

# How to incorporate Patient-Reported Outcomes (PROs) in Cochrane reviews?

Caroline Terwee, Donald Patrick,  
Gordon Guyatt, Riekje de Vet



# Aim of the workshop

To provide some guidelines for how PROs can be incorporated in Cochrane reviews and how information can be gathered about the quality of the PROs.

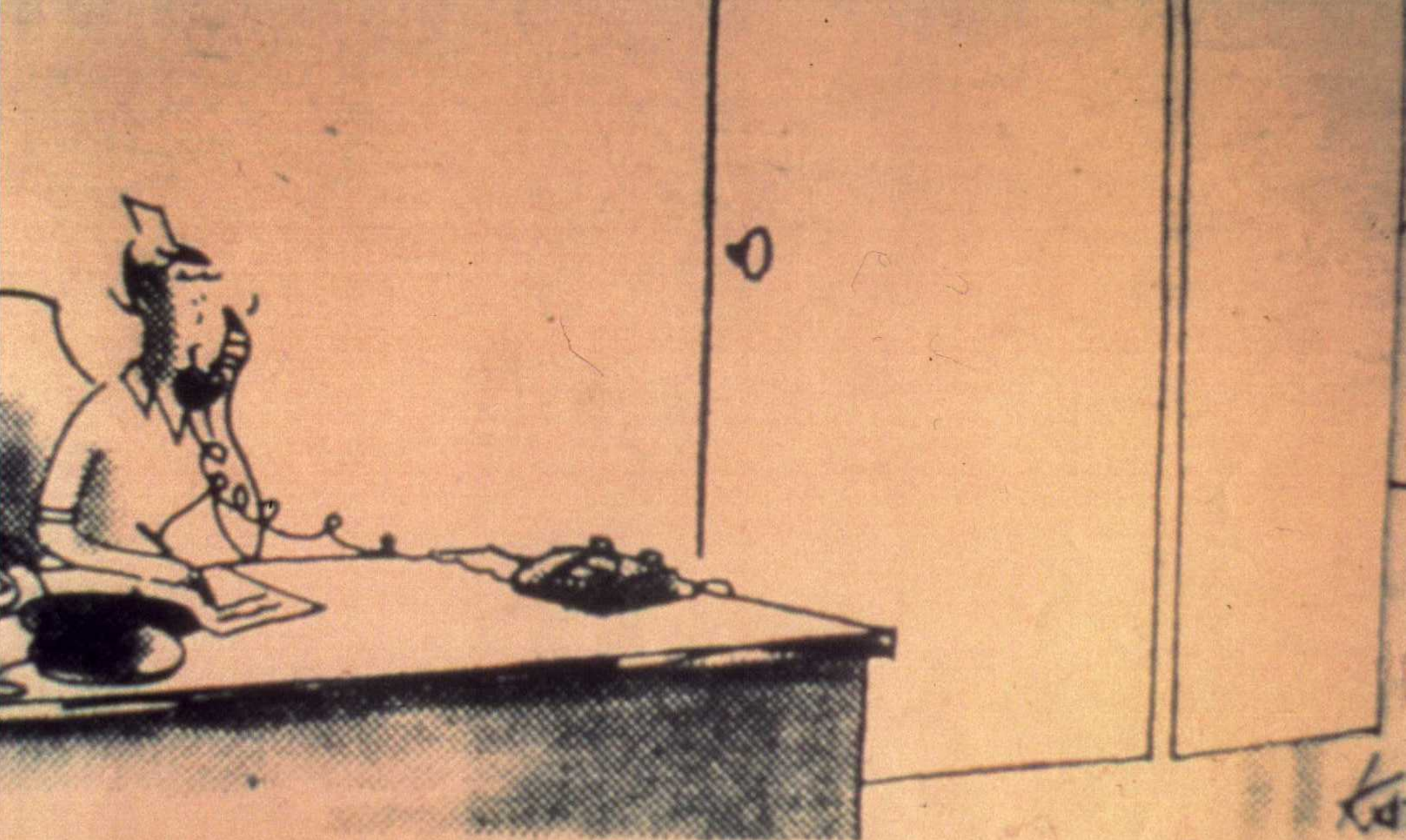
# Content of the workshop

## Example Cochrane review

1. Are PROs considered relevant outcomes in the review?
2. Were PROs used as outcome measures in the trials?
3. Is the quality of the PROs taken into account in the review?
4. How are PROs handled in the analyses of the review?
5. How are PROs included in the conclusions of the review?

What has been done? What could be better?

# Example: Urinary incontinence



“Urology Department. Can you hold?”

# Patient-Reported Outcomes in Urinary Incontinence

- ❖ Urinary episodes diary –common in most urinary studies
- ❖ Symptoms
- ❖ Symptom bothersomeness
- ❖ Impact on functional status
- ❖ Perceptions
- ❖ “Quality of life”

# How do you know what is being measured in a clinical trial when doing a Cochrane review?

- ❖ Read the article and list what the authors of the study say is being measured
- ❖ Find a copy of the instruments referenced
- ❖ Analyse the content of the instrument: is it a symptom, it is a perception, is it a functional impact?

# What is the I-QOL?

- ❖ a 22-item measure for the subjective evaluation of quality of life in incontinence and its treatment
- ❖ self-administered (approximately 5 minutes)
- ❖ responses made on 5-point Likert scale (extremely / quite a bit / moderately / a little / not at all)
- ❖ scored as three domain scores (Limiting Situations, Psycho-social and Social Embarrassment) and one total score
- ❖ easy to score, scores transformed on a zero to 100-pt scale (higher score = better quality of life)



# Sample items from the I-QOL

I worry about being able to get to the toilet on time (*Limiting*)

I feel depressed because of my incontinence (*Psychosocial*)

I feel frustrated because my incontinence prevents me from doing what I want (*Psychosocial*)

Because of my incontinence, I don't feel free to leave my home for long periods of time (*Limiting*)

I worry about others smelling urine on me (*Social Embarrassment*)

I worry about being embarrassed or humiliated because of my incontinence (*Social Embarrassment*)

# What makes a good instrument for evaluating treatment?

- ❖ Generated by persons with the condition and in their own language
- ❖ Supported by a clearly defined and tested conceptual and measurement model
- ❖ Has good measurement properties
- ❖ Is able to detect change when change occurs used.
- ❖ Scores can be interpreted (i.e., percent improved).

# Example Cochrane review

## **OBJECTIVES**

To determine the effects of pelvic floor muscle training in the management of female urinary (stress, urge, and mixed) incontinence.

The following hypothesis was tested:

- that pelvic floor muscle training is better than no treatment, placebo, sham, or any other form of inactive control treatment.

# 1. Are PROs considered relevant outcomes?

Do you consider PROs relevant outcomes for the hypothesis?

# 1. Are PROs considered relevant outcomes?

“The primary outcomes of interest were:

1) symptomatic cure or improvement (reported by the woman and not the clinician)

2) symptom and condition specific quality of life assessment (for example Incontinence Impact Questionnaire, Kings Health Questionnaire)”

Are these PROs?

Are these outcomes relevant for patients?

## 2. Were PROs used as outcome measures in the trials?

### **Included studies**

“Thirteen trials were included.

Four trials reported data on patient-perceived cure or improvement.

Two trials used psychometrically robust questionnaires for assessment of incontinence symptoms and/or the impact of these symptoms on quality of life, or both.”

### 3. Is the quality of the PROs taken into account in the review?

What was measured in the trials?

"Many different scales were used to measure patient response to treatment, including Likert scales, visual analogue scales and percent reduction in symptoms.

Bø and colleagues (Bø 1999) used the Bristol Female Lower Urinary Tract Symptoms Questionnaire (B-FLUTS).

Schagen van Leeuwen and co-workers (van Leeuwen 2004) reported the Quality of Life in Persons with Urinary Incontinence (I-QoL) score."

### 3. Is the quality of the PROs taken into account in the review?

#### **Assessment of methodological quality**

Assessment of methodological quality was undertaken by both review authors using the Cochrane Incontinence Group's criteria, which includes assessment of quality of random allocation and concealment, description of dropout and withdrawal, analysis by intention to treat, and blinding during treatment and at outcome assessment. Any disagreements were resolved as previously described.

(how) was the quality of PROs taken into account?



### 3. Is the quality of the PROs taken into account in the review?

“Bø and colleagues (Bø 1999) used the B-FLUTS, which has established validity, reliability and responsiveness to change for evaluation of urinary incontinence symptoms in women (Donovan 2005).

“Schagen van Leeuwen and co-workers (van Leeuwen 2004) reported I-QoL score; I-QoL has established validity, reliability and responsiveness to change for assessing quality of life impact of urinary incontinence (Donovan 2005).

What is established validity, reliability, and responsiveness?

# How to evaluate the quality of PROs?

Ideal situation:

1. collect all studies on the development and measurement properties of the PROs
2. evaluate the quality of the studies on measurement properties
3. evaluate the measurement properties of the PROs

# How to evaluate the quality of PROs?

## Alternative:

1. use high quality systematic reviews of measurement properties of PROs
2. [www.cosmin.nl](http://www.cosmin.nl)  
215 reviews of health status measurement instruments

3 reviews - most recent one:

Avery KN, Bosch JL, Gotoh M, Naughton M, Jackson S, Radley SC et al. Questionnaires to assess urinary and anal incontinence: review and recommendations. J Urol 2007; 177(1):39-49.

# How to evaluate the quality of PROs?

If systematic reviews of measurement properties are not available:

1. Perform a quick literature search

- name of the PRO [ti]
- AND precise search filter for studies on measurement properties

[www.cosmin.nl](http://www.cosmin.nl)

# Minimal requirements for PROs

1. Test-retest reliability, validity, and responsiveness studied
2. Sample size adequate ( $n \geq 50$ )
3. Measurement properties should have been evaluated in the population of interest
4. Results adequate

# Minimal requirements for PROs

## Adequate results:

1. test-retest reliability  $> 0.70$ ;
2. validity: at least moderate correlation with comparable instrument or significant difference between two groups that are expected to differ in the construct of interest;
3. responsiveness: at least moderate correlation between change in the instrument and change in comparable instrument or significant difference between change in two groups that are expected to change differently in the construct of interest;

## 4. How are PROs handled in the analyses of the review?

What was reported in the trial?

Bø 1999:

No information of B-FLUTS (wrong reference)

Bø 2000:

Percentage of women with problems in lifestyle and sex-life variables before and after intervention, and the difference between the groups after intervention

## 4. How are PROs handled in the analyses of the review?

Bø 2000:

Table III. Percentage of women with 'little', 'some', and 'much' problems in sex-life variables before and after 6 months intervention measured by B-FLUTS

		PFM exercise	Control	Difference between groups after intervention	Cochran-Mantel- Haenszel
Sex-life spoilt by urinary symptoms	Before	40.0%	46.2%	0.03	0.9
	After	16.7%	50.0%		
Problem with sex-life spoilt by urinary symptoms	Before	33.3%	52.2%	0.02	0.02
	After	11.1%	50.0%		
Problem with pain in intercourse	Before	33.4%	20%	0.1	<0.05
	After	10.5%	33.3%		
UI with intercourse	Before	20.0%	45.8%	0.02	
	After	10.5%	41.7%		



## 4. How are PROs handled in the analyses of the review?

What was reported in the review?

Bø 1999:

“Fewer women in the PFMT group reported that urinary incontinence symptoms interfered with activity, or were problematic.”

What could be better?

## 4. How are PROs handled in the analyses of the review?

What was reported in the trial?

Van Leeuwen 2004:

Table 2

Secondary efficacy variables by treatment group

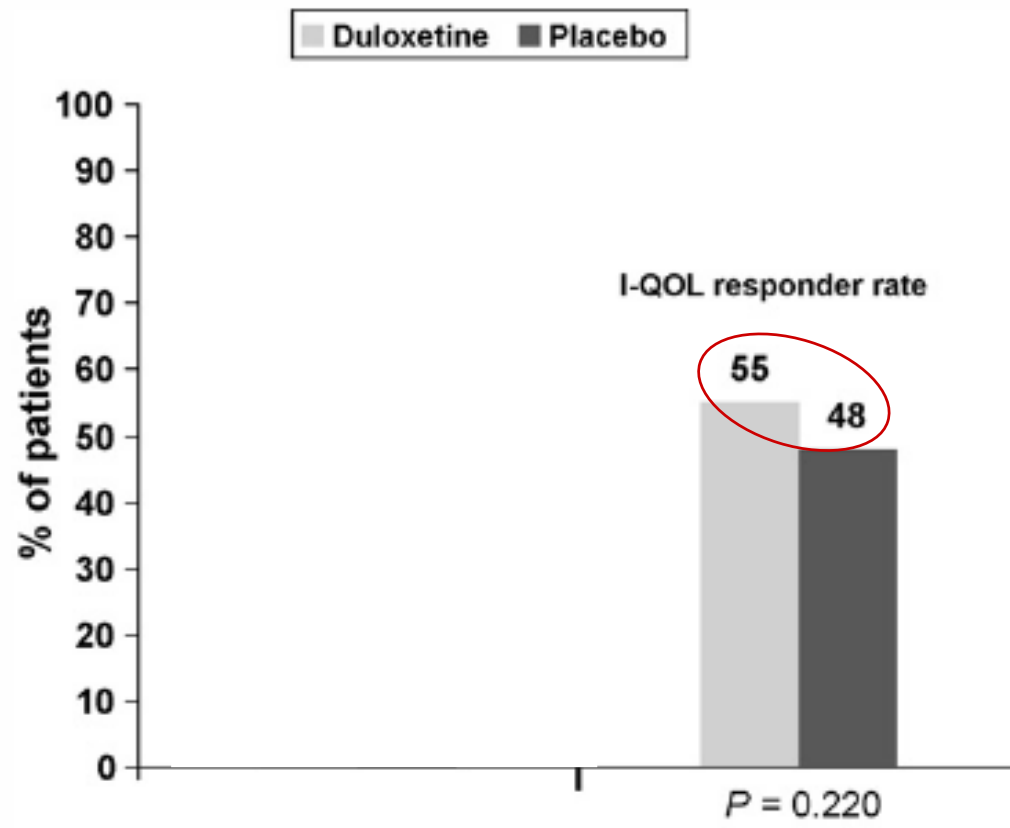
Efficacy variable	N placebo/duloxetine	Placebo	Duloxetine
Mean change in I-QOL	120/128		
Total score		6.84	12.81 <sup>†</sup>
Avoidance and limiting behaviour		8.78	14.60 <sup>†</sup>
Psychological impacts		3.08	10.50 <sup>†</sup>
Social embarrassment		10.47	14.11 <sup>*</sup>

\* Significantly different from placebo with  $P < 0.05$ .

† Significantly different from placebo with  $P \leq 0.001$ .

## 4. How are PROs handled in the analyses of the review?

Van Leeuwen 2004: responder analysis: responder being defined as a patient who has an increase in I-QOL total score of  $\geq 6.3$  points



## 4. How are PROs handled in the analyses of the review?

What was reported in the review?

Van Leeuwen 2004:

“Although quality of life was better in the PFMT group, it was not clear if there were important differences between PFMT and control groups; the means were presented without a measure of dispersion.”

What could be better?

## 4. How are PROs handled in the analyses of the review?

What was reported in the review?

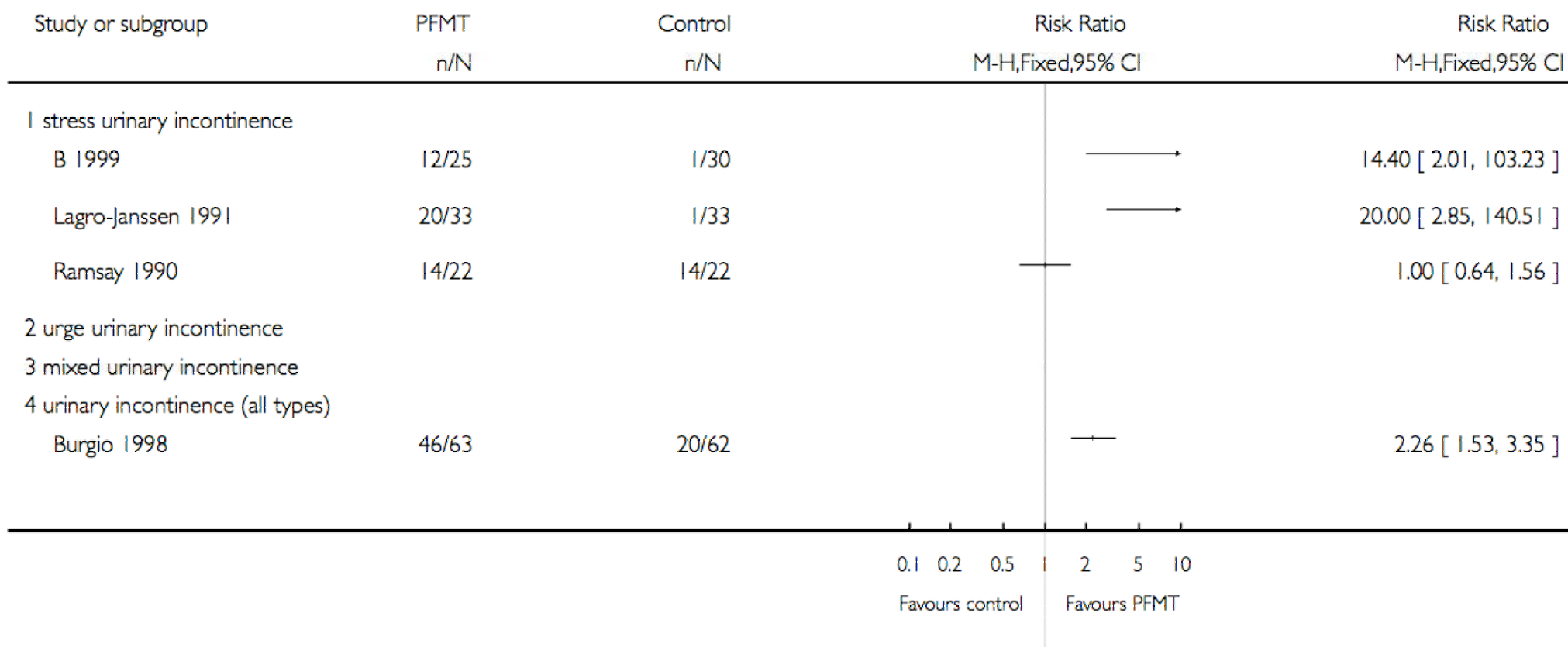
Overall there was no consistency in the choice of outcome measures by trialists. This limited the possibilities for considering results from individual studies together. It was disappointing that half the eligible trials did not contribute any data to the main analyses because they did not measure any of the pre-specified outcomes of interest, or did not report their outcome data in a usable way (for example mean without a measure of dispersion, P values without raw data).

## Analysis 1.2. Comparison 1 PFMT versus no treatment, placebo or control, Outcome 2 Patient perceived cure or improvement.

Review: Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women

Comparison: 1 PFMT versus no treatment, placebo or control

Outcome: 2 Patient perceived cure or improvement



# Making results interpretable

Control event rate (cure or improve) about 30%

Relative risk approximately 2 (60% event rate)

Absolute benefit 30% or 300 in 1,000

NNT 100/30 or 3

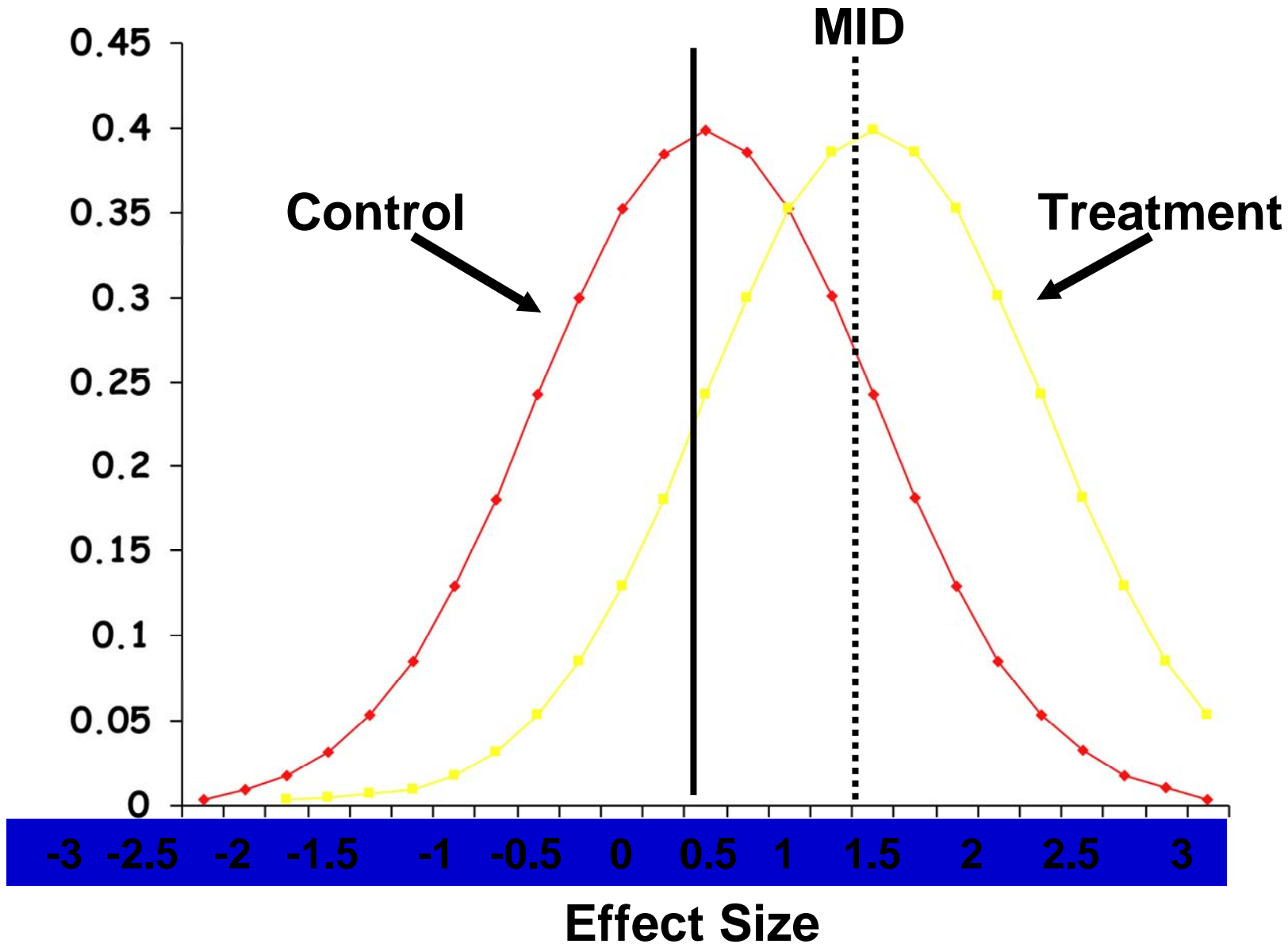
But only 3 studies, 125 patients

Possible to do more?

Pick patient-important outcomes with mean and SD

- urinary incontinence score, incontinence episode frequency, general health questionnaire.

# Change score difference





# What if you cannot find a way to change to binary in a consistent way across trials?

Assume binomial and equal variance assumptions generate numbers needed to treat (Furukawa, Lancet 1999; 353:680)

Control group response rate	10	20	30	40	50	60	70	80	90
SMD = 0.2	25	17	14	13	13	13	15	20	33
SMD = 0.5	9	6	5	5	5	5	7	9	16
SMD = 0.8	5	4	3	3	4	4	5	7	12
SMD = 1.0	4	3	3	3	3	4	4	6	11

Fancier methods available

Chinn (Statistics in Medicine 2000; 19: 3127-31)

# Summary of Findings

Outcomes	Illustrative Comparative Risks (95% CI)		Number of participants (studies)	Quality of the Evidence (GRADE)	Comments
	Proportion improved with Control	Proportion improved with exercise			
Symptoms measured as continuous variable	20%	50%	600 (8)	High (moderate)	This approach uses binomial and equal variance assumptions and baseline risks from need for rescue analgesia to translate SMD of 1.0 to risk difference

## 5. How are PROs included in the conclusions of the review?

### DISCUSSION

“It seems there might be improved condition specific quality of life (lifestyle and sex-life) in women treated with PFMT compared to controls, but there might be less or no effect on generic quality of life.”

Do you agree?

# Take home messages

## Trials

1. PROs need to be included in more trials

## Systematic reviews

2. The quality of the PROs needs to be taken into account
3. Results of PROs needs to be reported quantitatively
4. PROs can and should to be included in the SoF
  - use all available information !