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

COMET meeting, Rome, November 19 2014

Outcomes and GRADE Summary of Findings Tables: old and new

Disclosure

- Co-chair GRADE Working Group
- World Health Organization: various committees
  - Co-director, WHO collaborating center on evidence informed policy making
- Cochrane Collaboration – Steering group
- GIN – Board of Directors
- No direct financial COI

GRADE working group

Cochrane

Content

- **Summary of Findings Tables**
  - **Prior work**
  - **Updates**

Self management for patients with chronic obstructive pulmonary disease

Patient or population: patients with chronic obstructive pulmonary disease
Settings: primary care, community, outpatient
Intervention: self management[1]
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk **usual care**	Corresponding risk **self management**				
Quality of Life St George's Respiratory Questionnaire. Scale from: 0 to 100. (follow-up: 3 to 12 months)	The mean quality of life ranged across control groups from **38 to 60 points**	The mean quality of Life in the intervention groups was **2.58 lower** (5.14 to 0.02 lower)		698 (7)	⊕⊕⊕O **moderate**[2]	Lower score indicates better quality of life. A change of less than 4 points is not shown to be important to patients.
Dyspnoea Borg Scale. Scale from: 0 to 10. (follow-up: 3 to 6 months)	The mean dyspnoea ranged across control groups from **1.2 to 4.1 points**	The mean dyspnoea in the intervention groups was **0.53 lower** (0.96 to 0.1 lower)		144 (2)	⊕⊕OO **low**[3,4]	Lower score indicates improvement
Number and severity of exacerbations[5]	See comment	See comment	Not estimable[5]	591 (3)	See comment	Effect is uncertain
Respiratory-related hospital admissions (follow-up: 3 to 12 months)	**Low risk population**[6]		OR 0.64 (0.47 to 0.89)	966 (8)	⊕⊕⊕O **moderate**[7]	
	10 per 100	**7 per 100** (5 to 9)				
	High risk population[6]					
	50 per 100	**39 per 100** (32 to 47)				
Emergency department visits for lung diseases (follow-up: 6 to 12 months)	The mean emergency department visits for lung diseases ranged across control groups from **0.2 to 0.7 visits per person per year**	The mean emergency department visits for lung diseases in the intervention groups was **0.1 higher** (0.2 lower to 0.3 higher)		328 (4)	⊕⊕⊕O **moderate**[4]	
Doctor and nurse visits (follow-up: 6 to 12 months)	The mean doctor and nurse visits ranged across control groups from **1 to 5 vists per person per year**	The mean doctor and nurse visits in the intervention groups was **0.02 higher** (1 lower to 1 higher)		629 (8)	⊕⊕⊕O **moderate**[8]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; OR: Odds ratio;

Evidence profiles

Question and source of evidence (systematic review)

Population, intervention, comparator, outcomes

Out Methds and evaluation

Effect estima

Confidence/quality by outcome:
High
Moderate
Low
Very low

Serious Adverse Events during investigational 24 week treatment phase (C208 Stages 1 and 2: ITT) 7 (assessed through clinical and laboratory results)												
2[8]	randomized trials	no serious risk of bias	no serious inconsistency	Serious[9]	very serious[5]	none	7/102[10] (6.9%)	2/105 (1.9%)	RR 3.6 (0.77 to 14.00)			
Mortality up to end of study at 120 weeks (C208 Stage 2: ITT) (deaths reported)												
1[11]	randomized trials	no serious risk of bias	no serious inconsistency	serious[12]	very serious[3]	none	9/79[11] (12.7%)	1/81[11] (2.5%)	RR 9.23 (1.20 to 72.95)[13,14]	10 more per 100 (from 0 more to 53 more)	+OOO Very Low	Critical
Time to conversion over 24 weeks (C208 Stage 2: mITT1) (measured with microbiological endpoints - MGIT960)												
1[15]	randomized trials	no serious risk of bias[4]	no serious inconsistency	serious[16]	serious[5]	none	n=66[1] median=83 days	n=66[1] median=125 days		median 42 days lower[17]	++OO Low	Critical

1 The mITT modified intention to treat population in C208 trial consisted of 66 subjects in each randomization group after excluding 13 subjects (16.5%) treated with bedaquiline and 15 subjects (18.5%) with placebo who did not have MDR or pre-XDR-TB at baseline or for whom MGIT results were considered not evaluable.
2 Cure defined as 5 consecutive negative cultures from samples collected at least 30 days apart in the final 12 months of treatment , OR if only 1 culture is reported positive during that period, then a further 3 consecutive negative cultures from samples taken at least 30 days apart.
3 End of study data slide supplied by Janssen subsequent to US-FDA meeting. In this slide, mention is made of 'treatment success', but the company further clarified that the strict WHO definition of 'cure' was being used.
4 Representativeness of the mITT population (assumptions made for ITT population).
5 Small sample size and resulting large confidence interval limits precision: few (= serious) or very few (= very serious) observations.
6 This difference is statistically significant (Fisher p=0.005; Pearson p=0.003).

McMaster University
Inspiring Innovation and Discovery

First study of SoF

- Number of outcomes ≤ 7
- Presenting information on all important outcomes; ordering of outcomes
- 20 Cochrane review authors participated (20 new or updated reviews)
  - spent an additional 4 h (2 to 40 h)
  - layout clear
  - 11/17 increased accessibility
  - 5/17 improved quality
  - 1/17 rephrased conclusions

RCTs

ELSEVIER

Journal of Clinical Epidemiology 63 (2010) 620–626

Journal of Clinical Epidemiology

Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information

Sarah E. Rosenbaum[a,*], Claire Glenton[b], Andrew D. Oxman[a]

- RCTs
  - 1 EBCP workshop (N 72); 2 Cochrane entities meeting (N 33)
- RCT 1: easy to find results, SoF versus no: 68 vs. 40% (p = 0.02)
- RCT 2: SoF more correct answers to two questions re results
  - 93% vs 44% (p = 0.003) and 87% vs. 11% (p < 0.001)
- SoF participants spent average of 90 seconds to find key information vs 4 minutes without SoF table

Probiotics as an adjunct to antibiotics for the prevention of pediatric antibiotic-associated diarrhea in children

Patient or population: children given antibiotics
Settings: inpatients and outpatient
Intervention: probiotics
Comparison: no probiotics

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk No probiotics	Corresponding risk Probiotics	Relative effect (95% CI)	No of particip (studi	
Incidence of Diarrhea: Probiotic dose (equal to/greater than) 5 billion CFU/day Follow-up: 10 days to 3 months	**Children < 5 years**		**RR 0.4**[1] (0.29 to 0.5	14	134 fewer
	223 per 1000[1]	**89 per 1000** (65 to 122)			
	Children > 5 years		**RR 0.8**[1] (0.53 to 1.21)	624	22 fewer
	112 per 1000[1]	**90 per 1000** (59 to 136)			
Adverse events Follow-up: 10 to 44 days	**18 per 1000**[1]	**23 per 1000** (8 to 38)	Not estimable[4]	1575 (11 studies)	5 more
Duration of diarrhea Follow-up: 10 days to 3 months	The mean duration of diarrhea in control groups was **4 days**	**0.6 fewer days** (1.18 to 0.02 fewer days)			

...miting, increased phlegm, chest pain, constipation, taste ...sturbance, and low ...petite

Optimal format Evidence Profiles

ELSEVIER

Journal of Clinical Epidemiology ■ (2012) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial

Per Olav Vandvik[a,b,*], Nancy Santesso[c], Elie A. Akl[c,d], John You[c,e], Sohail Mulla[c], Frederick A. Spencer[c,e], Bradley C. Johnston[c], Jan Brozek[c], Julia Kreis[f,g], Linn Brandt[b], Qi Zhou[c], Holger S. Schunemann[c,e], Gordon Guyatt[c,e]

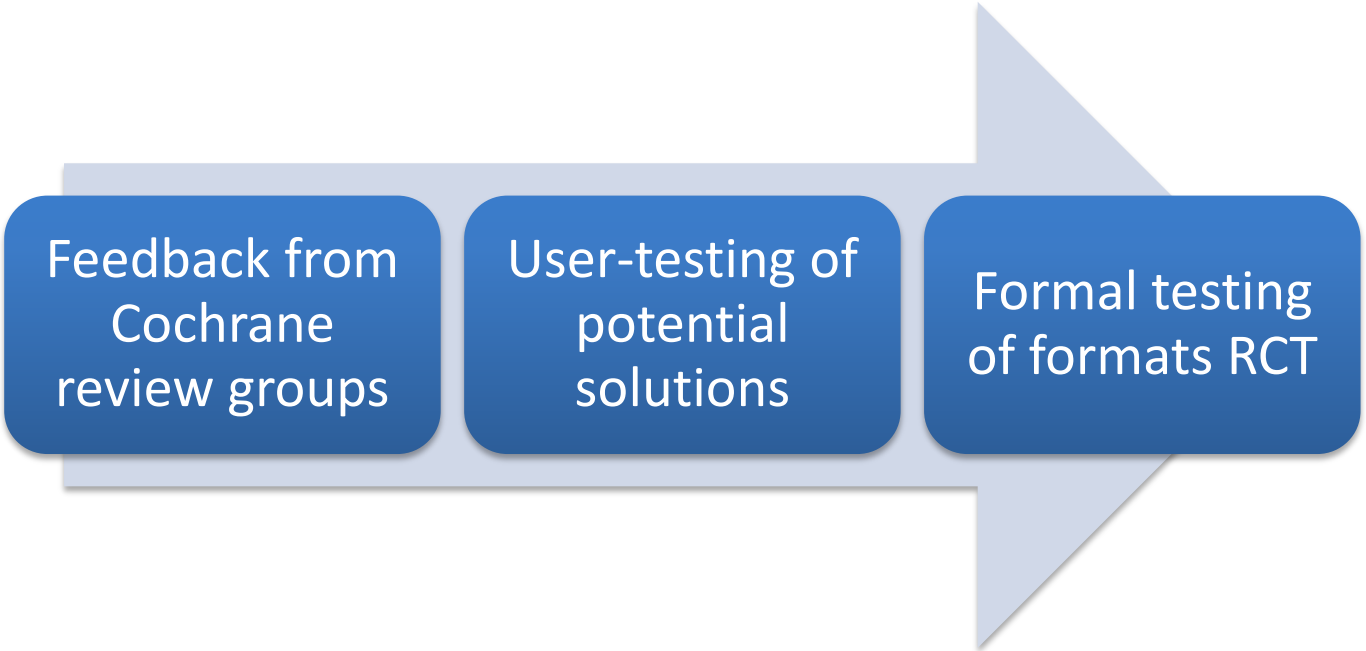
- RCT alternative testing formats of Evidence Profiles in 88 guideline panelists
- Preferences for:
  - Study event rates over no (small/moderate effect)
  - Absolute risk differences over absolute risks (large effect)
    - more likely to correctly interpret CI around difference (58 vs 11%, $p < 0.0001$)
  - Information in table cells vs footnotes (small/moderate effect)
    - In cells more likely to get time frame right (58% vs 11%, $p < 0.0001$)
- Information EP judged easy to find, comprehend, helpful both (median 6 of 7)

Content

- **Summary of Findings Tables**
  - Prior work
  - **Updates**

Cochrane method innovation fund project

- Enhancing the acceptance and implementation of SoF tables in Cochrane reviews
- Initiated in 2012



Enhancing the acceptance and implementation of SoF tables in Cochrane reviews

User testing

- 40 participants
- Cochrane review users (clinicians, guideline developers, researchers)
- Participants prefer simple, less crowded SoF tables
- Dichotomous: NNTs, Risk Difference over natural frequencies
- Continuous: Minimal important difference units over MD and SMD
- "what happens" column:
  - statement of presence/direction of effect and qualitative statement of confidence

Table 1. Comparison between items included in the current and alternative SoF tables formats

	Current formats (Table B)	**Alternative formats (Table A)**
1	Inclusion of the Nº of participants and studies column	Exclusion of the Nº of participants and studies column. Information presented in the outcomes column
2	Quality of evidence presented with symbols and labeled as High, moderate, low, or very low. Reasons for downgrading presented in the footnotes	Quality of evidence presented with main reasons for downgrading in the same column (e.g. MODERATE due to imprecision)
3	"Footnotes" label	"Explanations" label
4	Baseline risk and corresponding risk expressed as natural frequencies	Baseline risk and corresponding risk expressed as percentages
5	No column presenting absolute risk reduction (risk difference) or mean difference	Inclusion of a column presenting absolute risk reduction (risk difference) or mean difference
6	Comments column included	Comments column deleted
7	No "what happens" column*	"What happens" column included*
8	Description of the GRADE Working Group grades of evidence definitions below the table	No description of the GRADE Working Group grades of evidence definitions

Probiotics as an adjunct to antibiotics for the prevention of pediatric antibiotic-associated diarrhea in children

Patient or population: children given antibiotics
Settings: inpatients and outpatient
Intervention: probiotics
Comparison: no probiotics

Outcomes No of Participants (studies)	Relative effects (95% CI)	Anticipated absolute effects* (95% CI) Without probiotics	 With probiotics	 Difference	Quality of the evidence (GRADE)	What happens
Incidence of Diarrhea: Probiotic dose 5 billion CFU/day Follow-up: 10 days to 3 months		**Children < 5 years**				
Children <5 years 1474 (7 studies)	**RR 0.4**[1] (0.29 to 0.55)	**22.3%**[1]	**8.9%** (6.5 to 12.2)	**13.4% fewer children**[1] (10.1 to 15.8 fewer)	⊕⊕⊕⊖ **moderate**[2] Due to risk of bias	Probably decreases the incidence of diarrhea
		Children > 5 years				
Children >5 years 624 (4 studies)	**RR 0.8**[1] (0.53 to 1.21)	**11.2%**[1]	**9%** (5.9 to 13.6)	**2.2% fewer children**[1] (5.3 fewer to 2.4 more)	⊕⊕⊖⊖ **low**[2, 3] Due to risk of bias and imprecision	May decrease the incidence of diarrhea
Adverse events[4] Follow-up: 10 to 44 days	-	**1.8%**[1]	**2.3%** (0.8 to 3.8)	**0.5% more adverse events**[5] (1 fewer to 2 more)	⊕⊕⊖⊖ **low**[6, 7] Due to risk of bias and inconsistency	There may be little or no difference in adverse events
1575 (11 studies)						

RCT Design

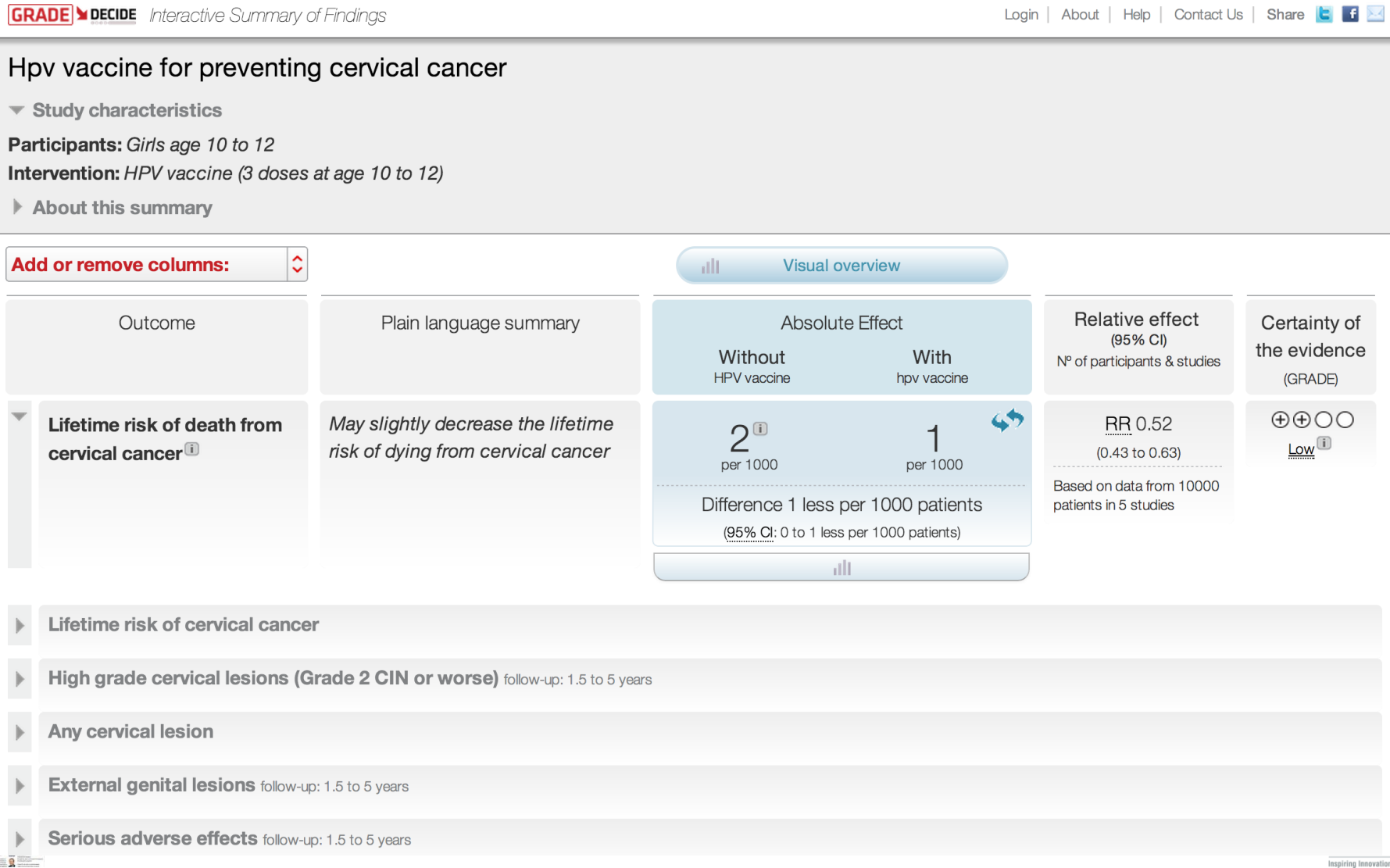
- 300 Clinicians, guideline developers, researchers
- Alternative vs current formats
- Understanding, accessibility, satisfaction, preference

RCT Results

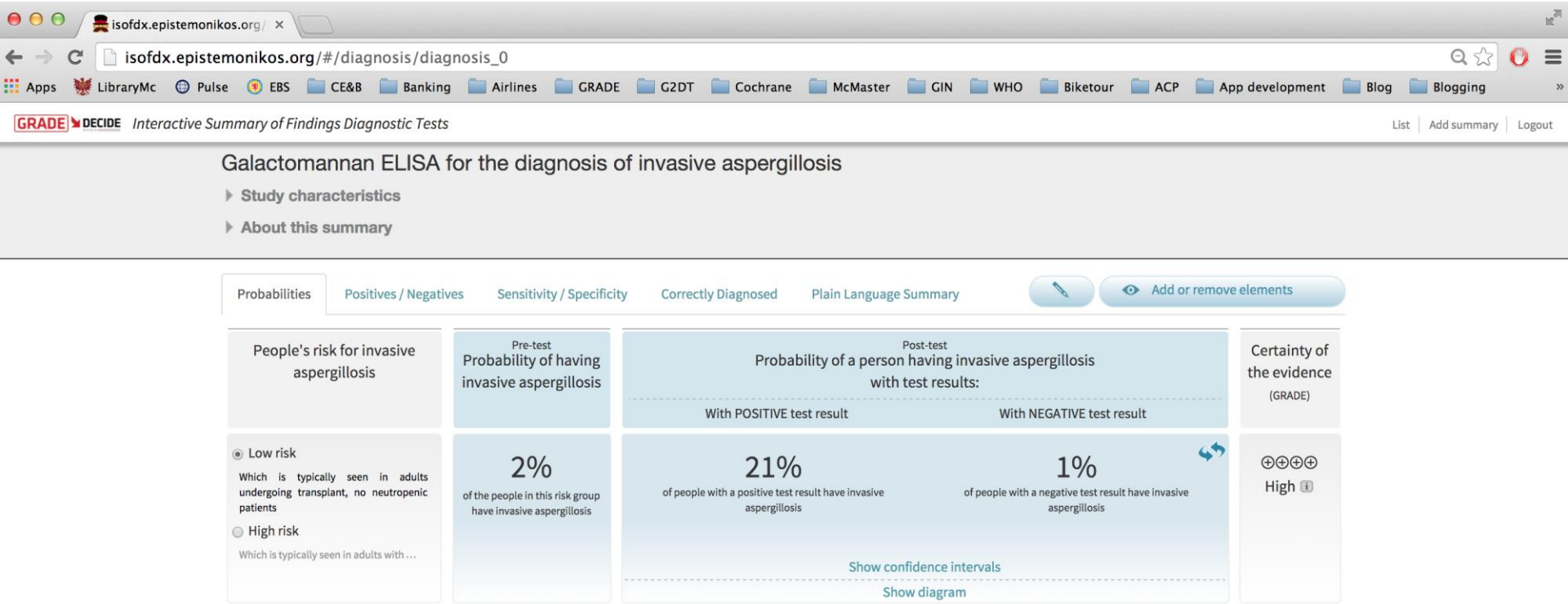
Percentage of participants that answered correctly understanding questions

Concept	Question asked	Alternative formats (N=122)	Current formats (N=168)	Difference	P value
Ability to determine risk difference	How many fewer children <5 years will have diarrhea if they have probiotics than if they do not?	98%	35%	63%	<0.001
Understanding of quality of evidence and treatment effect	Which of the following statements best represents the results informing the outcome adverse events?	88%	26%	62%	<0.001

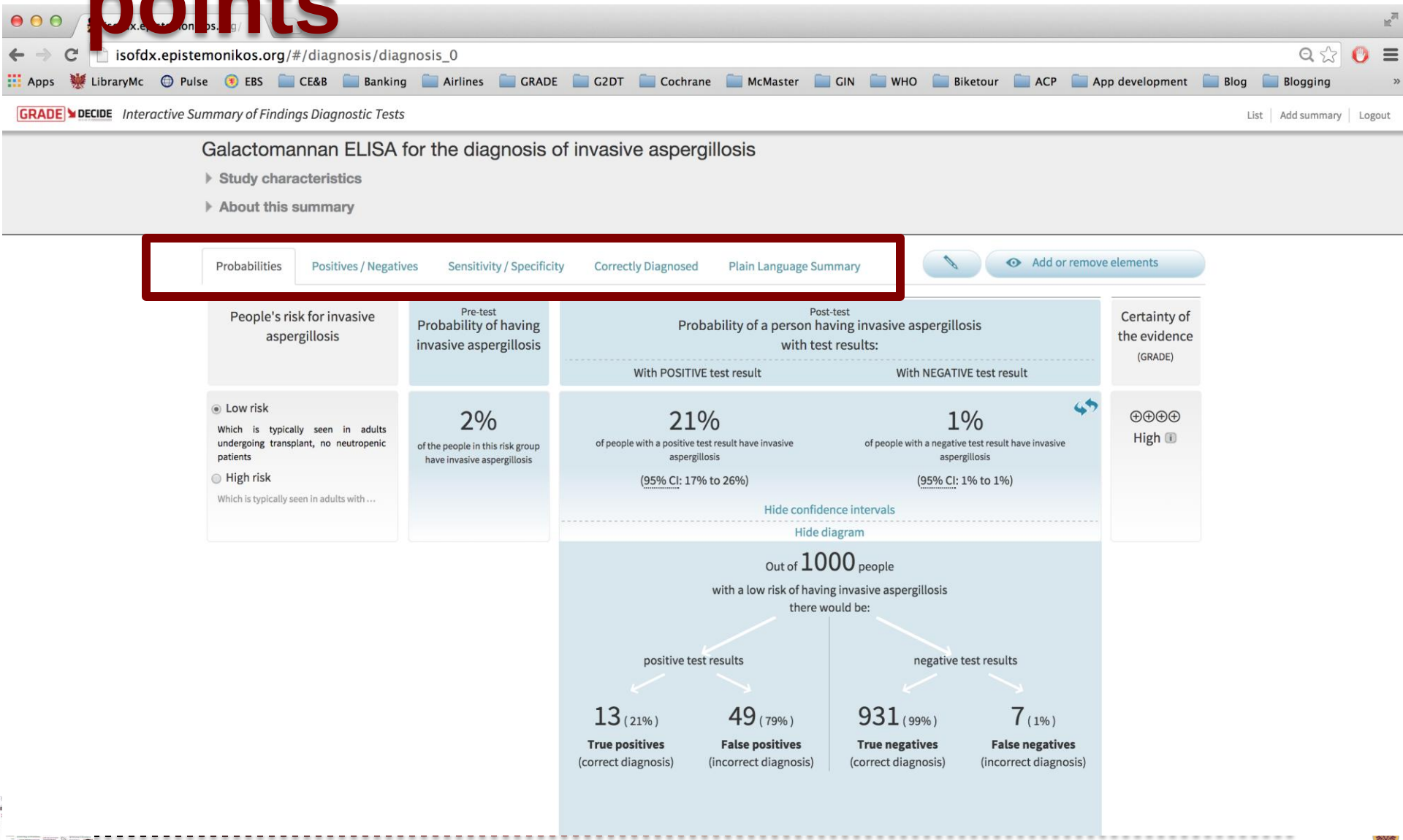
DECIDE: Interactive Summary of Findings tables

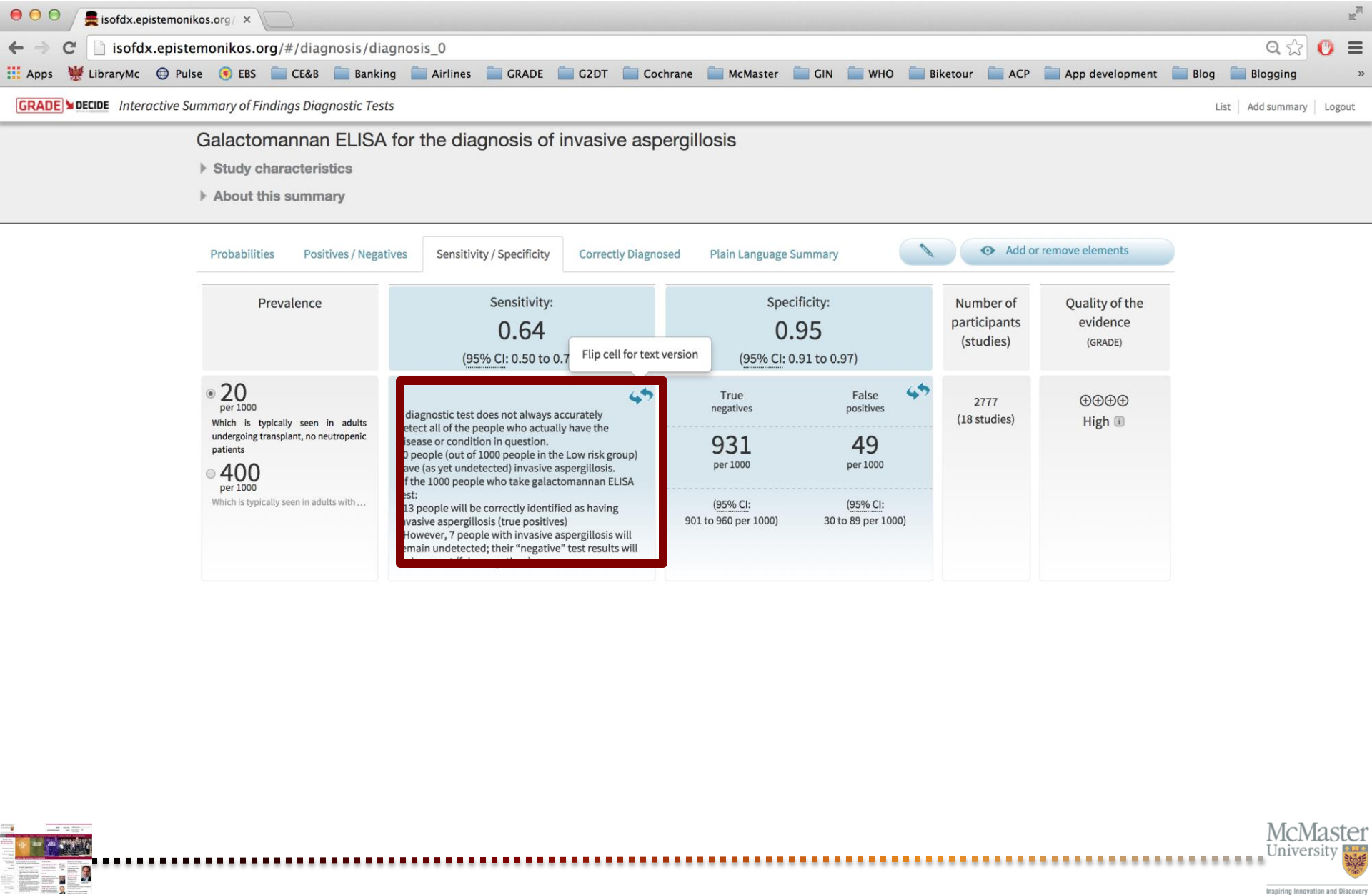


Diagnostic SoF Tables



Testing of different entry points




isofdx.epistemonikos.org/
isofdx.epistemonikos.org/#/diagnosis/diagnosis_0
Apps
LibraryMc
Pulse
EBS
CE&B
Banking
Airlines
GRADE
G2DT
Cochrane
McMaster
GIN
WHO
Biketour
ACP
App development
Blog
Blogging
GRADE DECIDE
Interactive Summary of Findings Diagnostic Tests
List
Add summary
Logout
Galactomannan ELISA for the diagnosis of invasive aspergillosis
Study characteristics
About this summary
Probabilities
Positives / Negatives
Sensitivity / Specificity
Correctly Diagnosed
Plain Language Summary
Add or remove elements
Prevalence
Sensitivity:
0.64
(95% CI: 0.50 to 0.7
Flip cell for text version
Specificity:
0.95
(95% CI: 0.91 to 0.97)
Number of participants (studies)
Quality of the evidence (GRADE)
20
per 1000
Which is typically seen in adults undergoing transplant, no neutropenic patients
400
per 1000
Which is typically seen in adults with …
diagnostic test does not always accurately
etect all of the people who actually have the
isease or condition in question.
0 people (out of 1000 people in the Low risk group)
ave (as yet undetected) invasive aspergillosis.
f the 1000 people who take galactomannan ELISA
st:
13 people will be correctly identified as having
vasive aspergillosis (true positives)
However, 7 people with invasive aspergillosis will
main undetected; their "negative" test results will
True negatives
931
per 1000
(95% CI:
901 to 960 per 1000)
False positives
49
per 1000
(95% CI:
30 to 89 per 1000)
2777
(18 studies)
⊕⊕⊕⊕
High

isofdx.epistemonikos.org/

isofdx.epistemonikos.org/#/diagnosis/diagnosis_0

Apps | LibraryMc | Pulse | EBS | CE&B | Banking | Airlines | GRADE | G2DT | Cochrane | McMaster | GIN | WHO | Biketour | ACP | App development | Blog | Blogging

GRADE DECIDE *Interactive Summary of Findings Diagnostic Tests*

List | Add summary | Logout

Galactomannan ELISA for the diagnosis of invasive aspergillosis

- Study characteristics
- About this summary

Probabilities | Positives / Negatives | Sensitivity / Specificity | Correctly Diagnosed | Plain Language Summary | Add or remove elements

Researchers reviewed studies comparing one/two tests to invasive aspergillosis: the composite of EORT/MSG clinical and histological criteria test and the commercial Platelia ¸sandwich ELISA detecting galactomannan in serum test. They searched for all relevant studies up to [date] and found 18 relevant studies.

What are composite of EORT/MSG clinical and histological criteria and commercial Platelia ¸sandwich ELISA detecting galactomannan in serum tests?

The composite of EORT/MSG clinical and histological criteria and the commercial Platelia ¸sandwich ELISA detecting galactomannan in serum tests are tests that a clinician performs to check for invasive aspergillosis. This disease can .

The composite of EORT/MSG clinical and histological criteria test checks if a person has invasive aspergillosis. The test is done in the following way: [description of how the test is done].

The commercial Platelia ¸sandwich ELISA detecting galactomannan in serum test also checks if a person has invasive aspergillosis. The test is done in the following way:

What the research says about the tests

What are commercial Platelia ¸sandwich ELISA detecting galactomannan in serum and composite of EORT/MSG clinical and histological criteria? The composite of EORT/MSG clinical and histological criteria and commercial Platelia ¸sandwich ELISA detecting galactomannan in serum tests check for invasive aspergillosis.

What the research says about the tests

A positive test should mean that the person has invasive aspergillosis. A negative test should mean the person does not have invasive aspergillosis. But very few tests are perfect and two problems can occur. A positive test could incorrectly say that a person has invasive aspergillosis when in fact s/he does not (called a "false positive"). As a consequence, this person may have more testing, be worried or treated for no reason. A negative test could incorrectly say that a person does not have invasive aspergillosis when in fact s/he does have invasive aspergillosis (called a "false negative"). In this person, invasive aspergillosis would be missed by the test and s/he may not receive the necessary treatment. When the quality of the evidence is low or very low as opposed to moderate or high, the size of this problem can be considerably larger or smaller than what the numbers indicate.

(For frequencies use)

Probabilities | Positives / Negatives | Sensitivity / Specificity

Add or remove columns

Prevalence	People with POSITIVE test result		People with NEGATIVE test result	Pooled Sensitivity/ Specificity	Number of participants (studies)	Quality of the evidence (GRADE)
	True positives	False positives	True negatives			
20 per 1000 400 per 1000	13 per 1000 (95% CI: 9 to 15 per 1000)	49 per 1000 (95% CI: 30 to 89 per 1000)	When the galactomannan ELISA has a NEGATIVE test result: • 931 out of 1000 people do not have invasive aspergillosis However, the test incorrectly says that 7 out of 1000 people do not have invasive aspergillosis but in reality they do ("false negatives")	**Sensitivity** 0.64 (95% CI: 0.50 to 0.77) **Specificity** 0.95 (95% CI: 0.91 to 0.97)	Based on data from 2777 individuals in 18 studies.	⊕⊕⊕⊕ High

Flip cell for text version

McMaster University
Inspiring Innovation and Discovery

Multiple comparisons

Outcome	Events in the screen-treat strategies for patient important outcomes (numbers presented per 1000000 patients)						
Mortality	HPV +/- CVK	HPV +/- Cyro	HPV +/- LEEP	VIA +/- CVK	VIA +/- Cyro	VIA +/- LEEP	No screening
HPV +/- CVK		10 LESS 7 LESS ⊕⊕⊕◯ Moderate	10 LESS 4 LESS ⊕⊕◯◯ Low	61 LESS 7 LESS ⊕⊕⊕⊕ High	68 LESS 7 LESS ⊕⊕◯◯ Low	50 LESS 7 LESS ⊕◯◯◯ Very low	230 LESS 87 LESS ⊕⊕⊕◯ Moderate
HPV +/- Cyro	10 MORE 7 MORE ⊕◯◯◯ Very low		0 MORE 3 MORE ⊕⊕◯◯ Low	51 LESS 0 MORE ⊕⊕⊕⊕ High	58 LESS 0 MORE ⊕⊕◯◯ Low	40 LESS 0 MORE ⊕◯◯◯ Very low	220 LESS 80 LESS ⊕◯◯◯ Very low
HPV +/- LEEP	10 MORE 4 MORE ⊕◯◯◯ Very low	0 MORE 3 LESS ⊕⊕⊕◯ Moderate		51 LESS 3 LESS ⊕⊕⊕⊕ High	58 LESS 3 LESS ⊕⊕◯◯ Low	40 LESS 3 LESS ⊕◯◯◯ Very low	220 LESS 83 LESS ⊕⊕⊕◯ Moderate
VIA +/- CVK	61 MORE 7 MORE ⊕◯◯◯ Very low	51 MORE 0 MORE ⊕⊕⊕◯ Moderate	51 MORE 3 MORE ⊕⊕◯◯ Low		7 LESS 0 MORE ⊕⊕◯◯ Low	11 MORE 0 MORE ⊕◯◯◯ Very low	169 LESS 80 LESS ⊕⊕⊕◯ Moderate
VIA +/- Cyro	68 MORE 7 MORE ⊕◯◯◯ Very low	58 MORE 0 MORE ⊕⊕⊕◯ Moderate	58 MORE 3 MORE ⊕⊕◯◯ Low	7 MORE 0 MORE ⊕⊕⊕⊕ High		18 MORE 0 MORE ⊕◯◯◯ Very low	162 LESS 80 LESS ⊕⊕⊕◯ Moderate
VIA +/- LEEP	50 MORE 7 MORE ⊕◯◯◯ Very low	40 MORE 0 MORE ⊕⊕⊕◯ Moderate	40 MORE 3 MORE ⊕⊕◯◯ Low	11 LESS 0 MORE ⊕⊕⊕⊕ High	18 LESS 0 MORE ⊕⊕◯◯ Low		180 LESS 80 LESS ⊕⊕⊕◯ Moderate

www.gradepro.org

A new version of GRADEpro proudly engineered by: Evidence Prime
The tools for health care decisions

The official tool of GRADE and DECIDE

A new quality in guideline development

Brought to you by the creators of GRADEpro

Guideline Development Tool is an easy to use all-in-one web solution for summarizing and presenting information for healthcare decision making. It supports creating concise summary tables for systematic reviews and health technology assessments as well as facilitates development

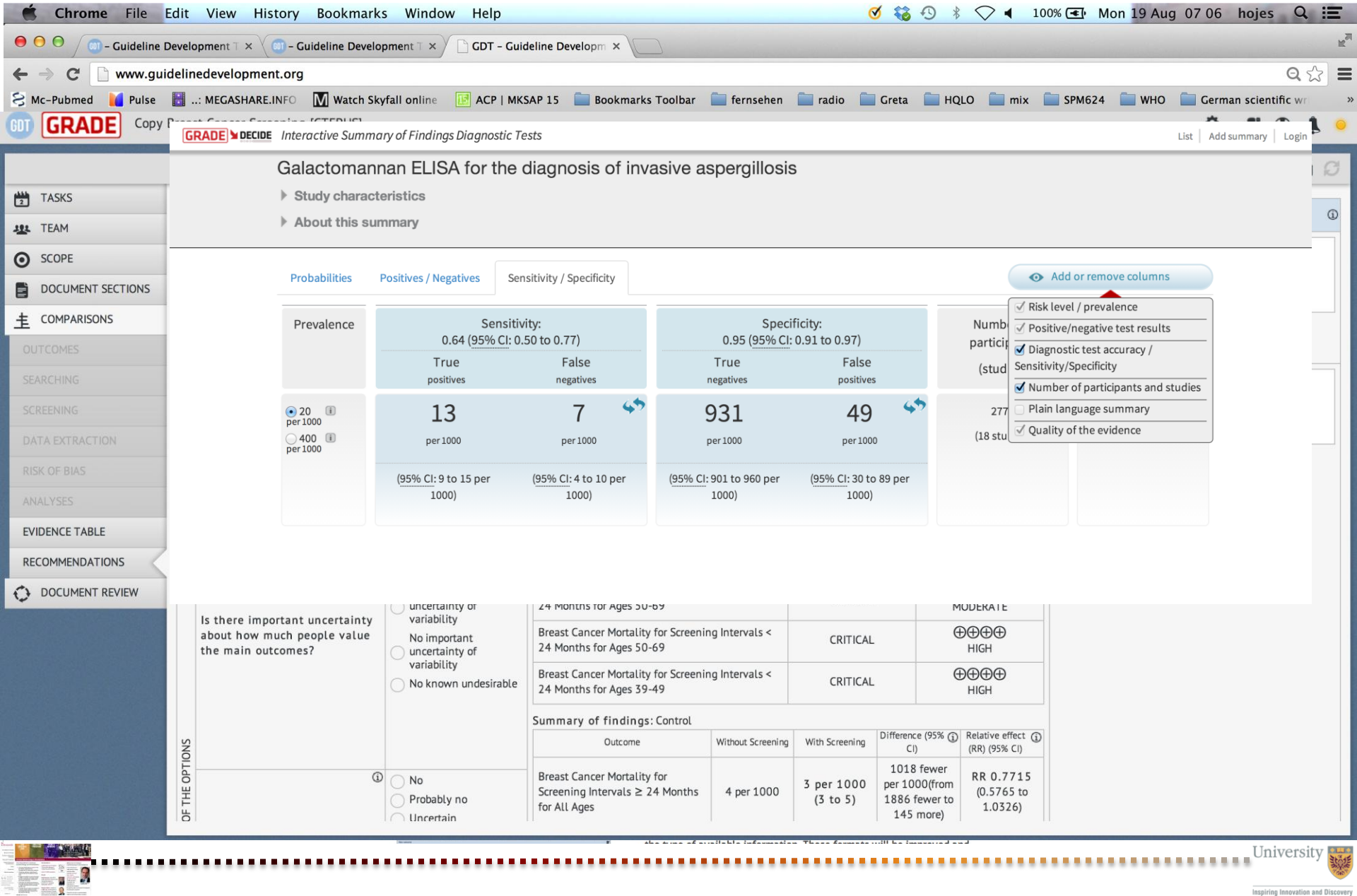
Best for both worlds

Guideline developers and authors of systematic reviews

Anyone preparing a systematic review will benefit from a simple online solution that assists creating summaries of evidence. Anyone developing clinical guidelines or other recommendations in healthcare will also benefit from a tool that smoothens the way of following a systematic and

McMaster University
Inspiring Innovation and Discovery

Tools


Chrome File Edit View History Bookmarks Window Help
www.guidelinedevelopment.org
GRADE
GRADE DECIDE Interactive Summary of Findings Diagnostic Tests
List | Add summary | Login
Galactomannan ELISA for the diagnosis of invasive aspergillosis
Study characteristics
About this summary
Probabilities
Positives / Negatives
Sensitivity / Specificity
Add or remove columns
Risk level / prevalence
Positive/negative test results
Diagnostic test accuracy / Sensitivity/Specificity
Number of participants and studies
Plain language summary
Quality of the evidence
Prevalence
Sensitivity: 0.64 (95% CI: 0.50 to 0.77)
Specificity: 0.95 (95% CI: 0.91 to 0.97)
True positives
False negatives
True negatives
False positives
20 per 1000
400 per 1000
13 per 1000
7 per 1000
931 per 1000
49 per 1000
(95% CI: 9 to 15 per 1000)
(95% CI: 4 to 10 per 1000)
(95% CI: 901 to 960 per 1000)
(95% CI: 30 to 89 per 1000)
TASKS
TEAM
SCOPE
DOCUMENT SECTIONS
COMPARISONS
OUTCOMES
SEARCHING
SCREENING
DATA EXTRACTION
RISK OF BIAS
ANALYSES
EVIDENCE TABLE
RECOMMENDATIONS
DOCUMENT REVIEW
Is there important uncertainty about how much people value the main outcomes?
No important uncertainty of variability
No known undesirable
No
Probably no
Uncertain
Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 50-69
CRITICAL
HIGH
Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 39-49
Summary of findings: Control
Outcome
Without Screening
With Screening
Difference (95% CI)
Relative effect (RR) (95% CI)
Breast Cancer Mortality for Screening Intervals ≥ 24 Months for All Ages
4 per 1000
3 per 1000 (3 to 5)
1018 fewer per 1000(from 1886 fewer to 145 more)
RR 0.7715 (0.5765 to 1.0326)

Why is this relevant for COMET

- Summary of Findings Tables report information for patient or population important outcomes
- Outcomes are determined by decision-makers (e.g. panels): over and over again

Formulate question

Select outcomes

Rate importance

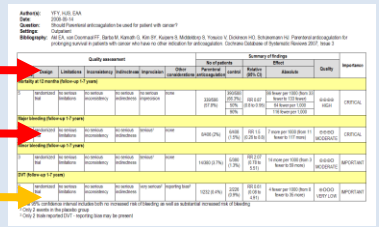
Outcomes across studies

Create evidence profile with GDT

Rate quality of evidence for each outcome

Randomization raises initial quality
RCTs: high
Observational: low

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



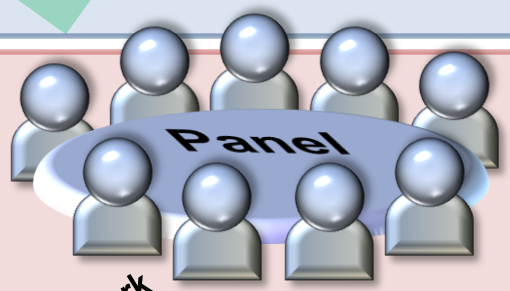
High
Moderate
Low
Very low

Summary of findings & estimate of effect for each outcome

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

Evidence synthesis



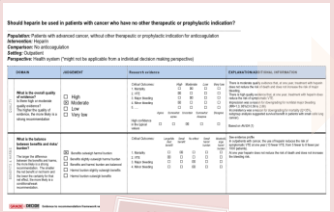
Recommendation

Grade recommendations (Evidence to Recommendation)
- For or against (direction) ↓↑
- Strong or conditional/weak (strength)

Grade overall quality of evidence across outcomes based on lowest quality of ***critical*** outcomes

EtD framework

By considering balance of consequences (evidence to recommendations):
- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences
- ❑ Feasibility, equity and acceptability
- ❑ Resource use (if applicable)



Guideline

Formulate Recommendations (↓↑ | ⊕…)
"The panel recommends that ….should..."
"The panel suggests that ….should..."
"The panel suggests to **not** ..."
"The panel recommends to **not**..."
Transparency, clear, actionable
Research?


longer life
fewer symptoms
fewer complications
better quality of life
surrogate
sensitivity
specificity

Why is this relevant for COMET

- Summary of Findings Tables report information for patient or population important outcomes
- Outcomes are determined by decision-makers (e.g. panels): over and over again
- But we wouldn't have to (if there was a core outcome set)

In summary

- SoF tables improve systematic review users' understanding and rapidity of grasping information
- Study showed this was using suboptimal formats – big improvement in formats with testing
- There is no one results presentation that fits all potential users' need – interactive SoF
- SoFs are essential and should be compulsory

Contact

Holger.Schunemann@McMaster.ca

@schunemann_mac