

Fondamentaux à l'Evidence Based Santé

CM4 – Appraise

Vendredi 12 Septembre 2014

Sommaire

- 1- Les différents types de recherche
- 2- Evaluation de la fiabilité d'une recherche
 - 2.1 niveau de preuve
 - 2.2 évaluation critique
 - 2.3 Peer Review et impact Factor

Partie 1: Les différents types de recherche

Copyright 2003 by Randy Glasbergen.
www.glasbergen.com



**"I found 1837 web sites about 'alternative medicine'
but none of them recommend pizza or chocolate
for lowering our cholesterol."**

Typologie des études cliniques

- Descriptives
 - séries de cas
 - études transversales
 - série temporelle
- Analytiques
 - Etudes d'observations
 - cohorte
 - cas-témoins
 - Etudes expérimentales
 - essais cliniques
 - revues systématiques

Différents types de recherche en santé

Type d'études

```
graph TD; A[Type d'études] --> B[Etudes Descriptives]; A --> C[Etudes Analytiques]; B --> D[Description<br/>Décrire<br/>Générer des hypothèses]; C --> E[Observation<br/>Trouver relations<br/>présomption]; C --> F[Expérimental<br/>Causalité<br/>évaluer];
```

Etudes Descriptives

Description

Décrire
Générer des
hypothèses

Etudes Analytiques

Observation

Trouver relations
présomption

Expérimental

Causalité
évaluer

Partie 2: Evaluation de la fiabilité d'une recherche

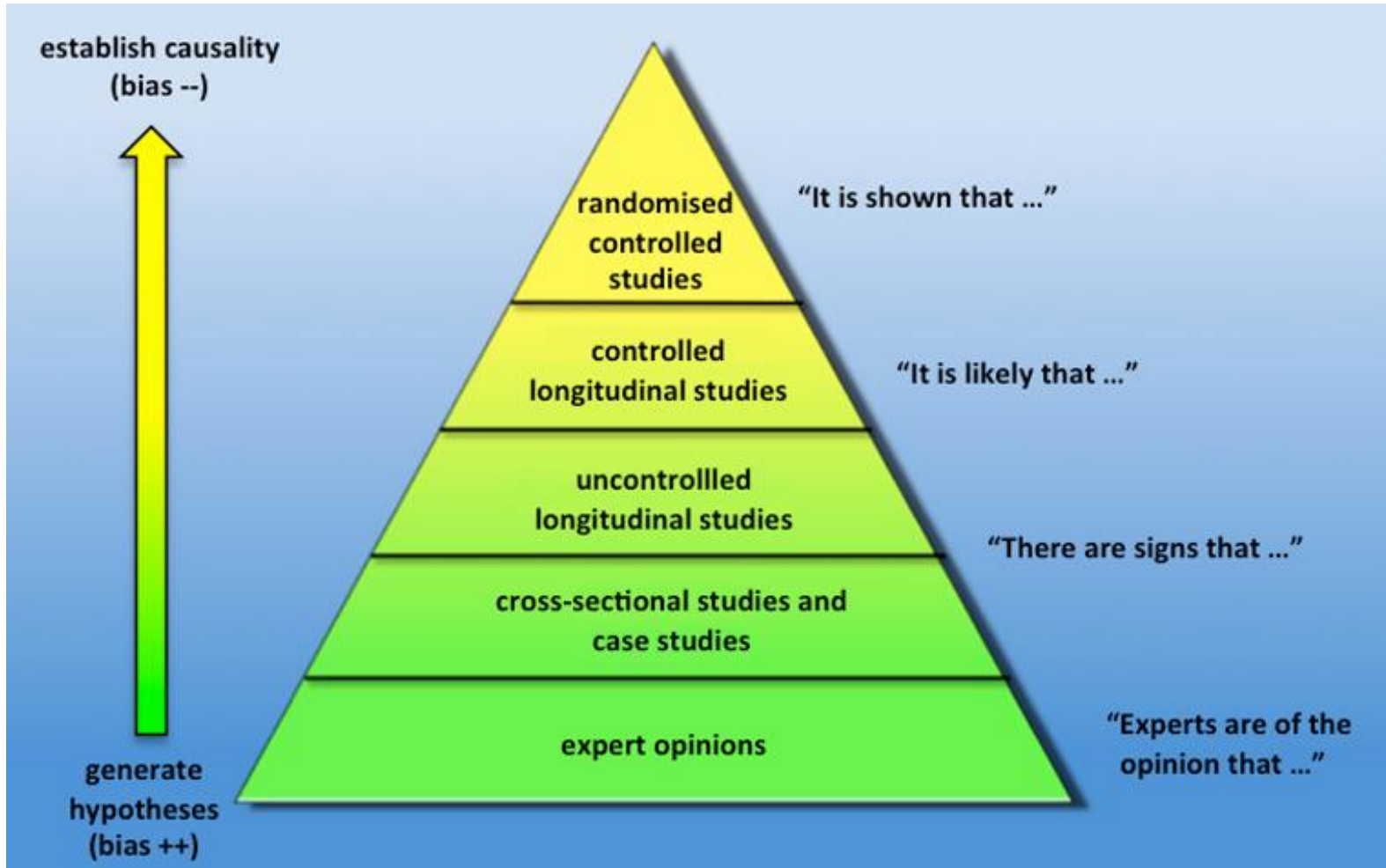
FRANK & ERNEST® by Bob Thaves



2.1 niveau de preuve

- toute information empirique concernant une relation entre des éléments est potentiellement une information factuelle
- mais toutes les informations n'ont pas la même validité (force pour convaincre)
- Il existe une hiérarchie
- qui varie selon le type de question

Différents niveaux de fiabilité de l'information



Hiérarchie pour des questions traitements, prévention, étiologie ou effets délétères

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard**	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or <i>n</i> -of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, <i>n</i> -of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or <i>n</i> -of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non-randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

<i>Research question</i>	<i>Randomised controlled studies</i>	<i>Controlled longitudinal studies</i>	<i>Cross- sectional surveys</i>	<i>Qualitative research</i>
Effectiveness: does it work?, does A work better than B?	++	+	-	--
Explanation : how does it work, why does it work?	--	-	+	++
Context: in what circumstances does it work, for whom?	--	-	+	++
Safety: will it do more good than harm?	++	+	+	+
Acceptability: will the target group accept the intervention / new method of working?	--	-	+	++
Prevalence: how often is this intervention / method applied / implemented?	--	--	++	--
Appropriateness: is this the right intervention / method for this target group?	--	-	+	++

D'après <http://www.cebma.org/frequently-asked-questions/what-are-the-levels-of-evidence/>

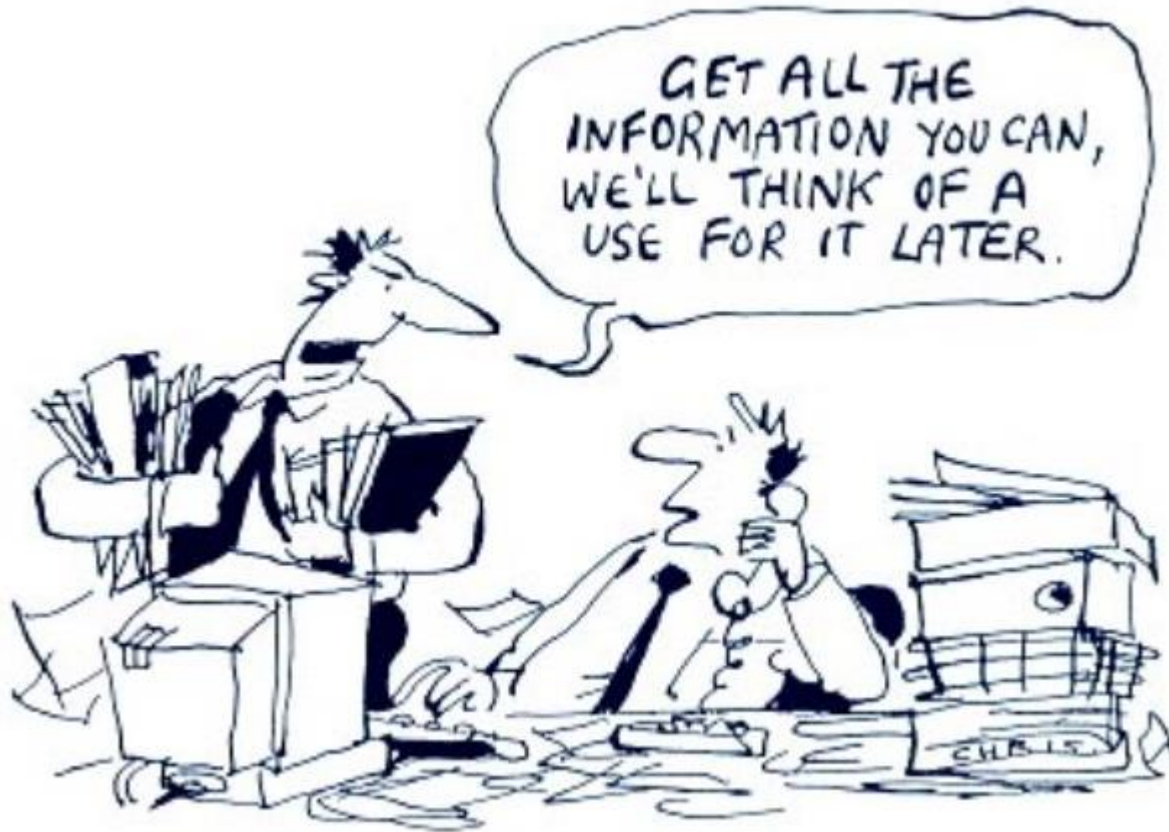
2.2 évaluation critique

- revues systématiques
- essai clinique randomisé
- recommandation de pratique clinique



**“If a vegetarian diet is good for losing weight,
how come they use grain to fatten pigs and cows?”**

Revue Systématique



Revue Systématique

Contexte:

- Information pour éclairer la prise de décision
- **filtrer** et **résumer** l'information validée avec une démarche scientifique
- avec une méthode assurant un haut niveau de validité (Egger, BMJ 1997; Koretz, Curr Opin Clin Nutr Metab Care. 2002)
- se distingue de revue narrative où auteurs considérés comme experts utilisent méthodes non formalisées, non systématiques, subjectives, pour rassembler et interpréter informations.

Objectifs d'un RS:

- collecter toutes les informations valides selon des critères préfinis d'inclusion pour répondre à une question précise de recherche
- réduire les biais en utilisant des méthodes claires et systématiques

Principes:

- **réaliser** recherche systématique pour identifier toutes études qui répondent aux critères d'inclusion
- **Évaluer** validité résultats études inclues,
par exemple les risques de biais
- **présenter** de manière systématique et **synthétiser** les caractéristiques et les résultats

Les étapes d'une revue systématique Cochrane

1. formuler une question (choix de la question)
2. déterminer des critères d'éligibilité (en terme de PICO)
3. recherche les études à inclure
4. sélectionner les études à inclure (consensus)
5. extraire les données (formulaire)
6. évaluer les risques de biais (choix de la grille)
7. analyser les résultats (exploration l'hétérogénéité, méta analyse)
8. réaliser la synthèse

- préférable de chercher une RS qui inclut une méta analyse (Murad, JAMA 2014)
- un EC est rarement suffisant pour répondre à une question clinique et peut souvent être biaisé (Ioannidis 2008; Born JCE 2009; Murad JAMA 2013)

Méta analyses = méthodes statistiques qui **combinent** les **résultats individuels** de plusieurs études (Kicinski, PLOSone 2013)

La méta analyse:

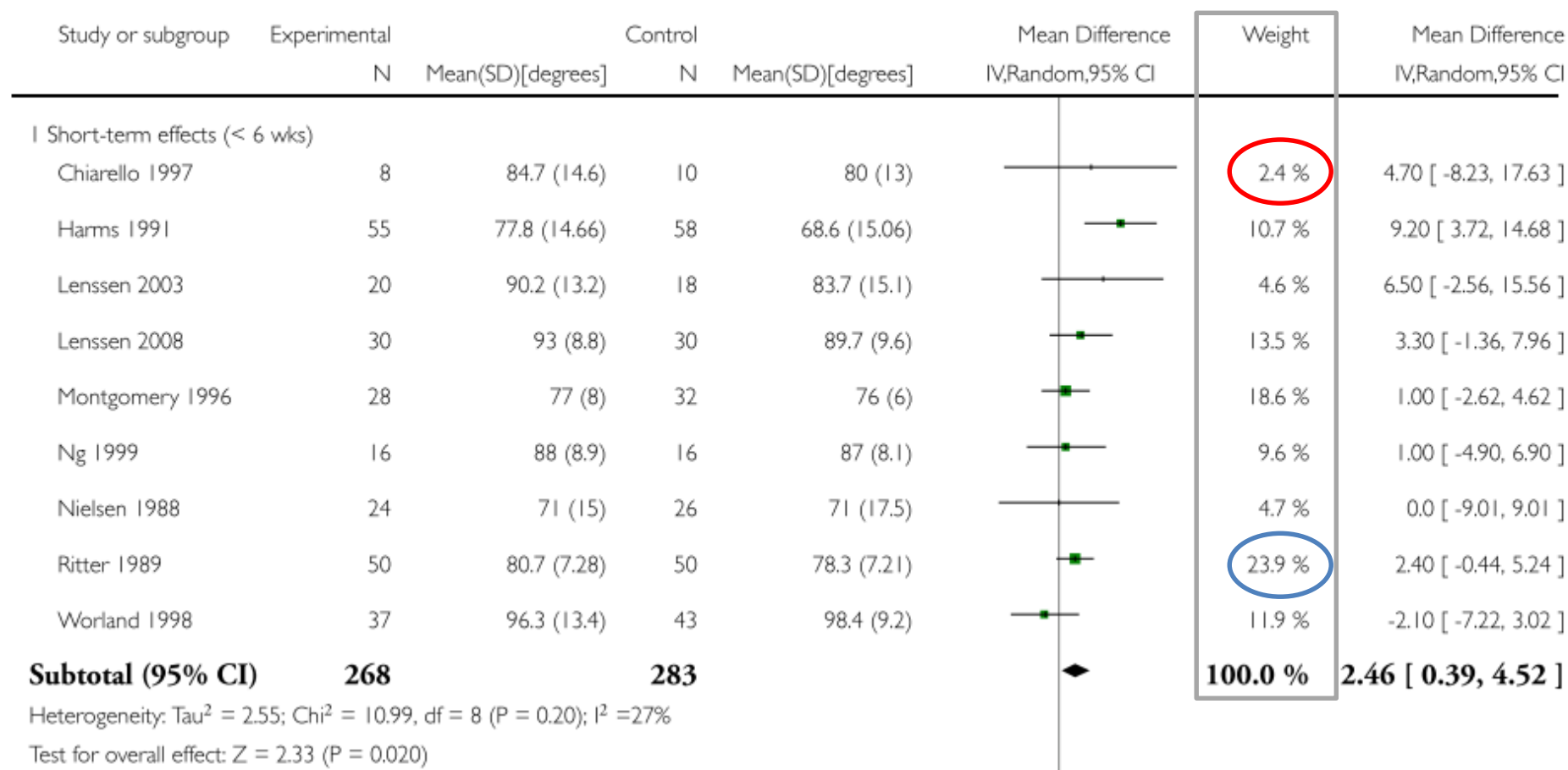
- fournit une **estimation** de **l'effet de taille** (magnitude d'une différence entre 2 traitements). Le type d'effet dépend du critère de jugement (moyenne, risque relatif, OD etc.)
- Les effets standardisés sont exprimés en multiples d'écart-type
- résultats exprimés généralement = **forest plot**
 - estimé individuel de chaque étude = carré dont la taille est proportionnel au poids de l'étude
 - IC représenté par une ligne horizontale
 - estimé combiné représenté par un diamant dont la largeur représente l'IC (valeurs possibles contenant l'effet véritable)

La méta analyse:

- la MA pondère l'estimation des effets individuels (weigh) selon la **précision** qui dépend du **nombre** d'événements ou du nb de sujets et de la **variabilité**
- Calcul variabilité entre études ou **hétérogénéité**
 - inspection visuelle (recouvrement IC, disparité des estimés) (**Hatala, CMAJ 2005**)
 - Statistique (Q^2 de cochrane, I^2 : 0-40%, 30-60%, 50-90%, 75-100%) (**Cochrane Handbook, Chap 9.5**)

Continuous passive motion following total knee arthroplasty in people with arthritis

Lisa A Harvey¹, Lucie Brosseau², Robert D Herbert³



Éléments de validité: lecture critique

- les résultats sont ils valides?
 - question justifiée?
 - stratégie de recherche cohérente et exhaustive?
 - sélection et évaluation reproductible?
 - qualité des études inclues?

Éléments de validité

- quels sont les effets?

- sont-ils comparables entre études?

inspection de l'hétérogénéité

- quels sont les résultats?

calcul de l'effet de taille selon type critères évaluation
(dichotomique ou continue)

- comment sont-ils précis? intervalle de confiance.

+ précision, + de contribution

Éléments de validité

- peut-on les appliquer ?
 - critères de jugement importants utilisés?
 - crédibilité des analyses en sous groupe ? par chance ou effet vrai
 - confiance dans les effets estimés?
- **Cotation GRADE** (-: risque de biais élevé, inconsistance, imprécision, indirecte ou biais de publication; + : taille,, relation dose /effet)
- balance bénéfice/risques?

http://www.gradeworkinggroup.org/

File Add View Options Help

☐ New ☐ Open ☐ Print ☐ Save ☐ Undo all changes ☐ Add profile group ☐ Add profile ☐ Add outcome ☐ Import from RevMan ☐ Preview SoF table

Profiles tree << Edit

- ☒ High-intensity versus low-intensity ph
- ☒ High-intensity versus low-intensity ph
- ☒ High versus Low intensity exercise
- ☐ Pain (at study completion)
- ☐ Pain (at study completion) - St
- ☐ Physical Function (at study co
- ☐ Physical Function (at study cor
- ☐ Quality of Life (at study compl
- ☐ Adverse events

Name of outcome: Pain (at study completion) Short name: Importance: ---

Assessed/measured with: WOMAC (VAS or likert) from: 0 to 20

No of studies: 4 Study design: randomised trials

Length of follow up: range 8 to 24 weeks

Downgrade quality of evidence

Risk of bias very serious (-2)

Inconsistency no

Indirectness no

Imprecision no

Publication bias undetected

Upgrade quality of evidence:

Large_effect no

Plausible confounding would change the effect no

Dose response gradient no

Quality of evidence: **LOW**

Profile: High versus Low intensity exercise for

Pain (at study completion) (follow-up 8 to 24 weeks; measured with: WOMAC (VAS or likert) from: 0 to 20; Better indicated by lower values) | 4 studies

Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Importance
randomised trials	very serious ¹	no serious inconsistency	no serious indirectness	no serious imprecision	none	
Patients (High versus Low intensity exercise)	Control	Relative effect		Absolute effect		Quality
162	151	-		MD 0.84 lower (1.63 to 0.04 lower)		LOW

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Multidisciplinary compared to usual care for chronic low back pain						
Patient or population: Patients with chronic low back pain Intervention: Multidisciplinary Biopsychosocial Rehabilitation Comparison: Usual care						
Outcomes	Baseline	Comparative effect (95% CI)		No. of participants (studies)	Quality of the evidence (GRADE)	Comments
Back pain long term 0-10 Numerical or visual scale, where 0 equals no pain at all and 10 is the worst pain imaginable. Follow-up: median 12 mth	# The baseline for the most representative study is 5.8 out of 10	The mean back pain long term in the MBR groups was 0.21 standard deviations lower (0.37 to 0.04 lower)		821 (7 studies)	⊕⊕⊕○ moderate ¹	This is a small effect that may be clinically relevant in this patient group
Disability long term Mostly Roland Morris 24-point scale where 0 equals no disability at all and 24 is seriously disabled. Follow-up: median 12 mth	# The baseline for the most representative study is 11.4 out of 24	The mean disability long term in the MBR groups was 0.23 standard deviations lower (0.4 to 0.06 lower)		722 (6 studies)	⊕⊕⊕○ moderate ¹	This is a small effect that may be clinically relevant in this patient group
	Assumed risk* Usual care	Corresponding risk MBR	Relative effect (95% CI)			
Work long term Proportion working Follow-up: median 12 mth	744 per 1000	751 per 1000 (679 to 810)	OR 1.04 (0.73 to 1.47)	1360 (7 studies)	⊕⊕⊕○ moderate ¹	This difference is not statistically or clinically relevant
Adverse events	not estimable	not estimable	not estimable	0	No evidence	

AUTHORS' CONCLUSIONS

Implications for practice

Choosing an MBR intervention over usual care or a physical treatment program for chronic low back pain is likely to result in a positive effect on pain and disability outcomes. It is also likely that MBR will have a beneficial effect on work outcomes compared to physical treatment. However, given the moderate size of these effects and the potentially high cost of an intensive intervention, in terms of both the monetary and time burden, the decision to refer to MBR requires some consideration. While our subgroup analyses were inconclusive regarding the influence of higher or lower symptom intensity at baseline, it would appear there is little to gain by referring those without substantial physical and psychosocial impacts of their condition to such an intervention. Clinical practice guidelines ([Dagenais 2010](#)) commonly recommend assessment and treatment of physical and psychosocial factors and referral to appropriately trained clinicians for management of these factors where present. This recommendation would seem more appropriate than a recommendation of MBR simply based on chronicity of symptoms.

Outils de lecture critique

- registre: PROSPERO, Cochrane
- reporting: PRISMA, MECIR
- qualité : AMSTAR

Sources spécifiques:

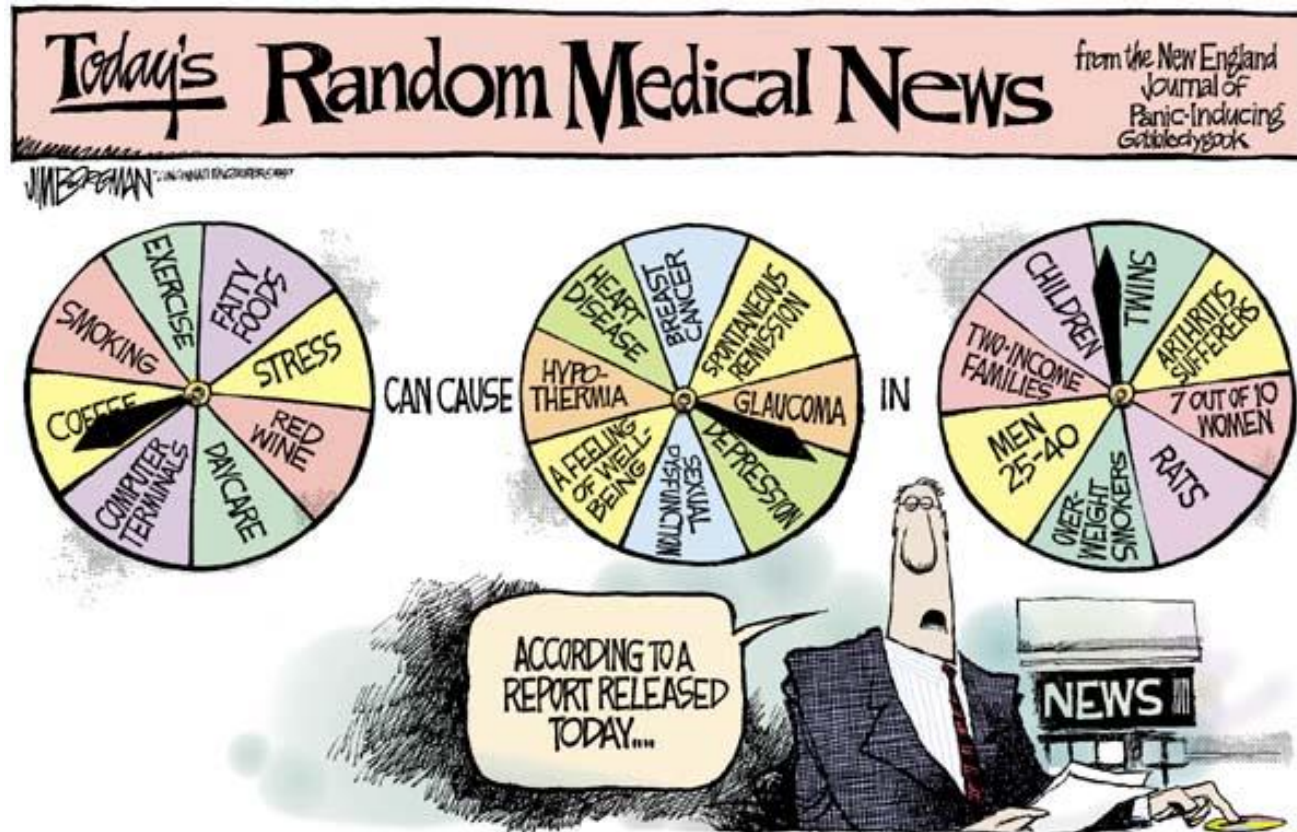
- Cochrane library
- Centre for Reviews and Dissemination databases (inclut DARE)
- Joanna Briggs institute
- The Campbell Collaboration Library of Systematic Reviews
- The Evidence for Policy and practice Information and Co-ordinating Centre (EPPI-Centre)

Les autres bases : Medline, Pubmed, CINAHL, PEDro, OTseeker, PsycBITE, SpeechBITE, SCOPUS, Web of Science

Limites et perspectives

- biais de publication
- small studies effect (effet + important que grandes études)
- inconclusivness
- influencée par qualité de la recherche: garbage in, garbage out
- méthodes qualitatives ou mixtes peu considérées
- implémentation: peu ou pas lues ([Wallace, BMJopen 2012](#))

Essai Clinique



Buts

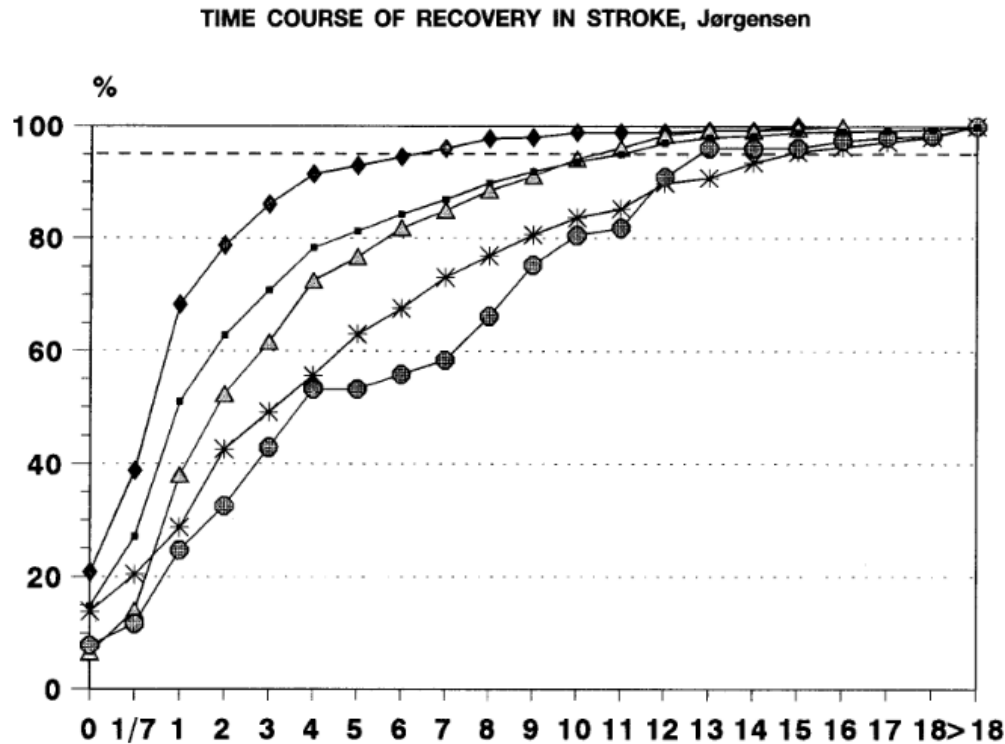
- s'assurer que la comparaison des groupes (de taille suffisante) diffère seulement par l'exposition à l'intervention concernée
- déterminer le rôle d'un ou plusieurs facteurs dans l'étiologie ou le pronostic d'une maladie

Facteurs qui influent sur la réponse du traitement

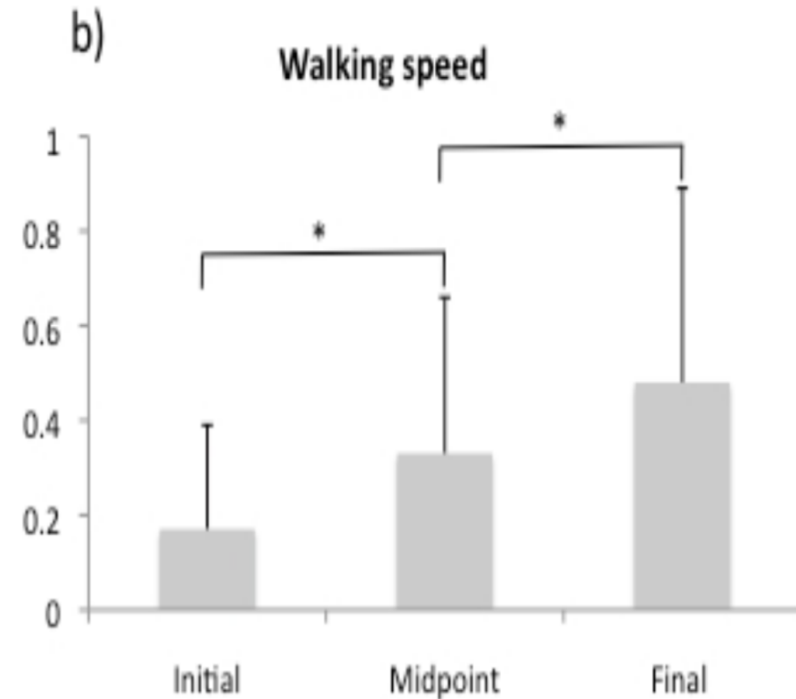
- facteurs de confusion
- biais
- effet vrai du traitement

- 3 composantes non spécifiques de l'effet d'un traitement (Sedgwick, BMJ 2012)
 - *L'évolution spontanée*: amélioration, progresser, stabiliser, disparaître sans recevoir aucun traitement
 - *Effet Hawthorne*: changement comportement ou critère d'évaluation attribuable à participation à une recherche – attention accordée
 - *L'effet placebo* : lorsqu'un patient reçoit n'importe quel traitement (perçu comme un traitement) mais aussi en réponse à un rituel ou interaction entre patient-soignant

Exemple de récupération motrice chez l'AVC



D'après Jørgensen, Arch Phys Med Rehabil 1995



D'après Conesa J Neuroeng Rehabil. 2012 Feb 13;9(1):13.

Placebo interventions for all clinical conditions

Asbjørn Hróbjartsson¹, Peter C Gøtzsche¹

Larger effects of placebo interventions were associated with physical placebo interventions (e.g. sham acupuncture), patient-involved outcomes (patient-reported outcomes such as pain, and observer-reported outcomes involving patient cooperation, such as depression rating scales), small trials, and trials with the explicit purpose of studying placebo. Larger effects of placebo were also found in the trials that falsely informed patients that the study compared two active treatments with no-treatment.

Citation: Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews* 2010, Issue 1. Art. No.: CD003974. DOI: 10.1002/14651858.CD003974.pub3.

- Traitements concomitants

Risque de contamination

Difficile à maîtriser, maladie chronique (+++)

Un biais = erreur systématique

- un biais est un écart de la réalité
- s'oppose à la chance . Biais = erreur systématique
- sens: - sous estimation de l'effet du traitement
 - sur estimation de l'effet du traitement
- nb types de biais possibles : sélection, réalisation, performance, attrition, publication

exemple études non randomisées

L'effet vrai du traitement

Il faut contrôler les autres facteurs pour l'évaluer:

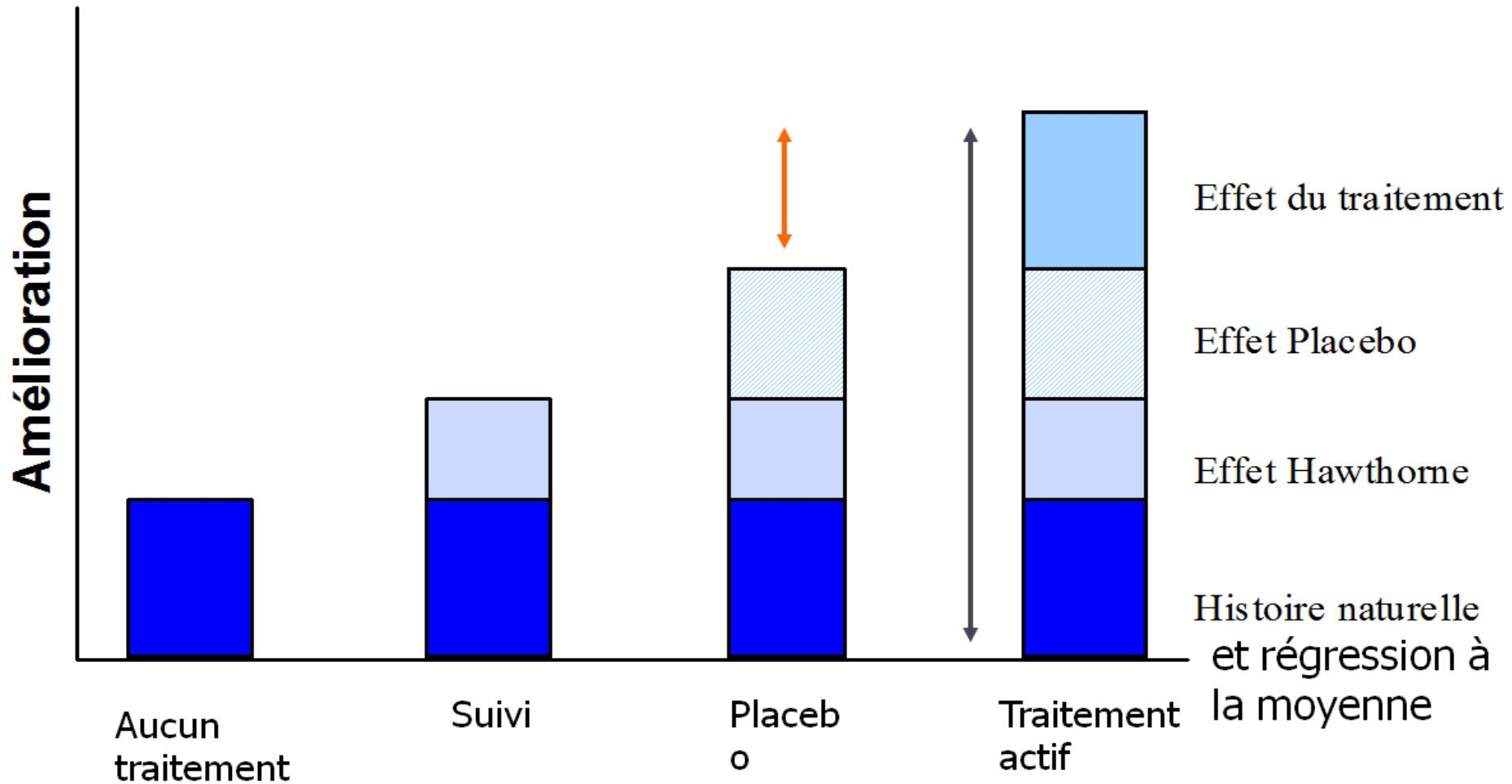
Qui dit évaluation dit comparaison

observations avant / après, causalité



changement provoqué par d'autres facteurs (confusion)

Il faut un **groupe contrôle** pour comparer



Les grands principes de l'évaluation thérapeutique

Principes de bases

3 Principes essentiels:

- Comparaison —————> Groupe témoin
- Causalité —————> Effet du Tx
- Significativité —————> Tests statistiques

Éléments de validité: lecture critique pour une étude évaluant l'efficacité d'un traitement

- les résultats sont ils valides? validité interne
 - sélection des patients
composition des groupes, randomisation, allocation cachée?
 - performance
aveugle
 - évaluation
Exclus de l'analyse? ITT? Données manquantes?

Éléments de validité

- quels sont les effets?
 - importance de l'effet traitement?
- Calcul effet de taille (ET)
 - précision de l'effet traitement?
- intervalle de confiance? taille de l'essai

Éléments de validité

- peut-on les appliquer ? Validité externe
 - patients compatibles avec ceux de ma pratique?
critères inclusion, exclusion
 - critères d'évaluation importants et pertinents?
critères de substitution ou surrogates? points
temporels de l'évaluation? centrés sur le patient?
 - rapports bénéfice / risques?
calcul NNT

Outils de lecture critique

- registre: WHO ICTRP
- reporting: CONSORT et autres
- qualité : ROB tool, JADAD scale, PEDro scale etc.

<http://www.cebm.net/critical-appraisal/>

items du Risk of Bias Tool proposé par la Cochrane

☐ Risk of bias table 🗨

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk ▼	Quote: "using a random numbers table from a random integer generator"
Allocation concealment (selection bias)	Unclear risk ▼	Comment: unclear if the list of randomisation is concealed
Blinding of participants and personnel (performance bias)	High risk ▼	Comment: participants and staff were not blinded
Blinding of outcome assessment (detection bias) Subjective	High risk ▼	Comment: participant not blinded
Blinding of outcome assessment (detection bias) Objective	Low risk ▼	Quote: "All evaluations were performed by the same examiner, who was unaware of the subjects' group assignments."
Incomplete outcome data (attrition bias)	High risk ▼	Quote: "3 subjects in HR group discontinued the exercise intervention due to severe knee pain. We speculate that the intensity and repetitions of the resistance training might be 2 factors influencing the training effect in patients with knee OA" Comment: despite an ITT analysis, the dropout rate was unbalanced between the 2 groups (3 vs 0) and seems related to the intervention. No imputation technique was described
Selective reporting (reporting bias)	Unclear risk ▼	Comment: Insufficient information. No protocol registration. Primary and secondary outcomes are not specified; unclear what they are.
Other bias	Low risk ▼	Comment: power sample size calculation. No difference at baseline between groups. From the data provided, no indication for other important risks of bias.

Sources électroniques :

publiques (Pubmed, Medline) , éditeurs (Ebscohost, Elsevier, Wiley, Thomson Reuters), filtrées (NHS evidence, Tripdatabase, UpToDate, ClinicalEvidence, PEDro, Otseeker etc.), journaux “médicaux”

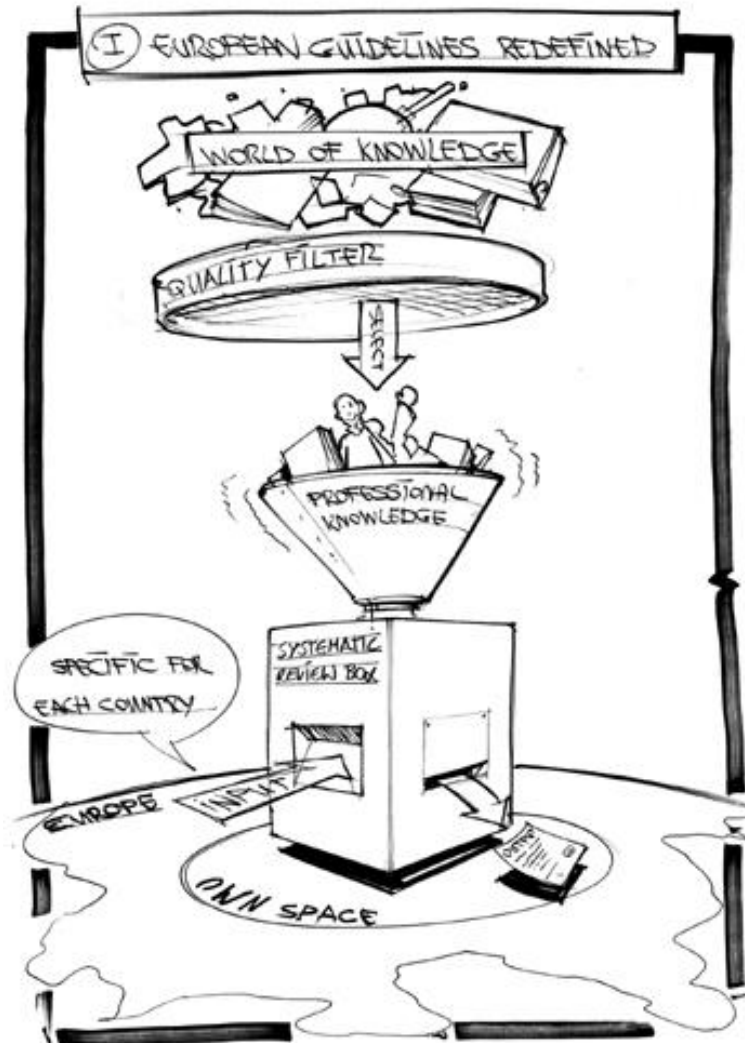
Littérature grise

- google / google scholar
- congrès / sociétés savantes
- thèses
- électroniques
- données personnelles d’auteurs
- agences de recherche

Limites et perspectives

- éthique?
- répond mal aux problématiques des patients (validité externe. Ex femmes enceintes)
- faisabilité (difficile à conduire, nb de sujets à inclure)
- coûts
- qualité - validité des résultats ([Ioannidis, PlosMedicine 2005](#))
- mal adapté aux interventions non médicamenteuses (NPT) (complexes)

Recommandation de bonne pratique



Définition:

« propositions développées méthodiquement pour aider le praticien et le patient à rechercher les soins les plus appropriés dans des circonstances cliniques données »
(Field, National Academy Press 1992)

- ne fournissent pas de quantification précise
- reposent sur consensus groupe de décideurs qui examinent les informations validées et décident de leurs implications (Guyatt, JAMAevidence, chap 21 2008)

Différents méthodes de recommandations de bonne pratique:

- recommandation pour la pratique clinique (RPC)

réalise des recommandations basées sur propositions élaborées à partir d'une **revue systématique**

- par consensus formalisé (RCF)

méthode de consensus des participants pour formaliser degré d'accord de propositions élaborées à partir d'une **analyse critique de la littérature**

Les étapes pour réaliser un RPC

- Préparer la réalisation
 - Formuler une question, définir les critères d'évaluation
- Examen systématique de la littérature
 - Revue systématique, inclusion/exclusion, évaluation critique
- Rédaction de la RPC
 - Proposition, consultation, formulation de recommandations
- Evaluer la RPC
 - Evaluer son impact, adaptation locale

Éléments de validité

- composition du groupe d'élaboration
 - inclure différents participants (professionnels, méthodologistes , patients, experts du domaine)
- procédure de décision
 - indiquer la méthode de décision avant de débiter
- conflit d'intérêt
 - déclaration des conflits financiers et non financiers
- objectif de la recommandation
 - indiquer clairement le(s) objectif(s) et champ d'application

Éléments de validité

- méthode
 - description en détail de la méthode d'élaboration
- revue de littérature
 - décrire la façon pour identifier et évaluer les informations
- recommandations
 - affirmations claires sur les bénéfices, effets délétères et si possible les coûts
- Système de notation
 - utilisation d'un système de notation pour informer sur la qualité et fiabilité des informations et du niveau de preuve

Éléments de validité

- Évaluation externe
 - relecture par des parties prenantes avant publication
- Expiration et mise à jour
 - doit inclure une date d'expiration et les prévisions de mise à jour
- Aide financière et sponsoring
 - déclaration du soutien financier pour la recherche de littérature et la réalisation de la recommandation

Outils de lecture critique

- registre: ?
- reporting: COGS checklist (Shiffman, Ann Intern Med 2003)
- qualité : AGREE II

<http://www.iom.edu/cpgstandards>

Sources électroniques :

régionales: agences gouvernementales (HAS, Inca, INPES, ANSM etc.) sociétés savantes,

internationales: NICE, G-I-N, National Guideline Clearinghouse

bases de données électroniques: Medline etc.

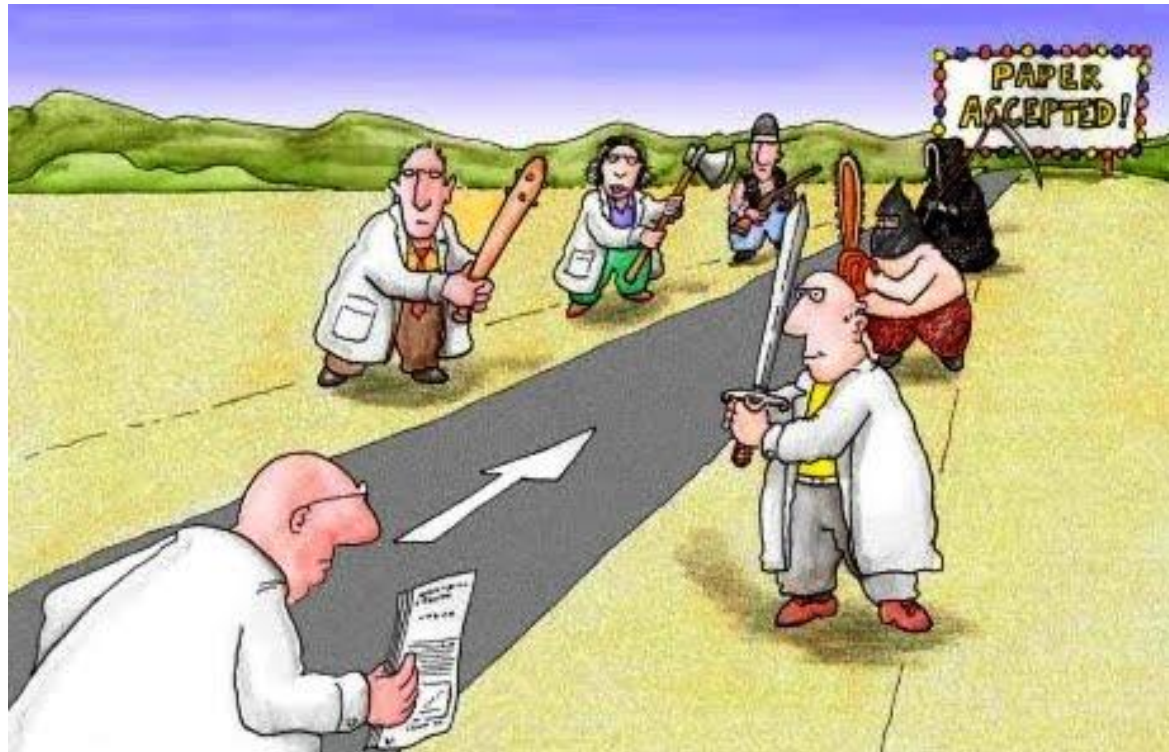
Limites et perspectives

- variabilité des recommandations (Burgers, Diabetes Care 2002)
- émet des avis pas de données quantifiées
- qualité insuffisante (Shaneyfelt, Arch Intern Med. 2012)
- recommandations basées sur des données quantitatives insuffisantes (Alexander, JCE 2014)
- valeur d'une recommandation basée sur opinion d'experts quand il n'y a pas de RCT? (ie conférence de consensus)
- confusion entre méthodologie et force de la recommandation (grade A, B, C ou I,II, III)

Limites et perspectives

- conflit d'intérêt majeur ([Lenzer, BMJ 2013](#))
- questions posées sont éloignées des données de la littérature
- elles ne sont pas suivies ([Lenzer, BMJ 2013](#))
- effets sur la santé modérés
- coût de réalisation élevés / durée de validité limitée

2.3 Peer Review et Impact Factor



le peer review

« évaluation de résultats de la recherche ou d'un appel à projet pour sa compétence, son intérêt et son originalité par des experts qualifiés qui effectue des recherches et soumettent des articles pour publication dans le même domaine » Morey, J Allied Health 2011

Objectifs:

- améliorer la transparence
- sélectionner les articles de meilleur qualité (Jackson, PLOSone 2011)

- peu d'information en faveur (Jefferson, Cochrane Database of Systematic Reviews 2008)
- faible impact mais positif sur le reporting (Hopewell, BMJ 2014)
- de nouvelles pratiques (open peer review, ...)

- **L'impact factor** est: rapport entre nombre de citations d'articles d'une revue par le nombre d'articles publiés par cette revue
- nombreuses limites à l'utilisation de l'impact factor comme marqueur de la qualité d'une publication (rapport de l'académie des sciences, 17 janvier 2011)
- La citation = mesure reflète manière dont une étude occupe l'esprit des scientifiques et retient l'attention – en bien ou en mal (Ioannidis., JAMA, 2005). Indicateur de notoriété d'une revue

En conclusion

- la fiabilité de l'information doit être évaluée
- méthode et qualité de l'étude sont des critères principaux pour évaluer cette fiabilité
- Ils doivent être adaptés au type de question, de recherche
- l'applicabilité et l'intérêt clinique des résultats doivent être considérés au regard de chaque personne à traiter

FIN

MERCI POUR VOTRE ATTENTION

DES QUESTIONS?