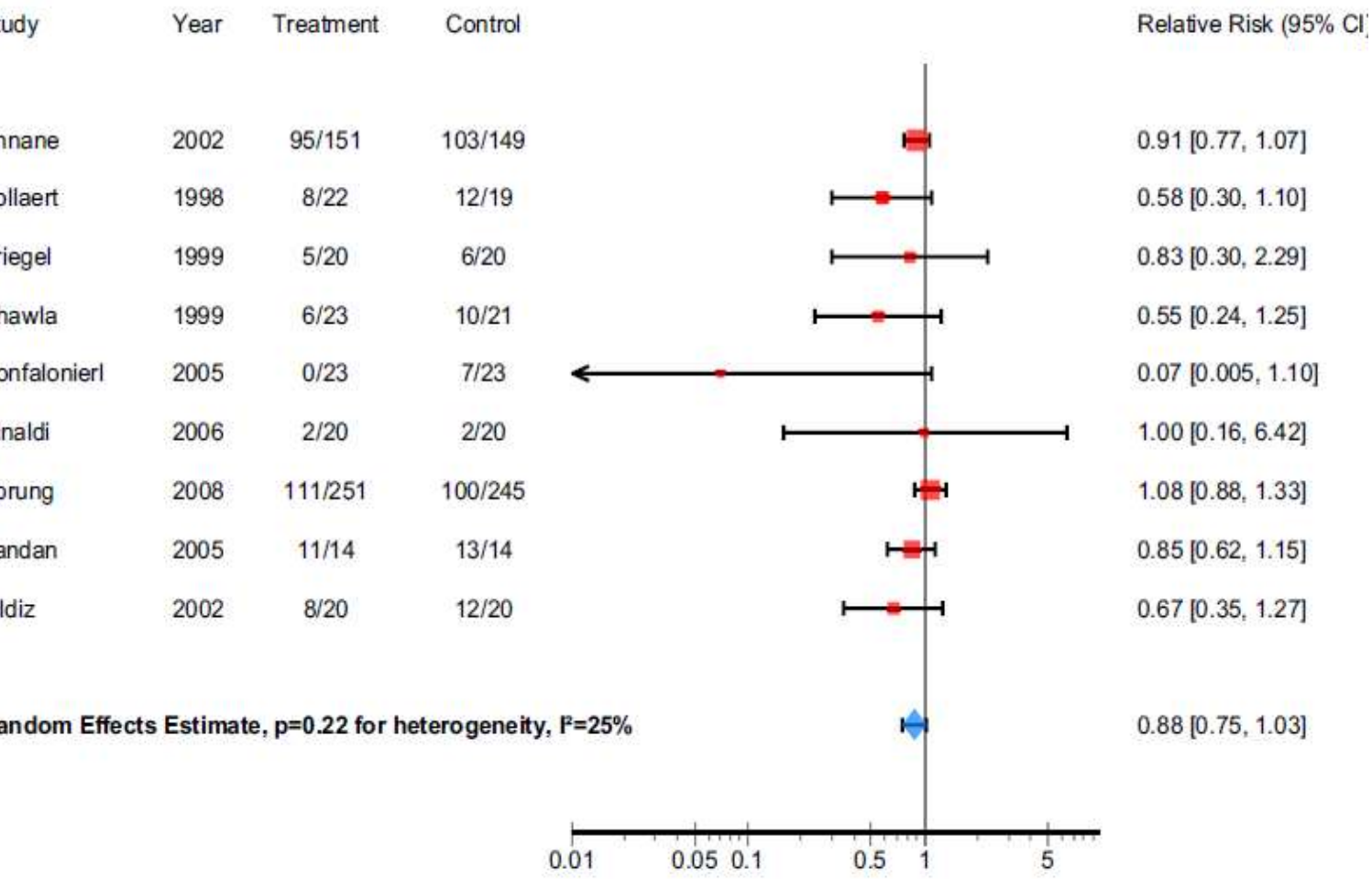
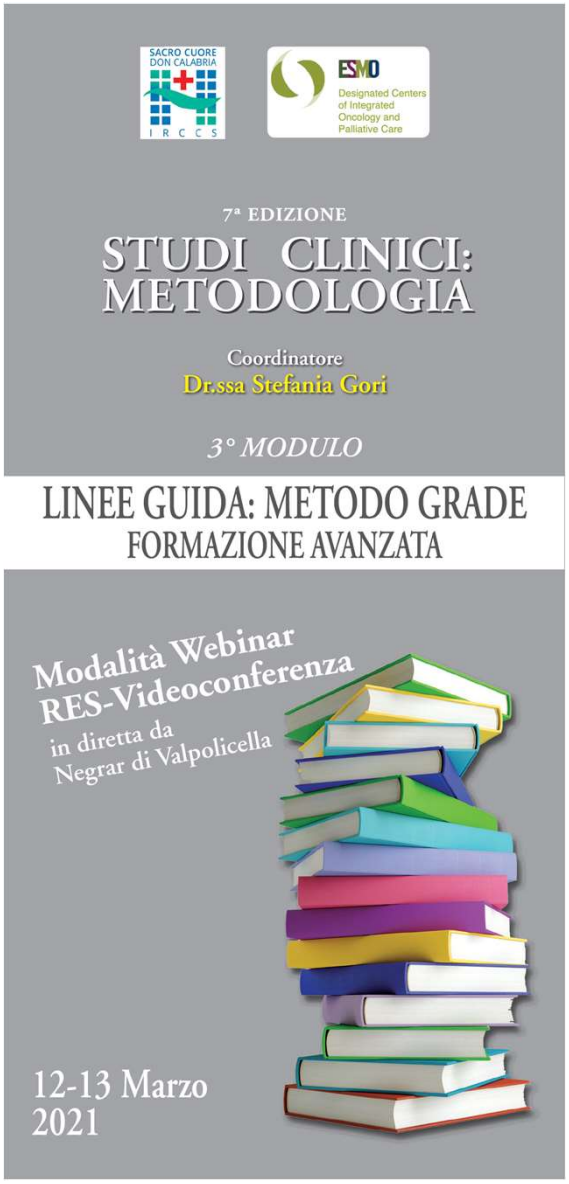


Fig. 2. Corticosteroids to reduce hospital mortality in septic shock.



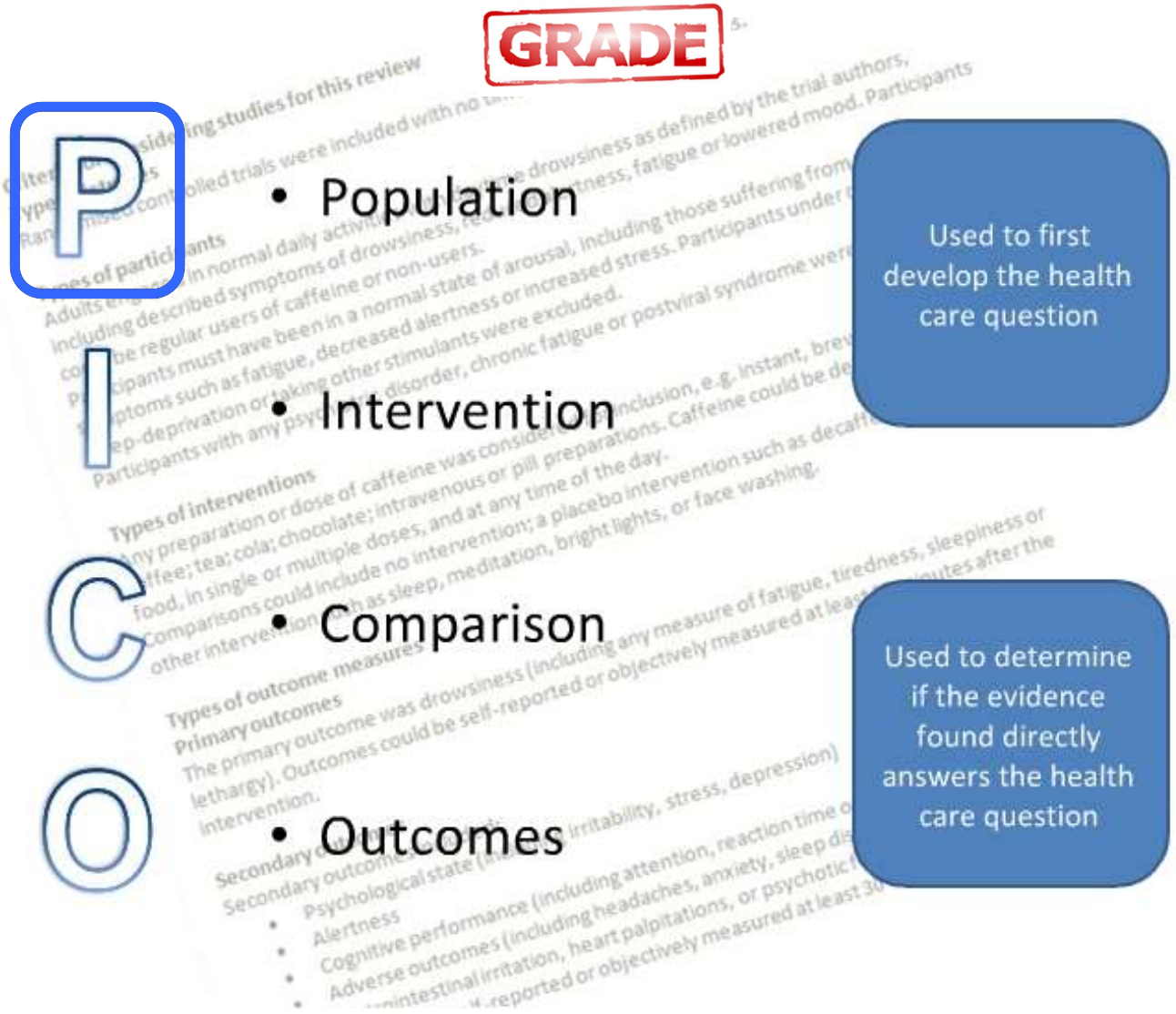
GRADE

Valutazione della qualità delle prove: *indirectness*

Direct evidence…

…comes from research that:

- is conducted in the **Population** that we are providing answers for;
- includes the **Intervention** that we are interested in…
- …and compares these interventions with the appropriate **Alternatives**;
- measures the **Outcomes** in which we are interested

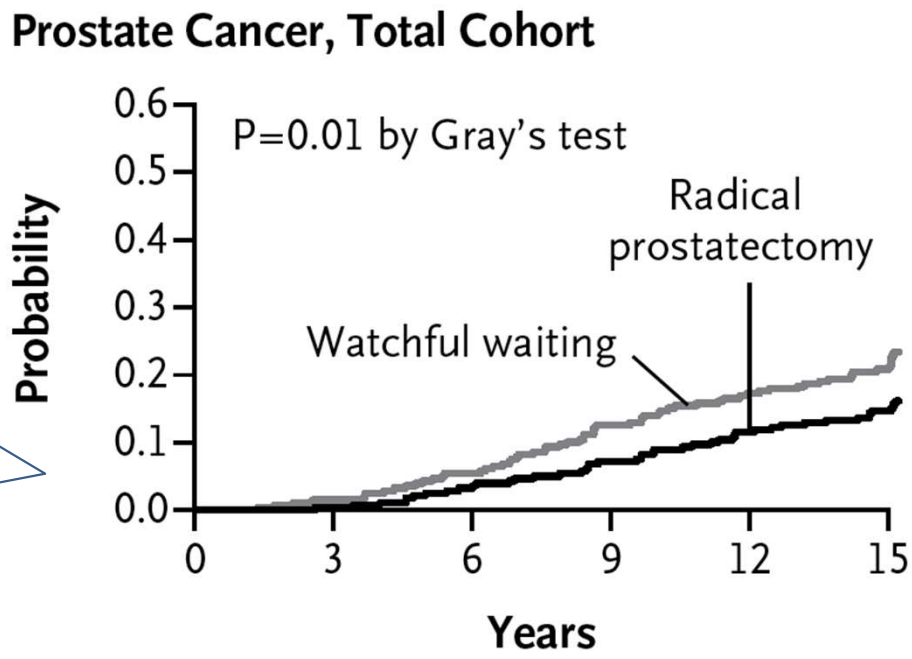

GRADE
P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

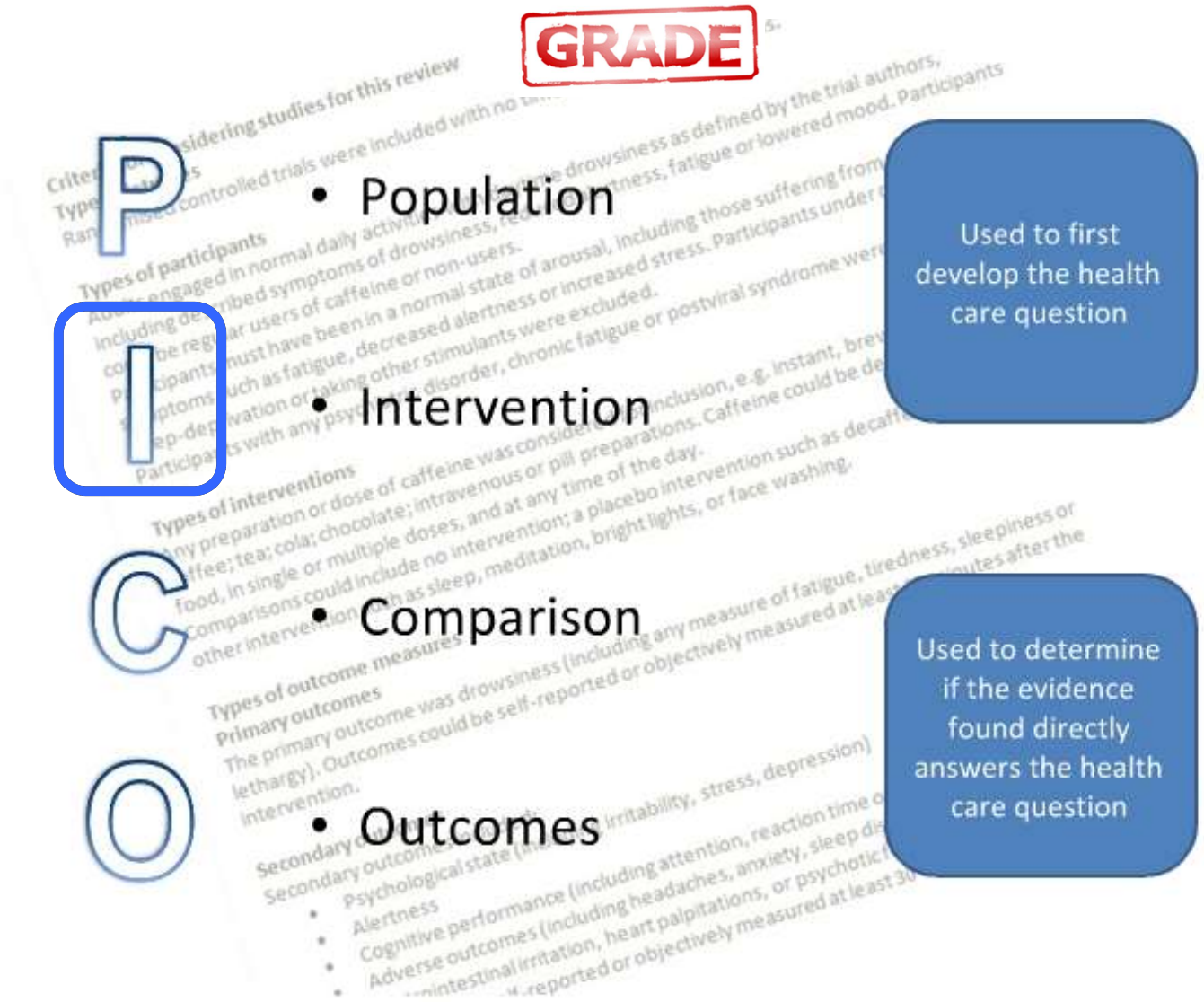
Radical Prostatectomy versus Watchful Waiting in Early Prostate Cancer

Anna Bill-Axelson, M.D., Ph.D., Lars Holmberg, M.D., Ph.D., Mirja Ruutu, M.D., Ph.D., Hans Garmo, Ph.D., Jennifer R. Stark, Sc.D., Christer Busch, M.D., Ph.D., Stig Nordling, M.D., Ph.D., Michael Häggman, M.D., Ph.D., Swen-Olof Andersson, M.D., Ph.D., Stefan Bratell, M.D., Ph.D., Anders Spångberg, M.D., Ph.D., Juni Palmgren, Ph.D., Gunnar Steineck, M.D., Ph.D., Hans-Olov Adami, M.D., Ph.D., and Jan-Erik Johansson, M.D., Ph.D., for the SPCG-4 Investigators*

N Engl J Med 2011;364:1708-17.

Pazienti in era pre-PSA



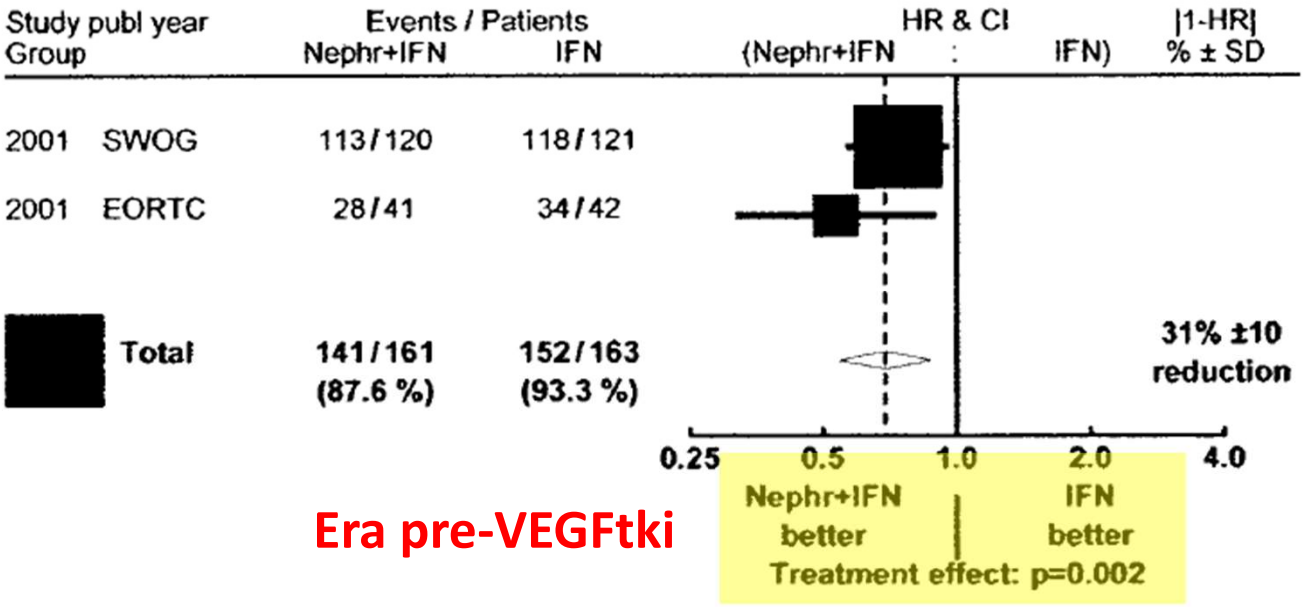

GRADE
P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

CYTOREDUCTIVE NEPHRECTOMY IN PATIENTS WITH METASTATIC RENAL CANCER: A COMBINED ANALYSIS

ROBERT C. FLANIGAN,* G. MICKISCH, RICHARD SYLVESTER, CATHY TANGEN,†
H. VAN POPPEL AND E. DAVID CRAWFORD

From the Southwest Oncology Group and European Organization for the Research and Treatment of Cancer Genitourinary Group, Loyola University Medical Center (RCF), Maywood, Illinois, Centrum Fuer Operative Urologie (GM), Bremen, Germany, European Organization for the Research and Treatment of Cancer Data Center (RS), Brussels and UZ Gasthuisberg (HVP), Leuven, Belgium, Southwest Oncology Group Statistical Center (CT), Seattle, Washington, and University of Colorado Medical Center (EDC), Denver, Colorado

THE JOURNAL OF UROLOGY® Vol. 171, 1071–1076, March 2004



Era pre-VEGFtki

Long-term oncologic outcomes of postoperative adjuvant versus salvage radiotherapy in prostate cancer: Systemic review and meta-analysis of 5-year and 10-year follow-up data

Ja Yoon Ku[1], Chan Ho Lee[1], Hong Koo Ha[1,2]

Korean J Urol 2015;56:735-741.

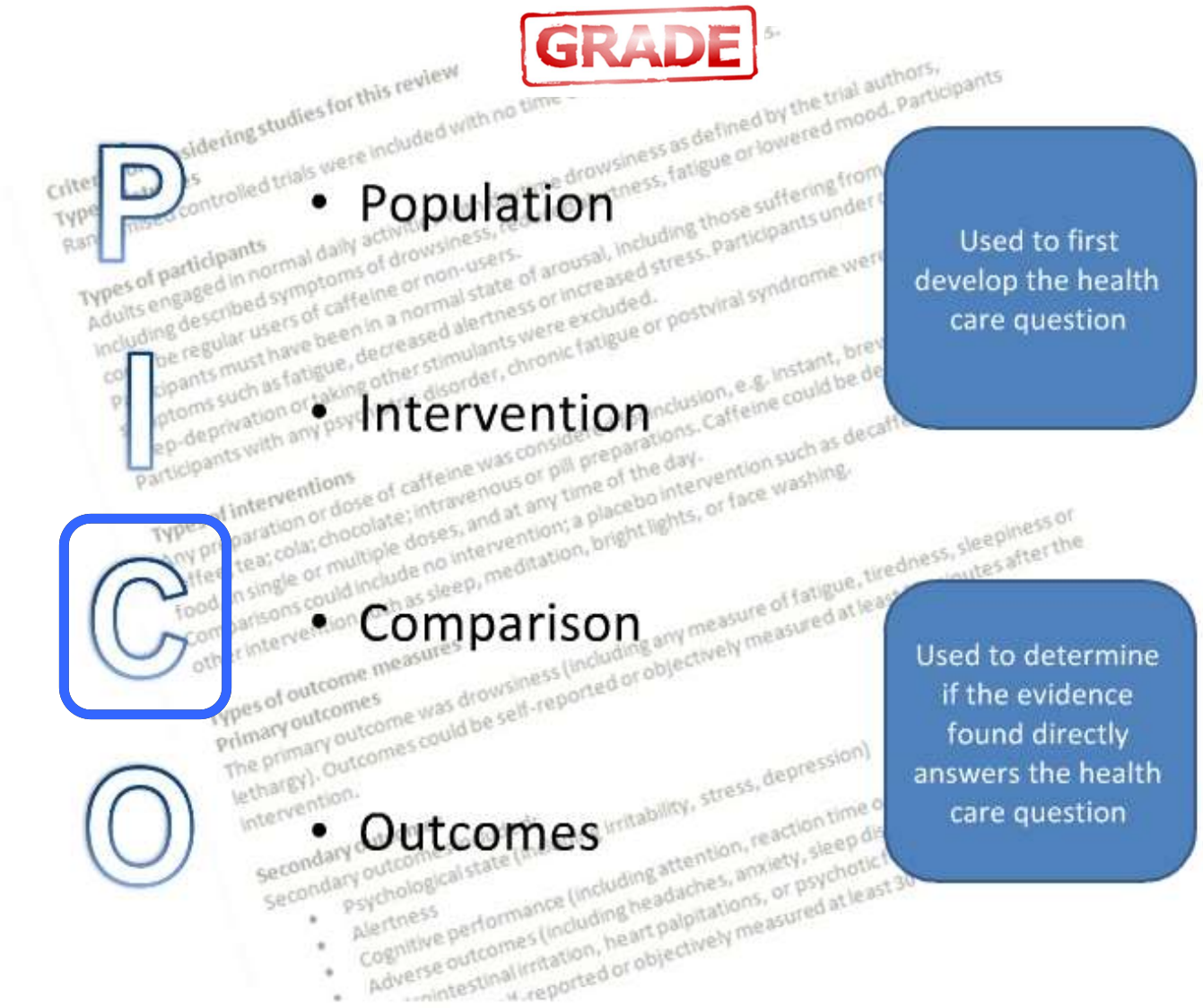
The present systemic review has the following limitations that must be taken into account.

The first limitation is that we heterogeneously recruit randomized controlled studies and retrospective studies: in retrospective studies, the initiation timing of radiotherapy is somewhat different in each study.

The second limitation is that there is a difference in dose and modality (2-dimensional, 3-dimensional, or intensity modulated radiation therapy) of radiation compared to recent practice, which may alter oncologic outcomes.

The third limitation [illegible] were different in ea[illegible] of long-term outcome be[illegible]

Indirectness per I. (di P.I.C.O.)


GRADE
P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

Phase III Trial of Vinflunine Plus Best Supportive Care Compared With Best Supportive Care Alone After a Platinum-Containing Regimen i Patients With Advanced Transitional Cell Carcinoma of t thelial Tract

Joaquim Bellmunt, Christine Théodore, Tomasz Demkov gelov, Gedske Daugaard, Armelle Caty, Joan Carles, Agnieszka Jag Patrick Hurteloup, Eric Winquist, Nassim

J Clin Oncol 27:4454-4461. © 2009

Second-line therapy in bladder cancer

Mark Bachner and Maria De Santis

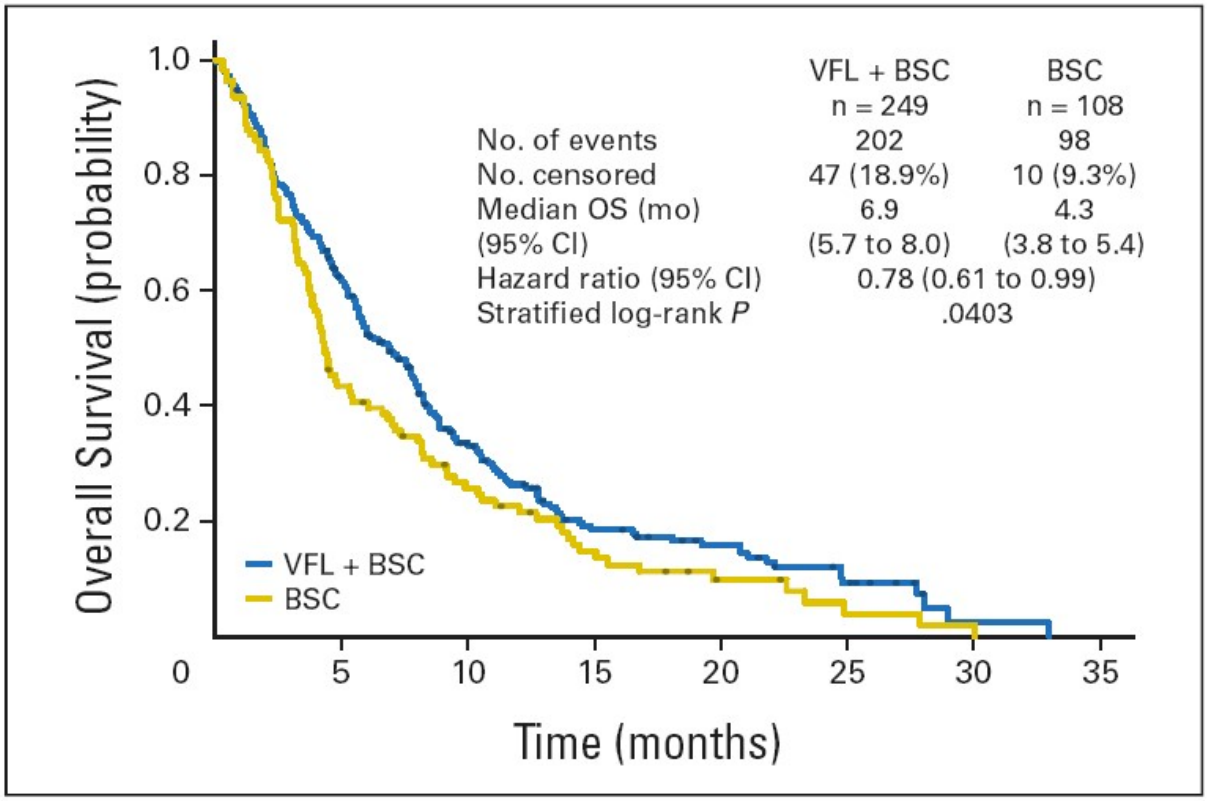
Current Opinion in Urology 2009, 19:533–539

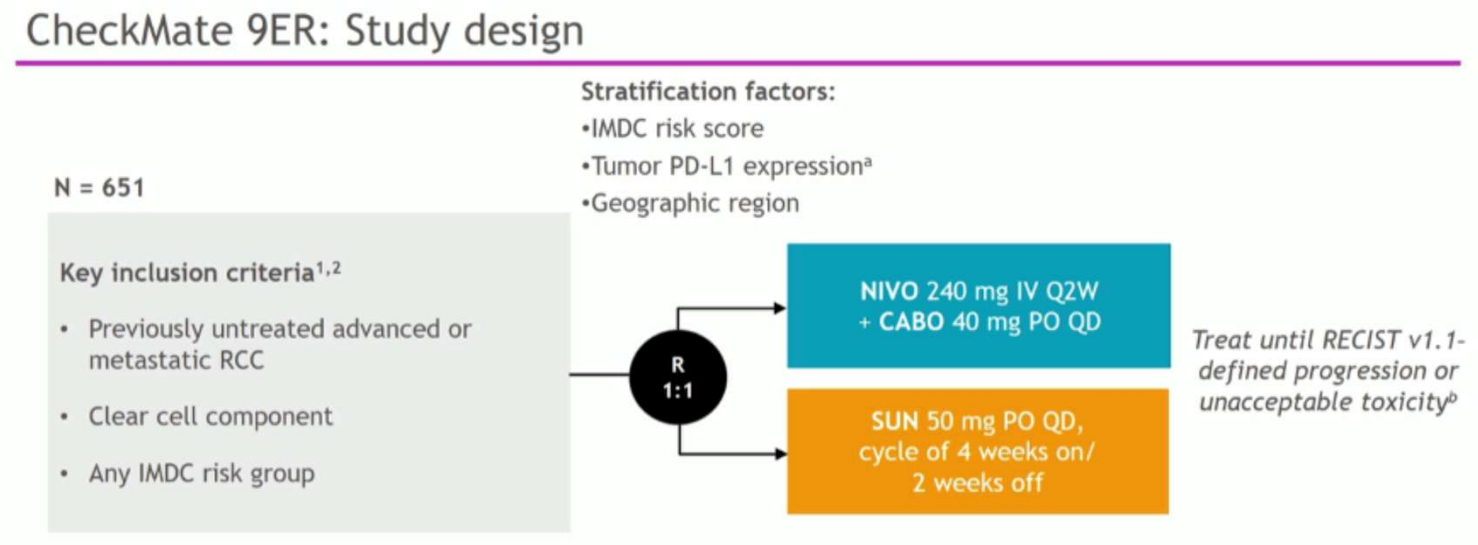
So far no standard therapy has been established for pretreated patients with transitional cell carcinoma.

Phase III Trial of Vinflunine Plus Best Supportive Care Compared With Best Supportive Care Alone After a Platinum-Containing Regimen in Patients With Advanced Transitional Cell Carcinoma of the Urothelial Tract

Joaquim Bellmunt, Christine Théodore, Tomasz Demkov, Boris Komyakov, Lisa Sengelov, Gedske Daugaard, Armelle Caty, Joan Carles, Agnieszka Jagiello-Gruszfeld, Oleg Karyakin, François-Michel Delgado, Patrick Hurteloup, Eric Winquist, Nassim Morsli, Yacine Salhi, Stéphane Culine, and Hans von der Maase

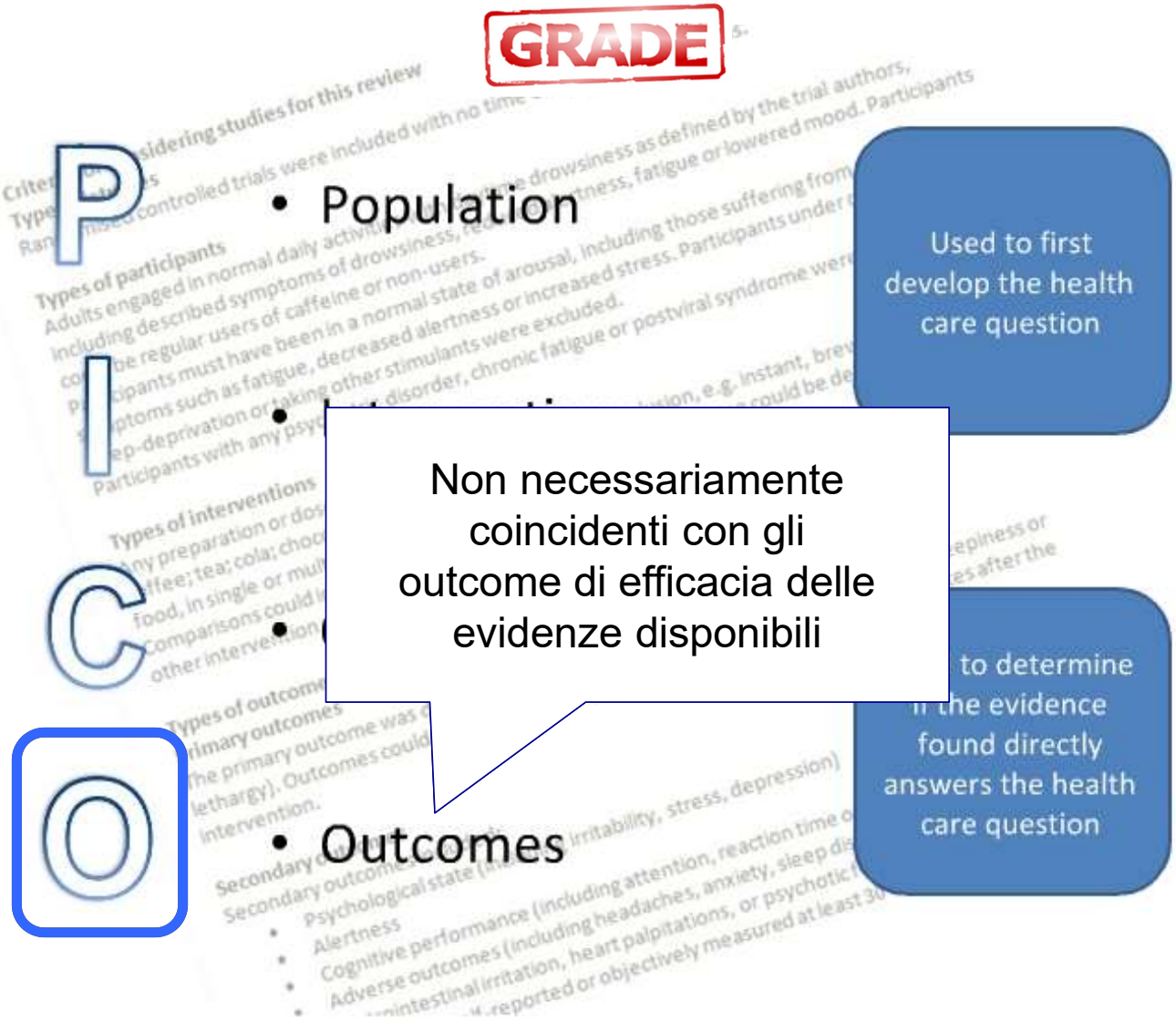
J Clin Oncol 27:4454-4461. © 2009 by American Society of Clinical Oncology





Pembrolizumab
Axitinib

Certezza globale delle prove	Raccomandazione clinica	Forza della raccomandazione clinica
BASSA	Nei pazienti affetti da carcinoma renale metastatico variante istologica a cellule chiare, rischio intermedio-alto sec. Heng, il trattamento di prima linea con pembrolizumab-axitinib dovrebbe essere preso in considerazione come approccio terapeutico di prima scelta	**Forte a Favore**


GRADE
P
• Population
I
C
O
• Outcomes
Used to first develop the health care question
Non necessariamente coincidenti con gli outcome di efficacia delle evidenze disponibili
to determine if the evidence found directly answers the health care question

Surrogate outcome markers in research and clinical practice

Scott Twaddell

(Aust Prescr 2009;32:47–50)

Table 1

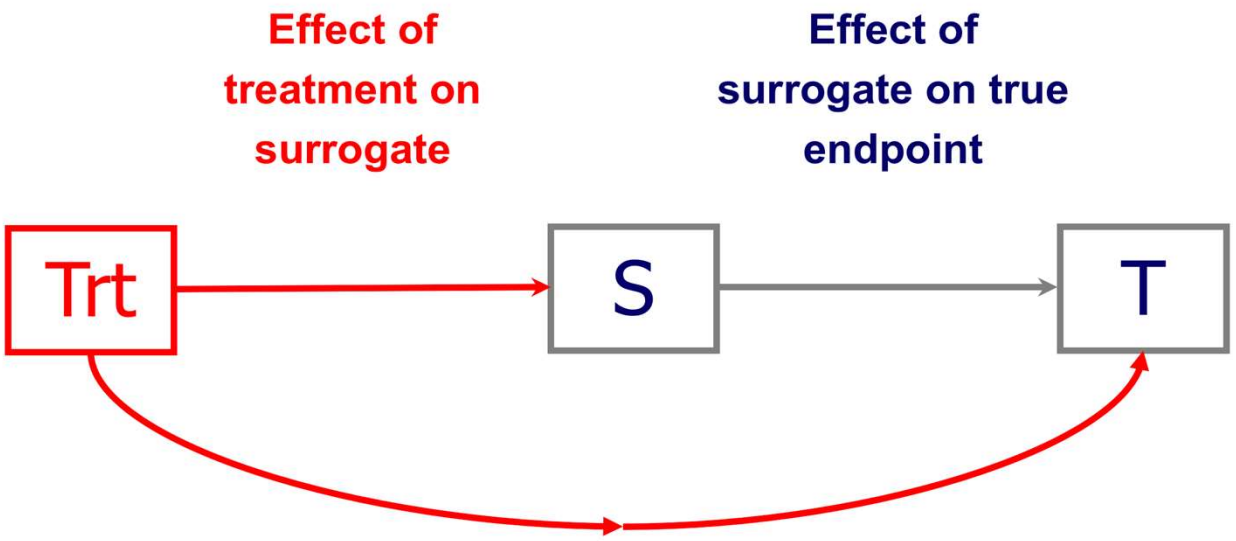
Surrogate markers often used in clinical practice

Generally accepted as valid		Doubt still exists about validity	
Surrogate marker	**Predicts**	**Surrogate marker**	**Predicts**
HbA1c	Diabetic microvascular complications	HbA1c	Diabetic macrovascular complications
FEV_1	Mortality in chronic obstructive pulmonary disease	Bone mineral density	Fracture risk
Blood pressure	Primary and secondary cardiovascular events	Prostate specific antigen	Prognosis of prostate cancer
Viral load	Survival in HIV infection	Suppression of arrhythmia	Long-term survival
Cholesterol concentration	Primary and secondary cardiovascular events	Carotid intima-media thickness	Coronary artery disease
Intraocular pressure	Visual loss in glaucoma	Albuminuria	Cardiovascular events

HbA1c glycated haemoglobin

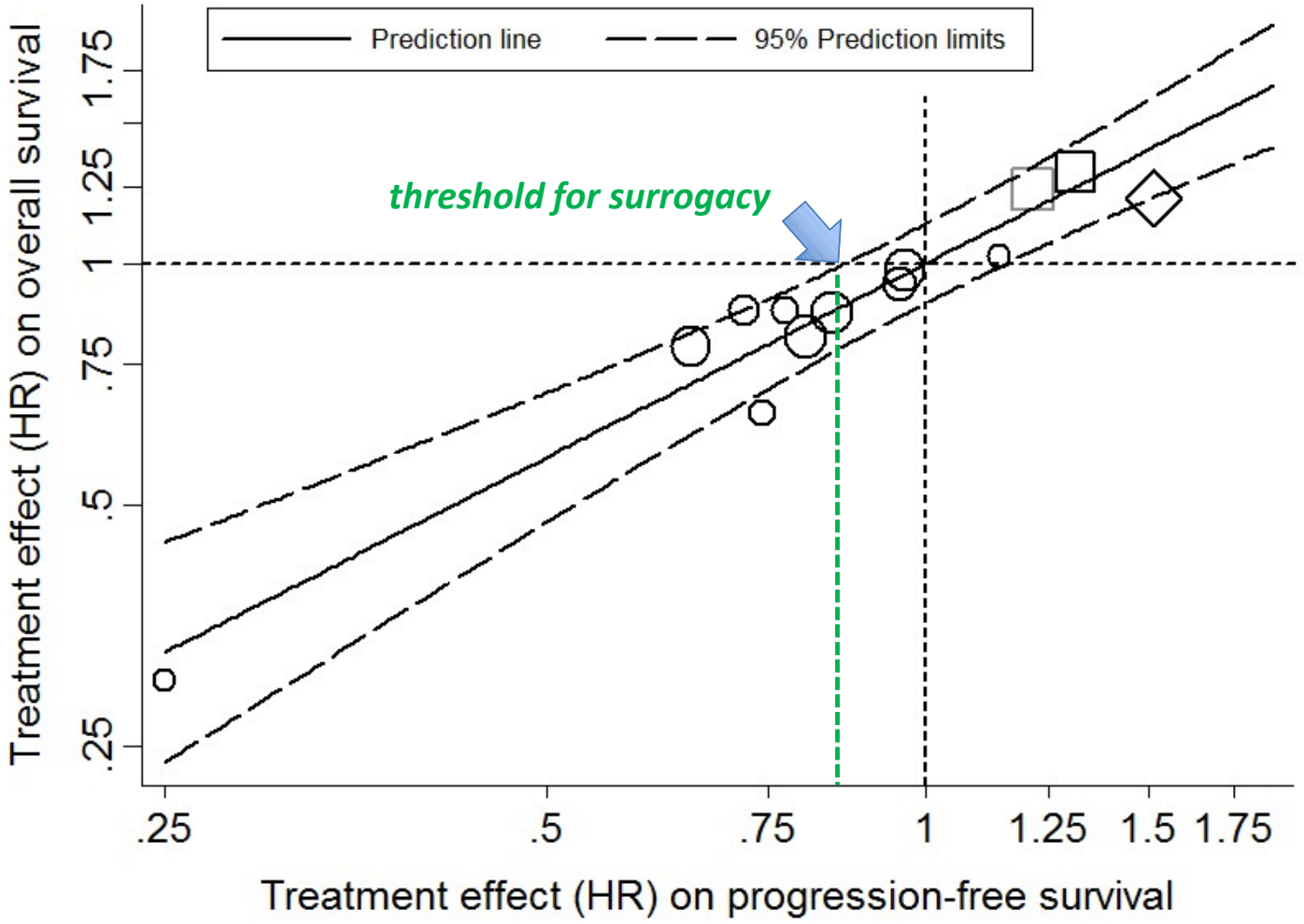
FEV_1 forced expiratory volume in one second

VALIDATION OF SURROGATE ENDPOINTS: "FULL CAPTURE OF EFFECT"



*Prentice, Statist Med 1989;**8**:431.*

TRIAL LEVEL CORRELATION BETWEEN EFFECTS

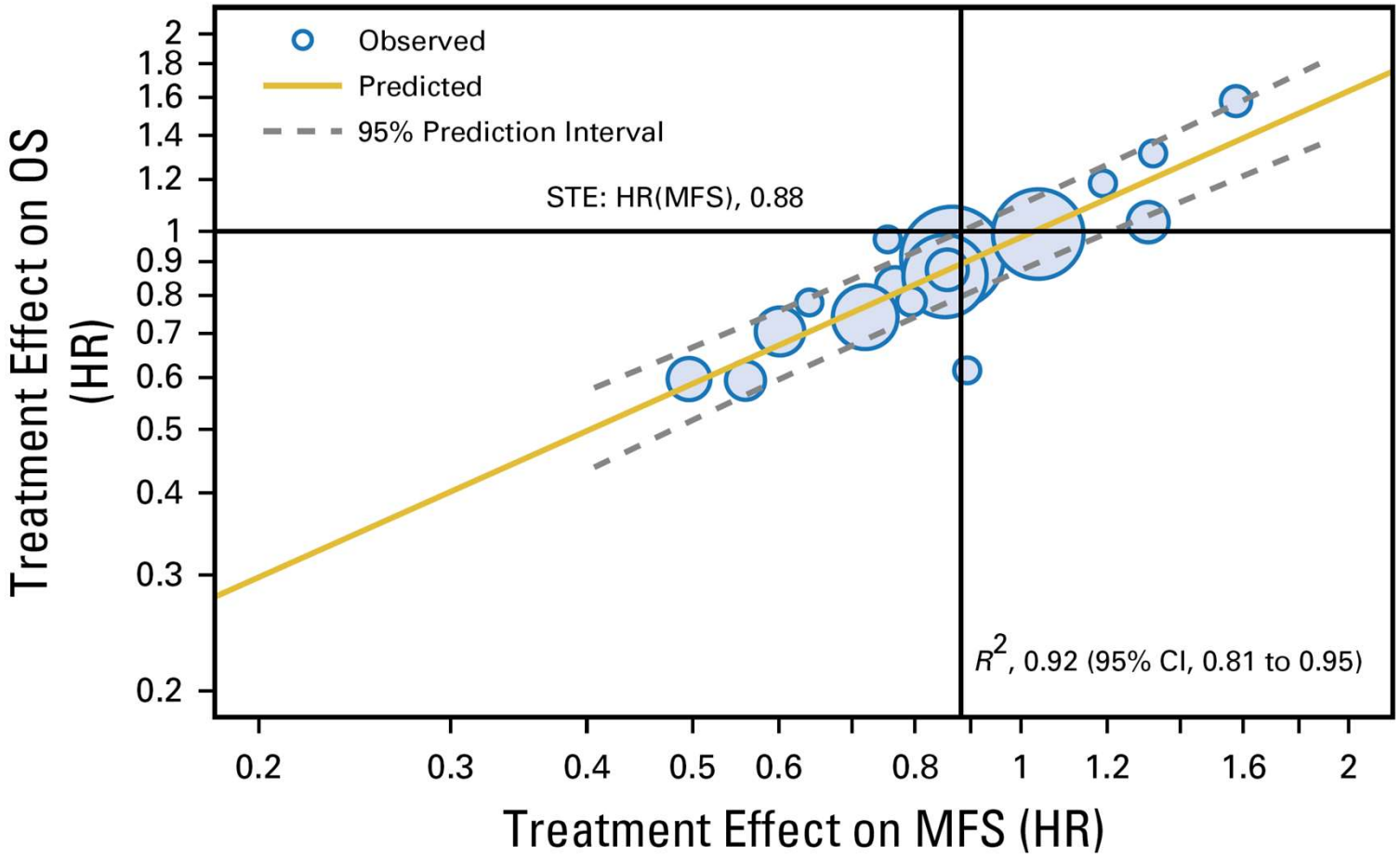


Burzykowski and Buyse, Pharmaceutical Statist 2006;5:173

Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer

Wanling Xie, Meredith M. Regan, Marc Buyse, Susan Halabi, Philip W. Kantoff, Oliver Sartor, Howard Soule, Noel W. Clarke, Laurence Collette, James J. Dignam, Karim Fizazi, Wendy R. Paruleker, Howard M. Sandler, Matthew R. Sydes, Bertrand Tombal, Scott G. Williams, and Christopher J. Sweeney, on behalf of the ICECaP Working Group

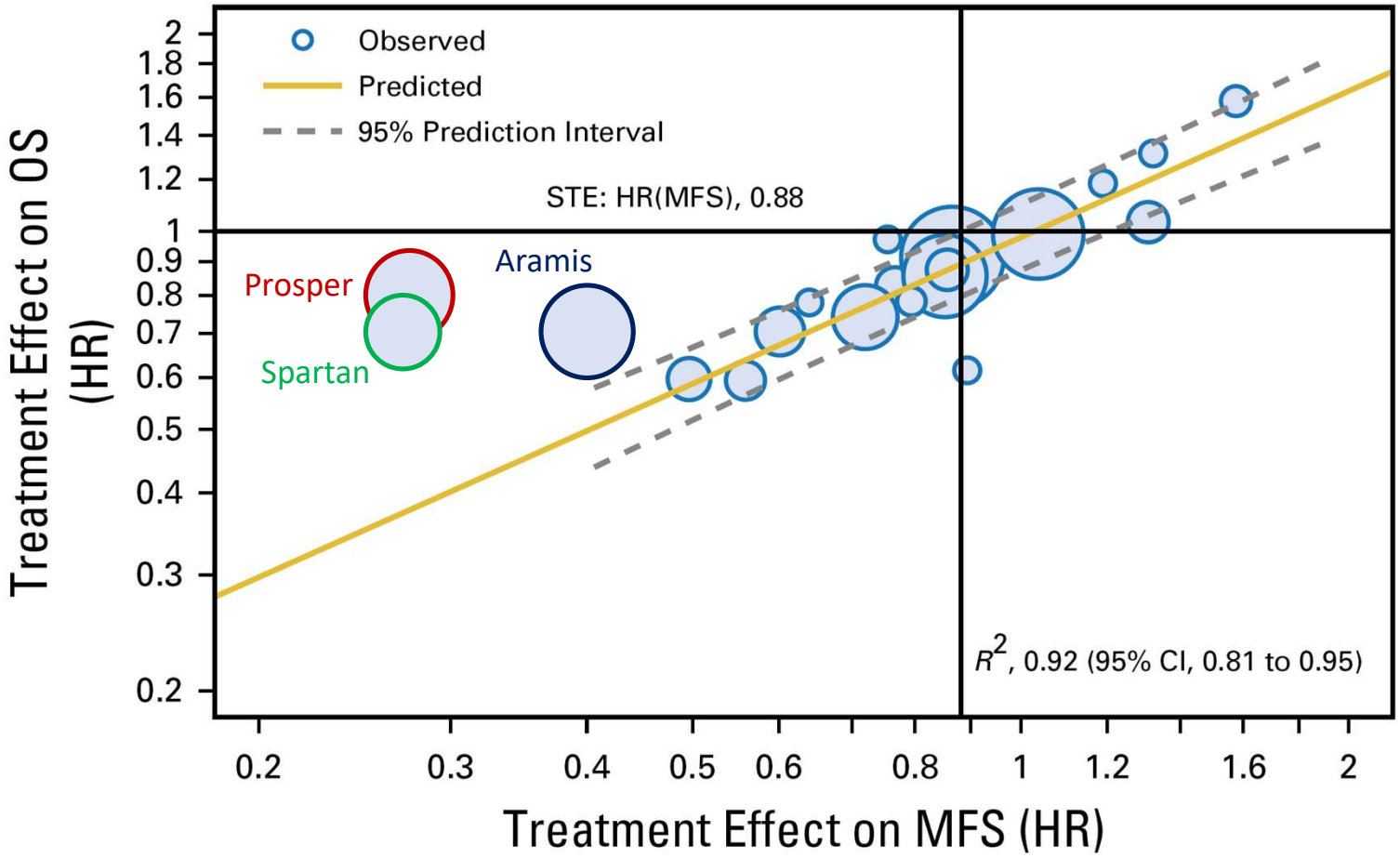
J Clin Oncol 35:3097-3104.



Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer

Wanling Xie, Meredith M. Regan, Marc Buyse, Susan Halabi, Philip W. Kantoff, Oliver Sartor, Howard Soule, Noel W. Clarke, Laurence Collette, James J. Dignam, Karim Fizazi, Wendy R. Paruleker, Howard M. Sandler, Matthew R. Sydes, Bertrand Tombal, Scott G. Williams, and Christopher J. Sweeney, on behalf of the ICECaP Working Group

J Clin Oncol 35:3097-3104. © 2017 by American Society of Clinical Oncology



Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

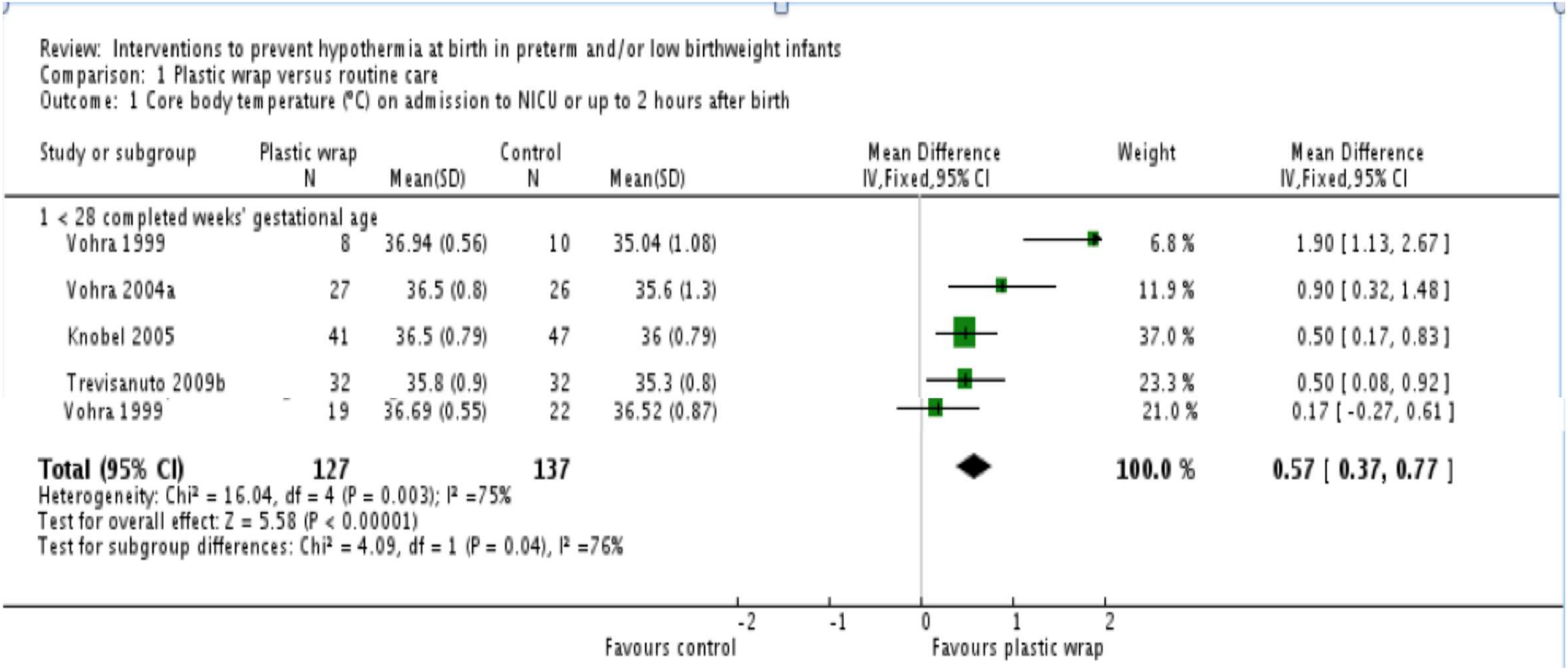
Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Valutazione della qualità delle prove: *inconsistency, publication bias*

Esempio di Metaview



What is heterogeneity?

Differences between studies with respect to:

Clinical heterogeneity (clinical diversity)

- *Participants*

  e.g. conditions under investigation, eligibility criteria for trials, geographical variation

- *Interventions*

  e.g. intensity / dose / duration, sub-type of drug, mode of administration, experience of practitioners, nature of the control (placebo/none/standard care)

- *Outcomes*

  e.g. definition of an event, follow-up duration, ways of measuring outcomes, cut-off points on scales

What is heterogeneity?

Differences between studies with respect to:

Methodological heterogeneity (methodological diversity)

- *Design*

  e.g. randomised vs non-randomised,
  crossover vs parallel group vs cluster randomised,
  pre-test and long follow up

- *Conduct*

  e.g. allocation concealment, blinding etc,
  approach to analysis, imputation methods for
  missing data

What is heterogeneity?

What do we do if there ***is*** statistical heterogeneity?

Variation in the *true effects* underlying the studies

...which may manifest itself in more observed variation than expected by chance alone

May be due to clinical diversity (different treatment effects) or methodological diversity (different biases)

Inconsistency is important only when it reduces confidence in results in relation to a particular decision.

Come si misura questa eterogeneità?

Confidence interval overlapping Eyeball test

Cochran's Q: to assess whether observed differences in results are compatible with change alone

χ^2 distribution; low power (small number of studies, small sample size)

p=<0.10 (heterogeneity)

I^2 quantifing heterogeneity (describes the percentage of variation across studies that is due to heterogeneity rather than chance)

0-40% might not be important

30-60% may represent moderate heterogeneity

50-90% may represent substantial heterogeneity

75-100% considerable heterogeneity

Tau….

Unexplained heterogeneity

Differenza fra effetto grande e piccolo.

Non importante se anche l''effetto piccolo è clinicamente significativo.

Rilevante se ci sono differenze clinicamente rilevanti (impatto sul paziente) fra effetto piccolo e effetto grande

Downgrade: dipende

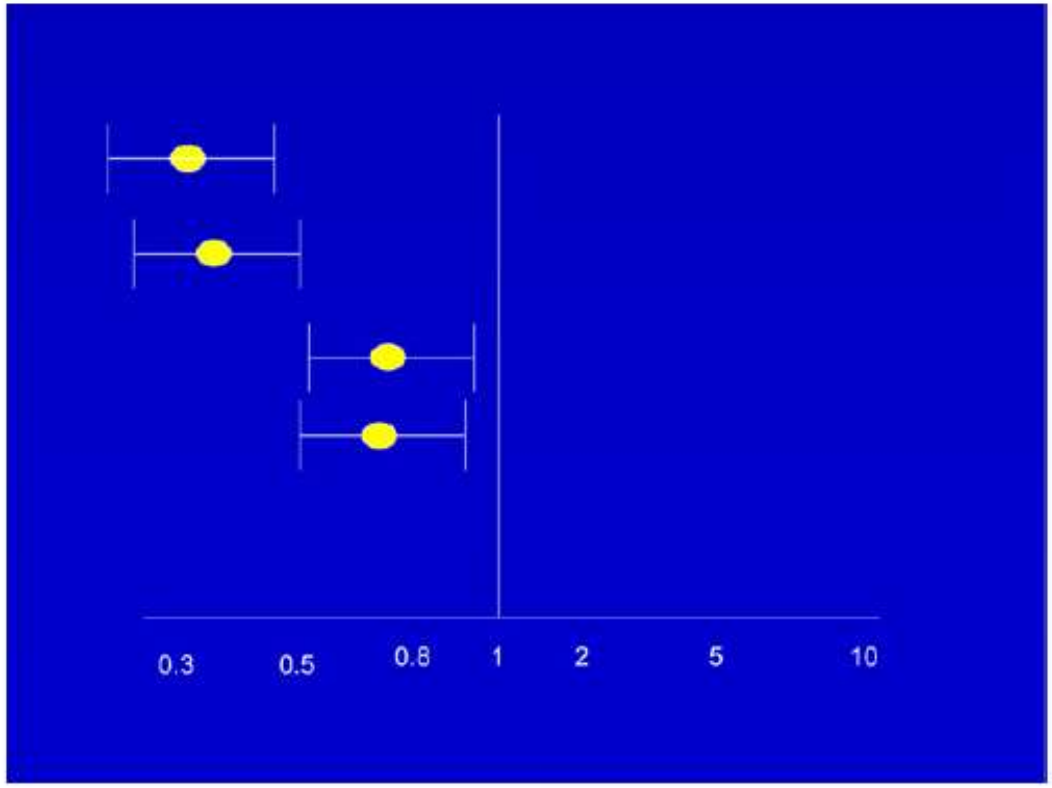


Fig. 2. Substantial heterogeneity, but of questionable importance.

Unexplained heterogeneity

La grandezza della variabilità è la stessa ma in questo caso due studi vanno in una direzione e due in un'altra.

Inconsistency importante

Pooled estimate di non effetto ma con grande eterogeneità

Downgrade: sì

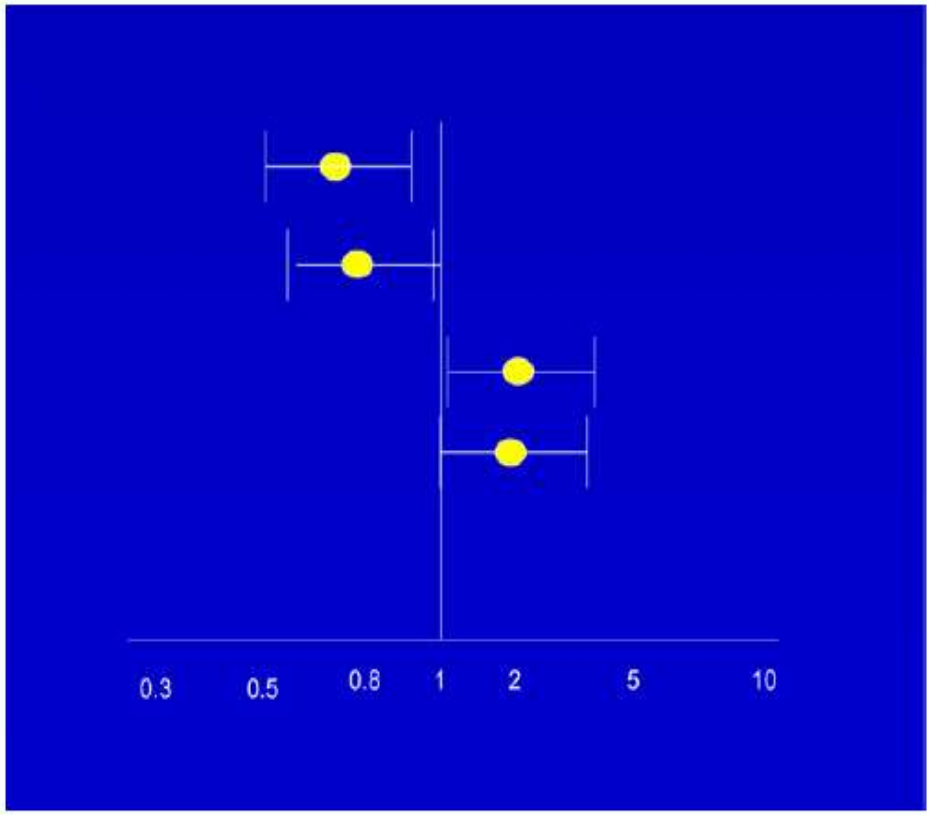


Fig. 3. Substantial heterogeneity, of unequivocal importance.

Unexplained heterogeneity

Pooled estimate di non effetto, come prima, ma in questo caso le differenze fra gli studi sono piccole , tutti concludono per differenze piccole e non significative

Downgrade: no

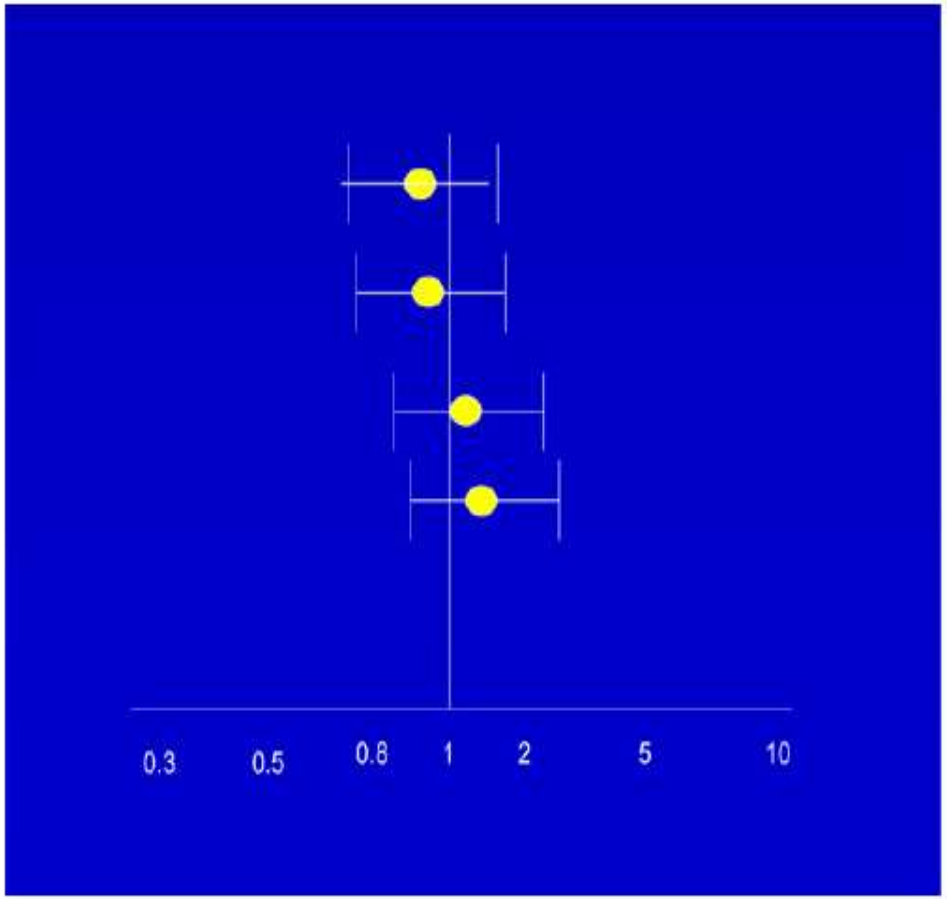


Fig. 1. Differences in direction, but minimal heterogeneity.

Publication bias

- Selective reporting bias: outcome dichiarato nel protocollo ma risultati non riportati: uno dei domini del risk of bias
- Publication bias: mancata pubblicazione di un intero studio

✓English language bias

✓Time lag bias

✓Citation bias

- Sovrastima dell'effetto del trattamento

Publication bias in context

Publication bias and other related biases can be summarised as statistically significant, 'positive' results being:

- more likely to be published (publication bias)
- more likely to be published rapidly (time lag bias)
- more likely to be published in English (language bias)
- more likely to be published more than once (multiple publication bias)
- more likely to be cited by others (citation bias)

[The Cochrane Collaboration]

Publication bias: come valutarne il rischio?

Sezione metodi della revisione:

Cercati studi non pubblicati?

Solo studi in lingua inglese?

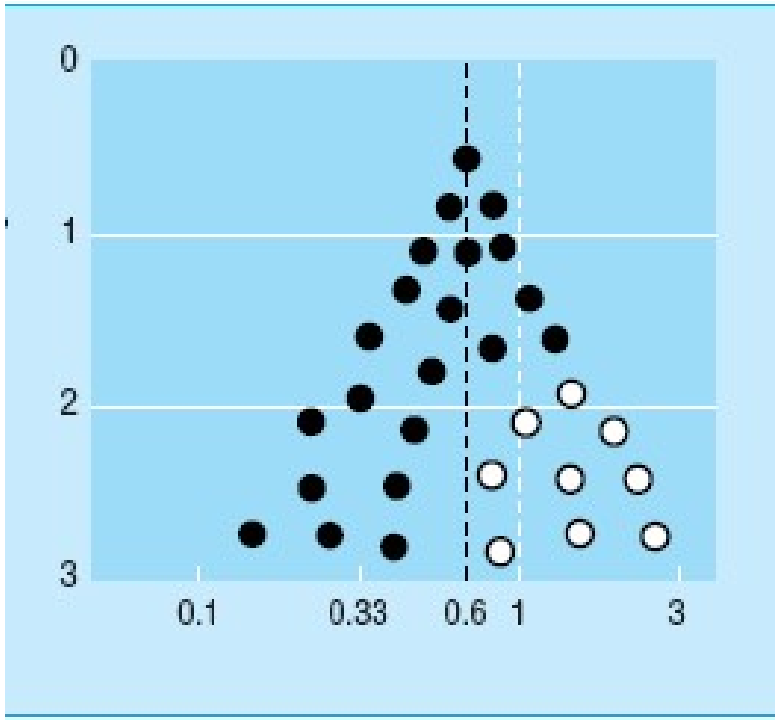
Revisione su nuova terapia introdotta da poco?

Trials piccoli e sponsorizzati da industria farmaceutica?

Funnel plot?

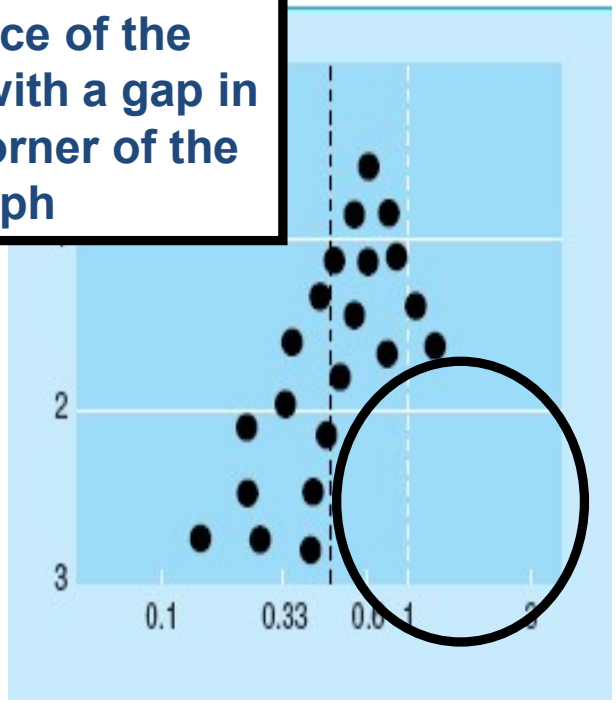
Studi osservazionali? Rischio maggiore perché non obbligatoria registrazione

Publication of All Trials

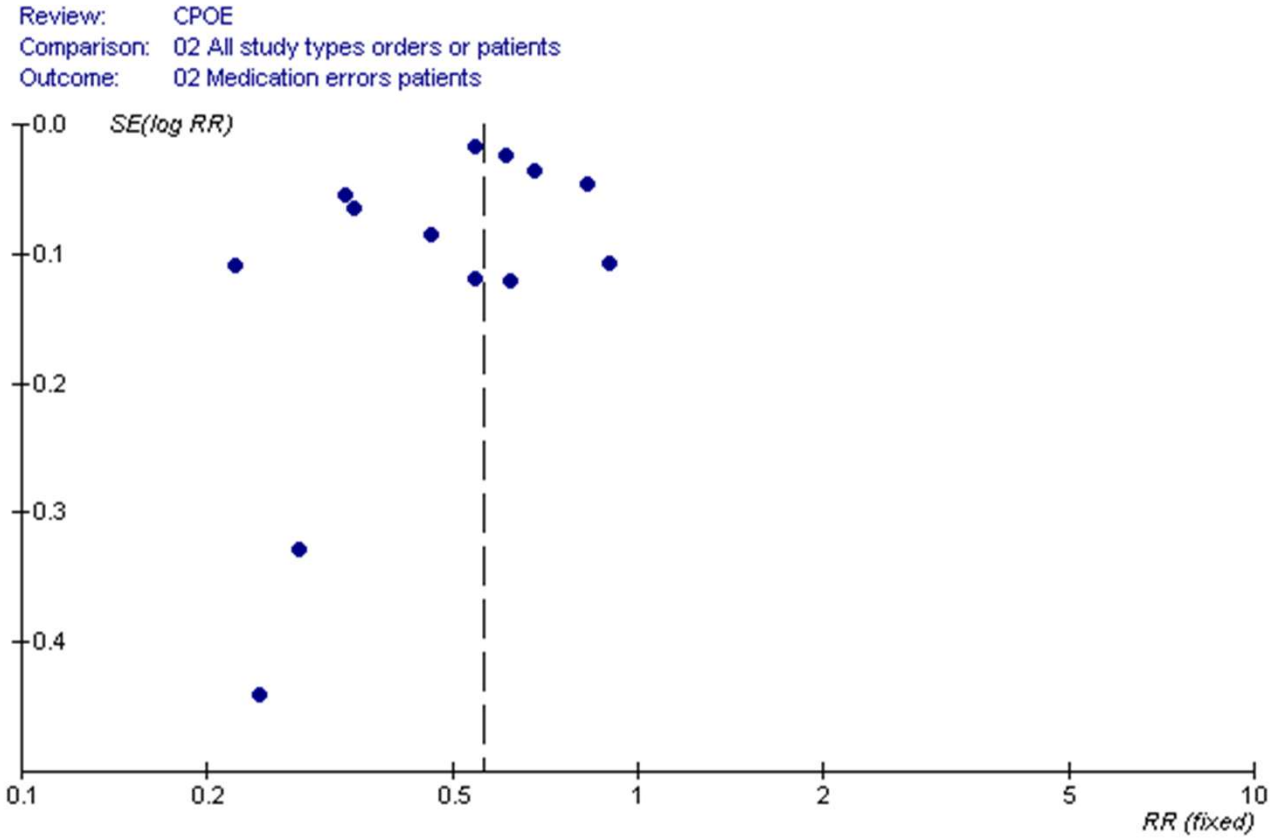


Publication Bias

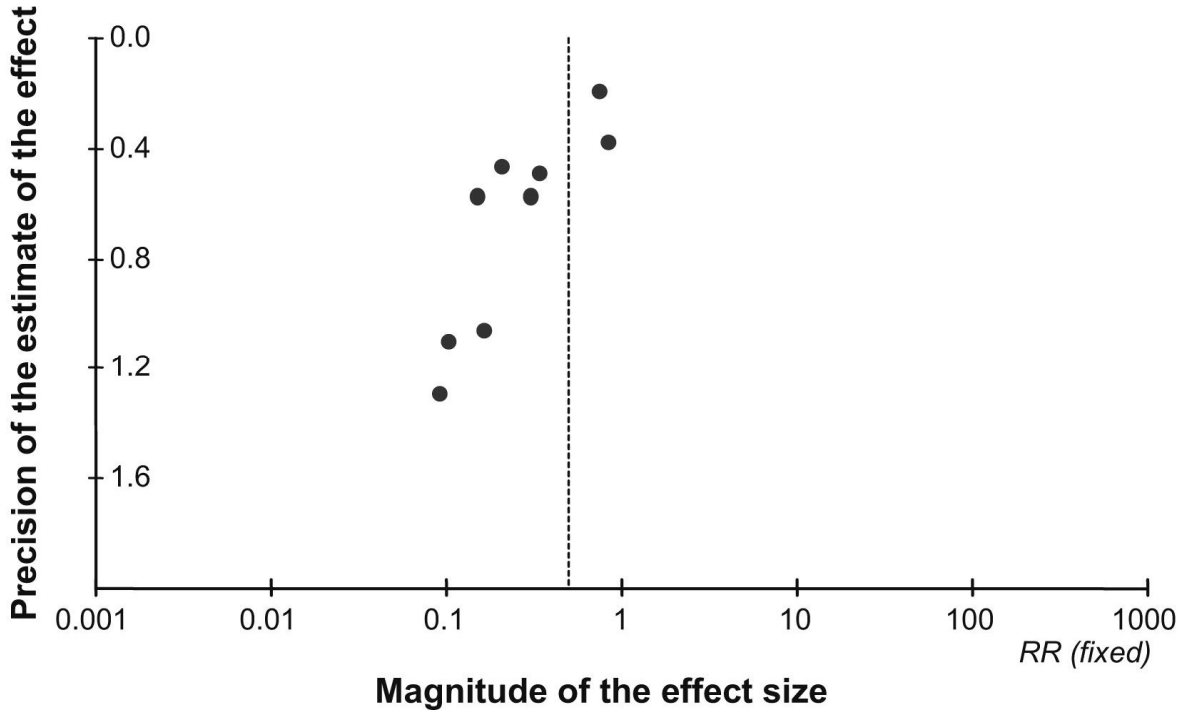
Asymmetrical appearance of the funnel plot with a gap in a bottom corner of the graph



Study 2: Effects of CPOE on medication errors / ADEs



Funnel plot of studies of flavonoids for ameliorating symptoms in patients with hemorrhoids



SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Tabella Sinottica delle Evidenze

Observational studies

Quality starts as low

Then it can be further downgraded for one of the reasons cited above (risk of bias, inconsistency, imprecision, indirectness, publication bias)

Then it can be upgraded ;

The circumstances under which we may wish to rate up the quality of evidence for intervention studies will likely occur infrequently and are primarily relevant to observational studies.

Indeed, although it is theoretically possible to rate up results from randomized control trials (RCTs), we have yet to find a compelling example of such an instance.

What can raise quality?

1. **large magnitude of effect** can upgrade (**RRR 50%/RR 2**)
   - very large two levels **(**RRR 80%/RR 5) ; modeling studies suggests that confounding (from nonrandom allocation) alone is unlikely to explain associations with a relative risk (RR) greater than 2 (or less than 0.5), and very unlikely to explain associations with an RR greater than 5 (or less than 0.2)
   - Es: relationship between infant sleeping position and sudden infant death syndrome (SIDS) found an odds ratio (OR) of 4.1 (95% confidence interval [CI]: 3.1, 5.5) of SIDS occurring with front vs. back sleeping positions

Large magnitude of effect

Further consideration for rating up:

- rapidity of treatment response, and the previous underlying trajectory of the condition

E.g.: we feel confident that hip replacement has a large effect on reducing pain and functional limitations in severe osteoarthritis not only because of the size of the treatment response, but because the natural history of hip osteoarthritis is a progressive deterioration that surgery rapidly and uniformly reverses

- indirect evidence that provides further support for large treatment effects.

E.g. : the effectiveness of antibiotic prophylaxis in a variety of other situations supports observational studies that suggest that antibiotic prophylaxis results in an 89% RR reduction in meningococcal disease in contacts of patients who have suffered the illness

What can raise quality?

2. dose response relation

e.g.: childhood lymphoblastic leukemia: risk for CNS malignancies 15 years after cranial irradiation

- no radiation: 1% (95% CI 0% to 2.1%)
- 12 Gy: 1.6% (95% CI 0% to 3.4%)
- 18 Gy: 3.3% (95% CI 0.9% to 5.6%)

E.g.: systematic review of observational studies on cyclooxygenase-2 inhibitors on cardiovascular events:

- RR 1.33 (95% CI: 1.00, 1.79) with doses less than 25 mg/d
- RR 2.19 (95% CI: 1.64, 2.91) with doses more than 25 mg/d

Residual confounding

- **3.** all **plausible residual confounding** may be working to reduce the demonstrated effect or increase the effect if no effect was observed (underestimate of the treatment effect)

- Es: effect of condom use on HIV infection among men who have sex with men RR: 0.34 [0.21, 0.54] (RRR: 66%) in favor of condom use compared with no condom use. Condom users were more likely to have more partners (but studies did not adjust for this confounding factor in their analyses). Considering the number of partners would, if anything, strengthen the effect estimate in favor of condom use.

Residual confounding

- Es: SR di studi osservazionali: mortalità per tutte le cause in ospedali privati for profit vs ospedali privati not – for profit: RR 1.020, 95% CI 1.003-1.038 : mortalità leggermente superiore in for profit
- Fattori confondenti per cui non hanno aggiustato: livello di reddito dei pazienti e disponibilità risorse ospedali (> in for profit) che dovrebbe essere associato a < mortalità. Quindi in caso di aggiustamento l'effetto (RR) sarebbe stato maggiore

La Table of evidence (ToEs)

Evidence profile: use of antibiotics (penicillin) versus no use of antibiotics in children with sickle cell disease. Source: Hirst et al. 4

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Penicillin prophylaxis	Standard care	Relative (95% CI)	Absolute (95% CI)		
Incidence of pneumococcal infection, for initiation of treatment - Initiation of penicillin												
2	Randomized trials	Not serious [1]	serious [2]	not serious [3]	serious [4]	none [5]	9/248 (3.6%)	19/209 (9.1%)	**OR 0.37** (0.16 to 0.86)	55 fewer per 1000 (from 12 fewer to 75 fewer)	⊕⊕◯◯ LOW	CRITICAL
Deaths, for initiation of treatment - Initiation of penicillin												
1	randomized trials	not serious [6]	not serious	not serious	serious [4]	none [5]	0/105 (0.0%)	4/110 (3.6%)	**OR 0.11** (0.01 to 2.11)	32 fewer per 1000 (from 36 fewer to 37 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Adverse drug effects - Nausea and vomiting												
1	randomized trials	not serious [6]	not serious	not serious	serious [4]	none [5]	2/201 (1.0%)	1/199 (0.5%)	**OR 1.99** (0.18 to 22.12)	5 more per 1000 (from 4 fewer to 95 more)	⊕⊕⊕◯ MODERATE	CRITICAL

1.blinding and concealment were not clear for one of the two studies
2.heterogeneity exists; p-value for testing heterogeneity is 0.07 and I2=69%
3.the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
4.total sample size is small, and the total number of events is <300 (a threshold rule-of-thumb value)
5.insufficient number of studies to assess publication bias
6.unclear allocation concealment

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Evidence to Decision Framework
Formulazione delle Raccomandazioni

The Evidence-to-Decision framework

GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction

Pablo Alonso-Coello,[1,2] Holger J Schünemann,[2,3] Jenny Moberg,[4] Romina Brignardello-Petersen,[2,5] Elie A Akl,[2,6] Marina Davoli,[7] Shaun Treweek,[8] Reem A Mustafa,[2,9] Gabriel Rada,[10,11,12] Sarah Rosenbaum,[4] Angela Morelli,[4] Gordon H Guyatt,[2,3] Andrew D Oxman[4] the GRADE Working Group

BMJ 2016;353:i2016

GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines

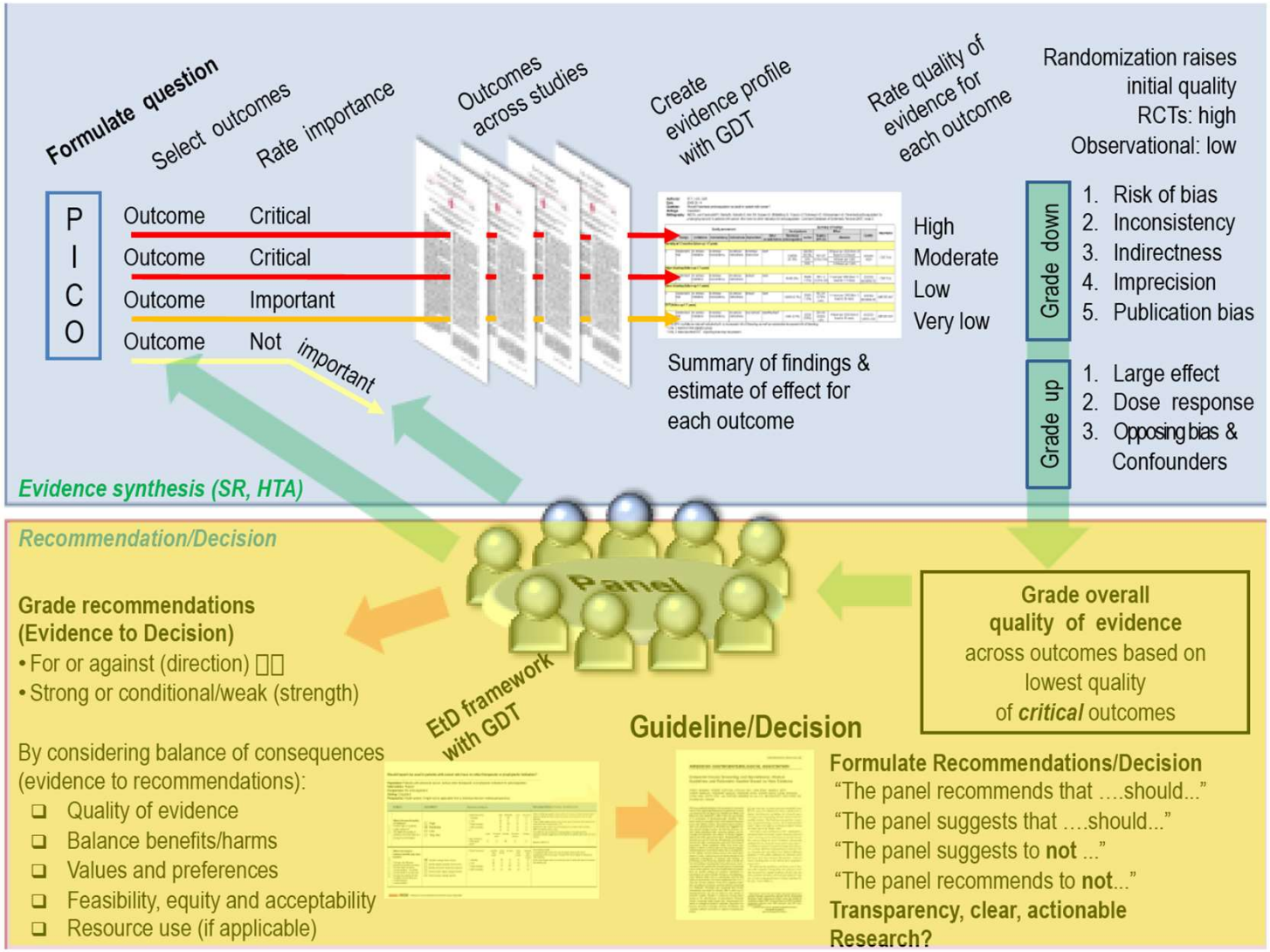
Pablo Alonso-Coello,[1,2] Andrew D Oxman,[3] Jenny Moberg,[3] Romina Brignardello-Petersen,[2,4] Elie A Akl,[2,5] Marina Davoli,[6] Shaun Treweek,[7] Reem A Mustafa,[2,8] Per O Vandvik,[3] Joerg Meerpohl,[9] Gordon H Guyatt,[2,10] Holger J Schünemann,[2,10] the GRADE Working Group

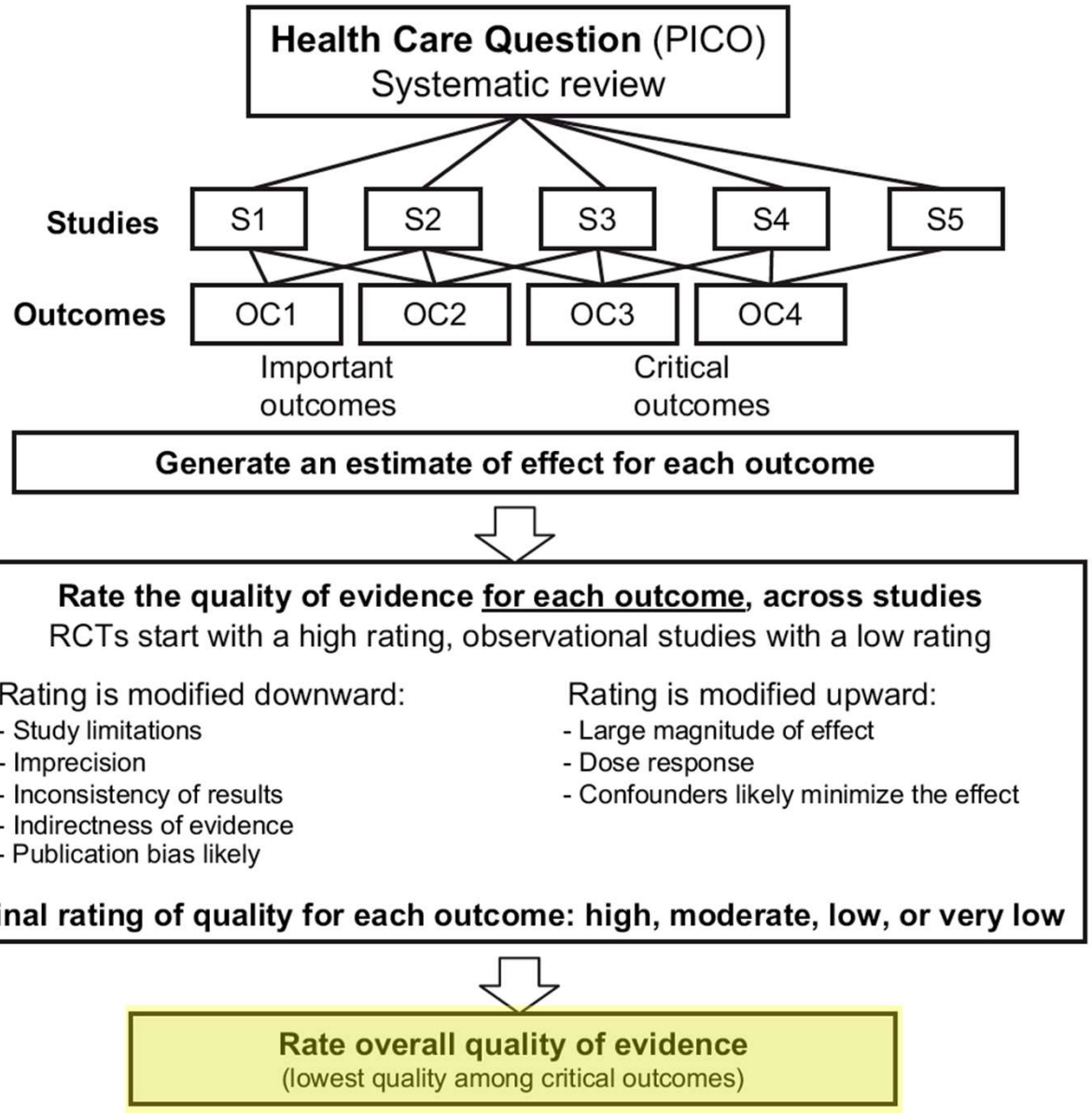
BMJ 2016;353:i2089

SUMMARY POINTS

- Explicit and transparent systems for decision making can help to ensure that all important criteria are considered and that decisions are informed by the best available research evidence
- The purpose of Evidence to Decision (EtD) frameworks is to help people use evidence in a structured and transparent way to inform decisions in the context of clinical recommendations, coverage decisions, and health system or public health recommendations and decisions
- EtD frameworks inform users about the judgments that were made and the evidence supporting those judgments by making the basis for decisions transparent to target audiences

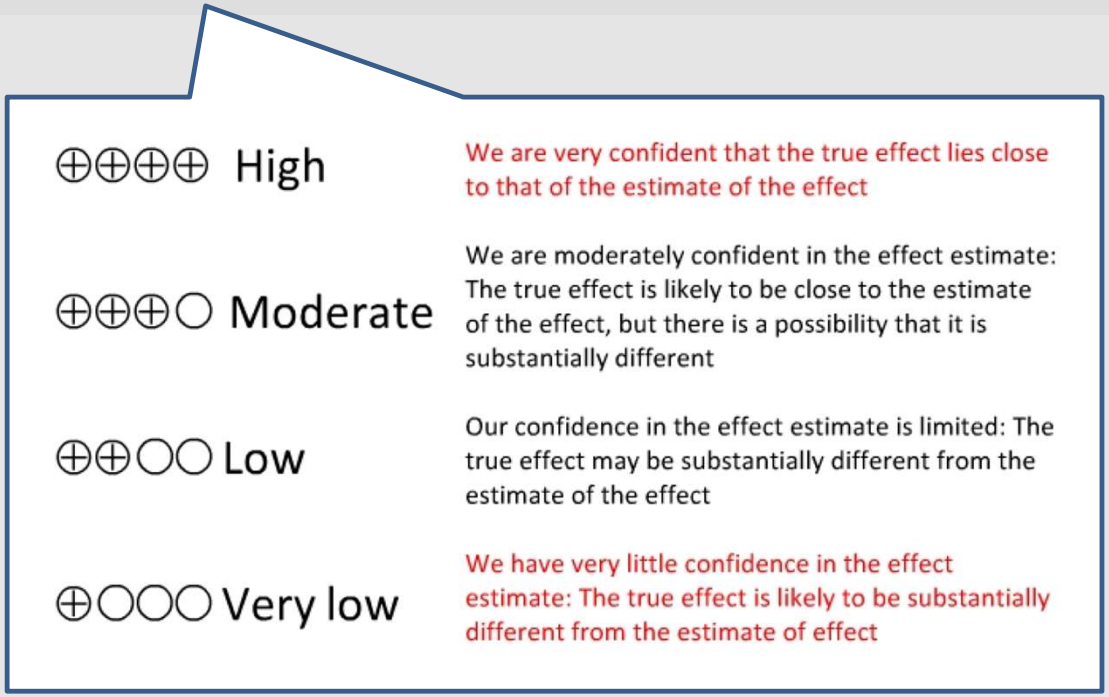
The GRADE process in developing guidelines




Health Care Question (PICO)
Systematic review
Studies
S1
S2
S3
S4
S5
Outcomes
OC1
OC2
OC3
OC4
Important outcomes
Critical outcomes
Generate an estimate of effect for each outcome
Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating
Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely
Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect
Final rating of quality for each outcome: high, moderate, low, or very low
Rate overall quality of evidence
(lowest quality among critical outcomes)

Box 4. Criteri considerati nel GRADE EtD framework

Qualità delle prove: Qual è la qualità globale delle evidenze?



Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

Grading quality of evidence and strength of recommendations

GRADE Working Group

Overall quality of evidence

Other systems have commonly based judgments of the overall quality of evidence on the quality of evidence for the benefits of interventions.

When the risk of an adverse effect is critical for a judgment, and evidence regarding that risk is weaker than evidence of benefit, ignoring uncertainty about the risk of harm is problematic.

We suggest that the lowest quality of evidence for any of the outcomes that are critical to making a decision should provide the basis for rating overall quality of evidence.

Valutazione della qualità globale delle prove

Dopo la valutazione della qualità per i singoli outcome importanti effettuata dall'ERT (Sezione 4.5) si deve formulare il giudizio complessivo di qualità. Il metodo GRADE suggerisce di procedere considerando soltanto gli outcome critici per la formulazione della raccomandazione relativa al quesito clinico.

Valutazione della qualità globale delle prove

Dopo la valutazione della qualità per i singoli outcome importanti effettuata dall'ERT (Sezione 4.5) si deve formulare il giudizio complessivo di qualità. Il metodo GRADE suggerisce di procedere considerando soltanto gli outcome critici per la formulazione della raccomandazione relativa al quesito clinico. Se la qualità è diversa fra i singoli outcome critici, il metodo indica la seguente linea di comportamento:

- se i risultati vanno in direzioni opposte (es. il trattamento oggetto della raccomandazione è migliore in termini di efficacia ma peggiore per quanto riguarda gli effetti indesiderati), la qualità globale viene attribuita basandosi sulla valutazione peggiore ossia assumendo come più rappresentativo l'outcome che ha ottenuto la più bassa valutazione di qualità;
- se i risultati vanno nella stessa direzione per tutti gli outcome (beneficio o danno), viene assunta come qualità globale delle prove la qualità di un singolo outcome critico che da solo basterebbe per formulare compiutamente la raccomandazione;

Valutazione della qualità globale delle prove

Dopo la valutazione della qualità per i singoli outcome importanti effettuata dall'ERT (Sezione 4.5) si deve formulare il giudizio complessivo di qualità. Il metodo GRADE suggerisce di procedere considerando soltanto gli outcome critici per la formulazione della raccomandazione relativa al quesito clinico. Se la qualità è diversa fra i singoli outcome critici, il metodo indica la seguente linea di comportamento:

- se i risultati vanno in direzioni opposte (es. il trattamento oggetto della raccomandazione è migliore in termini di efficacia ma peggiore per quanto riguarda gli effetti indesiderati), la qualità globale viene attribuita basandosi sulla valutazione peggiore ossia assumendo come più rappresentativo l'outcome che ha ottenuto la più bassa valutazione di qualità;
- se i risultati vanno nella stessa direzione per tutti gli outcome (beneficio o danno), viene assunta come qualità globale delle prove la qualità di un singolo outcome critico che da solo basterebbe per formulare compiutamente la raccomandazione;

Box 4. Criteri considerati nel GRADE EtD framework

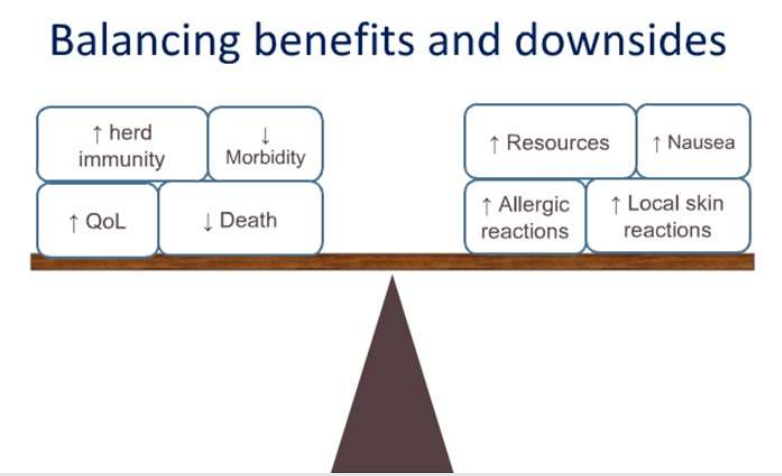
Qualità delle prove: Qual è la qualità globale delle evidenze?

Benefici attesi: Quanto sono consistenti i benefici attesi?

Effetti indesiderati: Quanto sono consistenti gli effetti indesiderati?

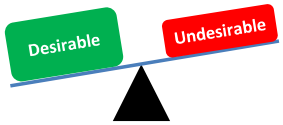
Valori: Esiste un importante grado di incertezza o variabilità rispetto al valore che le persone attribuiscono agli outcome principali?

Bilancio tra benefici e danni attribuibili all'intervento: Il bilanciamento tra benefici ed effetti indesiderati favorisce l'intervento proposto o il controllo?

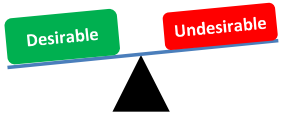


Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

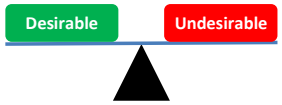
Based on the evidence presented, do the desirable consequences outweigh the undesirable consequences, or vice versa?



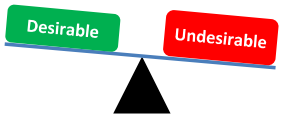
Desirable consequences *clearly outweigh* undesirable consequences in most settings



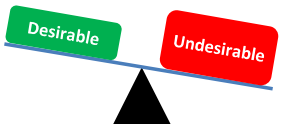
Desirable consequences *probably outweigh* undesirable consequences in most settings



The balance between desirable and undesirable consequences *is closely balanced or uncertain*



Undesirable consequences *probably outweigh* desirable consequences in most settings



Undesirable consequences *clearly outweigh* desirable consequences in most settings

Box 4. Criteri considerati nel GRADE EtD framework

Qualità d

Benefici

Effetti in

Valori: E che le persone attribuiscono agli outcome

Bilancio fici ed effetti indesiderati favorisce l'intervento proposto

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: Incorporating considerations of resources use into grading recommendations

Guideline panellists have differing opinions on whether resource use should influence decisions in individual patients. As medical care costs rise, resource use considerations become more compelling, but panellists may find dealing with such considerations challenging

Gordon H Guyatt

BMJ | 24 MAY 2008 | VOLUME 336

Risorse richieste:

- Quali sono i costi richiesti?
- Qual è il livello di evidenza relativo ai costi richiesti?
- L'analisi di costo-efficacia o di costo-utilità favorisce l'intervento proposto o il controllo?

Il metodo GRADE consente di tralasciare tale opzione quando non vi siano le evidenze a supporto di una corretta analisi costo/efficacia. Applicando l'EtD non verrà quindi valutata la dimensione costo-efficacia a meno che evidenze del genere siano trovate in fase di selezione degli studi.

Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

Box 4. Criteri considerati nel GRADE EtD framework

Qualità delle prove: Qual è la qualità globale delle evidenze?

Benefici attesi: Quanto sono consistenti i benefici attesi?

Effetti indesiderati: Quanto sono consistenti gli effetti indesiderati?

Valori: Esiste un importan… buiscono agli outcome principali?

Bilancio tra benefici e da… erati favorisce l'intervento proposto o il co…

Risorse richieste:

- Quali sono i costi richie…
- Qual è il livello di eviden…
- L'analisi di costo-efficacia o di costo-… favorisce l'intervento proposto o il controllo?

Equità: Quale potrebbe essere l'impatto sull'equità?

- ✓ Miglioramento di attuali situazioni di disparità
- ✓ Assicurazioni sul fatto che l'implementazione della raccomandazione non vada a creare disparità
- ✓ Possibile creazione di disparità

Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

Box 4. Criteri considerati nel GRADE EtD framework

Qualità delle prove: Qual è la qualità globale delle evidenze?

Benefici attesi: Quanto sono consistenti i benefici attesi?

Effetti indesiderati: Quanto

Valori: Esiste un importan agli outcome principali?

Bilancio tra benefici e dan sce l'intervento proposto o il con

Risorse richieste:

- Quali sono i costi richies
- Qual è il livello di eviden
- L'analisi di costo-efficacia

Equità: Quale potrebbe essere l'impatto su

Accettabilità: L'intervento proposto è accettabile da parte degli stakeholder?

- ✓ Non accettazione di costi o effetti indesiderati a breve termine Vs benefici futuri
- ✓ Non accordo con l'importanza (valore) dei benefici e degli effetti indesiderati (a causa di interessi personali o della percezione dell'importanza relativa degli effetti)
- ✓ Disapprovazione dell'intervento, per principi di etica o di giustizia

Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

Box 4. Criteri considerati nel GRADE EtD framework

Qualità delle prove: Qual è la qualità globale delle evidenze?

Benefici attesi: Quanto sono consistenti i benefici attesi?

Effetti indesiderati: Quanto s

Valori: Esiste un importante

outcome principali?

Bilancio tra benefici e dann

l'intervento proposto o il contro

Risorse richieste:

- Quali sono i costi richiesti?
- Qual è il livello di evidenza
- L'analisi di costo-efficacia

Equità: Quale potrebbe essere l'impatto sul

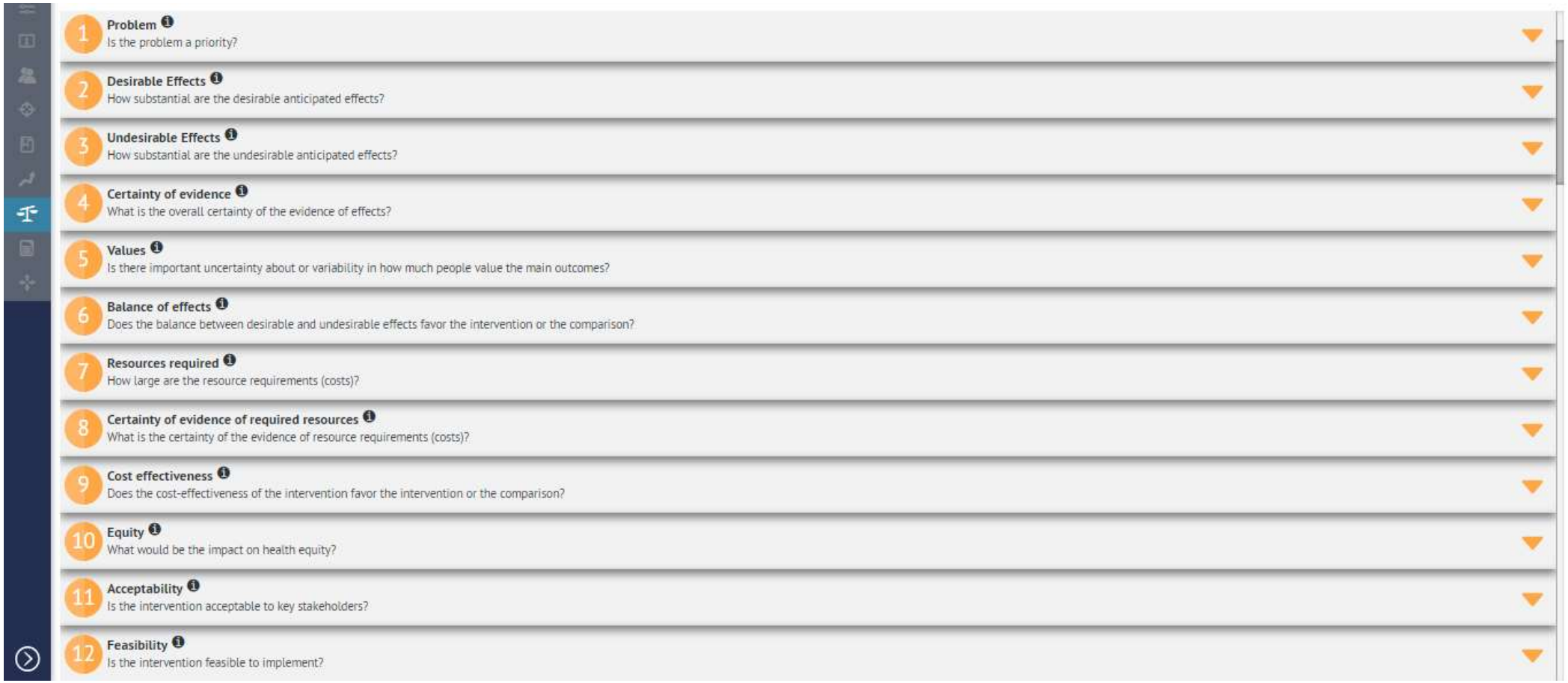
Accettabilità: L'intervento proposto è accett da parte degli stakeholder?

Fattibilità: L'intervento proposto può essere implementato?

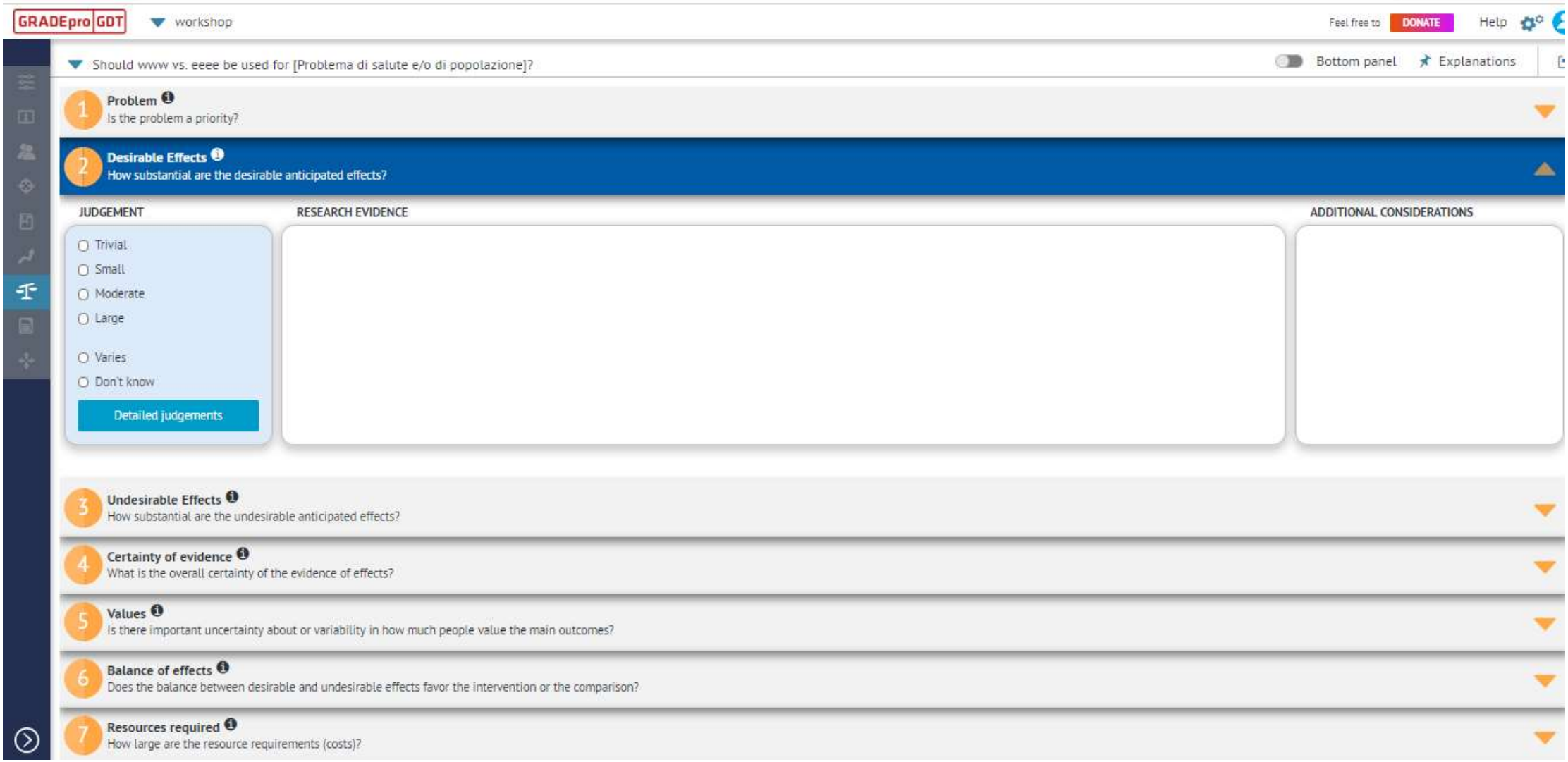
- ✓ Sostenibilità dell'intervento
- ✓ Esistenza di barriere tali da limitare la fattibilità dell'intervento
- ✓ Possibilità di verificare l'appropriatezza dell'intervento
- ✓ Preoccupazioni riguardo alla accessibilità all'intervento
- ✓ Leggi/regolamenti potenzialmente limitanti l'implementazione della raccomandazione

Centro nazionale per l'eccellenza clinica, la qualità e la sicurezza delle cure. Manuale metodologico per la produzione di linee guida di pratica clinica

GRADEpro GDT

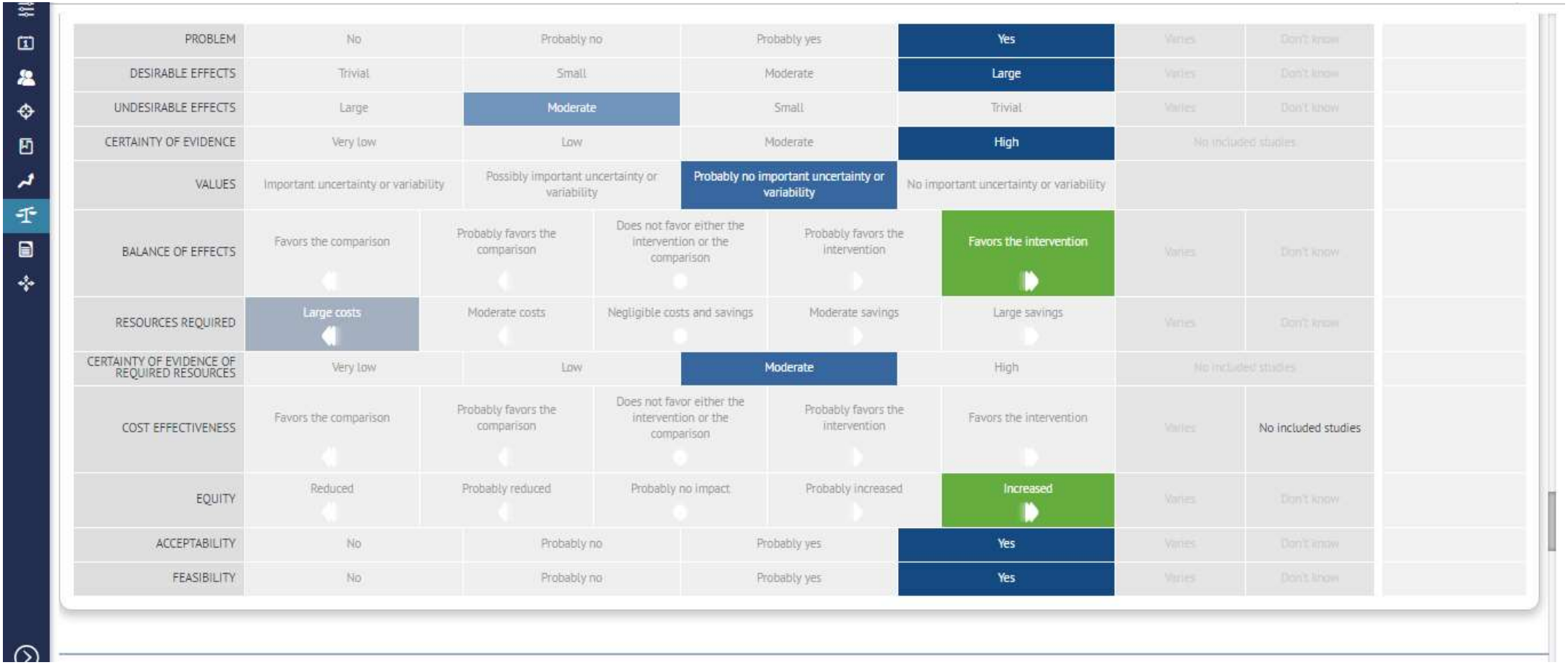


GRADEpro GDT



GRADEpro GDT

Immunoterapia nel melanoma



PROBLEM	No	Probably no	Probably yes	**Yes**	Varies	Don't know
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	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 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

GRADEpro GDT

Anti-EGFR vs chemio in NSCLC

CRITERIA	SUMMARY OF JUDGEMENTS						IMPORTANCE FOR DECISION
PROBLEM	No	Probably no	Probably yes	**Yes**	Varies	Don't know	
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	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 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	

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

La *Forza* delle Raccomandazioni

Defining Strength of recommendation

- The extent to which a guideline panel is confident that desirable effects of an intervention outweigh undesirable effects, or vice versa, across the range of patients for whom the recommendation is intended.

Strength of recommendation

- A recommendation can have one of 2 strength:
  - **Strong** : panel is confident that the desirable effects of adherence to the recommendation outweigh the undesirable effects (or vice versa).
  - **Conditional** : panel concludes that the desirable effects of adherence to the recommendation probably outweigh the undesirable effects (or vice versa), but is not confident.

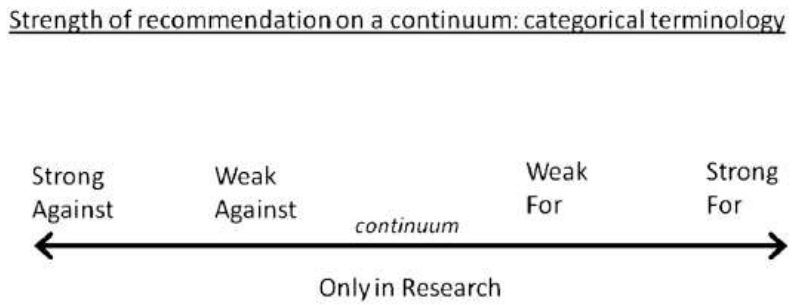


Fig. 1. Strength of recommendation: a continuum divided into categories.

Implications of strong and conditional recommendations

	Strong recommendation	Conditional recommendation
Patients	Most people in your situation would want the recommended course of action and only a small proportion would not	The majority of people in your situation would want the recommended course of action, but many would not
Clinicians		
Policy makers		

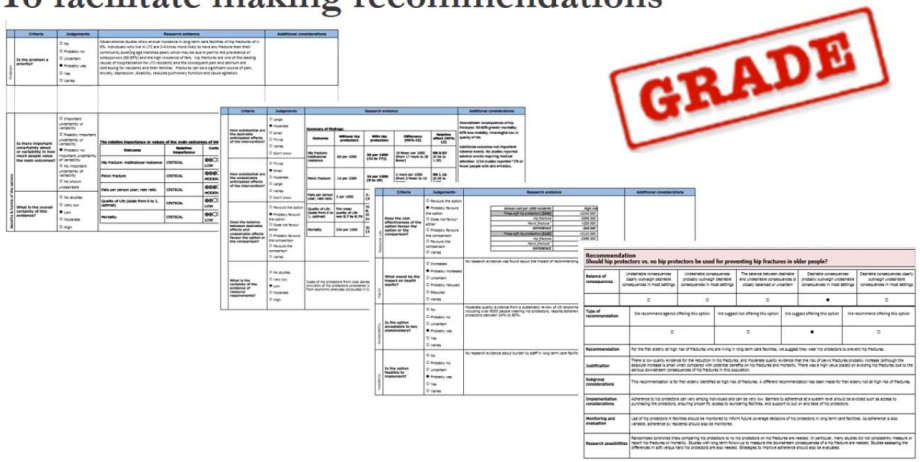
Implications of strong and conditional recommendations

	Strong recommendation	Conditional recommendation
Patients	Most people in your situation would want the recommended course of action and only a small proportion would not	The majority of people in your situation would want the recommended course of action, but many would not
Clinicians	Most patients should receive the recommended course of action	Be prepared to help patients to make a decision that is consistent with their own values
Policy makers		

Implications of strong and conditional recommendations

	Strong recommendation	Conditional recommendation
Patients	Most people in your situation would want the recommended course of action and only a small proportion would not	The majority of people in your situation would want the recommended course of action, but many would not
Clinicians	Most patients should receive the recommended course of action	Be prepared to help patients to make a decision that is consistent with their own values
Policy makers	The recommendation can be adapted as a policy in most situations	There is a need for substantial debate and involvement of stakeholders

Evidence to Decision Table:
To facilitate making recommendations



GRADE

Criteria

- Problem/priority?
- Benefits & harms of the options
- Certainty of evidence
- Values
- Resource use
- Equity
- Acceptability
- Feasibility

Evidence from Systematic Reviews

Formulazione delle Raccomandazioni LG AIOM (14 febbraio 2021)

Forza della raccomandazione clinica	Terminologia	Significato
Forte a Favore	"Nei pazienti con (criteri di selezione) l'intervento xxx dovrebbe essere preso in considerazione come opzione terapeutica di prima intenzione"	l'intervento in esame dovrebbe essere considerato come prima opzione terapeutica (**evidenza che i benefici sono prevalenti sui danni**)
Condizionata a Favore	"Nei pazienti con (criteri di selezione) l'intervento xxx può essere preso in considerazione come opzione terapeutica di prima intenzione, in alternativa a yyy"	l'intervento in esame può essere considerato come opzione di prima intenzione, consapevoli dell'esistenza di alternative ugualmente proponibili (**incertezza riguardo alla prevalenza dei benefici sui danni**)
Condizionata a Sfavore	"Nei pazienti con (criteri di selezione) l'intervento xxx non dovrebbe essere preso in considerazione come opzione terapeutica di prima intenzione, in alternativa a yyy"	l'intervento in esame non dovrebbe essere considerato come opzione di prima intenzione; esso potrebbe comunque essere suscettibile di impiego in casi altamente selezionati e previa completa condivisione con il paziente (**incertezza riguardo alla prevalenza dei danni sui benefici**)
Forte a Sfavore	"Nei pazienti con (criteri di selezione) l'intervento xxx non deve essere preso in considerazione come opzione terapeutica di prima intenzione"	l'intervento in esame non deve essere in alcun caso preso in considerazione (**evidenza che i danni sono prevalenti sui benefici**)

SACRO CUORE DON CALABRIA I R C C S

ESMO Designated Centers of Integrated Oncology and Palliative Care

7ª EDIZIONE

STUDI CLINICI: METODOLOGIA

Coordinatore
Dr.ssa Stefania Gori

3° MODULO

LINEE GUIDA: METODO GRADE
FORMAZIONE AVANZATA

Modalità Webinar
RES-Videoconferenza
in diretta da
Negrar di Valpolicella

12-13 Marzo
2021

GRADE

Sistema Nazionale Linee Guida: quali strumenti vengono utilizzati per valutare le Linee Guida Italiane?

Accreditamento delle linee guida

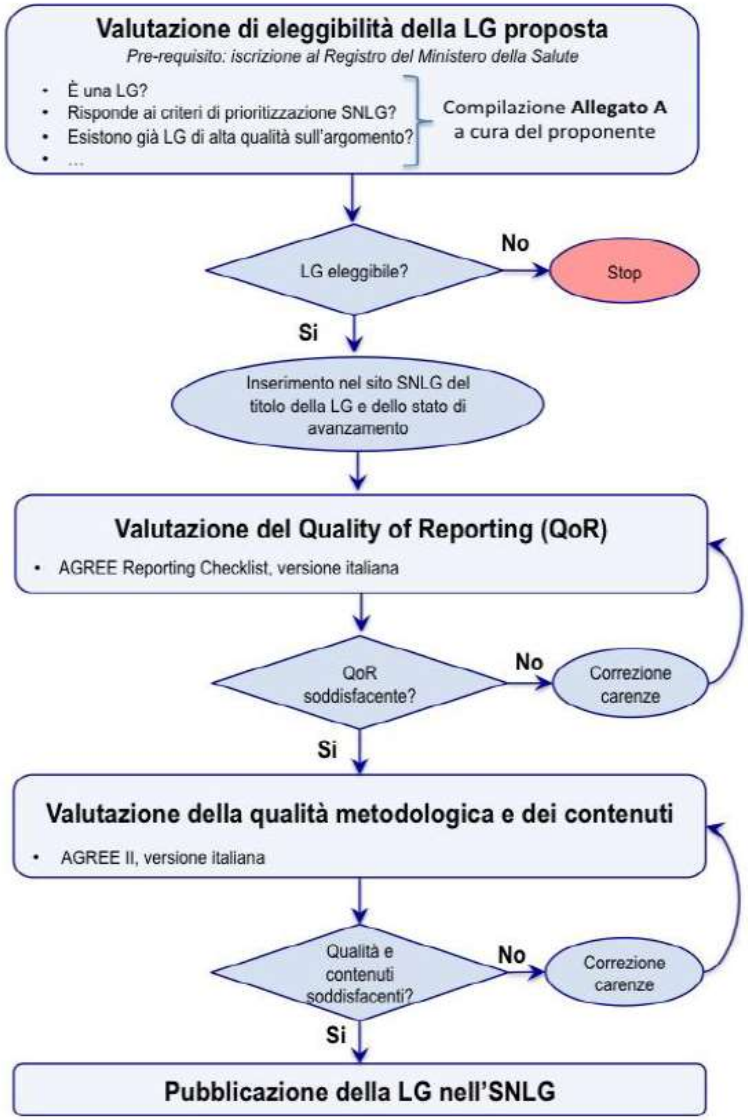
l'Italia si munisce quindi di un sistema di accreditamento, monitoraggio ed aggiornamento delle linee guida (2017)

- Attività di **verifica su linee guida** che, ai sensi del comma l, sono «**elaborate da**:
  - **enti** ed **istituzioni** pubblici e privati nonché dalle **società scientifiche** e dalle **associazioni tecnico-scientifiche** delle professioni sanitarie iscritte in **apposito elenco** istituito e regolamentato con decreto del Ministro della salute,
    - da aggiornare con cadenza biennale

Courtesy of Primiano Iannone

Processo di valutazione delle LG proposte da soggetti ex art.5 L. n.24/17 per la pubblicazione nell'SNLG

Courtesy of Primiano Iannone



Le richieste di valutazione vanno inviate online attraverso la **piattaforma SNLG**

Workflow sviluppato dal CNEC con il supporto del centro collaboratore **GIMBE**

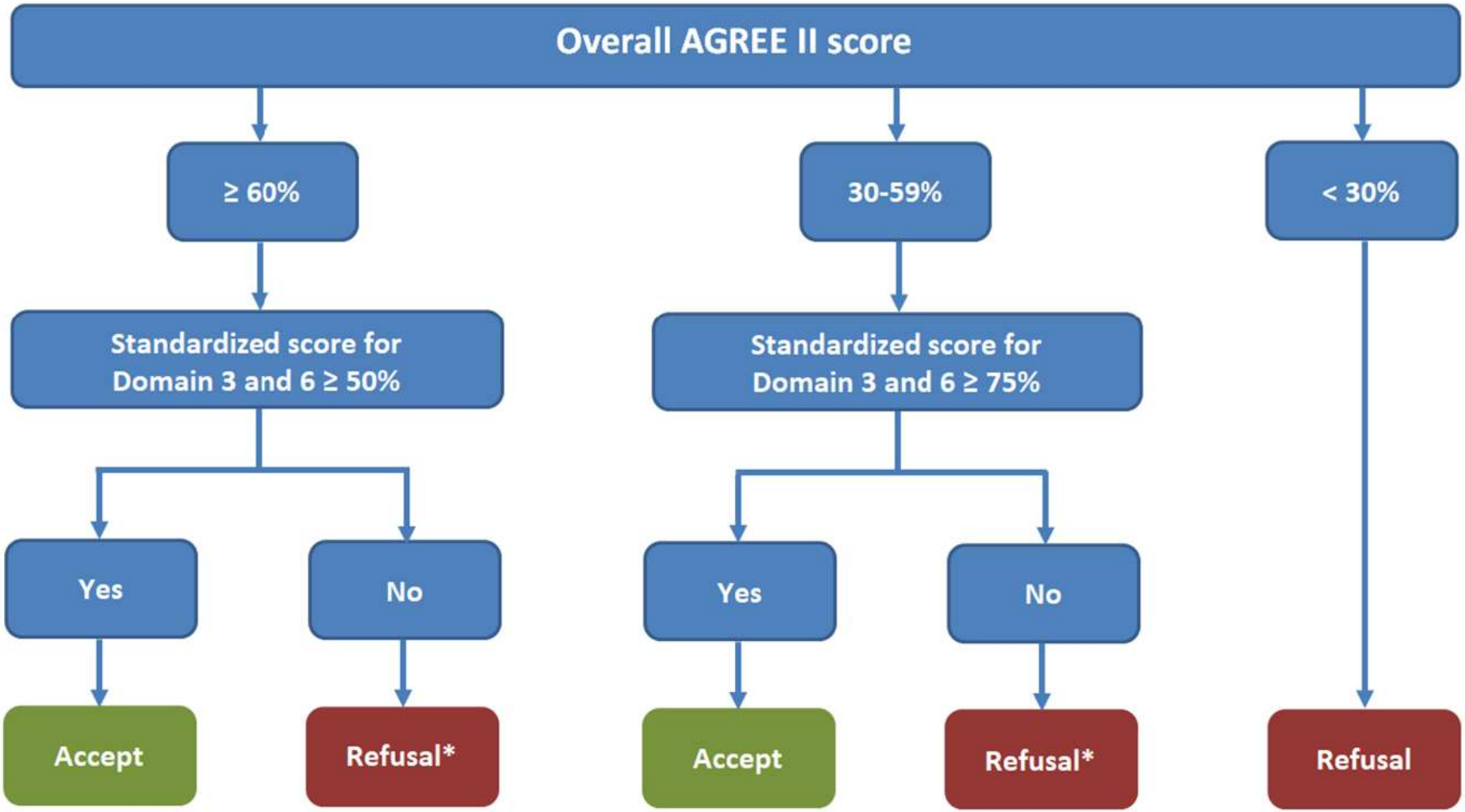
AGREE II e AGREE Reporting Checklist

APPRAISAL OF GUIDELINES FOR RESEARCH & EVALUATION

Dimensione 1	OBIETTIVI E AMBITI DI APPLICAZIONE (1-3)
Dimensione 2	COINVOLGIMENTO DEGLI STAKEHOLDER (4-6)
Dimensione 3	RIGORE METODOLOGICO (7-14)
Dimensione 4	CHIAREZZA ESPOSITIVA (15-17)
Dimensione 5	APPLICABILITÀ (18-21)
Dimensione 6	INDIPENDENZA EDITORIALE (22-23)

benefici potenziali LG proporzionali alla loro qualità → **strumento di valutazione standardizzato**

Courtesy of Primiano Iannone



* Accept with reserve if no other CPG on specific topic is available

Courtesy of Primiano Iannone

Archivi categoria: *Produzione*

Strumenti per gli sviluppatori di LG

In questa sezione sono riportati gli **standard metodologici di riferimento** che il CNEC raccomanda per la produzione e la valutazione critica di LG destinate a essere pubblicate nel sito SNLG.

Manuale metodologico ISS per la produzione di LG: rappresenta il nuovo manuale di riferimento per la produzione di LG dell'Istituto Superiore di Sanità. Incorpora la metodologia GRADE e tiene conto dell'esperienza maturata nella produzione di LG da parte delle maggiori organizzazioni internazionali dedicate allo scopo, adattata al contesto italiano.

E' disponibile una nuova versione del manuale [v. 1.3.2, scarica qui]. Rispetto alla v .1.3, sono state apportate delle modifiche, tra cui:

- aggiornamento Sezione 6.1

Il metodo GRADE [scarica file]: è il metodo adottato da un numero sempre maggiore di organizzazioni internazionali e agenzie di sanità pubblica quale standard di riferimento per la valutazione della qualità delle prove e la produzione di raccomandazioni cliniche che tengano conto in modo esplicito dei diversi fattori che, oltre alla qualità delle evidenze, condizionano la forza e la direzione delle raccomandazioni, incluso un chiaro bilancio dei rischi e dei benefici delle opzioni

Versione 2019

Sviluppato dal CNEC in collaborazione con il GRADE Working Group

Perché il metodo GRADE?

Vantaggi del GRADE rispetto precedenti sistemi di grading:

- netta separazione tra valutazione della qualità delle evidenze e forza delle raccomandazioni
- valutazione esplicita dell'importanza degli outcome di strategie alternative
- criteri espliciti per aumentare e ridurre il rating della qualità delle evidenze
- processo sistematico e trasparente che porta dalle evidenze alle raccomandazioni
- considerazione dei valori e delle preferenze dei pazienti e dei professionisti
- interpretazione chiara delle raccomandazioni forti e deboli per medici, pazienti e *policy maker*

Courtesy of Primiano Iannone

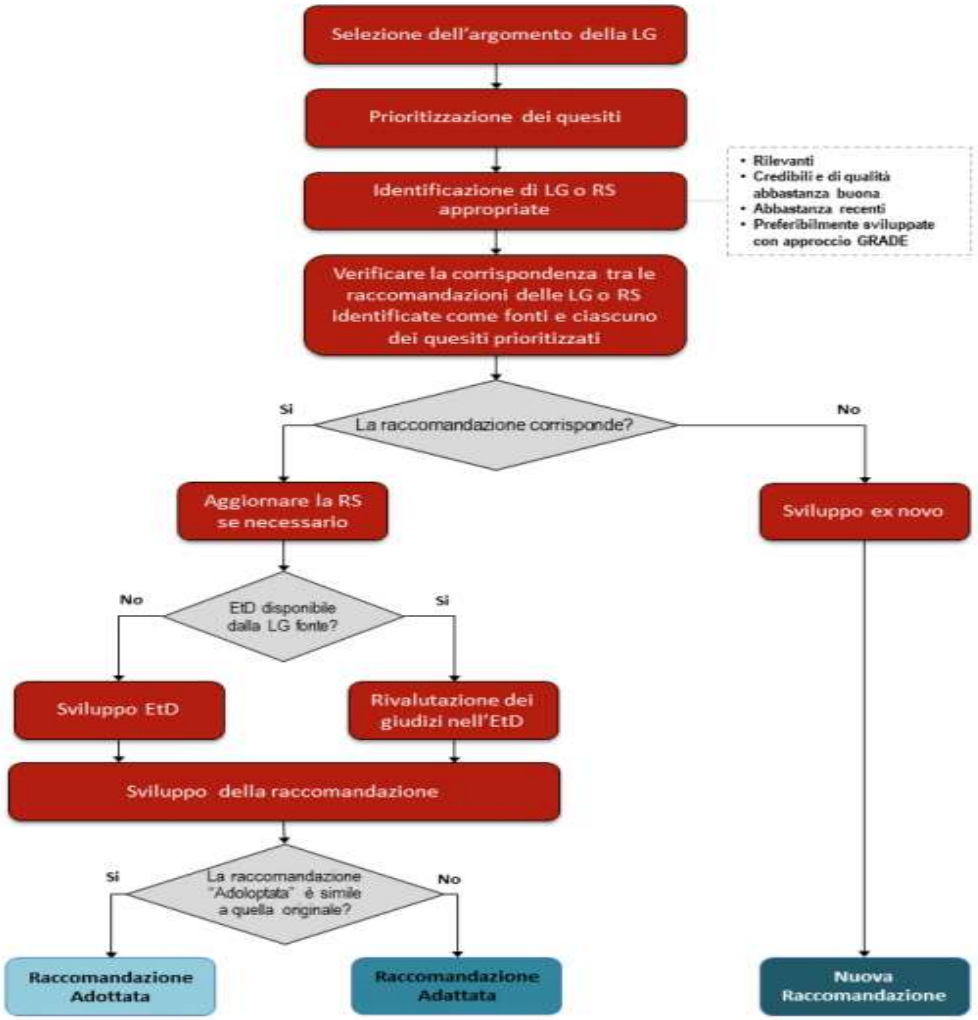


Figura 3 – Tradotta da: *Journal of Clinical Epidemiology* 2017 81, 101-110DOI: (10.1016/j.jclinepi.2016.09.009)

Istituto Superiore di Sanità

Piattaforma SNLG 2.0

Finalità

Inserimento e valutazione di linee guida per la pubblicazione nel Sistema Nazionale Linee Guida

Gli utenti abilitati* possono richiedere la valutazione di linee guida (LG) per la loro pubblicazione nel SNLG in linea con la legge n° 24/2017 e relativi decreti attuativi.

Il CNEC:

1. verifica l'eleggibilità della LG in base a requisiti di priorità e non ridondanza
2. verifica la presenza nella LG di una previsione di impatto delle raccomandazioni sui Livelli Essenziali di Assistenza (LEA)
3. valuta la LG con criteri espliciti in termini di qualità del reporting, metodologia adottata e rilevanza delle raccomandazioni rispetto alle evidenze citate
4. invia eventuali feedback al proponente per correzioni e revisioni
5. pubblica la LG nel SNLG in caso di esito positivo della valutazione

Per i dettagli riferirsi al manuale operativo.

**Enti e istituzioni pubbliche e private, società scientifiche e associazioni tecnico-scientifiche delle professioni sanitarie iscritte in apposito elenco istituito e regolamentato con DM 2 agosto 2017 (GU n. 186 del 10-8-2017).*

Un indirizzo e-mail può essere associato ad un solo account. Finché l'account non viene approvato dal CNEC è possibile registrarsi nuovamente indicando il medesimo indirizzo e-mail nel qual caso fa fede l'ultima registrazione effettuata

Login

Registrazione Produttore di linee guida

Registrazione Valutatori di linee guida

Area stakeholder per le LG prodotte dall'ISS

Gli stakeholder possono partecipare allo sviluppo delle LG ISS attraverso la procedura di consultazione pubblica

Procedure di invio e valutazione di Linee Guida per la pubblicazione nell'SNLG

Manuale operativo

Centro Nazionale Eccellenza
Clinica Qualità e Sicurezza delle Cure

Versione 3.01 – gennaio 2020