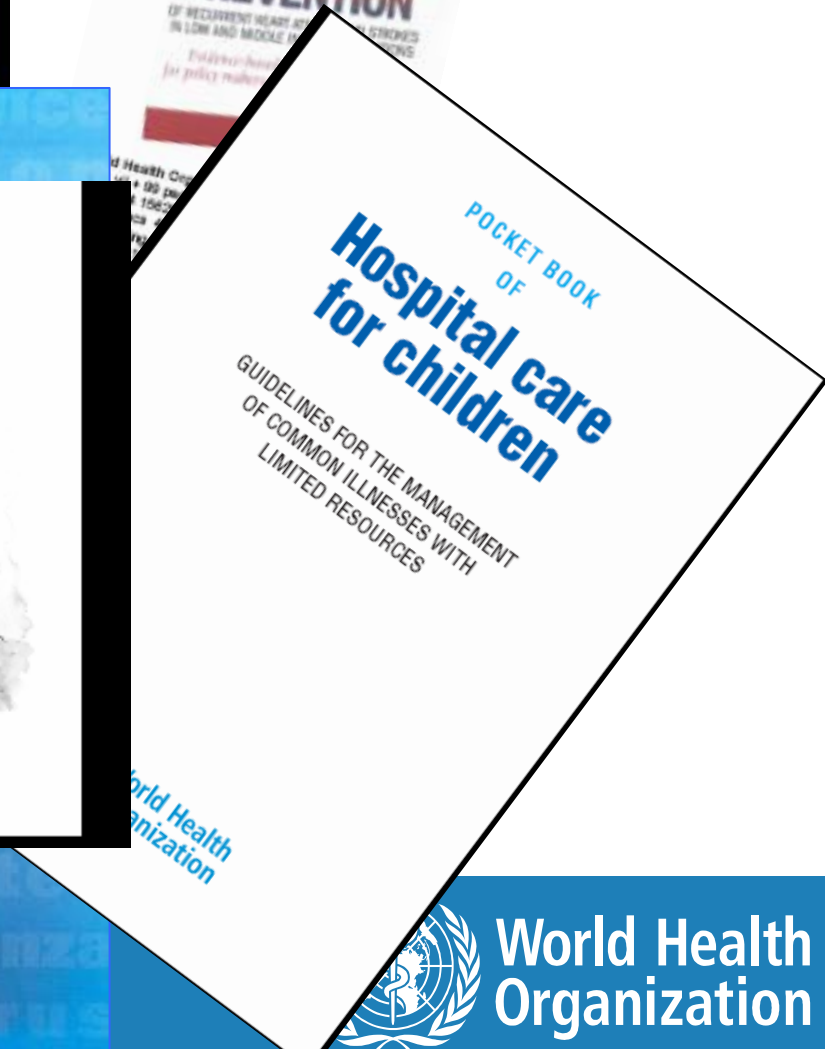
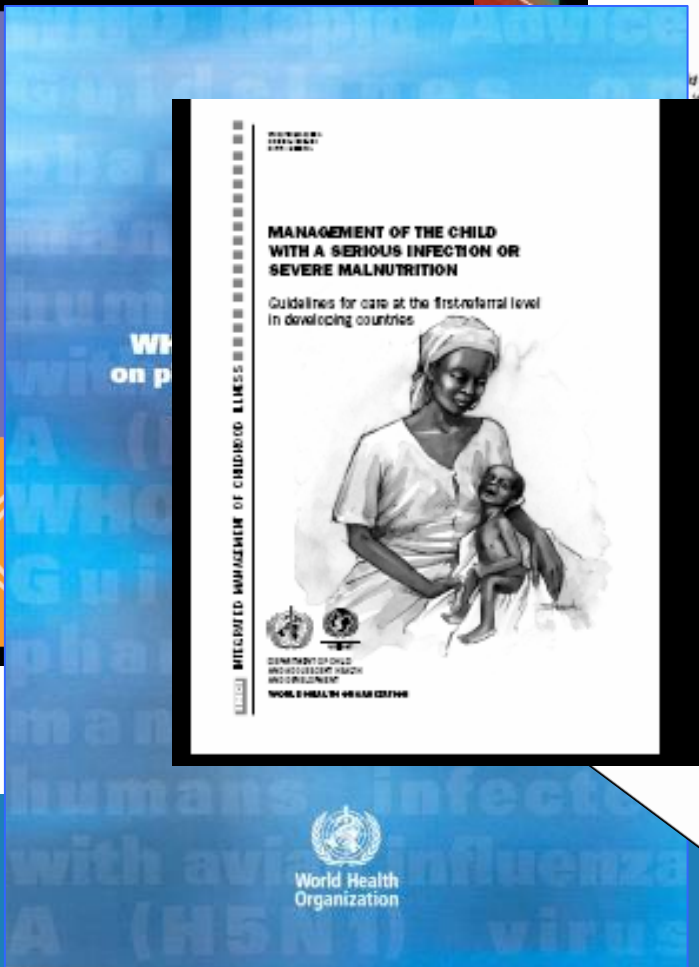
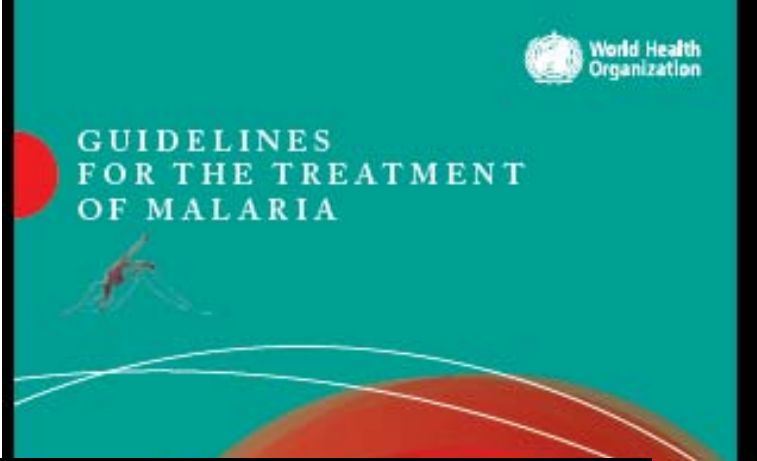




# **WHO, guidelines and evidence**

**Suzanne Hill**

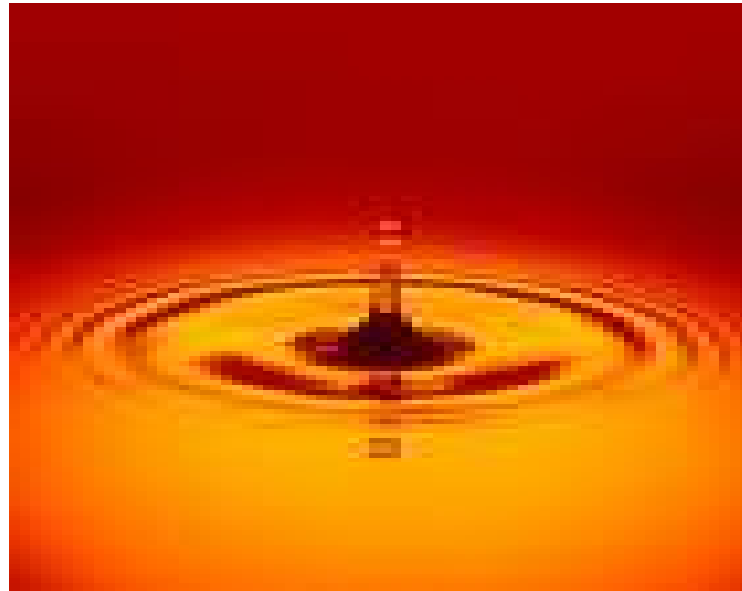
**March 2007**



World Health Organization



HONEYSETT



***'Critical appraisal of the JNC VI, WHO/ISH and BHS guidelines for essential hypertension.'***

**'These differing recommendations between JNC VI and BHS, and WHO/ISH cannot be reconciled and they are of such magnitude as to carry serious implications for clinical practice, not least among which is that acceptance of the WHO/ISH levels of 'normality' for blood pressure would result in some 45% of the population of all ages and nearly 60% of elderly people being classified as 'hypertensive'.'**

*O'Brien & Staessen, 2000*



**World Health  
Organization**

## ***Implementation of WHO/ISH Guidelines: role and activities of WHO.***

**'In order to increase its impact, however, an implementation strategy is needed that includes advocacy, dissemination, training and evaluation as its major components.'**

*Martin, Clin Exper Hypertens, 2000*

*'World health organisation-international society of hypertension  
(WHO/ISH) hypertension guidelines.'*

**'Since the publication of the 1999 WHO/ISH Guidelines for the  
Management of Hypertension, WHO determined in 2000 that in  
future the evidence base for all of its guidelines will be  
explicitly documented according to a defined methodology. '**

*Whitworth JA, Chalmers J. Clin Exper Hypertens  
2004*

2



- **Professional good intentions and plausible theories are insufficient for selecting policies and practices for protecting, promoting and restoring health.**
- **We will serve the public more responsibly and ethically when research designed to reduce the likelihood that we will be misled by bias and the play of chance has become an expected element of professional and policy making practice, not an optional add-on.**

- **Iain Chalmers**



**How can we judge how sure we are that adherence to a recommendation (our good intentions) will do more good than harm?**



**World Health  
Organization**

<b>Evidence</b>	<b>Recommendation</b>	<b>Organization</b>
<b>II-2</b>	<b>B</b>	<b>USPSTF</b>
<b>C+</b>	<b>1</b>	<b>ACCP</b>
<b>Strong</b>	<b>Strongly recommended</b>	<b>SIGN</b>



Recommendation for use of oral anticoagulation in patients with atrial fibrillation and rheumatic mitral valve disease

<b>Evidence</b>	<b>Recommendation</b>	<b>Organization</b>
<b>II-2</b>	<b>B</b>	<b>USPSTF</b>
<b>C+</b>	<b>1</b>	<b>ACCP</b>
<b>Strong</b>	<b>Strongly recommended</b>	<b>SIGN</b>

## Anxiety Guidelines (NICE)

- Shared decision-making should take place as it improves concordance and clinical outcomes. C
- Shared decision-making between the individual and healthcare professionals should take place during the process of diagnosis and in all phases of care. D
- Cognitive behavioural therapy (CBT) should be used. A
- For most people, CBT should take the form of weekly sessions of 1–2 hours and should be completed within a maximum of 4 months of commencement. B



## Levels of evidence

Level of evidence	Type 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 of RCTs, or RCTs with a low risk of bias
1-	Meta-analyses, systematic reviews of RCTs, 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, bias, or chance and a high probability that the relationship is causal
2+	Well-conducted case-control or cohort studies with a low risk of confounding, bias, or chance and a moderate probability that the relationship is causal
2-	Case-control or cohort studies with a high risk of confounding bias, or chance and a significant risk that the relationship is not causal
3	Non-analytic studies (for example, case reports, case series)
4	Expert opinion

## Grading of recommendations

Grade	Evidence
A	<ul style="list-style-type: none"> <li>• At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population, or</li> <li>• A systematic review of RCTs or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results</li> <li>• Evidence drawn from a NICE technology appraisal</li> </ul>
B	<ul style="list-style-type: none"> <li>• A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results, or</li> <li>• Extrapolated evidence from studies rated as 1++ or 1+</li> </ul>
C	<ul style="list-style-type: none"> <li>• A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results, or</li> <li>• Extrapolated evidence from studies rated as 2++</li> </ul>
D	<ul style="list-style-type: none"> <li>• Evidence level 3 or 4, or</li> <li>• Extrapolated evidence from studies rated as 2+, or</li> <li>• Formal consensus</li> </ul>
D (GPP)	A good practice point (GPP) is a recommendation for best practice based on the experience of the Guideline Development Group

# Problem

- **Too many systems**
- **Concentrating only on study design**
- **Not including other factors that influence judgements and recommendations**



# Why bother about grading?

- **People draw conclusions about the**
  - quality of evidence
  - strength of recommendations
- **Systematic and explicit approaches can help**
  - protect against errors
  - resolve disagreements
  - facilitate critical appraisal
  - communicate information
- **However, there is wide variation in currently used approaches**





# GRADE

## Grades of Recommendation Assessment, Development and Evaluation

# GRADE Working Group

David Atkins, chief medical officer<sup>a</sup>  
Dana Best, assistant professor<sup>b</sup>  
Peter A Briss, chief<sup>c</sup>  
Martin Eccles, professor<sup>d</sup>  
Yngve Falck-Ytter, associate director<sup>e</sup>  
Signe Flottorp, researcher<sup>f</sup>  
Gordon H Guyatt, professor<sup>g</sup>  
Robin T Harbour, quality and information director<sup>h</sup>  
Margaret C Haugh, methodologist<sup>i</sup>  
David Henry, professor<sup>j</sup>  
Suzanne Hill, senior lecturer<sup>k</sup>  
Roman Jaeschke, clinical professor<sup>k</sup>  
Gillian Leng, guidelines programme director<sup>l</sup>  
Alessandro Liberati, professor<sup>m</sup>  
Nicola Magrini, director<sup>n</sup>  
James Mason, professor<sup>d</sup>  
Philippa Middleton, honorary research fellow<sup>o</sup>  
Jacek Mrukowicz, executive director<sup>p</sup>  
Dianne O'Connell, senior epidemiologist<sup>q</sup>  
Andrew D Oxman, director<sup>r</sup>  
Bob Phillips, associate fellow<sup>r</sup>  
Holger J Schünemann, associate professor<sup>g,s</sup>  
Tessa Tan-Torres Edejer, medical officer/scientist<sup>t</sup>  
Helena Varonen, associate editor<sup>u</sup>  
Gunn E Vist, researcher<sup>f</sup>  
John W Williams Jr, associate professor<sup>v</sup>  
Stephanie Zaza, project director<sup>w</sup>

a) Agency for Healthcare Research and Quality, USA  
b) Children's National Medical Center, USA  
c) Centers for Disease Control and Prevention, USA  
d) University of Newcastle upon Tyne, UK  
e) German Cochrane Centre, Germany  
f) Norwegian Centre for Health Services, Norway  
g) McMaster University, Canada  
h) Scottish Intercollegiate Guidelines Network, UK  
i) Fédération Nationale des Centres de Lutte Contre le Cancer, France  
j) University of Newcastle, Australia  
k) McMaster University, Canada  
l) National Institute for Clinical Excellence, UK  
m) Università di Modena e Reggio Emilia, Italy  
n) Centro per la Valutazione della Efficacia della Assistenza Sanitaria, Italy  
o) Australasian Cochrane Centre, Australia  
p) Polish Institute for Evidence Based Medicine, Poland  
q) The Cancer Council, Australia  
r) Centre for Evidence-based Medicine, UK  
s) University of Buffalo, USA  
t) World Health Organisation, Switzerland  
u) Finnish Medical Society Duodecim, Finland  
v) Duke University Medical Center, USA  
w) Centers for Disease Control and Prevention, USA



## Definitions

Quality of evidence

**The extent to which one can be confident that an estimate of effect or association is correct.**

■ **Although the degree of confidence is a continuum, we use four categories:**

- **High**
- **Moderate**
- **Low**
- **Very low**

## Categories of quality

- High: **Further research is very unlikely to change our confidence in the estimate of effect. ++++**
- Moderate: **Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. +++**
- Low: **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: **Any estimate of effect is very uncertain. +**

# Judgements about the quality of evidence

**The quality of the evidence (i.e. our confidence) depends on:**

- **study design (e.g. RCT, case-control study)**
- **study quality/limitations (protection against bias; e.g. concealment of allocation, blinding, follow-up)**
- **consistency of results**
- **directness of the evidence including the**
  - **populations (those of interest versus similar; for example, older, sicker or more co-morbidity)**
  - **interventions (those of interest versus similar; for example, drugs within the same class)**
  - **outcomes (important versus surrogate outcomes)**
  - **comparison (A - C versus A - B & C - B)**



# Judgements about the quality of evidence

The quality of the evidence (i.e. our confidence) may also be **REDUCED** when there is:

- ↓ Sparse or imprecise data
- ↓ Evidence of reporting bias

The quality of the evidence (i.e. our confidence) may be **INCREASED** when there is:

- ↑ A strong association
- ↑ A dose response relationship

# Quality assessment criteria

Quality of evidence	Study design	Lower if	Higher if
High	Randomised trial	<b>Study quality:</b> -1 Serious limitations -2 Very serious limitations  -1 Important <b>inconsistency</b>  <b>Directness:</b> -1 Some uncertainty -2 Major uncertainty  -1 <b>Sparse or imprecise data</b>  -1 High probability of <b>reporting bias</b>	<b>Strong association:</b> +1 Strong, no plausible confounders +2 Very strong, no major threats to validity  +1 Evidence of a <b>Dose response gradient</b>
Moderate			
Low	Observational study		
Very low			



## Judgements about the overall quality of evidence

- **Most systems not explicit**
- **Options:**
  - **Benefits**
  - **Primary outcome**
  - **Highest**
  - **Lowest**
- **Based on lowest of all the critical outcomes**
- **Beyond the scope of a systematic review**



## Strength of recommendation

The degree of confidence that the desirable effects of adherence to a recommendation outweigh the undesirable effects.



### Desirable effects

- health benefits
- less burden
- savings

### Undesirable effects

- harms
- more burden
- costs



## Categories of recommendations

Although the degree of confidence is a continuum, we suggest using two categories: strong and weak.

- **Strong recommendation: the panel is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects.**
- **Weak recommendation: the panel concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, but is not confident.**



# Judgements about the strength of a recommendation

Reasons for not being confident can include:

- absence of high quality evidence
- imprecise estimates
- uncertainty or variation in how different individuals value the outcomes
- small net benefits
- uncertainty whether the net benefits are worth the costs (including the costs of implementing the recommendation)



# Judgements about the strength of a recommendation

- **No precise threshold for going from a strong to a weak recommendation**
- **The presence of important concerns about one or more of the above factors make a weak recommendation more likely.**
- **Panels should consider all of these factors and make the reasons for their judgements explicit.**
- **Recommendations should specify the perspective that is taken (e.g. individual patient, health system) and which outcomes were considered (including which, if any costs).**



## Implications of a strong recommendation

- **Patients: Most people in your situation would want the recommended course of action and only a small proportion would not**
- **Clinicians: Most patients should receive the recommended course of action**
- **Policy makers: The recommendation can be adapted as a policy in most situations**



## Implications of a weak recommendation

- **Patients: The majority of people in your situation would want the recommended course of action, but many would not**
- **Clinicians: Be prepared to help patients to make a decision that is consistent with their own values**
- **Policy makers: There is a need for substantial debate and involvement of stakeholders**



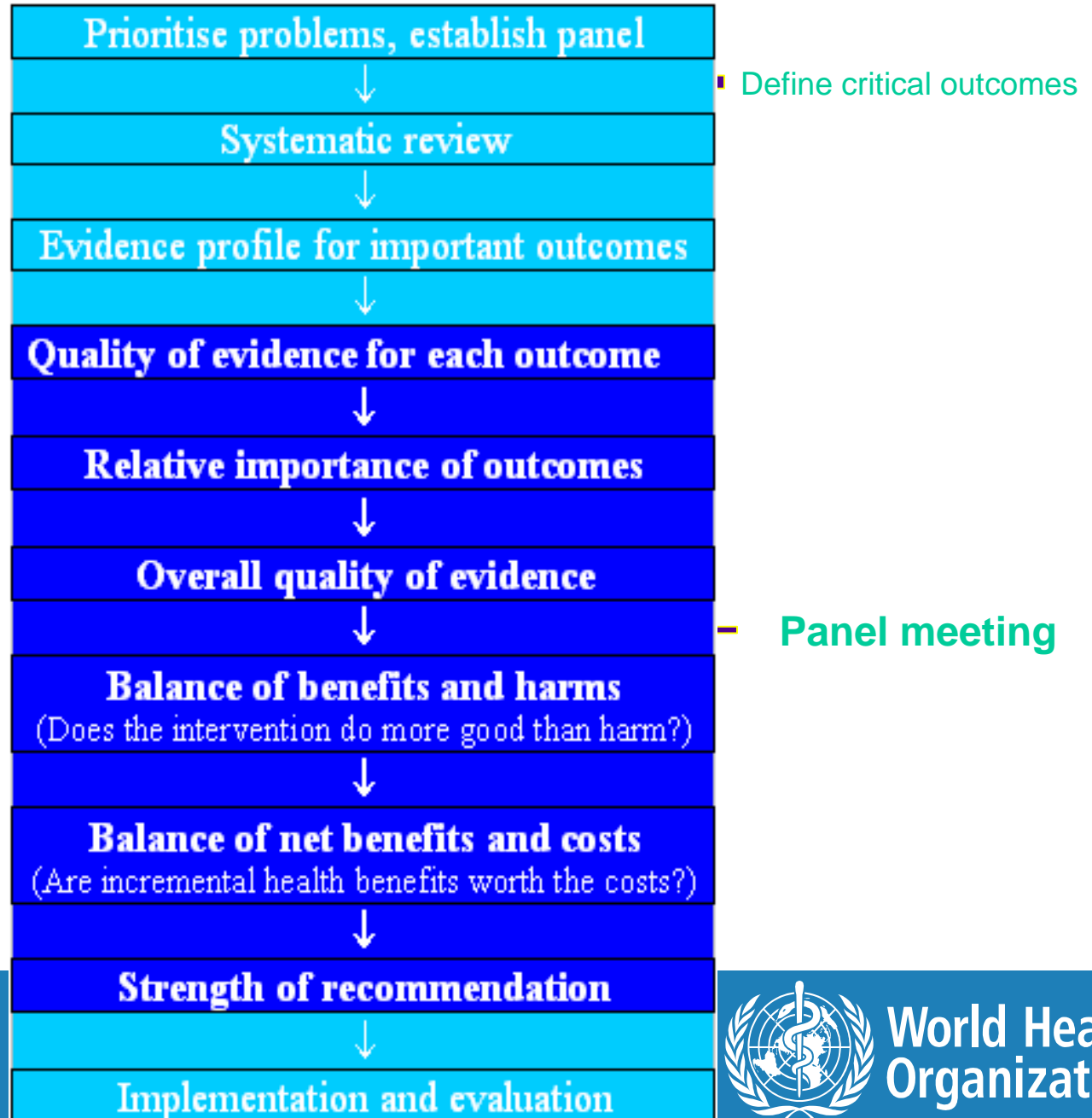
## **In practice....**

- **Post partum haemorrhage is the major cause of maternal mortality**
- **Effective interventions available –'active management'**
- **But...**
- **Which interventions?**
- **Who should use?**
- **Is one better than the other?**





# Guidelines development process



# QUESTION: Should active management of the third stage of labour be used by skilled providers for all women to prevent postpartum hemorrhage (PPH)?

Quality assessment						Summary of findings						
						No of patients		Effect			Quality	Importance
No of studies (Ref)	Design	Limitations	Consistency	Directness	Other considerations	Active management	Standard procedures	Baseline Risk (95%CI)	Relative risk (95%CI)	NNT (95%CI)		
<b>Benefits:</b>												
<b>Maternal deaths</b>												
0	-	-	-	-	-	-	-	-	-	-	-	8.5
<b>Admission to intensive care unit</b>												
0	-	-	-	-	-	-	-	-	-	-	-	6.4
<b>Blood loss ≥ 500 ml</b>												
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2</sup> ,3,17 -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none	3126	3158	min 8.3% (6.3, 10.3) max 17.9% (15.3, 20.5)	0.38 (0.32, 0.46)	min 8 (6.7, 11.2) max 16 (11.7, 24.7)	low quality ++oo	6.3
<b>Blood loss ≥ 1000 ml</b>												
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2</sup> ,3,17 -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none	3126	3158	min 1.5% (0.6-2.4) max 3.2% (2.0-4.4)	0.33 (0.21, 0.51)	min 41 (26.5, 90.1) max 73 (43.3, 225.5)	low quality ++oo	7.7



# QUESTION: Should active management of the third stage of labour be used by skilled providers for all women to prevent postpartum hemorrhage (PPH)?

## Quality assessment

No of studies (Ref)	Design	Limitations	Consistency	Directness	Other considerations
<b>Benefits:</b>					
<b>Maternal deaths</b>					
0	-	-	-	-	-
<b>Admission to intensive care unit</b>					
0	-	-	-	-	-
<b>Blood loss <math>\geq</math> 500 ml</b>					
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2</sup> ,3,17 -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none
<b>Blood loss <math>\geq</math> 1000 ml</b>					
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2</sup> ,3,17 -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none

# Quality assessment criteria

Quality of evidence	Study design	Lower if	Higher if
High	Randomised trial	<b>Study quality:</b> -1 Serious limitations -2 Very serious limitations  -1 Important <b>inconsistency</b>  <b>Directness:</b> -1 Some uncertainty -2 Major uncertainty  -1 <b>Sparse or imprecise data</b>  -1 High probability of <b>reporting bias</b>	<b>Strong association:</b> +1 Strong, no plausible confounders +2 Very strong, no major threats to validity  +1 Evidence of a <b>Dose response gradient</b>
Moderate			
Low	Observational study		
Very low			

# QUESTION: Should active management of the third stage of labour be used by skilled providers for all women to prevent postpartum hemorrhage (PPH)?

Quality assessment						Summary of findings						
						No of patients		Effect			Quality	Importance
No of studies (Ref)	Design	Limitations	Consistency	Directness	Other considerations	Active management	Standard procedures	Baseline Risk (95%CI)	Relative risk (95%CI)	NNM (95%CI)		
<b>Benefits:</b>												
<b>Maternal deaths</b>												
0	-	-	-	-	-	-	-	-	-	-	-	8.5
<b>Admission to intensive care unit</b>												
0	-	-	-	-	-	-	-	-	-	-	-	6.4
<b>Blood loss ≥ 500 ml</b>												
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2,3,17</sup> -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none	3126	3158	min 8.3% (6.3, 10.3) max 17.9% (15.3, 20.5)	0.38 (0.32, 0.46)	min 8 (6.7, 11.2) max 16 (11.7, 24.7)	low quality ++oo	6.3
<b>Blood loss ≥ 1000 ml</b>												
4 PW 00 <sup>1</sup> Ad 97 Br 88 Du 90 Hi 98	RCT	serious limitation <sup>2,3,17</sup> -1	no important inconsistency	some uncertainty about directness <sup>4,5</sup> -1	none	3126	3158	min 1.5% (0.6-2.4) max 3.2% (2.0-4.4)	0.33 (0.21, 0.51)	min 41 (26.5, 90.1) max 73 (43.3, 225.5)	low quality ++oo	7.7



By skilled providers								
	Critical outcomes	Studies n.	Patients n.	Baseline Risk without treatment (95%CI)	Relative effect (95%CI)	NNT	Quality	Notes
<b>Benefits</b>	Maternal deaths	No data available	-	-	-	-	-	-
	Blood loss $\geq 1000$ ml	<b>4</b> Ad 97 Br 88 Du 90 Hi 98	6284	min 1.5% (0.6-2.4) max 3.2% (2.0-4.4)	0.33 (0.21, 0.51)	min 41 (26.5, 90.1) max 73 (43.3, 225.5)	low quality ++00	1,2,3,4,5,17
	Need for blood transfusion	<b>5</b> Ad 97 Br 93 Br 88 Du 90 Hi 98	6477	5.7% (4.1-7.2)	0.34 (0.22, 0.53)	28 (18.7, 59,1)	moderate quality +++0	1,3,7,8,16
<b>Harms</b>	None judged critical							

# What would you recommend?

Although the degree of confidence is a continuum, we suggest using two categories: strong and weak.

- Strong recommendation: **the panel is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects.**
- Weak recommendation: **the panel concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, but is not confident.**



- **Active management of third stage of labour should be offered to by skilled attendants to all women (Strong recommendation, moderate quality).**



## What happens next?

- **Publication and dissemination**
- **Review of recommendations by countries – develop local treatment protocols**
- **Identification of indicators**
- **Implementation strategy**
- **.... Evaluation....**



## Summary

- Evidence is a tough taskmaster
- Systematic reviews and critical appraisal essential
- Content experts alone insufficient
- Transparent system required
- Judgements should be explicit
- To make it worth while, implementation and evaluation have to be integral to process



