



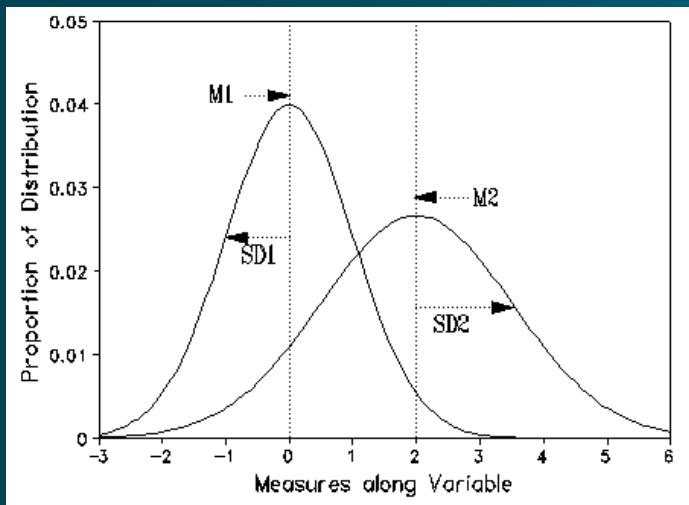
## Bridging the gap between aggregate data and individual patient management – a Bayesian approach

Gert J. van der Wilt PhD, J. Groenewoud MSc, P. van Riel, MD

Radboud University Medical Centre, Nijmegen, NL  
Department of Epidemiology, Biostatistics & HTA, and  
Department of Rheumatology



- Evidence from trials: essentially a public health perspective. Sometimes difficult to reconcile with the individual patient perspective.

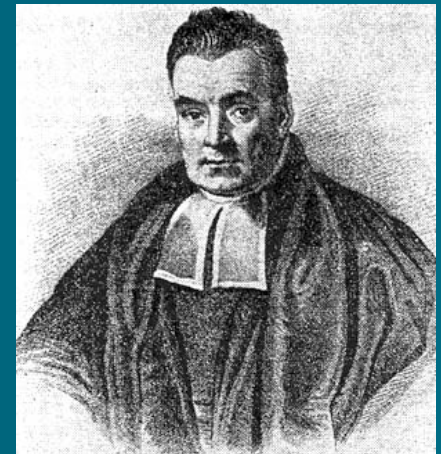


Frequency distribution of observed outcomes (e.g., change in blood pressure from baseline) in a single trial with two groups.



## Two schools of thought:

- Frequentist: what is the probability of the observed events under the null hypothesis? ( $p$  relates to a stochastic process)
- Bayesian: what is the probability that proposition  $X$  (which makes a claim about some aspect of reality) is true? ( $p$  = a measure of justified belief)





## Bayes: contentious, but widely accepted in the context of diagnostic research

Patient presents with clinical symptoms of acute appendicitis (AA).

*Q = what is the probability that the patient has AA?*

Further evidence: patient has leukocytosis.

*Q = how much more likely is leukocytosis to be found in persons with AA, as compared to persons without AA?  
(Likelihood ratio)*

Combine: how should the estimate of the probability that this patient has AA be revised, given the new evidence?



## Bayes applied to therapeutic issues

The relevant prior probability:

*what is the probability that this particular patient will respond to this therapy?*

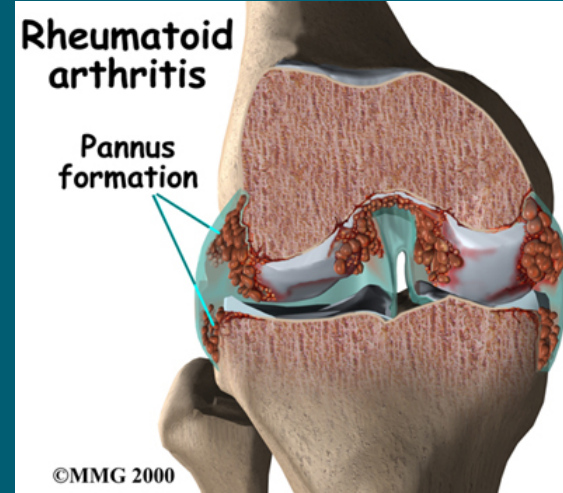
The relevant likelihood ratio:

*what is the probability of finding a specific 'marker' of success among respondents, as compared to non-respondents?*

Combine: given the presence (or absence) of the marker, how should we revise the estimate of the probability that the patient will respond to this therapy?



# Example: response to treatment for RA (MTX)



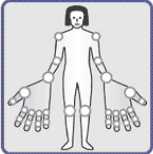
# Clinical response: Disease Activity Score (DAS)

**RISE**

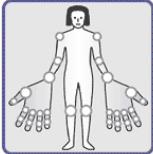
## Disease Activity Score (DAS 28)

### Joint Status - 28 Joint Count

**Tenderness**



**Swelling**




1 Joint Count TEN28     2 Joint Count SW28

3 ESR (after 1 hour in mm)

4 **General Health** or patient's global assessment of disease activity  
How active has your rheumatoid arthritis been during the last 7 days?\*

no activity highest activity possible



\*Please let patient assess this by drawing a vertical line.

Patient's assessment in mm

#### Formulas for DAS 28 calculation

$$0,56 \times \sqrt{1 \times \text{TEN28}} + 0,28 \times \sqrt{2 \times \text{SW28}}$$

$$+ 0,70 \times \ln(\text{ESR (after 1 hour in mm)}) + 0,014 \times (\text{Patient's assessment in mm})$$

=  DAS 28

#### Evaluation DAS 28

Current DAS 28	DAS 28: Difference to initial value		
	> 1,2	> 0,6 and ≤ 1,2	≤ 0,6
≤ 3,2 Inactive	Good improvement	Moderate Improvement	No Improvement
> 3,2 ≤ 5,1 Moderate	Moderate Improvement	Moderate Improvement	No Improvement
> 5,1 Very active	Moderate Improvement	No Improvement	No Improvement

International standard  
Validated

van Gestel AM, Haagsma CJ,  
van Riel PL. Arthritis Rheum  
1998;41:1845-50



## A suitable marker

- Easy to obtain
- Non-invasive
- Patient-reported
  
- E.g., score on Health Assessment Questionnaire (HAQ)



## HAQ Disability Index and Pain Scale

- Self-report functional status measure (Fries, 1980)
- Domains: disability, discomfort and pain during the past week
- Difficulties in performing dressing and grooming, arising, eating, walking, hygiene, reach, grip, and ADL
- Responses: without any difficulty (0), with some difficulty (1), with much difficulty (2), unable (3).
- Pain: VAS (convert cm to value between 0 and 3)



## Prior probability ('justified belief')

Months since start of treatment	% of patients in remission (DAS < 1.6)
6	13.6
12	14.0
24	15.8

Source: van der Heijde et al, Arthritis & Rheumatism 2006; 54 (4): 1063 – 74  
(2-year results from the TEMPO study)



## Calculation of Likelihood Ratio

- Poldas: clinical registry containing data from 3170 patients with RA
- First-time users of MTX, with a minimal follow up of 3 months (n = 1652)
- Define responders and non-responders on the basis of DAS (change in DAS > 1.2)
- Calculate: probability of specified improvement in HAQ among responders and non-responders

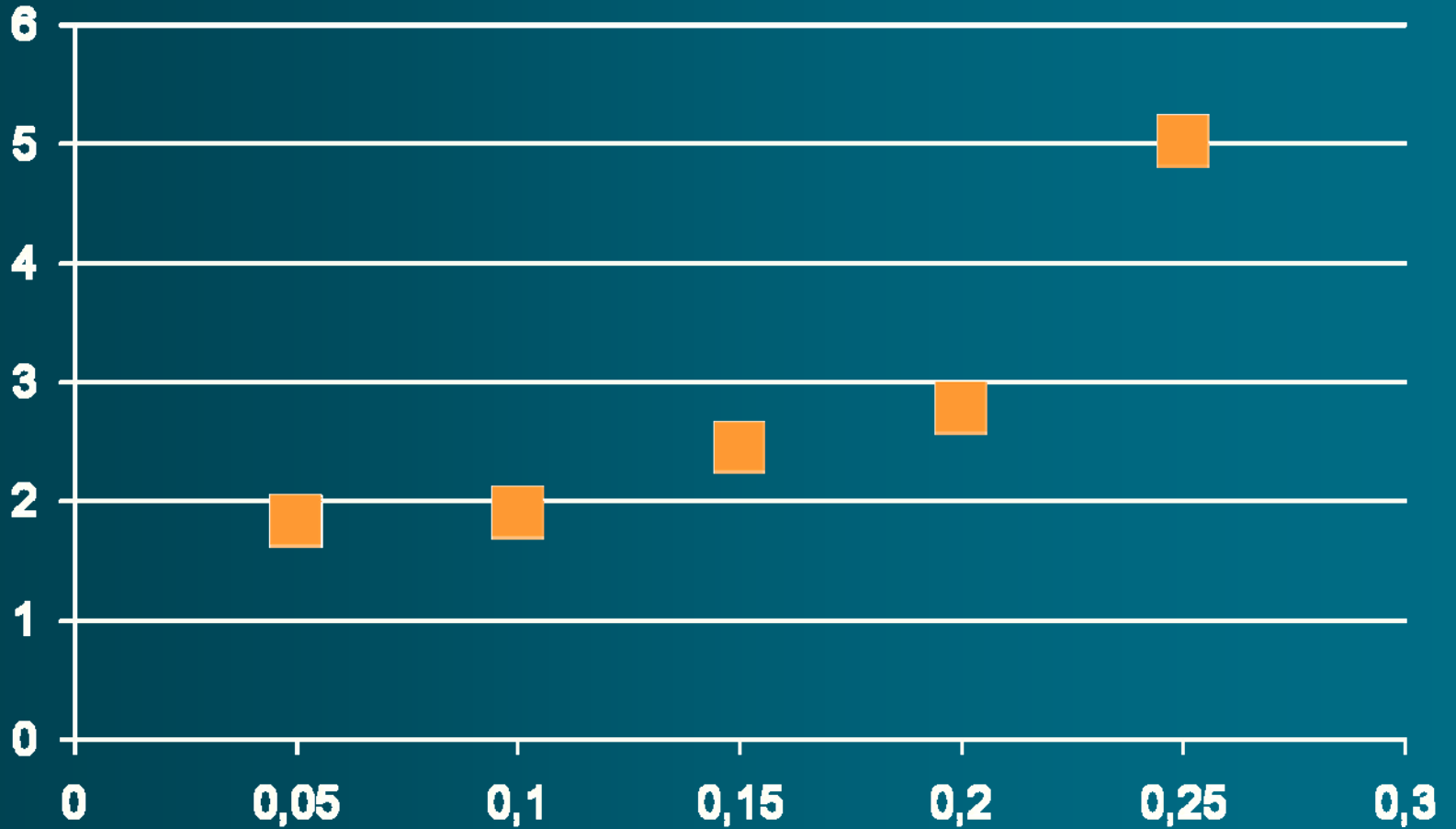


## Results: Likelihood Ratio

Improvement in HAQ	p (change in HAQ among responders)	p (change in HAQ among non-responders)	Likelihood Ratio
< 0.05	0.52	0.28	1.8
< 0.10	0.38	0.20	1.9
< 0.15	0.31	0.13	2.5
< 0.20	0.25	0.09	2.8
< 0.25	0.18	0.04	5.0



# LR as function of change in HAQ





## Results: Posterior probability

Improvement in HAQ score	Posterior probability of clinical response
$< 0.05$	0.23
$0.05 < x < 0.10$	0.24
$0.10 < x < 0.15$	0.28
$0.15 < x < 0.20$	0.31
$0.20 < x < 0.25$	0.46

Prior probability of clinical response: 0.14



## Interpretation (1)

- When, after 3 months of treatment with MTX, a patient reports substantially improved functional ability (c.q., improvement in HAQ  $> 0.25$ ), this patient is three times more likely to be a responder to the treatment (in the objective sense) as compared to his probability when treatment was initiated



## Interpretation (2)

- However, due to the relatively low prior probability of clinical response, the probability that the patient is responding still approximately equals the probability that he is not responding (0.46 vs. 0.54).
- Hence, not measuring clinical response at this time of follow up would lead to incorrect classification of patients in approximately half of the cases.



## Conclusions (1)

- Application of Bayesian analysis to therapeutic issues can be conceptually similar to its application to diagnostic issues
- Advantages of the Bayesian approach: ease of interpretation, relevance to clinical practice, flexibility (e.g., different prior probability of clinical response)
- A means of bridging the gap between aggregate evidence and individual, evidence-based patient management



# Requirements

- Reasonable estimate of prior probability (empirical basis; RCTs, or registries)
- Availability of a good marker (= 'diagnostic test result') and a 'gold standard'
- Empirical data to calculate likelihood ratios (e.g., a clinical registry)
- Threshold values (when to start with a particular treatment, when to (dis)continue or change treatment, what level of incorrect classification of (non)response is still acceptable?)



## Possible use

- In the communication with the patient: what may be expected from treatment? When to discontinue or change treatment?
- Self-management: teach the patient how to interpret experienced improvements or deteriorations in functional ability / when to consult your physician? (how does experienced improvement translate into objective response?)



## Conditions for wider use

- Tools to assist calculation
- Translate numbers into phrasings that can be readily interpreted by patients
- Markers with still higher discriminatory capacity



Thank you!

