Practice guidelines : overview of methodology with focus on GRADE

Benjamin Djulbegovic, M.D.,Ph.D.
Distinguished Professor
USF, Dpt of Internal Medicine
Chief, Division of Evidence-based Medicine and Health Outcome Research

CTSI, Director of
Center for Evidence-based Medicine and Health Outcome Research

Guideline Definition

"Systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances"

Institute of Medicine, 1990

Types of Guidelines

- What to do?
  - Pathways/Algorithms
  - Boundary Guidelines
- How to do it. . .
  - Critical Care Paths

Courtesy of Dr. R. Winn

Methods of Developing Guidelines

- **Informal consensus**
- **Formal consensus**
- **Evidence-based medicine approach**
- **Explicit approach**

Consensus

Although it may capture collective knowledge, it is also vulnerable to the possibility of capturing collective ignorance

-- Murphy, 1998

Courtesy of Dr. R. Winn

Importance of grading quality of evidence and strength of recommendations

- Patients and physicians using clinical practice and other recommendations need to know how much confidence they can place in the recommendations
- Clinical guidelines are only good as good as the evidence and judgments that are based on
- Systematic and explicit methods of making judgments can reduce errors and improve communication

BMJ 2004;328:1490-4

What does the patient and his physician need? Evidence-based principles

- **Evidence: selective citation vs. totality of evidence**
  - Need for systematic review
  - Simultaneous instead of se… …nefits and harms)
- **Assessing the quality**
  - Critical appraisal of the qua… …formed decision in health care
  - quantity, quality (internal v…
- **Benefits and harms**
  - Relative effect measures vs. abs… …e effect measures (NNT)
    - Minimizing framing effect
    - Probability vs. certainty
- **Patient-oriented evidence vs. disease-oriented evidence**
  - Evidence on survival, DFS, QOL is more important than evidence on tumor response, markers etc
- **Help with decisions**
  - Effective health-care recommendations vs. preference-sensitive health-care recommendations (decisions)

Mortality Rx1=10% Mortality Rx2=5%
RRR=(10-5)/10= 50%
ARD=5%
NNT=100/5%= 20

Mortality Rx1=1% Mortality Rx2=0.5%
RRR=(1-0.5%)/1= 50%
ARD=0.5%
NNT=100/0.5%=200

The need for research synthesis

- Health care decision makers need to access research evidence to make informed decisions on diagnosis, treatment and health care management for both individual patients and populations.
- There are few important questions in health care which can be informed by consulting the result of a single empirical study.

Systematic reviews of the **totality research evidence** represents a scientific foundation for development of clinical practice guidelines and health technology assessments.

The need for better methods of research synthesis: the rise of systematic reviews

- Systematic Review
  - "The application of strategies that limit bias in the **assembly**, **critical appraisal**, and **synthesis** of **all** relevant studies on a specific topic. Meta-analysis may be, but is not necessary, used as part of this process."
- Meta-Analysis
  - " The statistical synthesis of the data from separate but similar, i.e. comparable studies, leading to a quantitative summary of the pooled results."

Last JM. Dictionary of Epidemiology, 2001

Rational decision-making

- All major theories of choice agree that rational decision-making requires integrations of
  - **Benefits** (gains)
  - **Harms** (losses)
- Theories of decision-making
  - Differ in the proposal how benefits and harms should be integrated in a given decision

Evidence Profile

Patient or population: newly diagnosed (previously untreated) patients with multiple myeloma
Intervention: High-dose chemotherapy with single autologous transplant
Control: Chemotherapy
Outcomes: Overall Survival, Progression-free survival, Treatment related mortality

Denotes quantity of evidence | Denotes quality of evidence | Consistency | Generalizibility | Power of the evidence | Relative effect measures | Absolute effect measures | Denotes Quality

Quality assessment						Summary of findings				
						No of patients		Effect		
No of studies	Design	Limitations	Consistency	Directness	Other considerations	high-dose chemotherapy with single autologous transplant	chemotherapy	Relative (95% CI) Hazard ratio/Odds ratio (HR/OR)	Absolute (95% CI)	Quality
Benefits	**Overall Survival (presented as total mortality):**									
9	Randomized trials	No limitations	Important inconsistency (-1)[1]	No uncertainty	None	2411		HR 0.92[1] (0.74 to 1.13)	----	⊕⊕⊕○ Moderate
Benefits	**Progression-free survival:**									
9	Randomized trials	No limitations	Important inconsistency (-1)[2]	No uncertainty	None	2411		HR 0.75[2] (0.59 to 0.96)	9%[3]	⊕⊕⊕○ Moderate
Harms	**Treatment related mortality:**									
9	Randomized trials	No limitations	No important inconsistency	No uncertainty	None	2411		OR 3.01 (1.64 to 5.5)	3%	⊕⊕⊕⊕ High

EVIDENCE VS. ASSERTION VS. ADVERTISEMENT

EFFECT OF FRAMING: PROBABILITY VS.

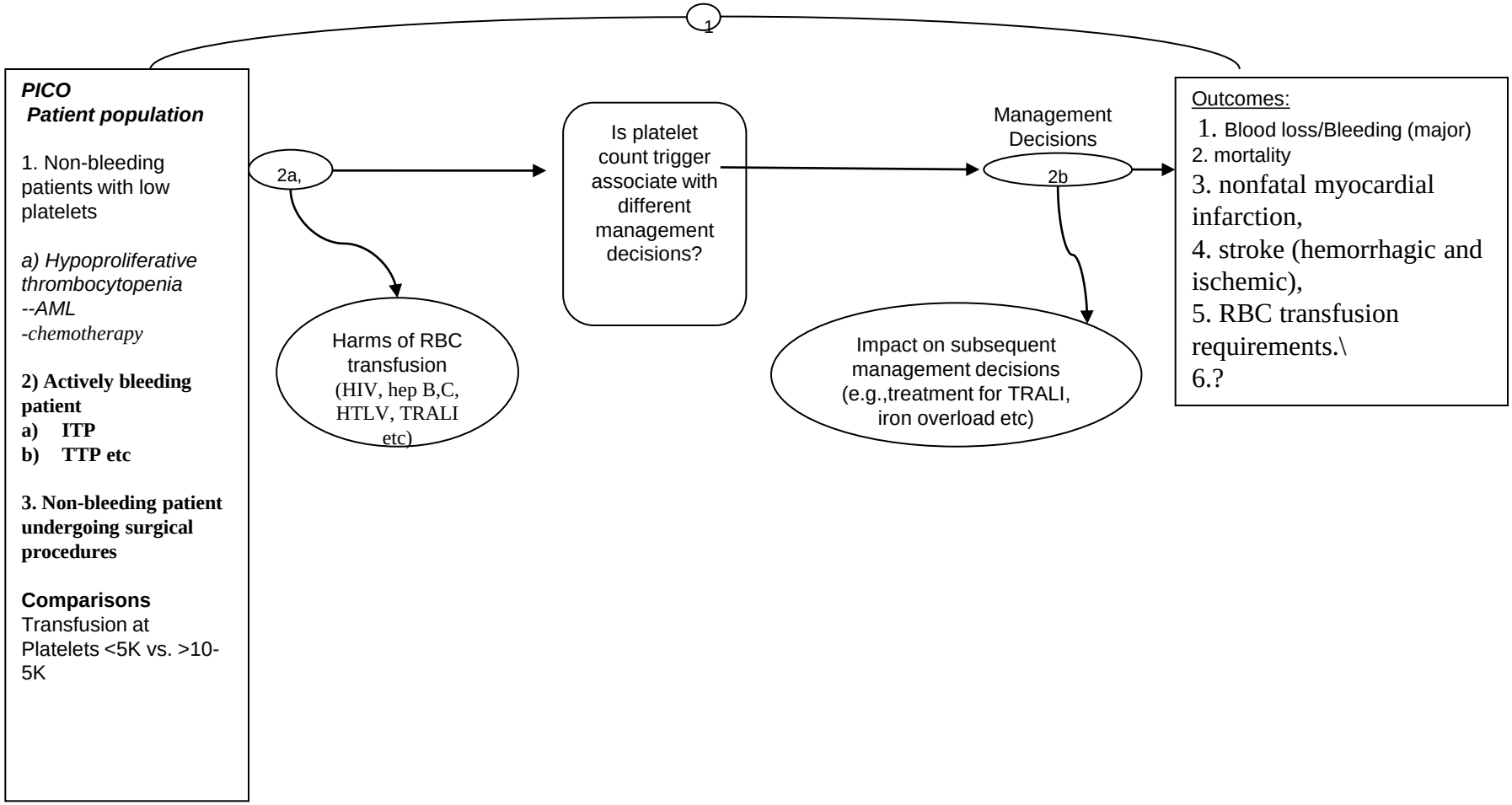
EFFECTIVE HEALTH-CARE RECOMMENATIONS VS. PREFERENCE BASED DECISION MAKING

Footnotes:

1. Significant heterogeneity among included studies (Q statistic of 27.65; p<0.01). All the exploratory sensitivity analysis results did not differ from the main result of excluding benefit for survival with high-dose therapy plus trasplant versus chemotherapy only.
2. Significant heterogeneity among included studies (q statistic of 51.57; p<0.01). However different sensitivity analysis based on either exclusing non standrad trials or source of stem cells or length of follow-up etc. did not result in any significant difference from the original results.
3. Assuming 50% PFS at 3 yrs
4. Assuming baseline risk of death of 1% in control group

Systematic review used to create this evidence profile: Koreth J, Cutler CS, Djulbegovic B, Behl R, Schlossman RL, Munshi NC, et al. High-dose therapy with single autologous transplantation versus chemotherapy ed multiple myeloma: A systematic review and meta-analysis of randomized controlled trials. Biol Blood Marrow Transplant 2007;13(2):183-96.

Analytic Framework : does platelet transfusion administered at different target values of platelet count result in different clinical outcomes? Linking guidelines/HTA to systematic reviews



Evidence & Decision making

- Evidence is necessary but not sufficient for optimal decision-making
- Making categorical recommendations

(considered judgments)

- Qualitative exercise
  - Occasionally is supplemented with quantitative (decision-analytic) modeling

GRADE: stressing explicitness and transparency and less reproducibility

- **Effective health-care recommendations**
  - Effective health-care (strong recommendations) when benefits >>>harms: candidate for quality criteria
- **Preference-sensitive recommendations**
  - Judgments about benefits/harm ratio uncertain, depend on patient values and preferences
  - May be based on **quantitative or qualitative judgments** about (explicitly) summarized evidence
- decision-making process must be **transparent** and **explicit** with **clarity** regarding the critical **criteria** that informed **recommendations;** based on **shared deliberation** & must include **appeal process**
- **"Accountability for reasonableness"**
  - which may help legitimize specific choices that may favor one set of stakeholders over others

After O'Connors, 2003; Teutsch et al, 2005 ; Daniels at al, 1997

McMaster University
Inspiring Innovation and Discovery

Formulate question

Select outcomes

Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

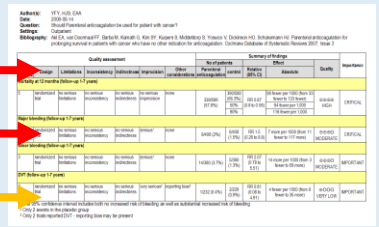
P
I
C
O

Outcome Critical

Outcome Critical

Outcome Important

Outcome Not important



Summary of findings & estimate of effect for each outcome

Grade down

High ⊕⊕⊕⊕

Moderate ⊕⊕⊕◯

Low ⊕⊕◯◯

Very low ⊕◯◯◯

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

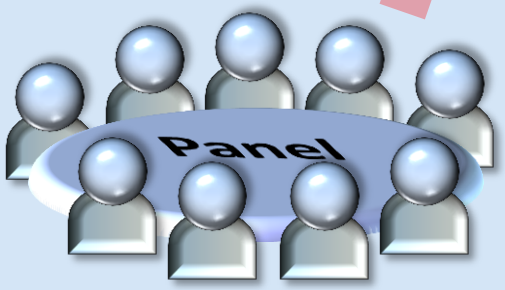
Grade up

1. Large effect
2. Dose response
3. Opposing bias & Confounders

Systematic review

Guideline development

Input?



Guideline

Grade overall quality of evidence
across outcomes based on lowest quality of ***critical*** outcomes

Grade recommendations
- For or against (direction) ↓↑
- Strong or conditional/weak (strength)

By considering balance of:
- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences

Revise if necessary by considering:
- ❑ Resource use (cost)


AMERICAN GASTROENTEROLOGICAL ASSOCIATION

Colorectal Cancer Screening and Surveillance: Clinical Guidelines and Rationale—Update Based on New Evidence

Formulate Recommendations (↓↑ | ⊕...)

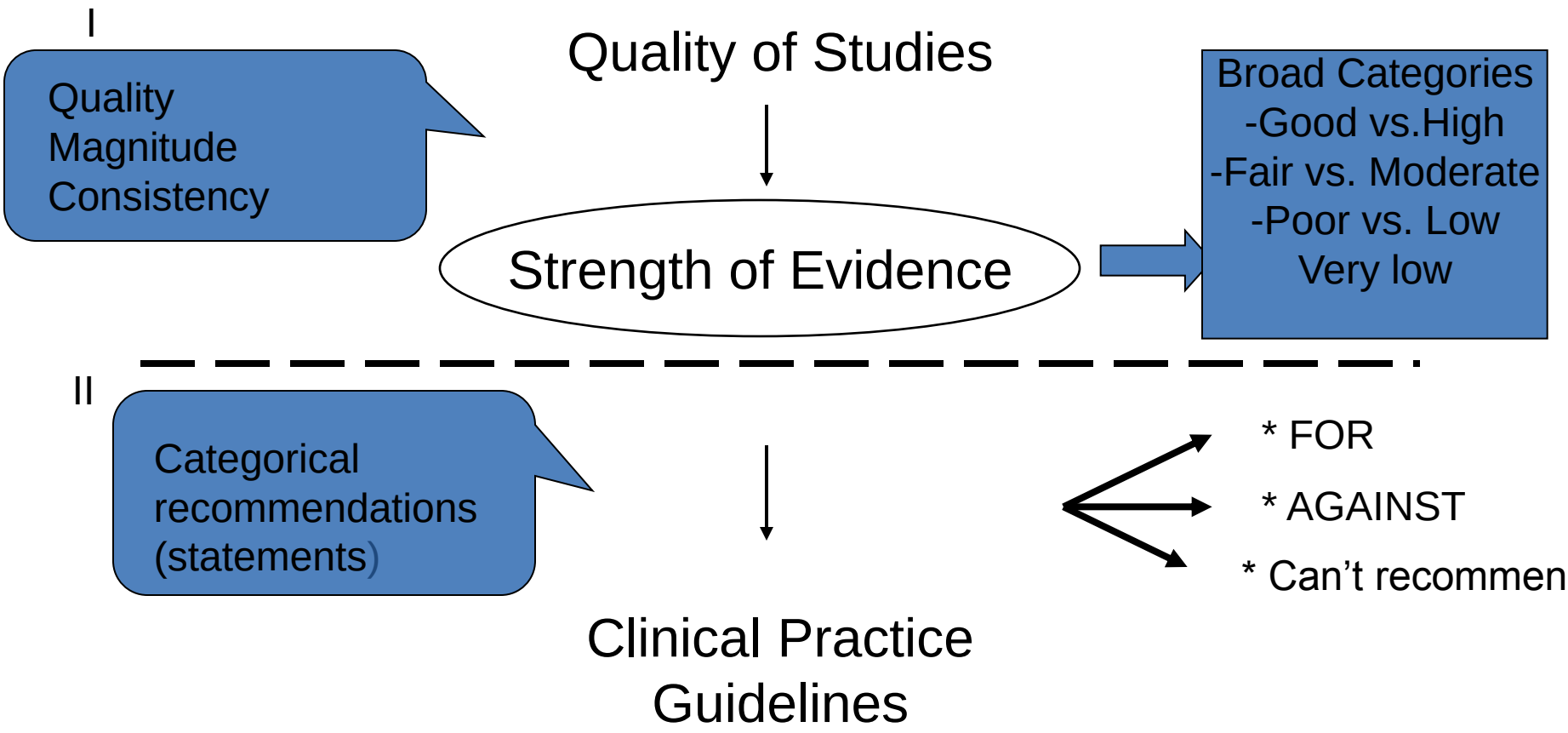
- "We recommend using..." | "Clinicians should..."
- "We suggest using..." | "Clinicians might..."
- "We suggest not using..." | "Clinicians ... not..."
- "We recommend not using..." | "Clinicians should not..."

Formulation of guidelines: main principles

- Separate evidence from decision-making
- **Quality of evidence** indicates the extent to which one can be confident that an estimate of effect is correct
  - represented on a **continuum scale** of credibility
- **Strength of recommendations** indicates the extent to which one can be confident that adherence to a recommendation will do more good than harm
  - Represent decision-making about choice and is **categorical exercise** (we recommend or do not)

From Evidence to Decision-making (recommendations)

Continuum from Study Quality Through Strength of Evidence to Guideline Development



Central role of evidence

Guidelines development process

Prior steps in developing guidelines

Prioritise problems, establish panel

Preparatory steps

Systematic review

↓

Evidence profile for important outcomes

Grading the quality of evidence and the strength of recommendations

Quality of evidence for each outcome

↓

Relative importance of outcomes

↓

Overall quality of evidence

↓

Balance of benefits and harms
(Does the intervention do more good than harm?)

↓

Balance of net benefits and costs
(Are incremental health benefits worth the costs?)

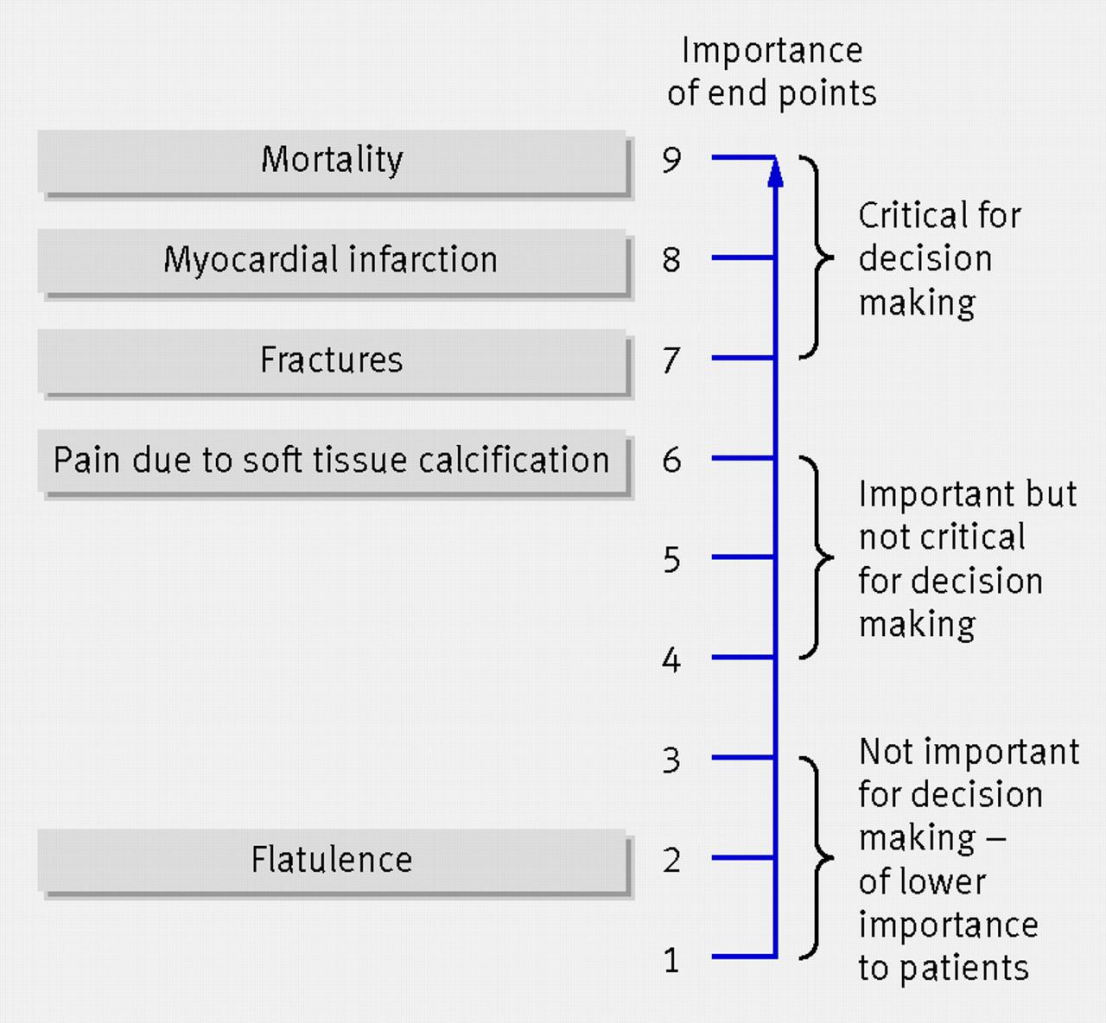
↓

Strength of recommendation

Subsequent steps

Implementation and evaluation

Hierarchy of outcomes according to importance to patients to assess effect of phosphate lowering drugs in patients with renal failure and hyperphosphataemia



Guyatt, G. H et al. BMJ 2008;336:995-998

BMJ

What are we assessing/grading?

- two components
- **quality of evidence**
  - extent to which confidence in estimate of effect adequate to support decision
    - high, moderate, low, very low
- **strength of recommendation**
  - The extent to which we can be confident that the desirable effects of an intervention outweigh the undesirable effects

The importance of context: conclusions vs. decisions

- **Quality of evidence** (="conclusions")
  - The extent of confidence that an estimate of effect is correct i.e. representing the "truth"
    - Important for systematic reviews
  - The extent to which confidence in an estimate of the effect is adequate to support recommendations
    - Importance for the guidelines panes
- **Making recommendations**
  - The extent to which we can be confident that the desirable effects of an intervention outweigh the undesirable effects
    - Important for guidelines panels
    - **NB as long as there is judgment that benefits>>>harms, recommendation can be strong even if the quality of evidence is low or very low**
      - **assumes that the error making a strong recommendation will be regretted less than the error making a weak recommendation**

GRADE: categories of quality

- **High**: Considerable confidence in the estimate of effect.
  - True effect likely lies close to our estimate of the effect
  - Further research unlikely to change our confidence in estimate
- **Moderate**: moderately confident that the estimate is close to the truth
  - Further research likely to have important impact on confidence in estimate, may change estimate.
- **Low**: confidence in the effect limited. True effect may be substantially different from the estimate
  - Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- **Very low**: little confidence in the effect estimate
  - Any estimate of effect is very uncertain.

GRADE quality assessment criteria: therapeutic studies

Quality of evidence	Study design	Lower if *	Higher if *
High	Randomised trial	**Risk of bias**: -1 Serious limitations -2 Very serious limitations **Inconsistency** -1 Serious -2 Very serious **Indirectness**: -1 Serious -2 Very Serious **-Imprecision** 1 Serious -2 Very serious **Reporting bias** 1 likely -2 Very likely	**Strong association**: +1 Large effect (Strong, no plausible confounders, consistent and direct evidence)** +2 Very large effect (Very strong, no major threats to validity and direct evidence)*** +1 Evidence of a **Dose response** gradient +1 All **plausible confounders** would have reduced the effect
Moderate	Quasi-randomised trial		
Low	Observational study		
Very low	Any other evidence		

* 1 = move up or down one grade (for example from high to intermediate)
2 = move up or down two grades (for example from high to low)

** A statistically significant relative risk of >2 (< 0.5), based on consistent evidence from two or more observational studies, with no plausible confounders

*** A statistically significant relative risk of > 5 (< 0.2) based on direct evidence with no major threats to validity

Sources of bias: Rx

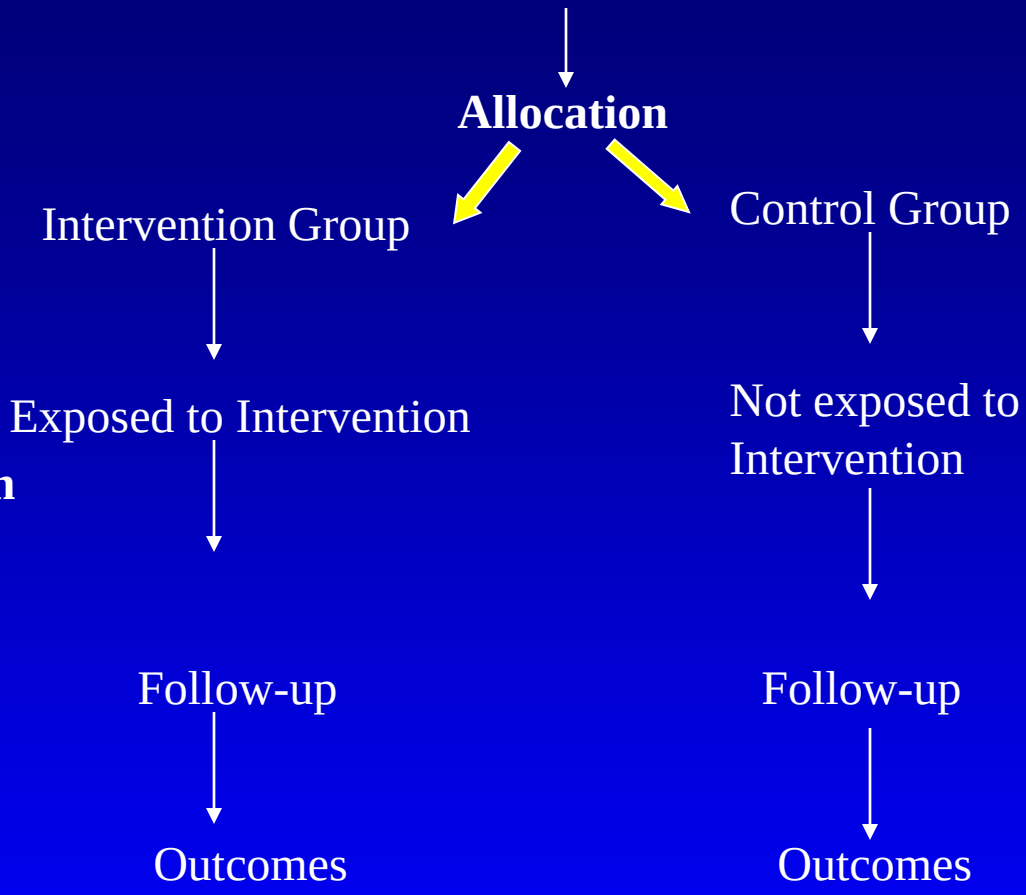
- **Choice of the control intervention**

Selection bias

systematic differences in comparison groups

- **Performance bias/information bias**
- **systematic differences in care provided apart from the intervention being evaluated**
- **systematic error in the measurement of information on exposure or outcome**
  - **Attrition bias (systematic differences in withdrawals from the trial**
  - **Detection bias/Recall bias (systematic differences in outcome assessment)**

- **Specimen handling bias (systematic differences in analysis of specimens)**



- **Analysis appropriateness (was analysis reflective of the problem at hand? ITT vs. PP)**

Controlling for selection bias

- **Randomized controlled trials**
  - **Generation** of allocation sequence
    - *In RCTs this is usually done by computer using any number of available methods (usually block randomization, etc)*
  - ***Concealing** treatment assignment until after the treatment has been allocated*
- **Observational research**
  - *In a cohort study: are participants in the exposed and unexposed groups similar in all important respects except for the exposure?*
    - *Control for confounders*
  - *In a case-control study: are cases and controls similar in all important respects except for the disease in question?*
    - *matching*

Controlling for performance bias

- **RCTs**
  - *Are we controlling for co-intervention/contamination?*
  - *a method to prevent that those who providing and receiving care do not know to which intervention group the recipients of care have been allocated*
  - *use of "blinding/masking"*
- **Observational research**
  - **Accounting for information (measurement) /recall bias**
  - *In a **cohort** study: is information about outcome obtained in the same way for those exposed and unexposed?*
  - *In a **case-control** study, is information about exposure gathered in the same way for cases and controls?*

Controlling for attrition bias

- **RCTs**
  - ***Complete follow-up***
  - *Baseline characteristics of participants lost to follow-up and those included in the analysis should be reported separately*
- **Observational research**
  - *Cohort/case-control studies: Completeness of follow-up*
  - *Baseline characteristics of participants lost to follow-up and those included in the analysis should be reported separately*

“Intention to treat” vs. ‘per protocol’ analysis

- All patients should be analysed in the arm to which they were allocated at randomisation, regardless of whether they receive the allocated treatment (‘Intention-to-treat’ analysis).

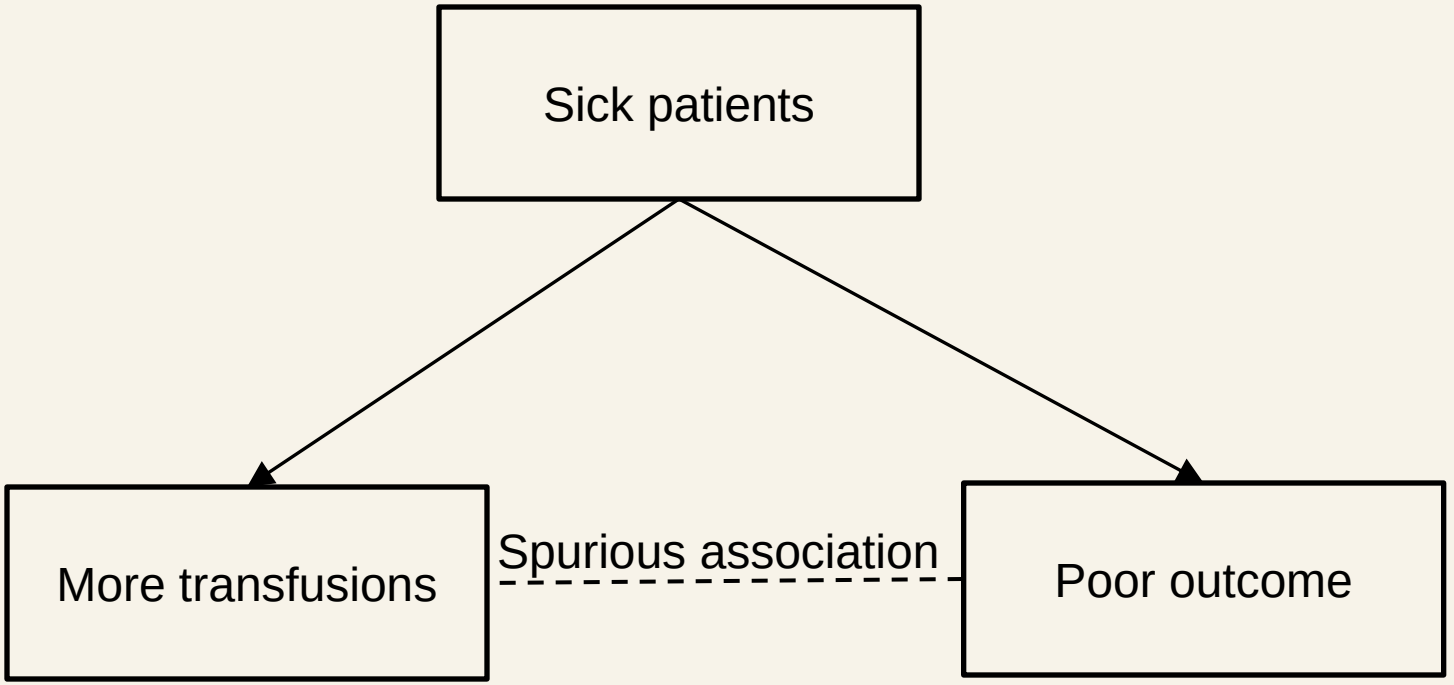
Observer bias

- The biases that lead to misperceptions that we have detected, seen or experienced something that actually isn't there
  - **Placebo/masking technique to control for observer bias**
    - Mesmerism and Franklin's commission appointed by Louis XVI in 1784 to investigate the medical claims of "animal magnetism", or "mesmerism".
  - The people being studied felt the effects of mesmerism only when they were "told" and felt no effects when they were not told, whether or not they were receiving the treatment.

Confounding by indication: important quality issue in transfusion medicine

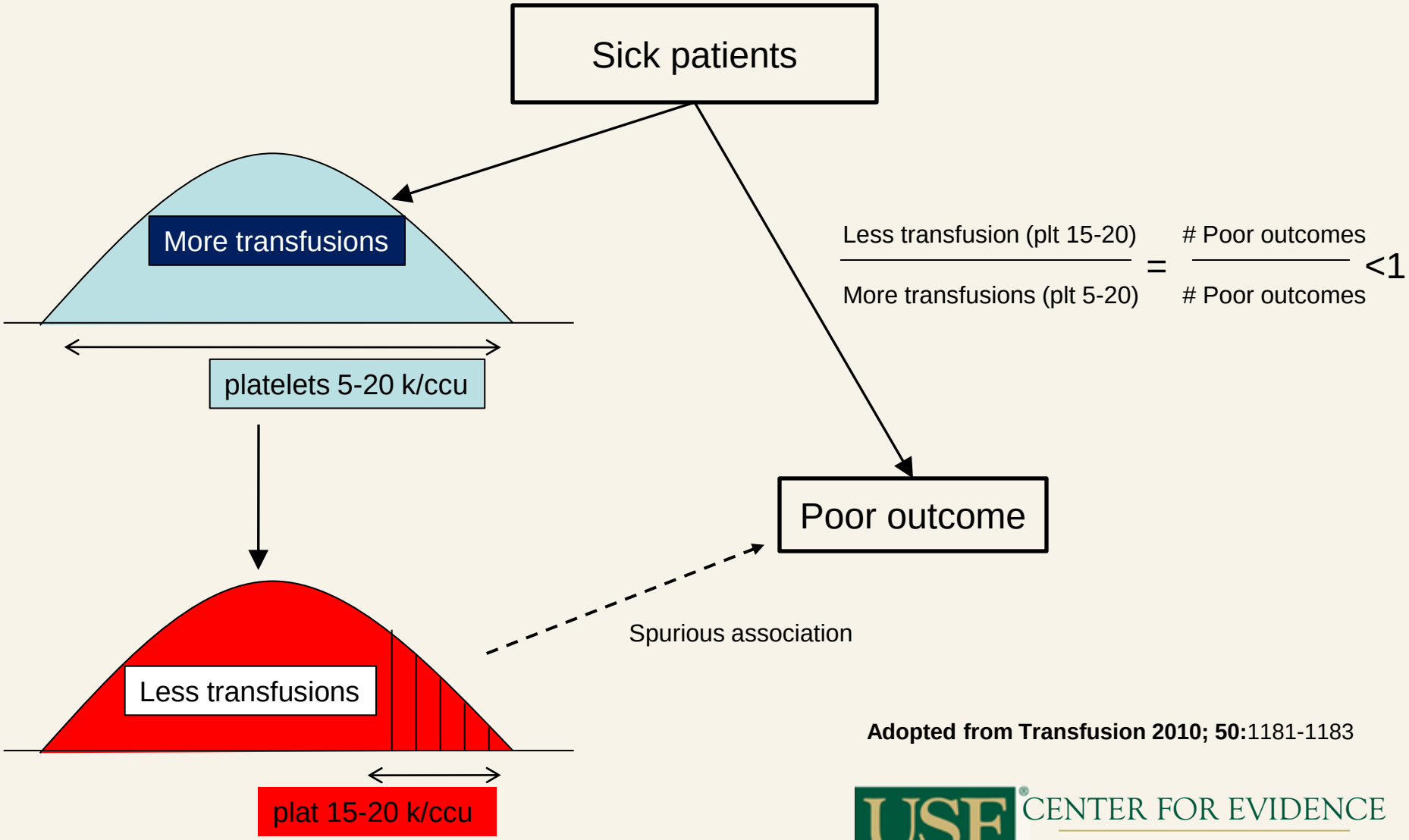
- Results from the conscious choice of different treatments for patients with different prognosis
  - According to severity of disease
- Probably the most important bias in clinical research
  - Observational studies
  - Can be avoided by performing a well designed RCT

Blood transfusions: Good or Bad?



Transfusion 2010; 50:1181-1183

Confounding by indications: apparent protective effect of liberal platelet transfusion strategy on poor outcomes (e.g., bleeding, etc)



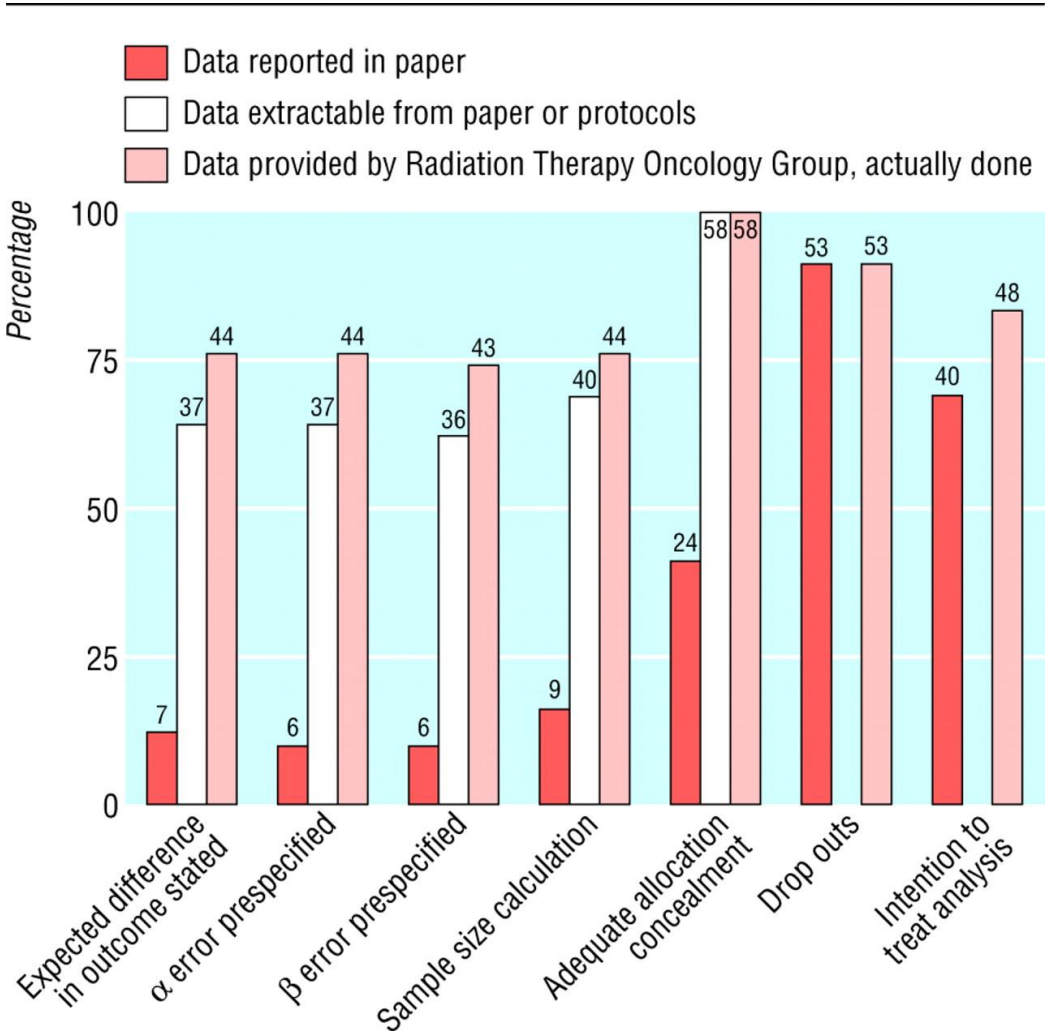
Adopted from Transfusion 2010; 50:1181-1183

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Information needed to judge whether our estimates are correct
GRADE methods for assessing quality of evidence: intervention studies (clinical utility)

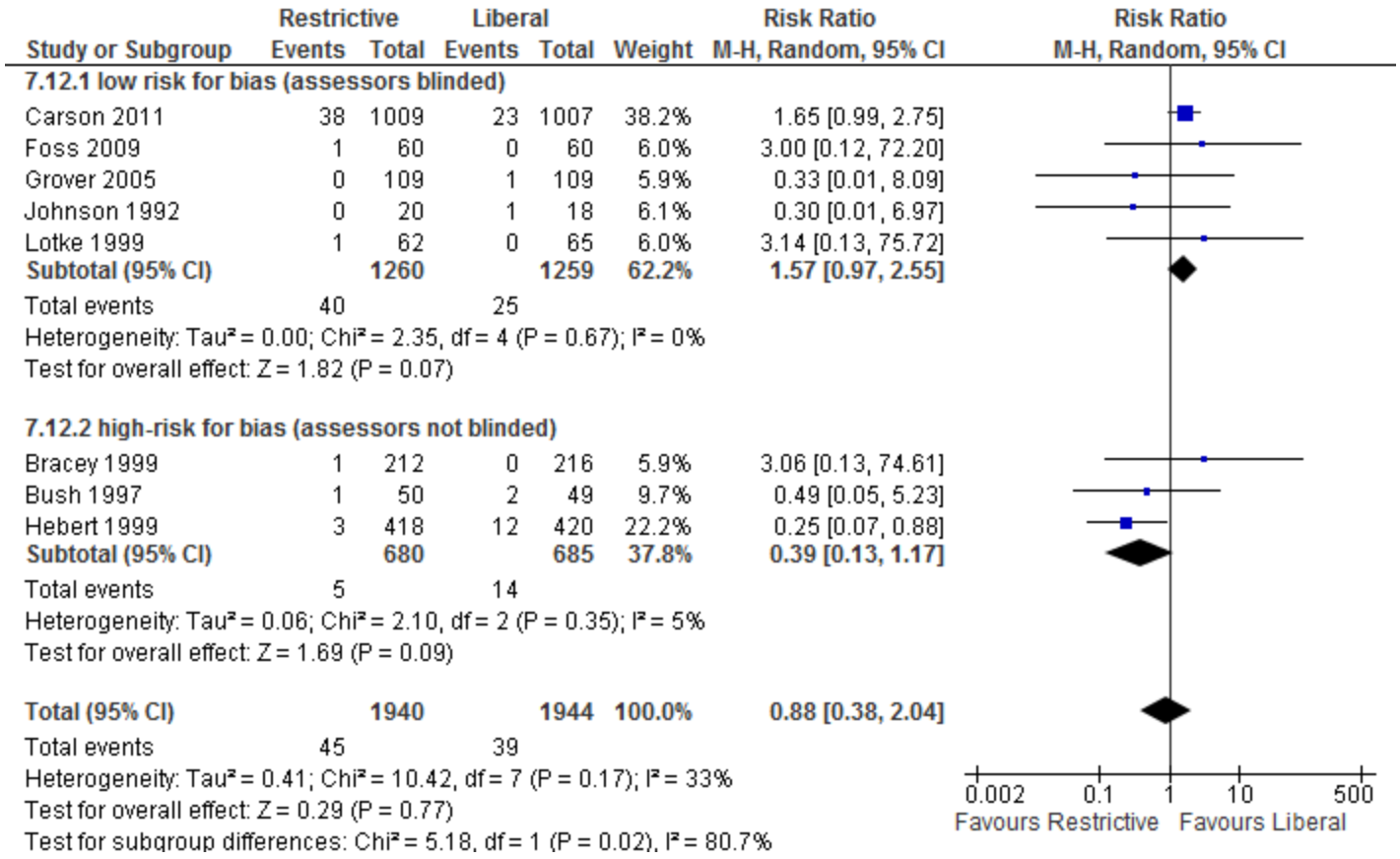
- ***Factors that might decrease quality of evidence***
  - **Study limitations (risk of bias)**
    - Inadequacy of allocation concealment; lack of blinding, large drop-outs, failure to perform ITT, failure to report outcomes,
    - **Inconsistency of results**
    - Variability or heterogeneity in results due true differences in treatment effect (due to ***P-I-C-O)***
    - **Statistical**: large I^2 (e.g.>50%); **clinical:** (PICO)
  - **Indirectness of evidence (2 types)**
    - Lack of head-to-head comparisons
    - differences in treatment effect (due to ***P-I-C-O)***
  - **Imprecision**
    - A few events (<200-300?), small studies (N<400), wide confidence intervals consistent with important differences in both directions or no effects or all)
  - Reporting (**publication**) bias
  - **Other factors**
    - Carryover effect in crossover trials, use of unvalidated outcome measures, recruitment bias in cluster RCT etc
- ***Factors that might increase quality of evidence***
  - **Large magnitude of effect**
    - A statistically significant relative risk of > 5 (< 0.2)
  - **Plausible confounding, which would reduce a demonstrated effect is accounted for without affecting treatment effect**
  - **Dose-response gradient**

Quality of reporting compared with actual methodological quality



Soares H P et al. BMJ 2004;328:22-24

Restrictive vs. liberal RBC transfusion: effect of assessors' blinding (MI)



Assessment of the quality of evidence in RCTs testing restrictive vs. liberal transfusion strategy

(based on Cochrane review by Carless et al

- **Use of RBC transfusion:**
  - High
    - None of the potential flaws appear to have significant effect on the results
- **30 days mortality**
  - High
    - The results between high-quality trials and those with the flaws consistent. Hence, none of the potential flaws appear to have significant effect on the results
- **Myocardial infarction/Cardiac events**
  - Very low
    - Trials in which outcome assessors were not blinded favored restricted strategy, which may have incorporated biased assessment of outcomes
- **Walking independently at 60 days**
  - Low
    - Based on self-reporting from one (high-quality) trial (sparse data)
- **Length of stay**
  - Moderate
    - Decision to discharge may be a function of knowledge of treatment group
- **CHF**
  - Very low
    - It is not clear how CHF was diagnosed; a few data provided in the Cochrane review
      - Pulmonary edema clinically can encompass many conditions including TRALI, CHF etc

Rating quality of evidence across outcomes

- GRADE recommends that the guideline developers consider the quality of evidence across outcomes as that associated with the

  ***critical outcome with the lowest quality evidence*.**

  - GRADE requires ***guideline developers, but not systematic review authors***, to make an overall rating of evidence quality ***across outcomes*** deemed critical for decision-making.
  - [NB The principle is that if there is higher quality evidence from some critical outcomes to support a decision in favour of an intervention (that is, benefits on critical outcomes clearly outweigh undesirable effects of the intervention, for which there is also high quality evidence) one needn't rate down the quality because of lower quality evidence regarding other critical outcomes that support the same recommendation]

BMJ 2008

Annals of Internal Medicine
Established in 1927 by the American College of Physicians

From: Red Blood Cell Transfusion: A Clinical Practice Guideline From the AABB*

Ann Intern Med. 2012;157(1):49-58. doi:10.7326/0003-4819-157-1-201206190-00429

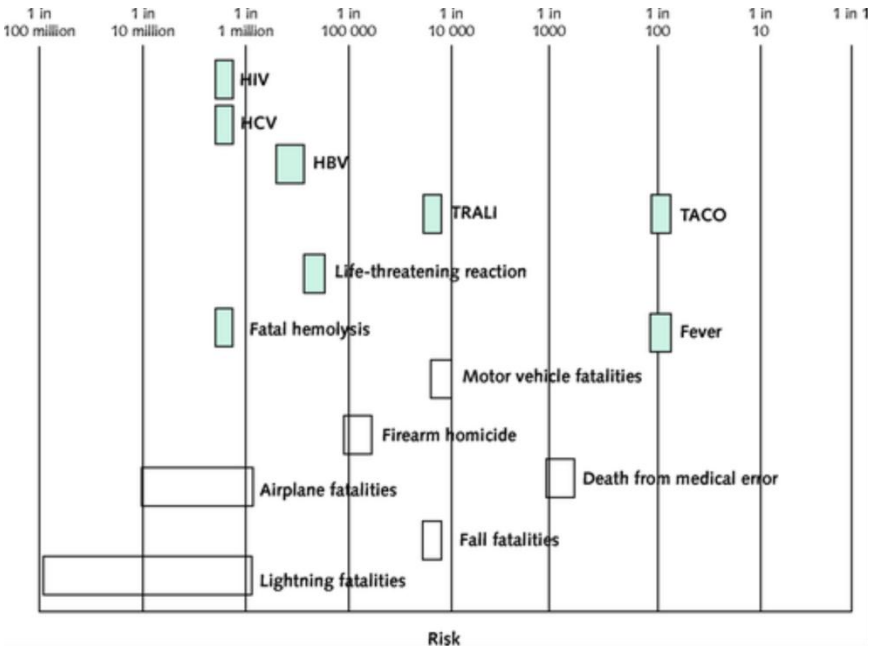


Figure Legend:

Adverse effects of RBC transfusion contrasted with other risks.

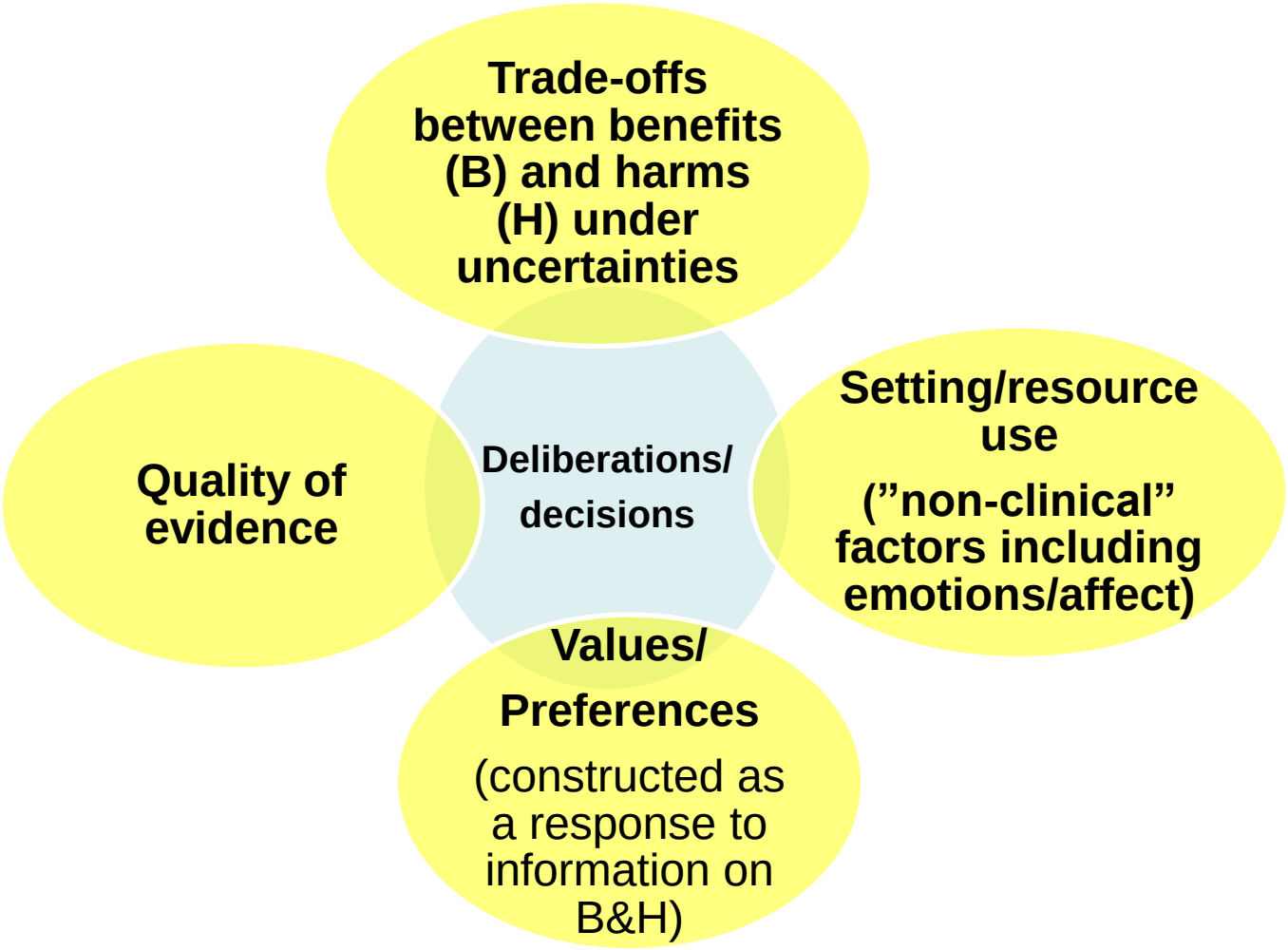
From evidence to recommendations

- Evidence is necessary but not sufficient for optimal decision-making
- Making categorical recommendations

(considered judgments)

- Qualitative exercise
  - Occasionally is supplemented with quantitative (decision-analytic) modeling
  - Driven by normative/prescriptive principles

Factors affecting decision-making

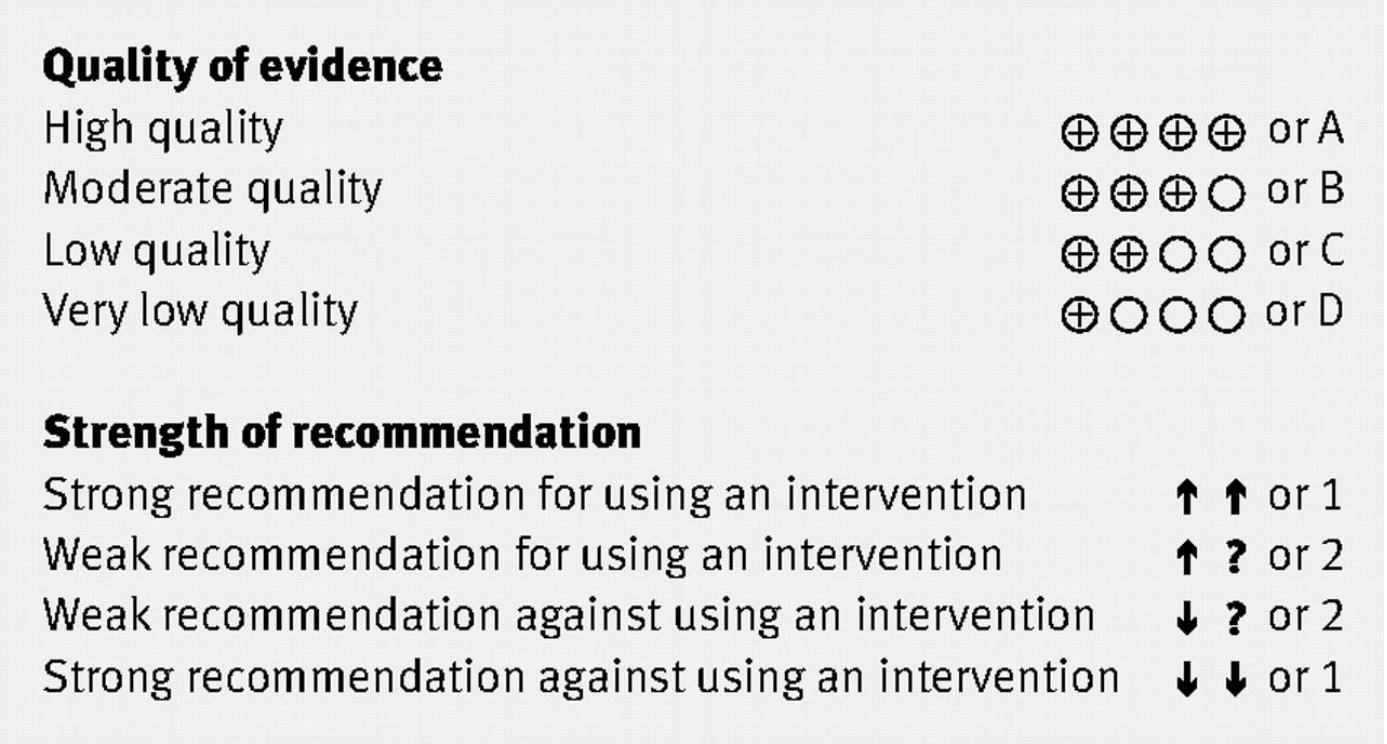


GRADE 2008

Determinants of strength of recommendation

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted
Costs (resource allocation)	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted

Representations of quality of evidence and strength of recommendations



Guyatt, G. H et al. BMJ 2008;336:1049-1051

BMJ

Factors that affect the strength of a recommendation

Factor	Examples of strong recommendations	Examples of weak recommendations
Quality of evidence	Many high quality randomised trials have shown the benefit of inhaled steroids in asthma	Only case series have examined the utility of pleurodesis in pneumothorax
Uncertainty about the balance between desirable and undesirable effects	Aspirin in myocardial infarction reduces mortality with minimal toxicity, inconvenience, and cost	Warfarin in low risk patients with atrial fibrillation results in small stroke reduction but increased bleeding risk and substantial inconvenience
Uncertainty or variability in values and preferences	Young patients with lymphoma will invariably place a higher value on the life prolonging effects of chemotherapy than on treatment toxicity	Older patients with lymphoma may not place a higher value on the life prolonging effects of chemotherapy than on treatment toxicity
Uncertainty about whether the intervention represents a wise use of resources	The low cost of aspirin as prophylaxis against stroke in patients with transient ischemic attacks	The high cost of clopidogrel and of combination dipyridamole and aspirin as prophylaxis against stroke in patients with transient ischaemic attacks

Has a mistake been made? Explicitly taking consequences into guidelines considerations

- We can **always** make a mistake
  - Recommend ineffective treatments
    - Regret of commission
  - Fail to recommend effective treatments
    - Regret of omission
- Sense of loss, or **regret**
  - **How many times regret of commission is worse than regret of omission**

Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive

If you need to vote: Insert the number of votes for the recommendation in each category				
Assessors' view of the balance between desirable and undesirable consequences of the intervention	**Desirable consequences clearly outweigh undesirable consequences**	**Desirable consequences probably outweigh undesirable consequences**	**Undesirable consequences probably outweigh desirable consequences**	**Undesirable consequences clearly outweigh desirable consequences**
Strength of recommendation	**Strong for an intervention**	**Conditional (weak) for an intervention**	**Conditional (weak) against an intervention**	**Strong against an intervention**
Wording of a recommendation	**We recommend to "do something"**	**We suggest (conditionally recommend) to "do something"**	**We suggest (conditionally recommend) not to "do something"**	**We recommend not to "do something"**
Number of votes				

NB typically one defines the rules advance. For example, one suggested rule is that recommendation for or against a particular intervention (compared with a specific alternative) will be made if at least 50% of the panel members vote in favor, with less than 20% preferring comparator. Failure to meet this criterion result in no recommendation (or "only in research"). For recommendations to be graded as strong vs. weak, at least 70% of the panel members should endorse it as "strong")

Making recommendations

Strength of the recommendation:

- Strong
- Conditional (weak)

Final recommendation:

Strength: **Quality of evidence:**

Assumptions about underlying values and preferences

Remarks

McMaster University
Inspiring Innovation and Discovery

Formulate question

Select outcomes

Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

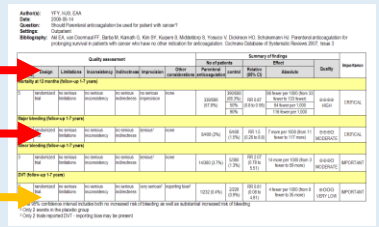
P
I
C
O

Outcome Critical

Outcome Critical

Outcome Important

Outcome Not important



Summary of findings & estimate of effect for each outcome

Grade down

High ⊕⊕⊕⊕

Moderate ⊕⊕⊕◯

Low ⊕⊕◯◯

Very low ⊕◯◯◯

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

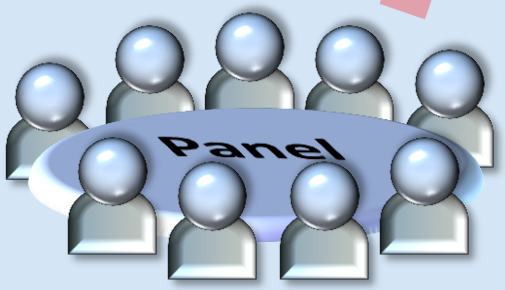
Grade up

1. Large effect
2. Dose response
3. Opposing bias & Confounders

Systematic review

Guideline development

Input?



Guideline

Grade overall quality of evidence
across outcomes based on lowest quality of ***critical*** outcomes

Grade recommendations
- For or against (direction) ↓↑
- Strong or conditional/weak (strength)

By considering balance of:
- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences

Revise if necessary by considering:
- ❑ Resource use (cost)


AMERICAN GASTROENTEROLOGICAL ASSOCIATION

Colorectal Cancer Screening and Surveillance: Clinical Guidelines and Rationale—Update Based on New Evidence

Formulate Recommendations (↓↑ | ⊕...)

- "We recommend using..." | "Clinicians should..."
- "We suggest using..." | "Clinicians might..."
- "We suggest not using..." | "Clinicians ... not..."
- "We recommend not using..." | "Clinicians should not..."

Evidentiary Standards: Clinical, Judicial , GRADE and FDA

Beyond reasonable doubt [= when both in the worst (Pworst) (skeptic) and best (Pbest) (enthusiast) case scenarios probability that intervention will exceed clinically important thresholds > 95%]	**Criminal cases**	**"Substantial Evidence"** (FDA marketing approval)	**GRADE: Strong recommendation for intervention** (high quality of evidence)	**Regret (of wrongly) recommending << regret of not recommending**
Clear and convincing evidence [Pworst <95%; Pbest>95%]	**Malpractice litigation**		**Strong recommendation** (moderate quality of evidence)	
Preponderance of evidence [Pworst>50%, Pbest<95%]	**Civil trials**	**AA (DOE→POE)**	**Strong vs. weak recommendation** (low quality of evidence) (context-dependent)	**Regret Rx << regret NoRx**
Reasonable to believe [Pworst<50%; Pbest>50%]	**Search warrants, reasonable suspicion**	**"Reasonable to believe" (EUA)**	**Weak recommendation** (very low quality of evidence)	**Regret Rx < regret NoRx**
Insufficient evidence [Pworst<50%, Pbest<50%]			**Do not recommend vs. "Only in research"**	**Regret Rx ≥ regret NoRx**