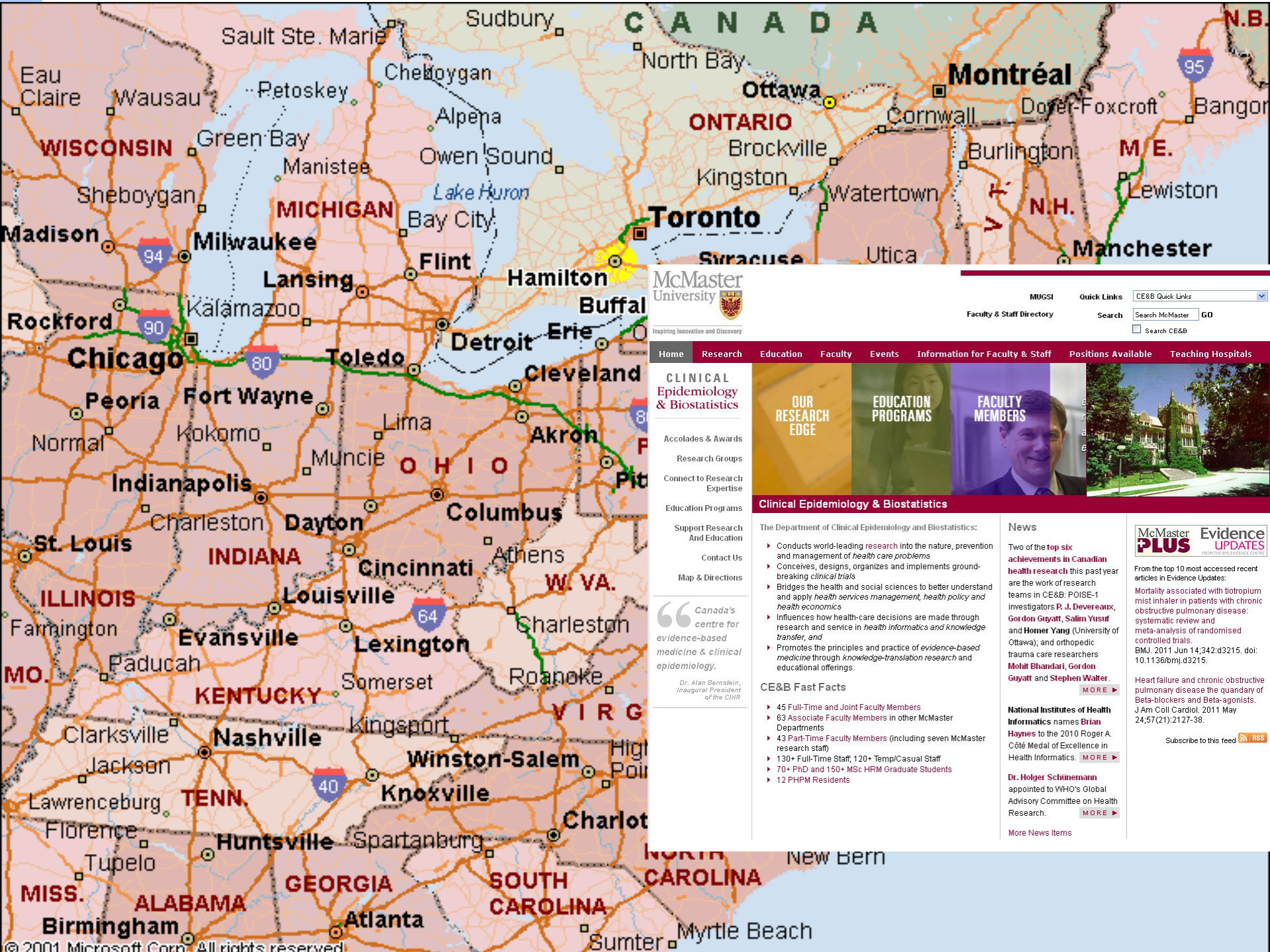
McMaster University
Inspiring Innovation and Discovery

Holger Schünemann, MD, MSc, PhD
Professor and Chair, Dept. of Clinical Epidemiology & Biostatistics
Professor of Medicine
Michael Gent Chair in Healthcare Research
McMaster University, Hamilton, Canada

IQWiG Herbsttagung, 25. Nov. 2011

DIAGNOSTIK UND LINKED EVIDENCE – WIE ROBUST MUSS DIE KETTE SEIN?



McMaster University
Inspiring Innovation and Discovery

MUGSI | Faculty & Staff Directory | Quick Links: CE&B Quick Links | Search: Search McMaster GO | Search CE&B

Home | Research | Education | Faculty | Events | Information for Faculty & Staff | Positions Available | Teaching Hospitals

CLINICAL Epidemiology & Biostatistics

- Accolades & Awards
- Research Groups
- Connect to Research Expertise
- Education Programs
- Support Research And Education
- Contact Us
- Map & Directions

Canada's centre for evidence-based medicine & clinical epidemiology.

Dr. Alan Bernstein, Inaugural President of the CIHR

Clinical Epidemiology & Biostatistics

The Department of Clinical Epidemiology and Biostatistics:

- Conducts world-leading research into the nature, prevention and management of *health care problems*
- Conceives, designs, organizes and implements ground-breaking *clinical trials*
- Bridges the health and social sciences to better understand and apply *health services management, health policy and health economics*
- Influences how health-care decisions are made through research and service in *health informatics and knowledge transfer, and*
- Promotes the principles and practice of *evidence-based medicine* through *knowledge-translation research* and educational offerings.

CE&B Fast Facts

- 45 Full-Time and Joint Faculty Members
- 63 Associate Faculty Members in other McMaster Departments
- 43 Part-Time Faculty Members (including seven McMaster research staff)
- 130+ Full-Time Staff, 120+ Temp/Casual Staff
- 70+ PhD and 150+ MSc HRM Graduate Students
- 12 PHPM Residents

News

Two of the **top six achievements in Canadian health research** this past year are the work of research teams in CE&B: POISE-1 investigators **P. J. Devereaux, Gordon Guyatt, Salim Yusuf** and **Homer Yang** (University of Ottawa); and orthopedic trauma care researchers **Mohit Bhandari, Gordon Guyatt** and **Stephen Walter**. MORE ▸

National Institutes of Health Informatics names **Brian Haynes** to the 2010 Roger A. Côté Medal of Excellence in Health Informatics. MORE ▸

Dr. Holger Schünemann appointed to WHO's Global Advisory Committee on Health Research. MORE ▸

More News Items

McMaster PLUS Evidence UPDATES FROM THE BMJ EVIDENCE CENTRE

From the top 10 most accessed recent articles in Evidence Updates:

Mortality associated with tiotropium mist inhaler in patients with chronic obstructive pulmonary disease: systematic review and meta-analysis of randomised controlled trials.
BMJ. 2011 Jun 14;342:d3215. doi: 10.1136/bmj.d3215.

Heart failure and chronic obstructive pulmonary disease the quandary of Beta-blockers and Beta-agonists.
J Am Coll Cardiol. 2011 May 24;57(21):2127-38.

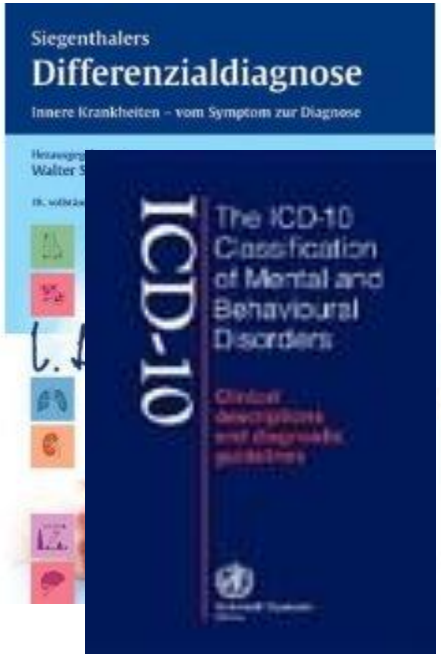
Subscribe to this feed RSS

Disclosure

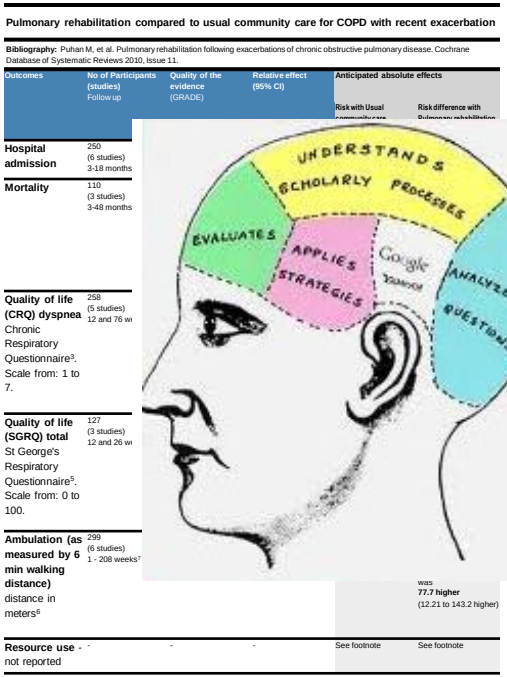
- Co-chair GRADE Working Group
- Leitlinienprojekte - GRADE
  - *American College of Physicians (ACP)* Clinical Practice Guidelines Committee
  - American College of Chest Physicians (ACCP)
- Weltgesundheitsorganisation (WHO): Advisory Committee for Health Research, Leitlinien, Drittmittel für systematische Übersichtsarbeiten
- Keine direktes Einkommen von profitorientierten Unternehmen/Organisationen
- Dank an Kollegen (Drs. Jan Brozek & Reem Mustafa)

Übersicht

Diagnostische Fragestellungen



Einführung



Pulmonary rehabilitation compared to usual community care for COPD with recent exacerbation

Bibliography: Puhan M, et al. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2010, Issue 11.

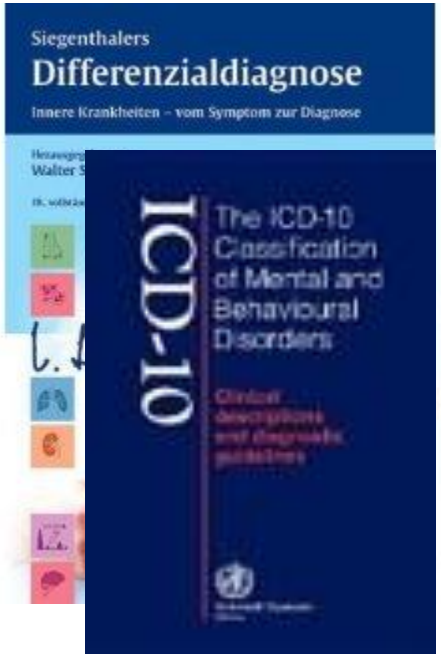
Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects: Risk with Usual community care	Anticipated absolute effects: Risk difference with Pulmonary rehabilitation
Hospital admission	250 (6 studies) 3-18 months				
Mortality	110 (3 studies) 3-48 months				
Quality of life (CRQ) dyspnea Chronic Respiratory Questionnaire[3]. Scale from: 1 to 7.	258 (5 studies) 12 and 76 we				
Quality of life (SGRQ) total St George's Respiratory Questionnaire[5]. Scale from: 0 to 100.	127 (3 studies) 12 and 26 we				
Ambulation (as measured by 6 min walking distance) distance in meters[6]	299 (6 studies) 1 - 208 weeks[7]				was **77.7 higher** (12.21 to 143.2 higher)
Resource use - not reported	-	-	-	See footnote	See footnote

Evidenz & Beurteilungen

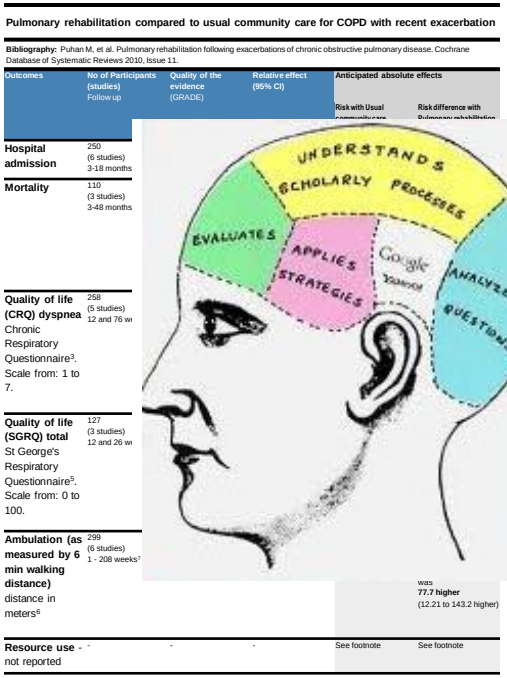
Empfehlungen & Implementierung

Übersicht

Diagnostische Fragestellungen



Einführung



Pulmonary rehabilitation compared to usual community care for COPD with recent exacerbation

Bibliography: Puhan M, et al. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2010, Issue 11.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects: Risk with Usual community care	Anticipated absolute effects: Risk difference with Pulmonary rehabilitation
Hospital admission	250 (6 studies) 3-18 months				
Mortality	110 (3 studies) 3-48 months				
Quality of life (CRQ) dyspnea Chronic Respiratory Questionnaire[3]. Scale from: 1 to 7.	258 (5 studies) 12 and 76 we				
Quality of life (SGRQ) total St George's Respiratory Questionnaire[5]. Scale from: 0 to 100.	127 (3 studies) 12 and 26 we				
Ambulation (as measured by 6 min walking distance) distance in meters[6]	299 (6 studies) 1 - 208 weeks[7]				was **77.7 higher** (12.21 to 143.2 higher)
Resource use - not reported	-	-	-	See footnote	See footnote

Evidenz & Beurteilungen

Empfehlungen & Implementierung

Anwendung eines Tests

- Kliniker benutzen eine Reihe von Tests (oder Teststrategien), die “diagnostisch” genannt werden:
  - Symptome und Zeichen, bildgebende Verfahren, Laborparameter, pathologische und psychologische Befunde
- Wenige Tests sind wirklich diagnostisch (positiv oder negativ - Schwangerschaftstest)
  - Im allgemeinen verbunden mit Wahrscheinlichkeiten
- Für diesen Vortrag: vereinfachtes Modell (positiv und negativ)

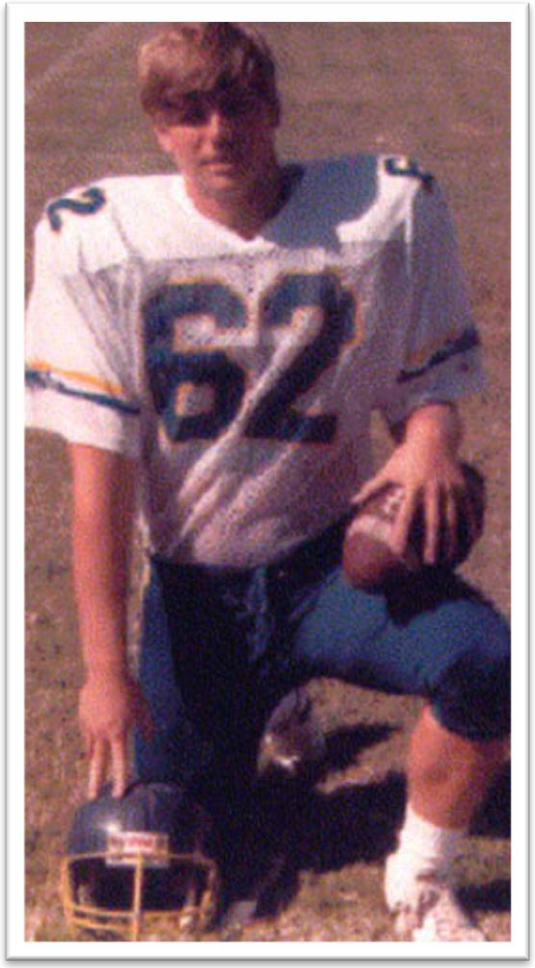
McMaster University
Inspiring Innovation and Discovery

HUNTINGTONs CHOREA

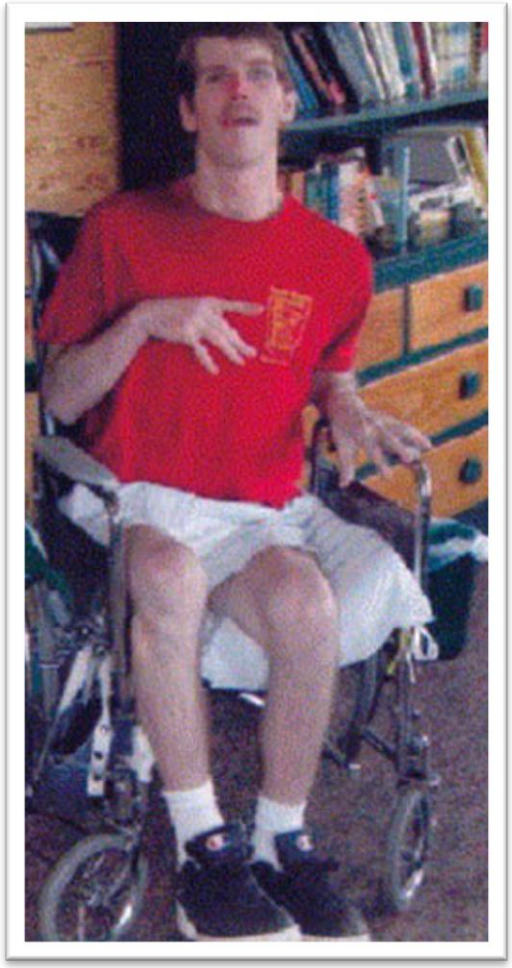
Morbus Huntington

McMaster University

Inspiring Innovation and Discovery



18 years

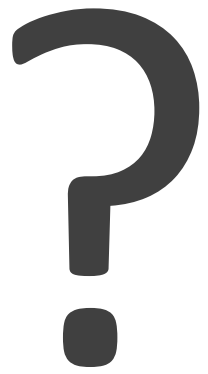


33 years

Scott Redford

Sensitivität = 98.8%
Spezifität = 100%
“pre-test” Wahrscheinlichkeit in Kindern = 50%

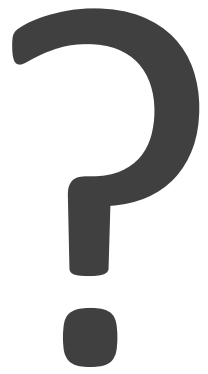
		Erkrankung vorliegend	Erkrankung nicht vorliegend
DNA Test	+	494	0
	–	6	500
		500	500



Würden Sie diesen genetischen Test für Kinder von betroffenen Patienten empfehlen?

Keine Prävention

Keine effektive Behandlung



Würden Sie diesen genetischen Test für Kinder von betroffenen Patienten empfehlen?


Sensitivität
Spezifität

Test accuracy ist ein Surrogatparameter für patientenrelevante Endpunkte

- Kliniker konzentrieren sich typischerweise auf ‘test accuracy’/Testgüte

- Annahme: Diagnose führt zu besserer Behandlung oder endpunktübergreifendem Zusatznutzen

Patientennutzen sollte vorliegen

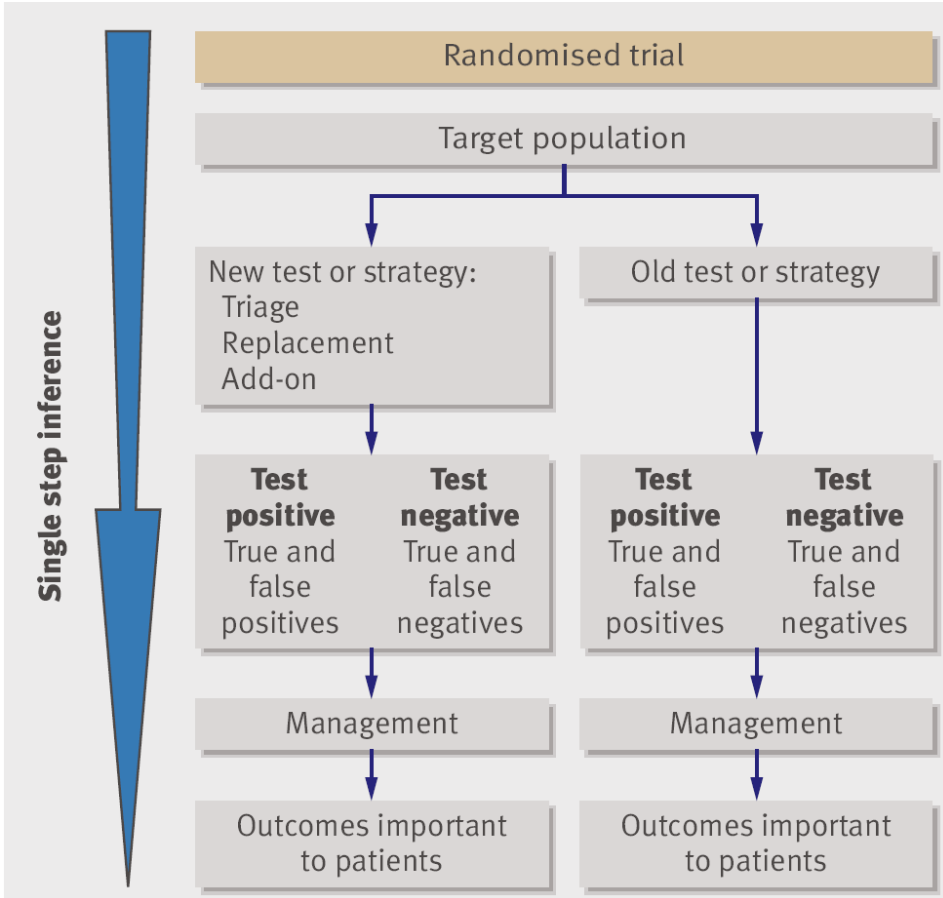
- Die Annahme auf der Basis von ‘accuracy’ Daten, dass ein Test patientenrelevante Endpunkte verbessert, erfordert das Vorhandensein von effektiven Behandlungsstrategien = linked evidence
- Inklusive:
  - Verringerung von testgebundenen Nebenwirkungen
  - Ausschluss von Erkrankungen oder Verminderung von Angst
  - Bestätigung einer Diagnose verbessert Lebensqualität durch die prognostische Information, die vermittelt wird

Studiendesigns in der Diagnoseerstellung

- Wenn ein Test patientenrelevante Endpunkte nicht verbessert, gibt es keinen Grund für seine Anwendung (unabhängig von seiner ‘accuracy’)

- Vernünftigste Verfahren, um ein Testverfahren zu evaluieren: randomisierte, kontrollierte Studien die Tests (mit Behandlung) gegeneinander vergleichen

Studiendesign I



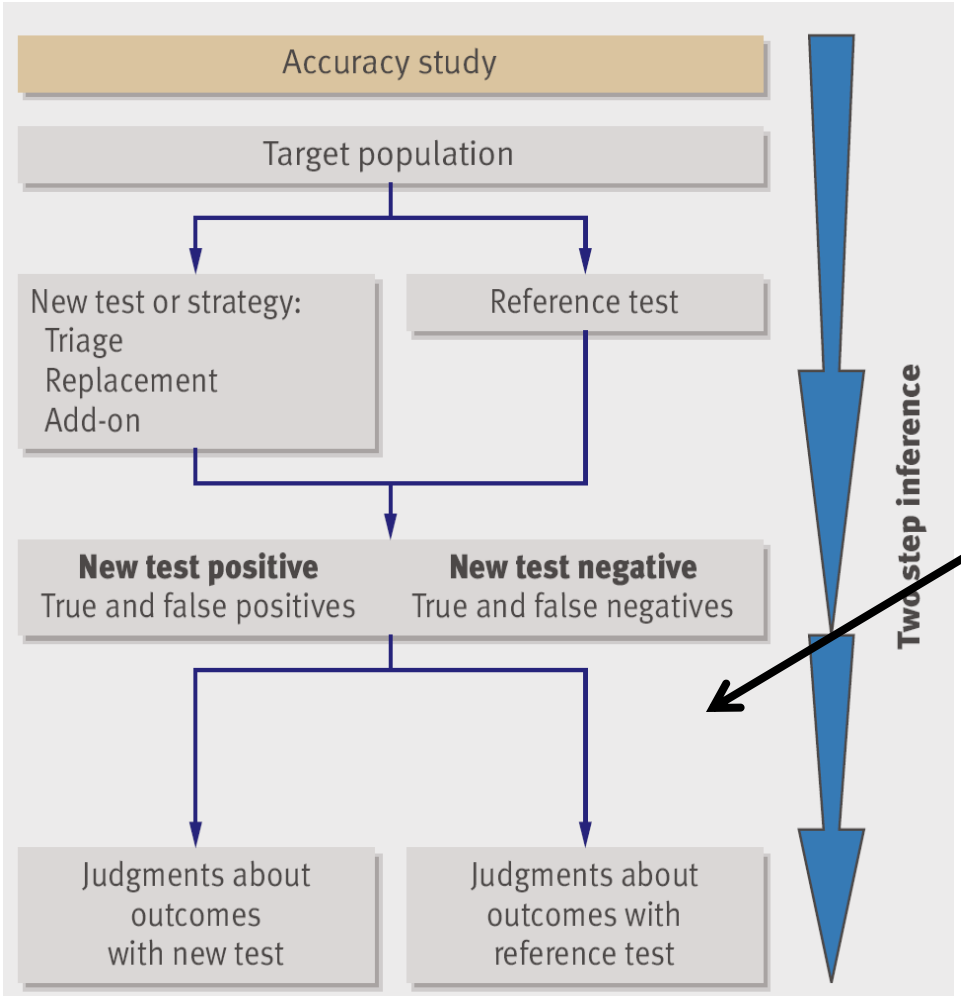
Example
Randomised control trials (RCTs) explored a diagnostic strategy guided by the use of B type natriuretic peptide (BNP)—designed to aid diagnosis of heart failure—compared with no use of BNP in patients presenting to the emergency department with acute dyspnoea.[8 9] As it turned out, the group randomised to receive BNP spent a shorter time in the hospital at lower cost, with no increased mortality or morbidity

Endpunkte:
- Mortalität
- Morbidität
- Nebenwirkungen
- QoL

GRADE für Interventionen und Behandlungen:
- Qualitätsbeurteilung
- Nutzen/Schaden/Werte/Ressourcen
- Empfehlung

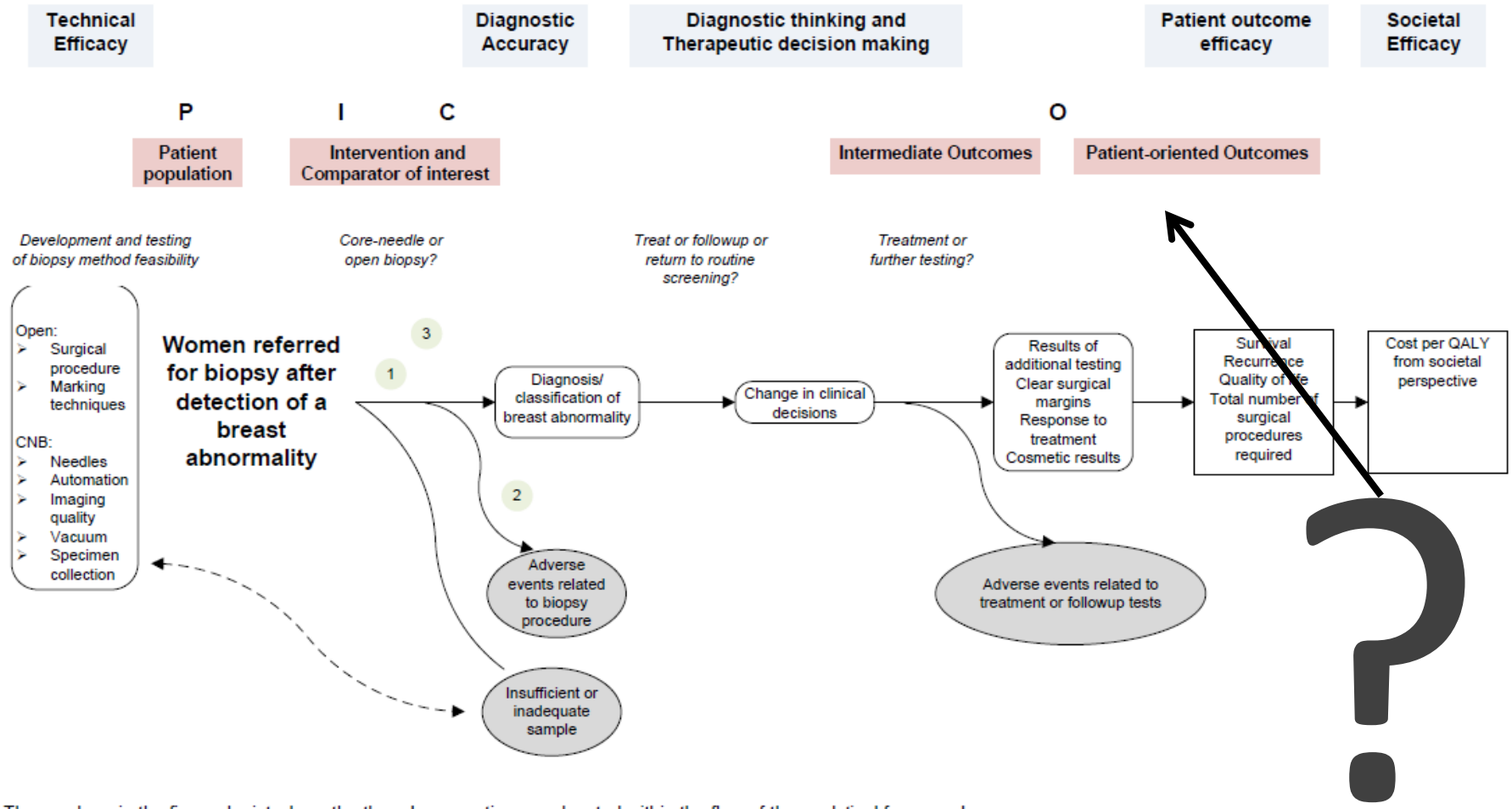
BMJ 17 May 2008
Volume 336 | Page 1106-1110

Studiendesign II



BMJ 17 May 2008
Volume 336 | Page 1106-1110

Figure 2-2. Example of an analytical framework within an overarching conceptual framework in the evaluation of breast biopsy techniques*



The numbers in the figure depict where the three key questions are located within the flow of the analytical framework.

AHRQ website – available upon request

‘accuracy’

Sensitivität & Spezifität

Patientenrelevante Konsequenzen

- TP (behandelt…)
- TN (vergewissert…)
- FP (unnötigerweise behandelt…)
- FN (nicht behandelt...)

Unklare Resultate

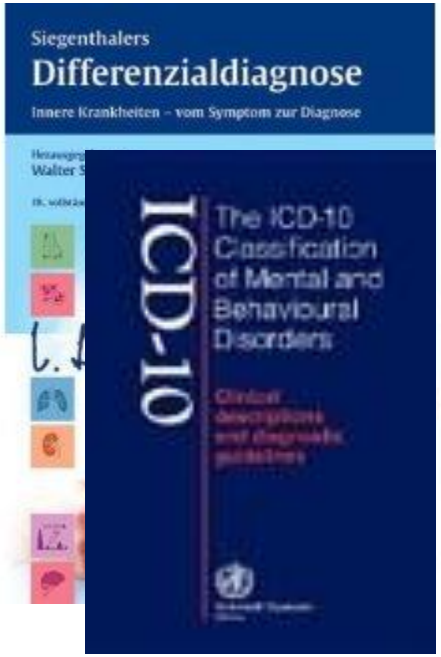
Komplikationen durch Test

Ressourcenverbrauch


Lebensverlängerung
Weniger Symptome
Komplikationen?
Lebensqualität?
Sensitivität
Spezifität

Übersicht

Diagnostische Fragestellungen

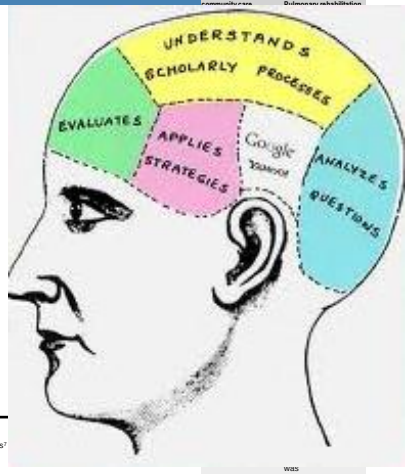


Einführung

Pulmonary rehabilitation compared to usual community care for COPD with recent exacerbation

Bibliography: Puhan M, et al. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2010, Issue 11.

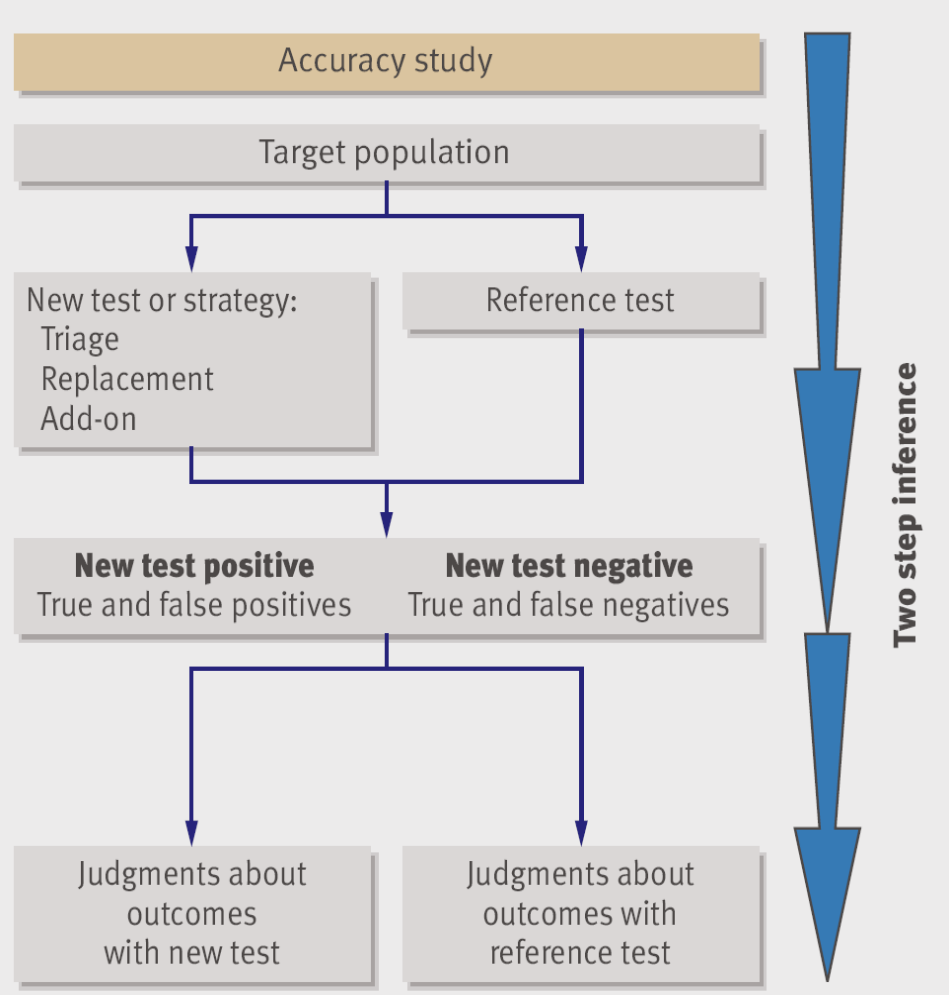
Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects: Risk with Usual community care	Anticipated absolute effects: Risk difference with Pulmonary rehabilitation
Hospital admission	250 (6 studies) 3-18 months				
Mortality	110 (3 studies) 3-48 months				
Quality of life (CRQ) dyspnea Chronic Respiratory Questionnaire[3]. Scale from: 1 to 7.	258 (5 studies) 12 and 76 we				
Quality of life (SGRQ) total St George's Respiratory Questionnaire[5]. Scale from: 0 to 100.	127 (3 studies) 12 and 26 we				
Ambulation (as measured by 6 min walking distance) distance in meters[6]	299 (6 studies) 1 - 208 weeks[7]				was **77.7 higher** (12.21 to 143.2 higher)
Resource use - not reported	-	-	-	See footnote	See footnote



Evidenz & Beurteilungen

Empfehlungen & Implementierung

‘Linked’ Evidenz



Systematische Übersichtsarbeiten

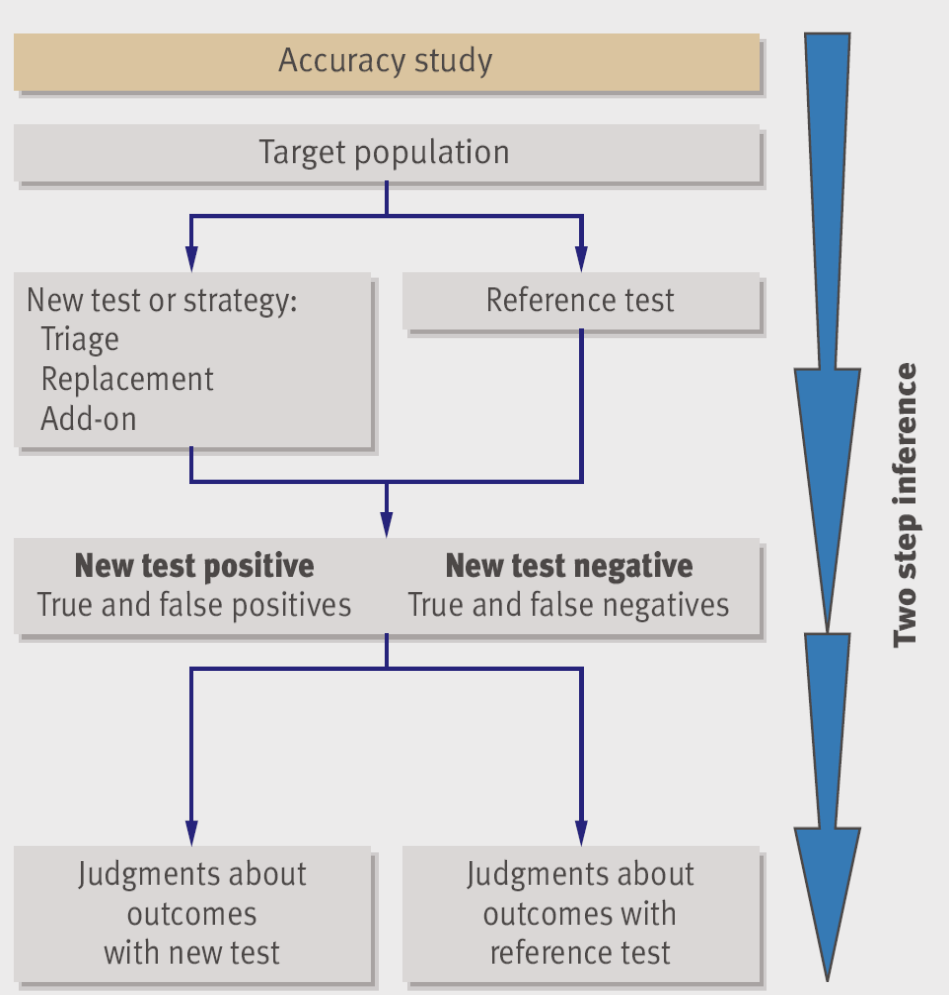
GRADE für ‘diagnostic accuracy’:

8 Qualitätsdomänen

Vertrauen in die Effektschätzer

Vertrauen in die Konsequenzen

‘Linked’ Evidenz



Hohe/gute Qualität

Directness: Surrogat – patientenrelevante Endpunkte?

Herunterstufen der Qualität?

Systematische Übersichtsarbeiten: Therapie, ‘natural history’

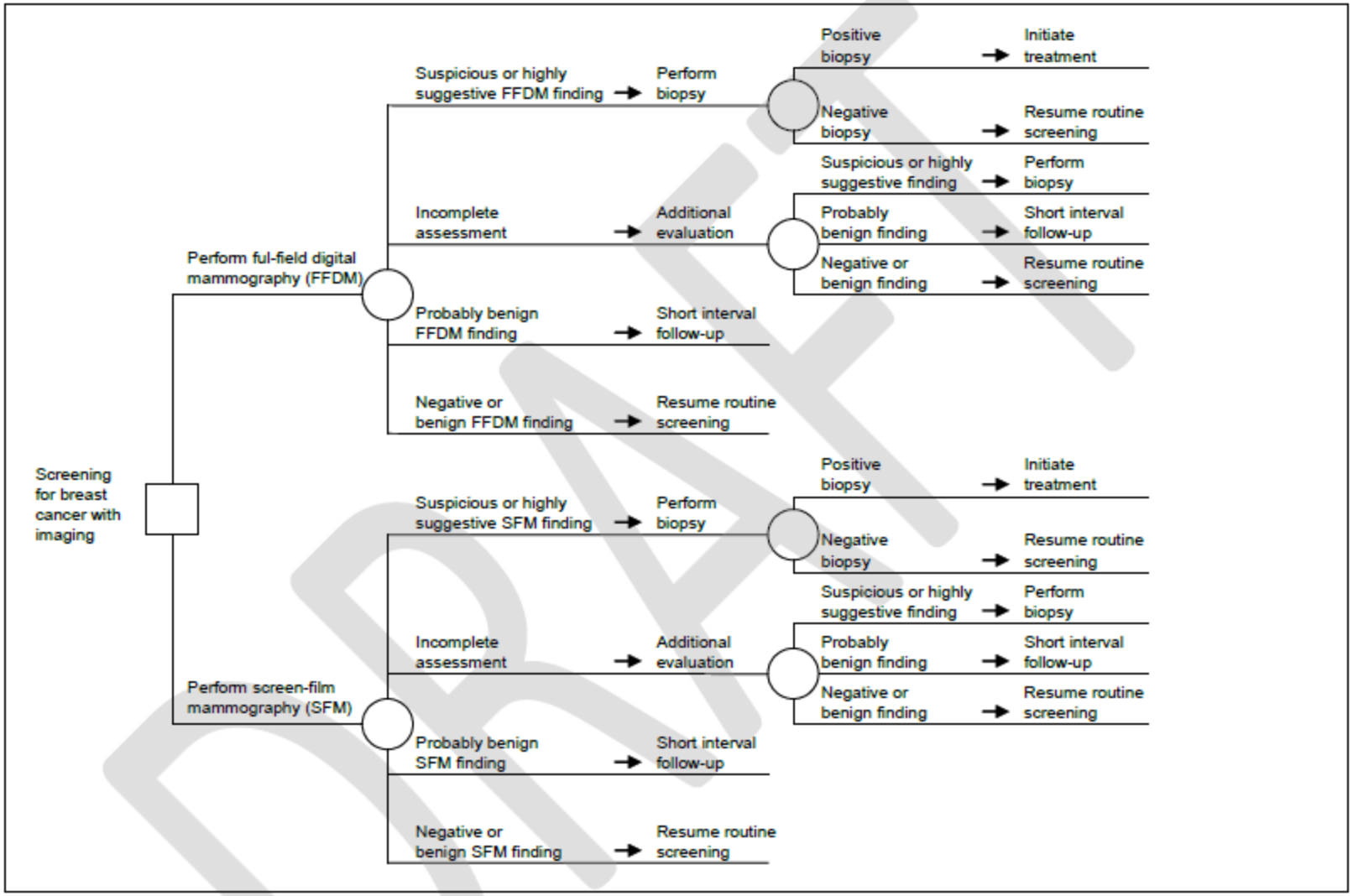
Domains, sub-domains and items in a conceptual framework for decision modeling in diagnostic test studies

McMaster University
Inspiring Innovation and Discovery

Domain	Sub-domain	Items *
Purpose	Triage Replacement Add-on	Screening Diagnosis Staging of disease Monitoring of treatment Monitoring of disease
Population		Pretest probability of a condition Any subgroups with different baseline risk or prevalence (co morbidities, patients' characteristics ..etc) Stage of the disease
Intervention (test of interest, aka index test)		Test's accuracy characteristics Test's side effects Test benefits Cut-off points Resources required Inconclusive results Values and preferences
Comparison (reference test or alternative test)		Test's accuracy characteristics Test's side effects Test benefits Cut-off points Resources required Inconclusive results Values and preferences
Diagnostic test accuracy	Test +ve, sensitivity	TP & FP
outcomes	Test –ve, specificity	TN & FN
Patient outcomes	Treatment 1 Treatment 2 Treatment 3 No treatment	Efficacy of available treatment Rate of side effects of available treatment Resource use with available treatment Values and preferences Prognosis/natural course of condition
Quality of evidence	Criteria for downgrading	Risk of bias Inconsistency Imprecision

Decision modelling

Figure 2-3. Replacement test example: full-field digital mammography versus screen-film mammography*



* Figure taken from Blue Cross and Blue Shield Association Technology Evaluation Center, 2002.[14]

World Allergy Organization

COWS MILK ALLERGY GUIDELINES

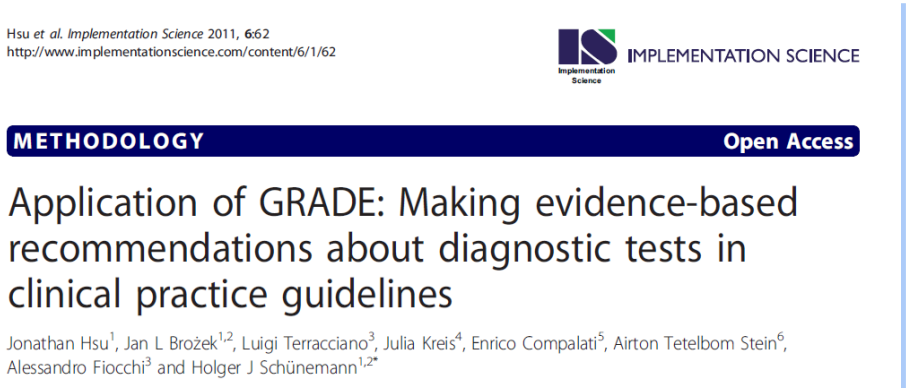
Workshop summary

Diagnosis and Rationale for Action against Cow's Milk Allergy (DRACMA): A summary report

Alessandro Fiocchi, MD,[a] Holger J. Schünemann, MD, PhD,[b] Jan Brozek, MD,[b] Patrizia Restani, PhD,[c] Kirsten Beyer, MD,[d] Riccardo Troncone, MD,[e] Alberto Martelli, MD,[f] Luigi Terracciano, MD,[f] Sami L. Bahna, MD,[g] Fabienne Rancé, MD,[h] Motohiro Ebisawa, MD,[i] Ralf G. Heine, MD, FRACP,[j] Amal Assa'ad, MD,[k] Hugh Sampson, MD,[l] Elvira Verduci, MD,[m] G. R. Bouygue, MSc,[f] Carlos Baena-Cagnani, MD,[n] Walter Canonica, MD,[o] and Richard F. Lockey, MD[p] *Milan, Naples, and Genoa, Italy, Hamilton, Ontario, Canada, Berlin, Germany, Shreveport, La, Toulouse, France, Kanagawa, Japan, Melbourne, Australia, Cincinnati, Ohio, New York, NY, Cordoba, Argentina, and Tampa, Fla*

World Allergy Organization (WAO) Diagnosis and Rationale for Action against Cow's Milk Allergy (DRACMA) Guidelines

Alessandro Fiocchi, (Chair), Jan Brozek, Holger Schünemann, (Chair), Sami L. Bahna, Andrea von Berg, Kirsten Beyer, Martin Bozzola, Julia Bradsher, Enrico Compalati, Motohiro Ebisawa, Maria Antonieta Guzman, Haiqi Li, Ralf G. Heine, Paul Keith, Gideon Lack, Massimo Landi, Alberto Martelli, Fabienne Rancé, Hugh Sampson, Airton Stein, Luigi Terracciano, and Stefan Vieths



Hsu *et al. Implementation Science* 2011, **6**:62
http://www.implementationscience.com/content/6/1/62

IMPLEMENTATION SCIENCE

METHODOLOGY **Open Access**

Application of GRADE: Making evidence-based recommendations about diagnostic tests in clinical practice guidelines

Jonathan Hsu[1], Jan L Brożek[1,2], Luigi Terracciano[3], Julia Kreis[4], Enrico Compalati[5], Airton Tetelbom Stein[6], Alessandro Fiocchi[3] and Holger J Schünemann[1,2*]

Formulate question

Select outcomes

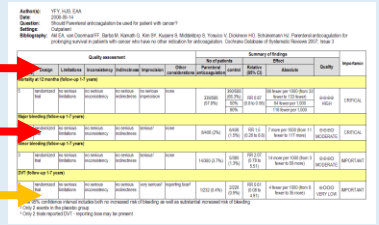
Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

P	Outcome	Critical
I	Outcome	Critical
C	Outcome	Important
O	Outcome	Not important



Summary of findings & estimate of effect for each outcome

High	⊕⊕⊕⊕
Moderate	⊕⊕⊕◯
Low	⊕⊕◯◯
Very low	⊕◯◯◯

Grade down

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

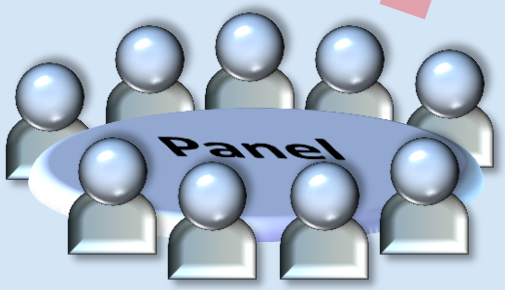
Grade up

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)

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..."

Darf sie Milch trinken?

McMaster University
Inspiring Innovation and Discovery

Food challenge test

Skin prick test

~~Food challenge test~~

Skin prick test

McMaster University
Inspiring Innovation and Discovery

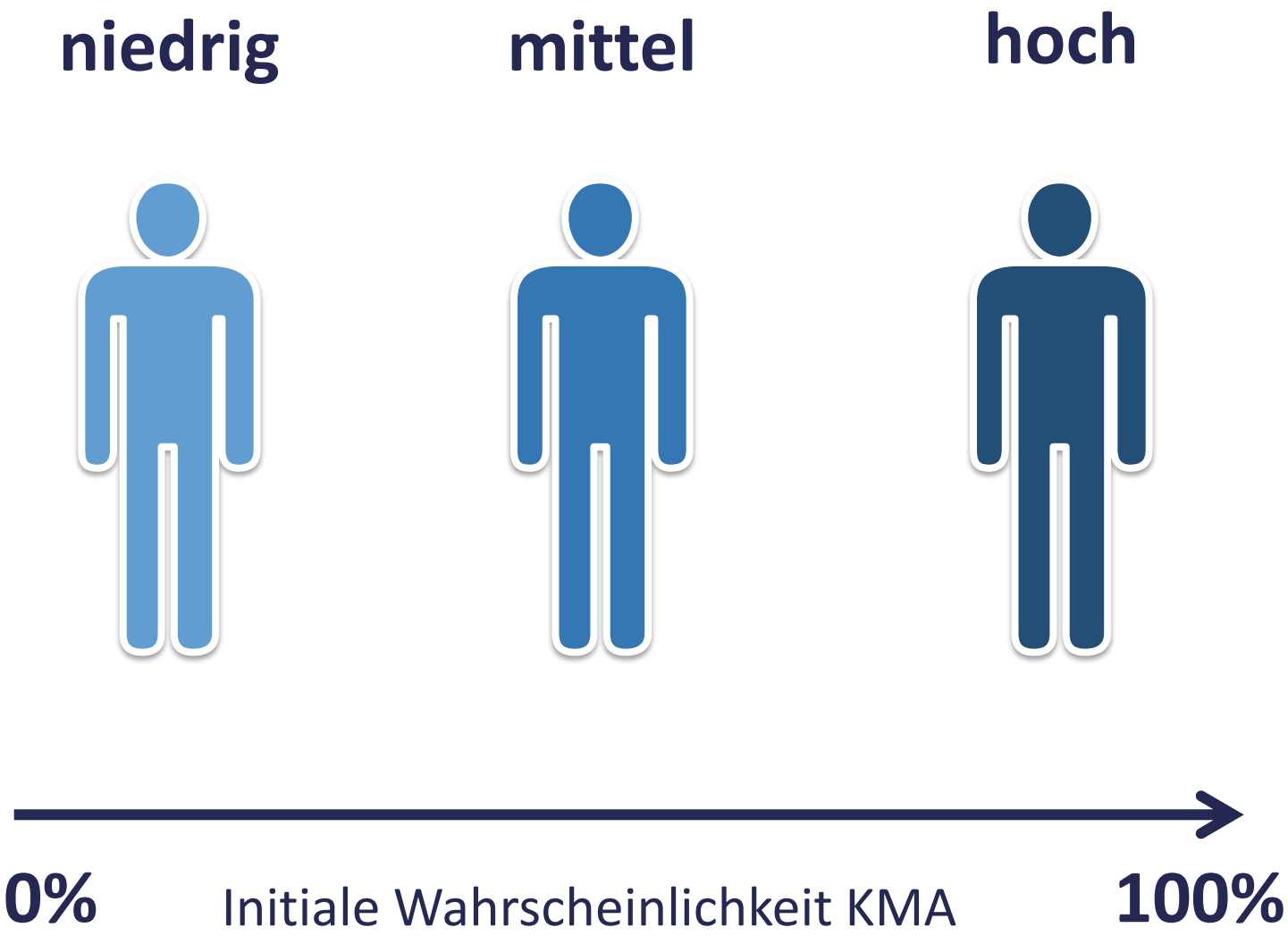
Sollten ‘skin prick tests’ zur Diagnose von Kuhmilchallergien (KMA) angewandt werden?

Population (Wer?)

Intervention (Welcher test)

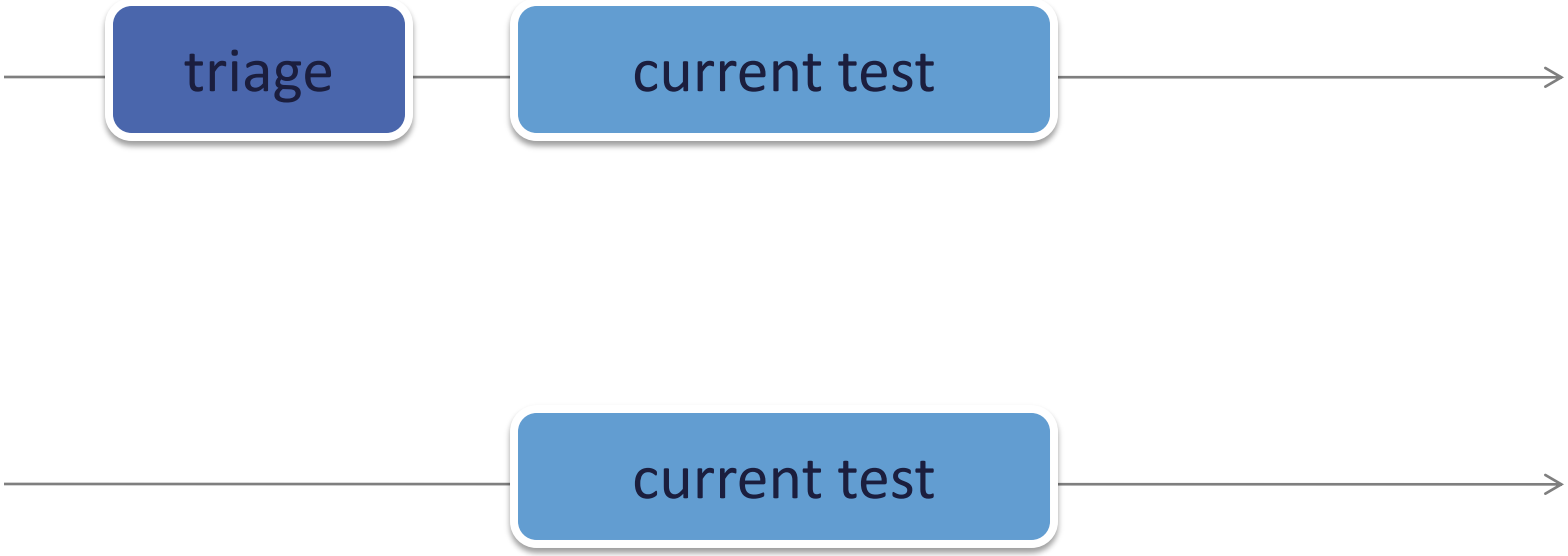
Comparison (Anstatt?)

Outcomes (Wofür?)

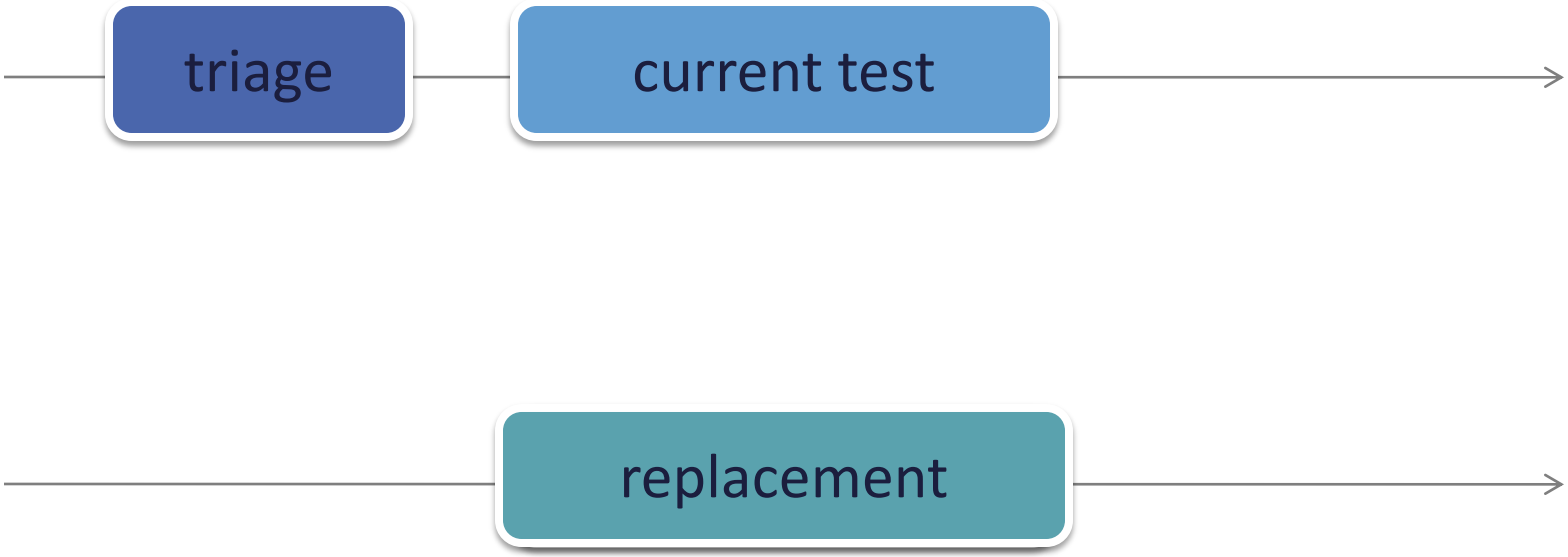


P. Bossuyt et al. BMJ 2006

P. Bossuyt et al. BMJ 2006



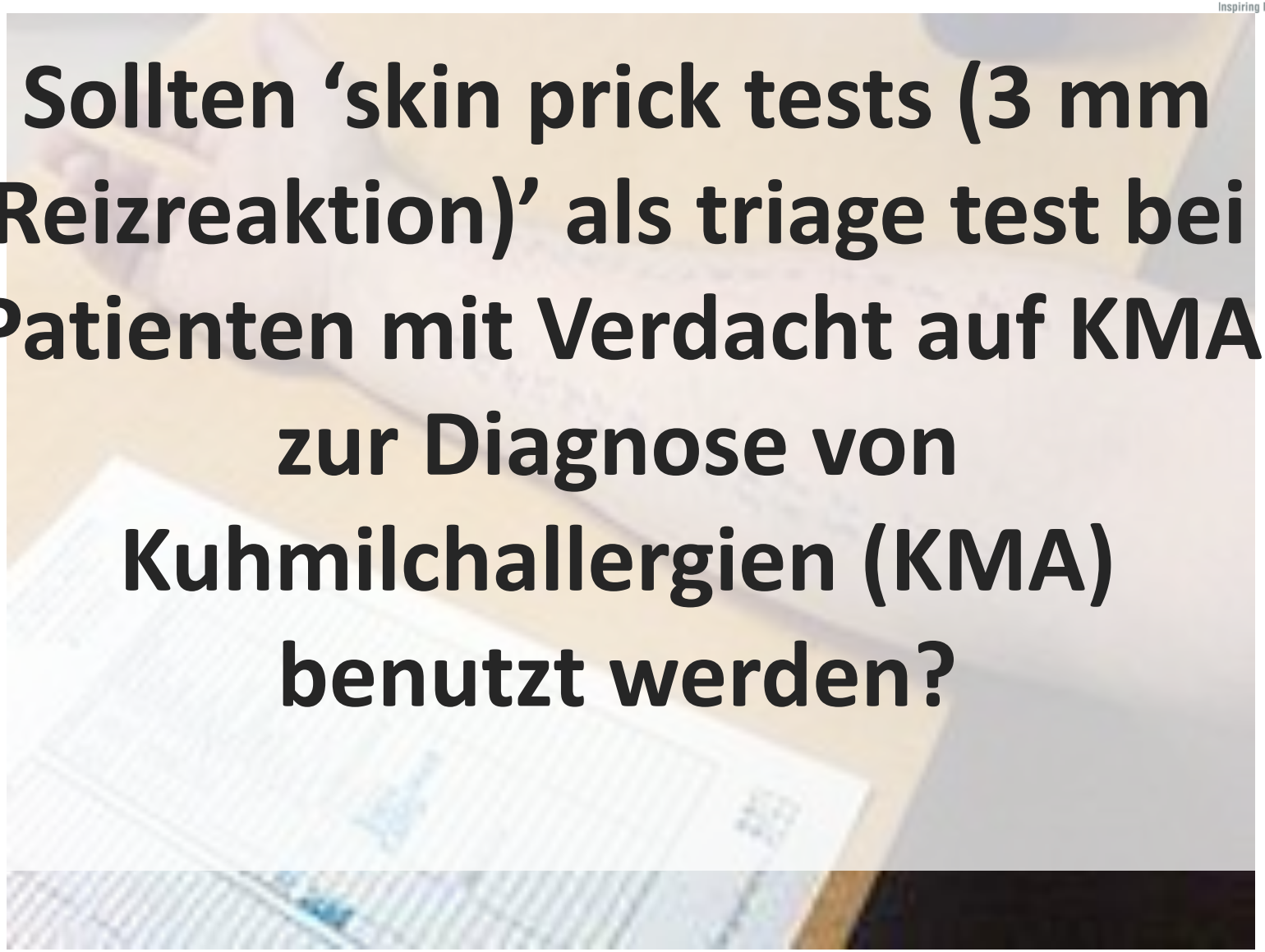
P. Bossuyt et al. BMJ 2006



P. Bossuyt et al. BMJ 2006

McMaster University
Inspiring Innovation and Discovery

Sollten ‘skin prick tests (3 mm Reizreaktion)’ als triage test bei Patienten mit Verdacht auf KMA zur Diagnose von Kuhmilchallergien (KMA) benutzt werden?



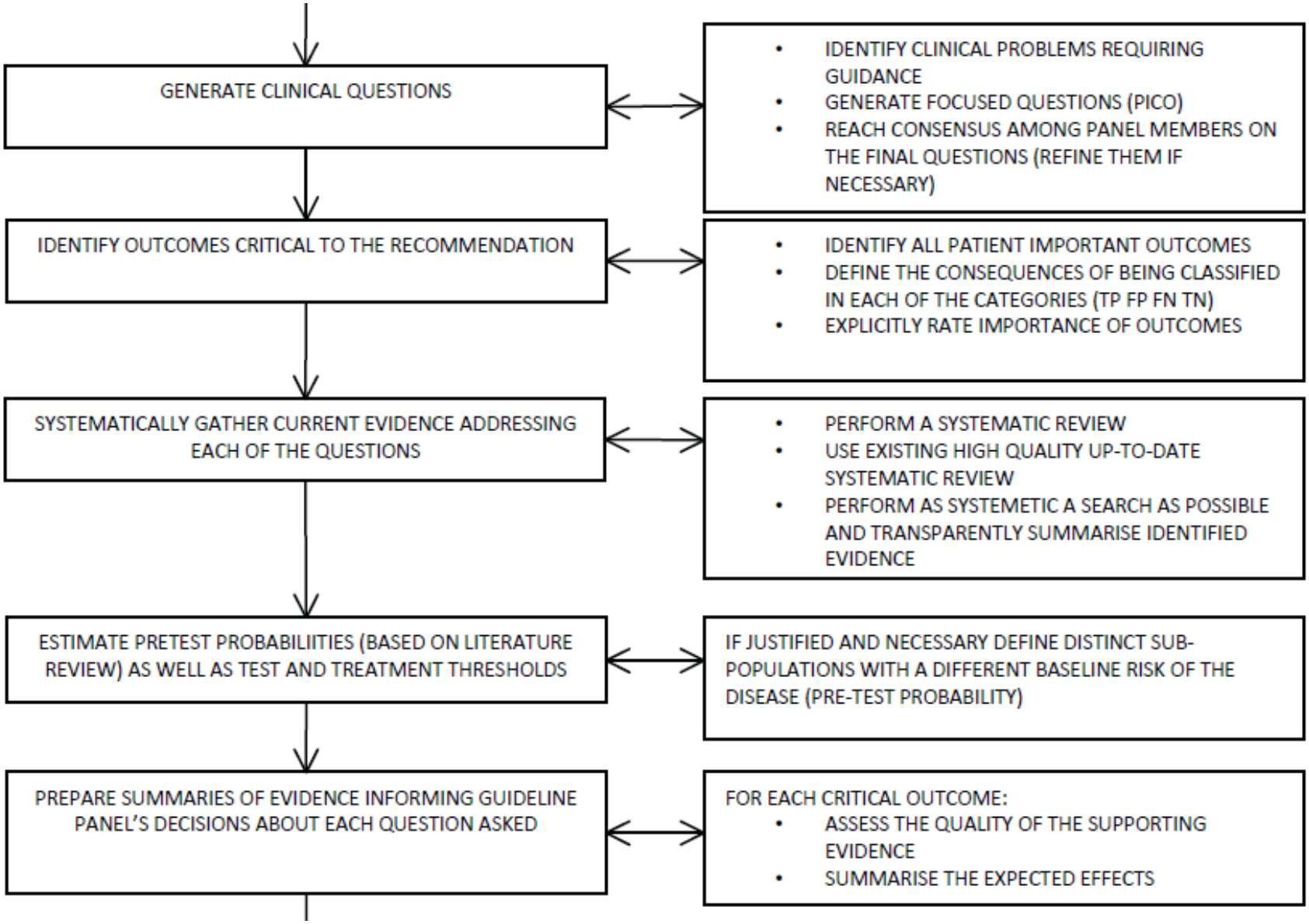
Population (Wer?)

Intervention (Welcher test)

Comparison (Anstatt?)

Outcomes (Wofür?)

- Anaphylaxis
- Umstände
- Benutzung von Kuhmilchersatz
- Korrekte Diagnose verzögert
- Ressourcen


GENERATE CLINICAL QUESTIONS
• IDENTIFY CLINICAL PROBLEMS REQUIRING GUIDANCE
• GENERATE FOCUSED QUESTIONS (PICO)
• REACH CONSENSUS AMONG PANEL MEMBERS ON THE FINAL QUESTIONS (REFINE THEM IF NECESSARY)
IDENTIFY OUTCOMES CRITICAL TO THE RECOMMENDATION
• IDENTIFY ALL PATIENT IMPORTANT OUTCOMES
• DEFINE THE CONSEQUENCES OF BEING CLASSIFIED IN EACH OF THE CATEGORIES (TP FP FN TN)
• EXPLICITLY RATE IMPORTANCE OF OUTCOMES
SYSTEMATICALLY GATHER CURRENT EVIDENCE ADDRESSING EACH OF THE QUESTIONS
• PERFORM A SYSTEMATIC REVIEW
• USE EXISTING HIGH QUALITY UP-TO-DATE SYSTEMATIC REVIEW
• PERFORM AS SYSTEMETIC A SEARCH AS POSSIBLE AND TRANSPARENTLY SUMMARISE IDENTIFIED EVIDENCE
ESTIMATE PRETEST PROBABILIITIES (BASED ON LITERATURE REVIEW) AS WELL AS TEST AND TREATMENT THRESHOLDS
IF JUSTIFIED AND NECESSARY DEFINE DISTINCT SUB-POPULATIONS WITH A DIFFERENT BASELINE RISK OF THE DISEASE (PRE-TEST PROBABILITY)
PREPARE SUMMARIES OF EVIDENCE INFORMING GUIDELINE PANEL'S DECISIONS ABOUT EACH QUESTION ASKED
FOR EACH CRITICAL OUTCOME:
• ASSESS THE QUALITY OF THE SUPPORTING EVIDENCE
• SUMMARISE THE EXPECTED EFFECTS

TP: the child will undergo oral food challenge which will turn out positive with risk of anaphylaxis, albeit in controlled environment; burden on time and anxiety for family; exclusion of milk and use of special formulae. Some children with high pre-test probability of disease and/or at high risk of anaphylactic shock during the challenge will not undergo challenge test and be treated with the same consequences of treatment as those who underwent food challenge.

TN: the child will receive cow's milk at home with no reaction, no exclusion of milk, no burden on family time and decreased use of resources (no challenge test, no formulae); anxiety in the child and family may depend on the family; looking for other explanation of the symptoms.

FP: the patient will undergo an oral food challenge which will be negative; unnecessary burden on time and anxiety in a family; unnecessary time and resources spent on oral challenge. Some children with high pre-test probability of CMA would not undergo challenge test and would be unnecessarily treated with elimination diet and formula that may led to nutritional deficits (e.g. failure to thrive, rickets, vit D or calcium deficiency); also stress for the family and unnecessary carrying epinephrine self injector which may be costly as well as delayed diagnosis of the real cause of symptoms.

FN: the child will be allowed home and will have an allergic reaction (possibly anaphylactic) to cow's milk at home; high parental anxiety and reluctance to introduce future foods; may lead to multiple exclusion diet. The real cause of symptoms (i.e. CMA) will be missed leading to unnecessary investigations & treatments.

Inconclusive results: (either negative positive control or positive negative control): the child would repeat SPT which may be distressing for the child and parent; time spent by a nurse and a repeat clinic appointment would have resource implications; alternatively child would have sIgE measured or undergo food challenge

Complications of a test: SPT can cause discomfort or exacerbation of eczema which can cause distress and parental anxiety; food challenge may cause anaphylaxis and exacerbation of other symptoms.

Resource utilization (cost): SPT adds extra time to clinic appointment however; oral food challenge has much greater resource implications.

Konsequenzen der Fehldiagnose: KMA nicht diagnostiziert

(False Negative SPT result)

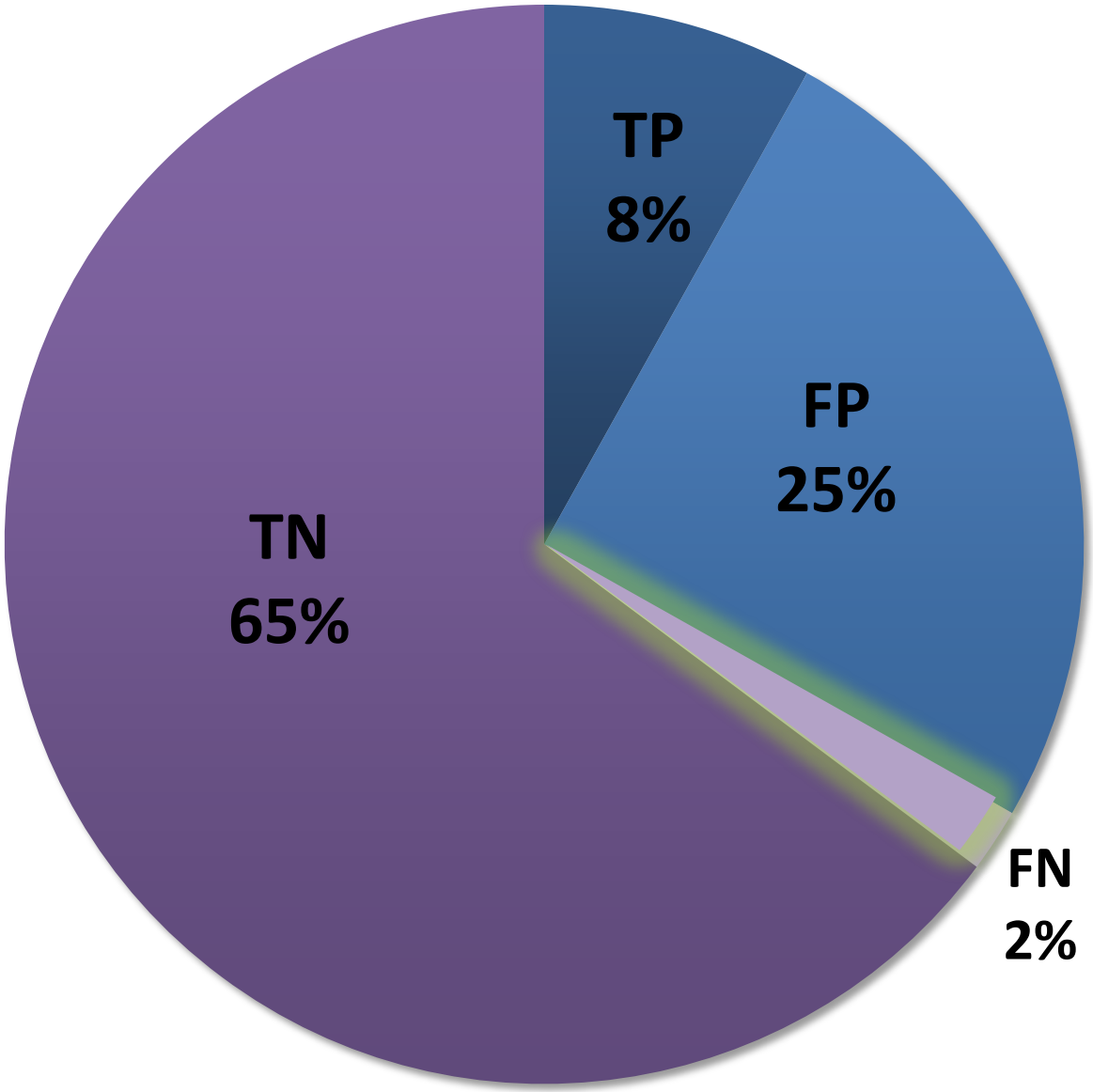
- Allergische (anaphylaktische) Reaktion auf Kuhmilch
- Angst der Eltern
- Verminderte Einführung von anderen Nahrungsstoffen
- Unnötige andere Untersuchungen und Behandlungen

McMaster University
Inspiring Innovation and Discovery

Sensitivität
0,81 (95% CI: 0,77 to 0,85)

Specifizität
0,72 (95% CI: 0,68 to 0,76)

SPT 3 mm
Initiale Wahrscheinlichkeit ~10%



SPT 3 mm

Initiale Wahrscheinlichkeit ~10%

Outcome	No. of studies	Study design	Factors that may decrease quality of evidence					Final quality	Effect per 1000[11]	Importance
			Limitations	Indirectness	Inconsistency	Imprecision	Reporting bias			
True positives (patients with CMA)	11 studies (1088 patients)	Consecutive or non-consecutive series	Serious[12]	None	Serious[13]	None	Unlikely	⊕⊕OO low	Prev 80%: 648 Prev 40%: 324 Prev 10%:81	CRITICAL
True negatives (patients without CMA)	11 studies (1088 patients)	Consecutive or non-consecutive series	Serious[2]	None	Serious[3]	None	Unlikely	⊕⊕OO low	Prev 80%: 144 Prev 40%: 432 Prev 10%: 648	CRITICAL
False positives (patients incorrectly classified as having CMA)	11 studies (1088 patients)	Consecutive or non-consecutive series	Serious[2]	Serious[14]	Serious[3]	None	Unlikely	⊕OOO very low	Prev 80%: 56 Prev 40%: 168 Prev 10%: 252	CRITICAL
False negatives (patients incorrectly classified as not having CMA)	11 studies (1088 patients)	Consecutive or non-consecutive series	Serious[2]	None	Serious[3]	None	Unlikely	⊕⊕OO low	Prev 80%: 152 Prev 40%: 76 Prev 10%: 19	CRITICAL
Inconclusive [15]	Not reported	–	–	–	–	–	–	–	–	IMPORTANT
Complications	Not reported	–	–	–	–	–	–	–	–	NOT IMPORTANT
Cost	Not reported	–	–	–	–	–	–	–	–	NOT IMPORTANT

Based on combined sensitivity of 81% (95% CI: 77 to 85) and specificity of 72% (95% CI: 68 to 76)

1,2 Most studies enrolled highly selected patients with atopic eczema or gastrointestinal symptoms, no study reported if an index test or a reference standard were interpreted without knowledge of the results of the other test, but it is very likely that those interpreting results of one test knew the results of the other; all except for one study that reported withdrawals did not explain why patients were withdrawn.

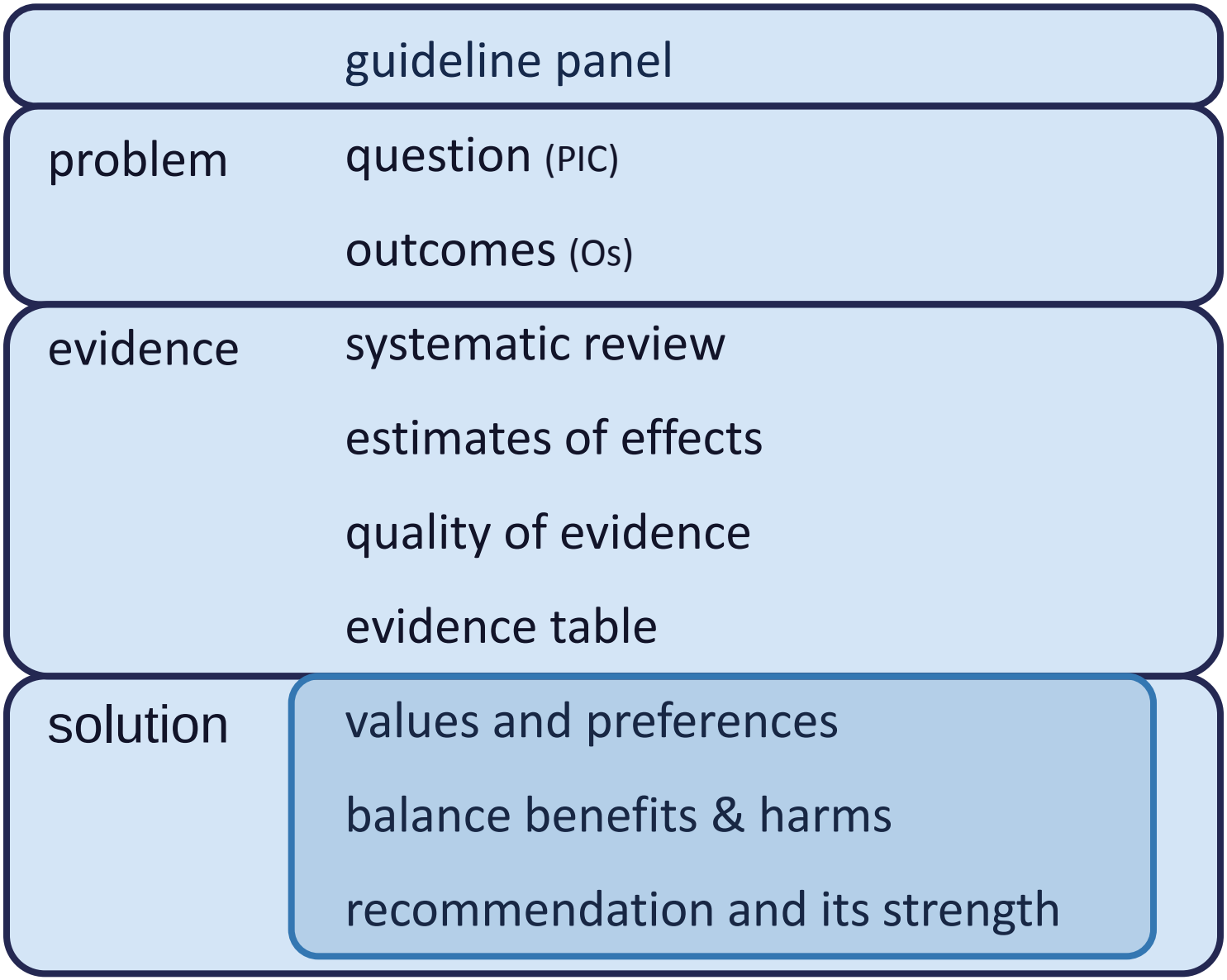
3 Estimates of sensitivity ranged from 10% to 100%, and specificity from 14% to 100%; we could not explain it by quality of the studies, tests used or included population

4 There is uncertainty about the consequences for these patients; in some a diagnosis of other potentially serious condition may be delayed

One study in a different population (children younger than 12 months) reported 8% inconclusive challenge tests but did not report number of inconclusive skin prick tests.

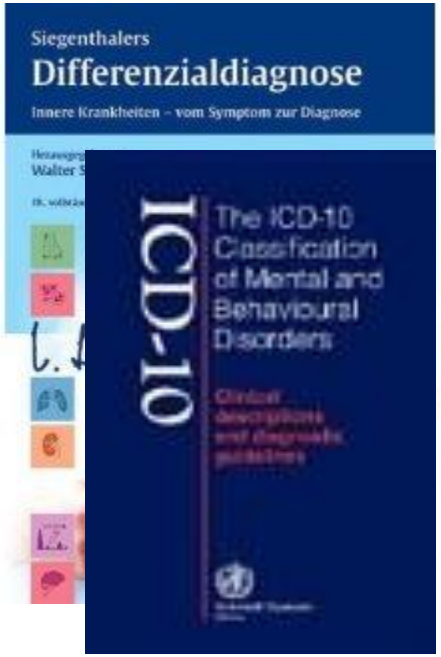
Outcomes	Illustrative Risks (95% CI) Assumed outcome with CT – prevalence of 20%	Number of participants (studies)	Quality of the Evidence[1]	Comments
True positives (Patients correctly classified as having coronary artery disease)	**192 per 1000**	1570 (21)	⊕⊕⊕◯ **Moderate**[2]	Benefit from treatment and fewer complications.* Some patients will have to undergo angiography.
True negatives (Patients correctly classified as not having coronary artery disease)	**592 per 1000**	1570 (21)	⊕⊕⊕◯ **Moderate**[2]	Benefit from reassurance and fewer complications
False positives (Patients incorrectly classified as having coronary artery disease)	**208 per 1000**	1570 (21)	⊕⊕⊕◯ **Moderate**[2]	Harm from unnecessary treatment
False negatives (Patients incorrectly classified as not having coronary artery disease)	**8 per 1000**	1570 (21)	⊕⊕◯◯ **Low**[2, 3]	Detriment from delayed diagnosis or myocardial insult
Complications (MI, allergic reactions, renal failure)	**99 per 1000**	1570 (21)	⊕⊕◯◯ **Low**[2]	There is a higher rate of rare complications (infarction and death) and higher cost with angiography – a full profile would be required.
Resource use* (cost of CT and Angiography)	**See comment**	See comment	See comment	Cost are higher for angiography,

1- Quality rated from 1 (very low quality) to 4 (high quality), 2- Cross sectional studies. Indirectness of outcomes in a wide spectrum of patients and indirect comparison of tests, 3– there is greater uncertainty whether these patients will have negative outcomes.
*Assumed efficacy of: 1) aspirin daily = 20% RRR; 2) beta-blockage = 18% RRR.

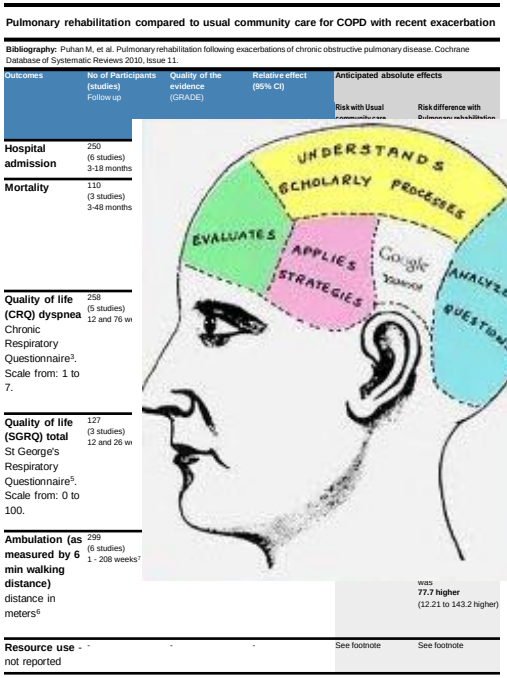


Übersicht

Diagnostische Fragestellungen



Einführung



Pulmonary rehabilitation compared to usual community care for COPD with recent exacerbation

Bibliography: Puhan M, et al. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2010, Issue 11.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects: Risk with Usual community care	Anticipated absolute effects: Risk difference with Pulmonary rehabilitation
Hospital admission	250 (6 studies) 3-18 months				
Mortality	110 (3 studies) 3-48 months				
Quality of life (CRQ) dyspnea Chronic Respiratory Questionnaire[3]. Scale from: 1 to 7.	258 (5 studies) 12 and 76 we				
Quality of life (SGRQ) total St George's Respiratory Questionnaire[5]. Scale from: 0 to 100.	127 (3 studies) 12 and 26 we				
Ambulation (as measured by 6 min walking distance) distance in meters[6]	299 (6 studies) 1 - 208 weeks[7]				was **77.7 higher** (12.21 to 143.2 higher)
Resource use - not reported	-	-	-	See footnote	See footnote

Evidenz & Beurteilungen

Empfehlungen & Implementierung

Recommendation 1.4.

In patients with low pre-test probability of CMA we suggest using a skin prick test with a cut-off value of ≥3 mm as a triage test to avoid oral food challenge in those in whom the result of a skin prick test turns out negative.
(weak recommendation | low quality evidence)

Underlying values and preferences

This recommendation places a relatively high value on avoiding risk of anaplylaxis, burden and resource use with an OFC test (~67% challenges avoided). It places a lower value on avoiding an allergic reaction in around 1 in 25–50 patients misclassified as not having CMA while they would actually be allergic to cow's milk (2–4% false negative results).

Other examples of GRADE in diagnostic reviews and guidelines

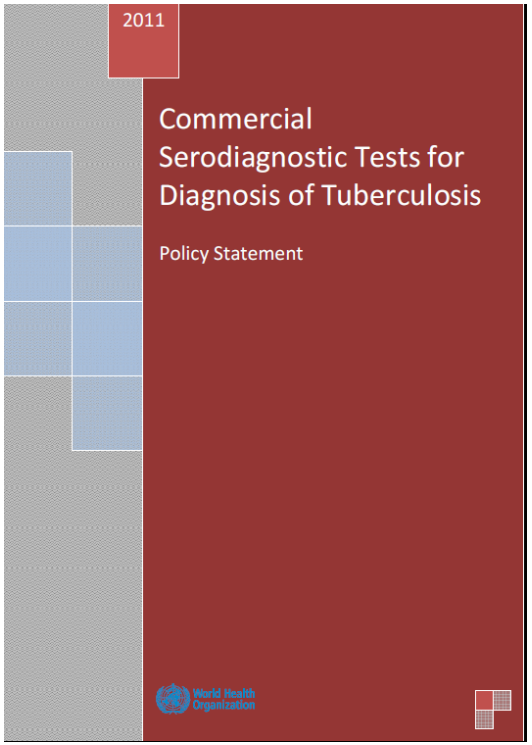
OPEN ACCESS Freely available online

PLOS MEDICINE

Commercial Serological Tests for the Diagnosis of Active Pulmonary and Extrapulmonary Tuberculosis: An Updated Systematic Review and Meta-Analysis

Karen R. Steingart[1], Laura L. Flores[2,3], Nandini Dendukuri[4], Ian Schiller[4], Suman Laal[5,6,7], Andrew Ramsay[8], Philip C. Hopewell[2,3], Madhukar Pai[4]*

1 Department of Health Services, University of Washington School of Public Health, Seattle, Washington, United States of America, **2** Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, University of California, San Francisco, California, United States of America, **3** Curry International Tuberculosis Center, University of California, San Francisco, California, United States of America, **4** Department of Epidemiology, Biostatistics, and Occupational Health, McGill University & Montreal, Chest Institute, Montreal, Quebec, Canada, **5** Department of Pathology, New York University Langone Medical Center, New York, New York, United States of America, **6** Department of Microbiology, New York University Langone Medical Center, New York, New York, United States of America, **7** Veterans Affairs Medical Center, New York, New York, United States of America, **8** UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, World Health Organization, Geneva, Switzerland



Annahmen und Beurteilungen

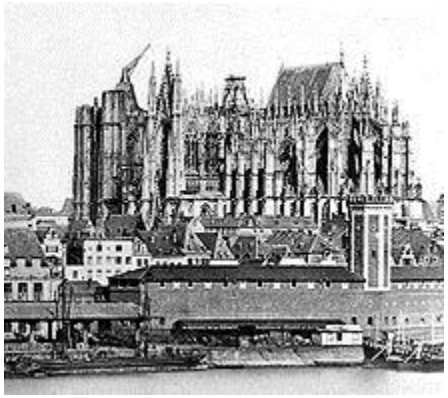
Example of new test and reference test or strategy	Putative benefit of new test	Diagnostic accuracy		Patient Outcomes and expected impact on management for the following test outcomes				Balance between presumed patient outcomes, test complications and cost
		Sensitivity	Specificity	True positives	True negatives	False positives	False negatives	
A shorter version of a dementia test compared with the original Mini Mental State Exam for diagnosis of dementia	Simpler test, less time	equal	equal	*Presumed influence on patient important outcomes*				Evidence of shorter time and similar test accuracy (and thus patient outcomes) would generally support the new test's usefulness
				Uncertain benefit from earlier diagnosis and treatment	Almost certain benefit from reassurance	Likely anxiety and possible morbidity from additional testing and treatment	Possible detriment from delayed diagnosis	
				Directness of the evidence (test results) for patient-important outcomes				
				Some uncertainty	No uncertainty	Some uncertainty	Major uncertainty	

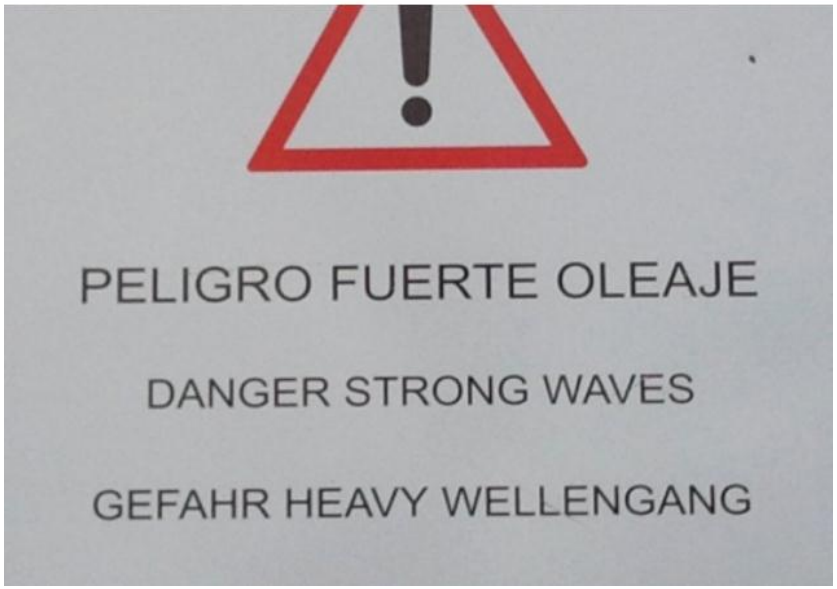
Annahmen und Beurteilungen

Example of new test and reference test or strategy	Putative benefit of new test	Diagnostic accuracy		Patient Outcomes and expected impact on management for the following test outcomes				Balance between presumed patient outcomes, test complications and cost
		Sensitivity	Specificity	True positives	True negatives	False positives	False negatives	
Helical CT for renal calculus compared with intravenous pyeolgram	Detection of more (but smaller) calculi	greater	equal	Presumed influence on patient important outcomes				Less complications and downsides compared to IVP would support the new test's usefulness, but the balance between desirable and undesirable effects is not clear in view of the uncertain consequences of identifying smaller stones.
				Certain benefit for larger stones, for smaller stones the benefit is less clear and unnecessary treatment can result	Almost certain benefit from avoiding unnecessary tests	Likely detriment from unnecessary additional invasive tests	Likely detriment for large stones, less certain for small stones, but a possible detriment from unnecessary additional invasive tests for other potential causes of complaints	
				Directness of the evidence (test results) for patient-important outcomes				
				Some	No	No	Major	

Zusammenfassung

- ‘Diagnostic accuracy’/Testgüte bedarf Evaluierung im Zusammenhang mit Konsequenzen
  - TP, FP, TN, FN, Ressourcen, Testnebenwirkungen
- Qualitätsbeurteilung muss sich auf alle Glieder in der Kette beziehen
  - Explizite Bewertung der Evidenz – Konsequenzen
    - Systematische Übersichtsarbeiten – Transparenz in den Annahmen
- Ansätze vorhanden, Pilotprojekte




PELIGRO FUERTE OLEAJE
DANGER STRONG WAVES
GEFAHR HEAVY WELLENGANG