



Provincia Religiosa di San Pietro
 dell'Ordine Ospedaliero di San Giovanni di Dio
 Ospedale "Sacro Cuore di Gesù", Fatebenefratelli
 UOC PEDIATRIA-NEONATOLOGIA-UTIN

Radiologia

Pillola di EBM

I  Renè

La valutazione di un nuovo
 test diagnostico

Maria Carmen Verga

11 giugno 2016

**JOURNAL CLUB
 of Pediatrics
 in Benevento**
Prima Parte
 incontri monotematici
 ospedale - territorio
 per la elaborazione
 di linee guida comuni
Febbraio 2016
Marzo



BENEVENTO
 Centro Congressi
 Ospedale Fatebenefratelli di Benevento
 Via Napoli 14/a
 82100 Benevento

Responsabili Scientifici
 Iride Dello Iacono
 Maria Carmen Verga



...???

[APAL] La confusione, l'EBM e il campione alle Olimpiadi 1 - Unicode (UTF-8)

File Messaggio

Posta indesiderata Elimina Elimina

Rispondi a tutti Rispondi Inoltra Rispondi

Messaggio istantaneo

Aggiungi al calendario Sposta in Copia in Azioni

Contrassegno Segui

Copia Trova testo Codifica

Precedente Successivo Esplora

[APAL] La confusione, l'EBM e il campione alle Olimpiadi 1

vergasa@virgilio.it [APAL]

A: mailing@siaip.it; APAL;

Il messaggio ha priorità alta.

15

Confusione 1.

- Siamo bombardati da richiami all'appropriatezza, ci rimandano alle evidenze scientifiche, alle raccomandazioni delle LG
- La Corte di Cassazione ha condannato un medico che si era attenuto rigorosamente alle LG

Confusione 2.

- quando dobbiamo rispondere ad un quesito clinico dobbiamo cercare la risposta nelle evidenze scientifiche, partendo dalle sintesi delle evidenze come le Revisioni Sistematiche
- Molto spesso le Revisioni Sistematiche, partendo da decine di lavori, ne includono solo 1 o 2, magari li valutano come metodologicamente deboli e concludono che "...sono necessari ulteriori studi....ecc..."

Confusione 3.

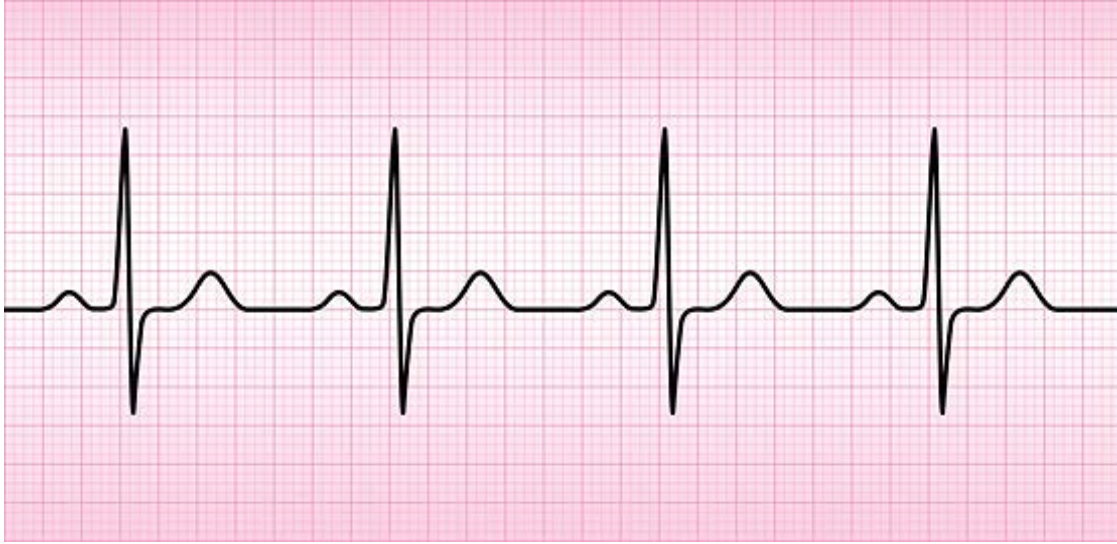
- L'EBM richiama, nella pratica clinica, a criteri di appropriatezza molto rigorosi
- quella terapia non è raccomandata perchè è poco efficace, ma io l'ho prescritta ed il mio paziente è stato bene.

Confusione 3 bis

L'EBM richiama a criteri di appropriatezza molto rigorosi

c. Io ed il mio collega abbiamo pazienti con lo stesso problema.

Siamo stati entrambi molto rigorosi, abbiamo entrambi tenuto conto delle stesse evidenze scientifiche, ma io sono arrivato ad una conclusione ed il mio collega ad un'altra ed alla fine abbiamo fatto 2 prescrizioni diverse....



Come ci siamo lasciati...

Technical Factors

Similar to many tests, the ECG is subject to issues related to consistency, reproducibility, and interobserver variability in interpretation. Indeed, performance of the ECG in a real-world mass screening setting will vary when readers and technicians

history and physical examination.^{11,12}

In the United States, the preparticipation screening process has lacked a measure of standardization, including in the selection of examiners, who have traditionally included

Negli USA l'organizzazione dei programmi di screening manca di misure di standardizzazione, anche sugli esaminatori (infermiere, MMG, specialista, spesso volontari)

Reliability of ECG Diagnostic Criteria

A major obstacle in screening is the overlap in measurements with the ECG between healthy subjects and those with

Un ostacolo importante nello screening è la sostanziale sovrapposizione di alcune misurazioni ECG (ampiezza, durata degli intervalli ecc...) tra soggetti sani e soggetti con patologie importanti



Come per altri test, anche per l'ECG bisogna considerare consistenza, riproducibilità e variabilità inter-operatori

as inherited channelopathies and cardiomyopathies is strongly dependent on test sensitivity, specificity, and positive and negative predictive values. However, even effective screening ECGs for the identification of diseases with low prevalence in the general population is not equivalent to identifying individuals who will actually experience clinically adverse outcomes, because the incidence of SD will be substantially lower than the prevalence of dis

disorders of function of predictive its screening ods can rel

ts, if it appou

L'accuratezza del test sulla popolazione generale in cui le patologie sono a bassa prevalenza non è la stessa che si ha su soggetti con sospetto diagnostico. Un PPV basso non basta per giudicare inutile uno screening se c'è utilità clinica ed il rapporto costo efficacia è vantaggioso

Sensitivity and Specificity (False-Positives and False-Negatives)

The sensitivity and specificity of the ECG for detection of cardiovascular disease will vary with the severity of phenotypic expression. In HCM, for example, specificity is reduced (and false-positive results increased) with only modestly increased

Sensibilità e specificità dell'ECG variano con la severità dell'espressione fenotipica

exceeding true-positive results.^{3,28,93,97} *However, even if ECGs with false-positive results could be reduced to only 5% the course of screening 10 million individuals (the estimated number of US competitive athletes), screening ECGs would nevertheless identify a formidable obstacle of 500 000 people who required further testing to exclude underlying heart disease and resolve eligibility for sports participants. Very few*

Anche se i falsi negativi fossero solo il 5%, se si screenassero 10.000.000 di individui avremmo 500.000 persone che sarebbero sottoposti ad esami di 2° livello

Con una prevalenza della patologia < 1%, forse < 0.1%, il PPV dell'ECG è *small* o *trivial*

With a prevalence of disease <1%, and perhaps <0.1% for even the most common channelopathies and inherited forms of cardiomyopathy (ie, HCM), the positive predictive value of the ECG is small or trivial. When the prevalence of rare causes of adverse outcome in young people is as low as 0.1%, the negative predictive value of the ECG nevertheless remains

detect (or raise suspicion) of those diseases that cause SD in young people and athletes.^{184,213} In 1 study, expert pediatric cardiologists blindly interpreted ECGs from young, nonathlete patients with HCM, LQTS, and other conditions.¹⁸⁴ False-positive rates in the range of $\approx 30\%$ were reported and were lowest for LQTS, Wolff-Parkinson-White, and HCM; 40% of the HCM patients were missed (ie, false-negatives) in the process. Indeed, in another study, correct classification of QT intervals from patients with or without LQTS was achieved by 96% of QT experts and 62% of arrhythmia experts but by <25% of general cardiologists and noncardiologists.²¹³

Race and Sex

It is becoming apparent that sex, race, and ethnicity are important determinants of pattern in the ECG.^{6,117,225-227,238-242} Notably,

La corretta classificazione dell'intervallo QT nei pazienti con e senza LQTS viene fatta dal

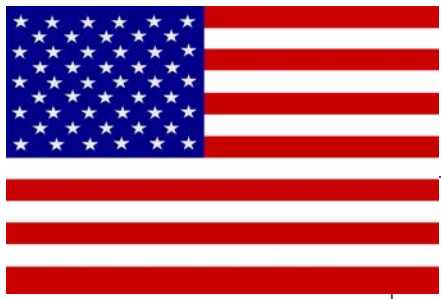
- 96% di esperti di QT
- 62% di esperti di aritmia
- < 25% di cardiologi generalisti e non cardiologi.

L'esperienza italiana

Tasso di mortalità prima dello screening (per 2 anni) e dopo lo screening (per 22 anni).

Riduzione da 4,3/100.000 py a 0.87/100.000 py >>> 55 morti/26 anni >>> 2/anno

Corrado et al⁴ compared mortality rates in athletes before the introduction of mandatory screening ECGs versus mortality evident after implementation of the program. The prescreening period was only 2 years in duration (1979–1981), whereas the postscreening period was much longer (22 years; 1982–2004). The authors reported an impressive 90% decline in SD rate, from 4.3 per 100 000 person-years in the prescreening period to 0.87 per 100 000 person-years with screening ECGs. This striking mortality reduction was attributed by the authors entirely to routine inclusion of the ECG into the screening process. However, these retrospective, nonrandomized data are based on a low event rate, with 55 cardiovascular deaths in 26 years (only 2 per year) as the numerator. The estimated denom-



United States Versus Italy: Global Considerations and Comparison of Prior Athlete Screening Demographics

Given the controversy generated over current screening strategies in Italy (as a proposed model) and in the United States, it is worth underscoring certain distinct differences between the 2 countries. **First**, the population of the United States is 5-fold larger than that of Italy (313 million versus 61 million) and 40 times that of Israel. **Furthermore**, the United States is 30 times larger in land mass than Italy, which presents an obstacle that impacts the practicality and durability of large-scale screening. **Second**, the US population is substantially more diverse ethnically and racially than that in Italy, a factor that may influence how one interprets patterns in ECGs in athletes. *for a variety of cardiovascular disease*

USA-Italia

- Popolazione USA 5 volte superiore di quella italiana
- Superficie degli USA è 30 volte maggiore
- Maggiori diversità etniche e razziali

Risoluzione dello scenario

- Le decisioni di politica sanitaria sono molto complesse, sono relative alla gravità ed alla prevalenza della patologia, all'utilità di una diagnosi precoce, alla disponibilità di un test di screening accurato, alla valutazione costo-efficacia
- Sono decisioni che richiedono competenze multidisciplinari
- Ad oggi, dai documenti pubblicati in Italia, lo screening ECGrafico non è supportato da robuste evidenze di appropriatezza

Elementi che influiscono sulla decisione clinica

- utilità di una diagnosi precoce,
- gravità e prevalenza della patologia,
- disponibilità di un test di screening accurato,
- valutazione costo-efficacia

E' quindi possibile che, pur partendo dalle stesse evidenze scientifiche, si arrivi a

decisioni diverse

TUTTE EBM-BASED!

Confusione 3 bis

L'EBM richiama a criteri di appropriatezza molto rigorosi

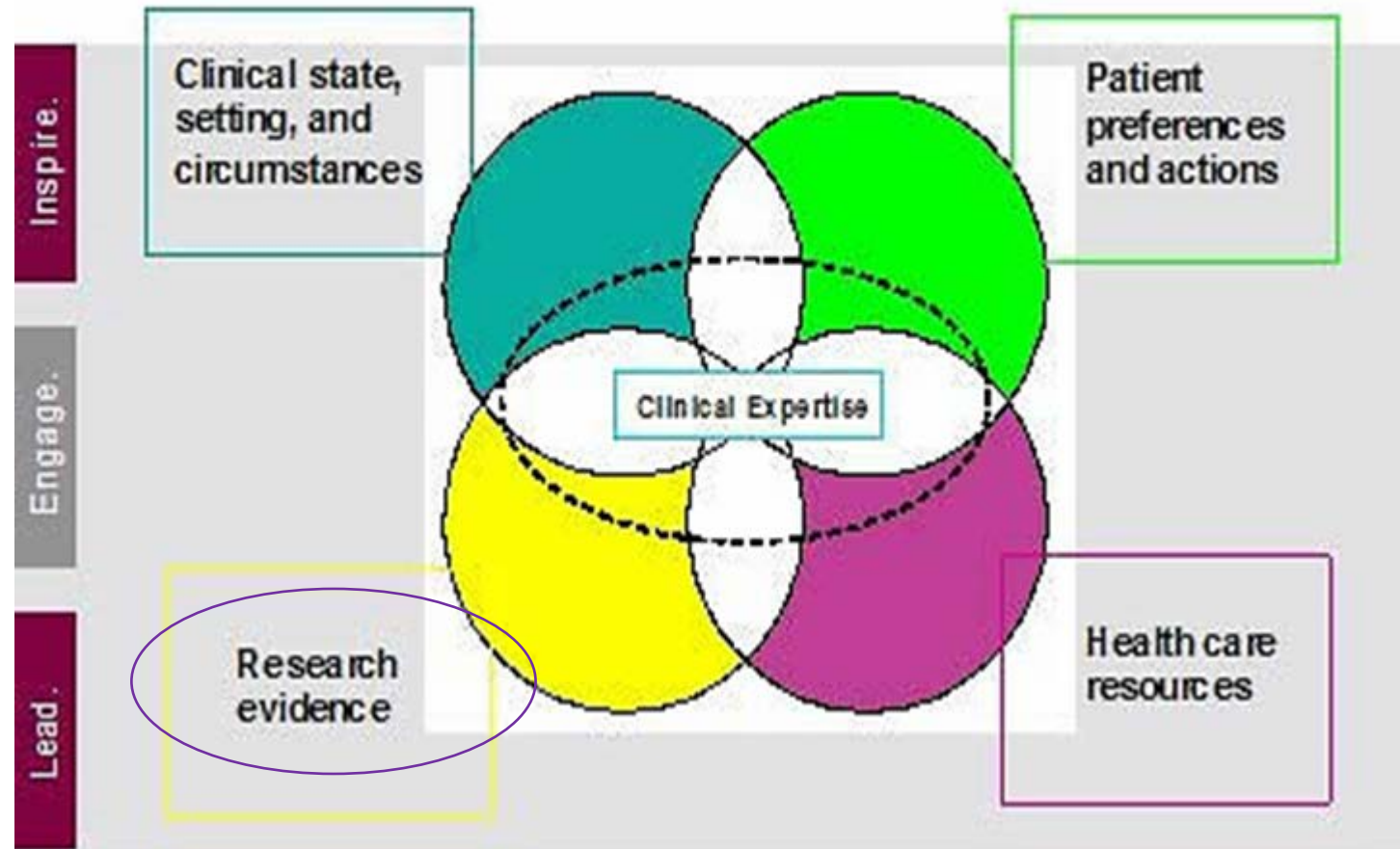
c. Io ed il mio collega abbiamo pazienti con lo stesso problema.

Siamo stati entrambi molto rigorosi, abbiamo entrambi tenuto conto delle stesse evidenze scientifiche, ma io sono arrivato ad una conclusione ed il mio collega ad un'altra ed alla fine abbiamo fatto 2 prescrizioni diverse....

...ma allora in che cosa consiste l'uniformità dei comportamenti a cui si richiama l'EBM???

Le evidenze scientifiche sono solo 1 degli elementi da considerare nel processo di decisione clinica

Clinical Decision Making



Research
evidence

Decisione 1

Patient
preferenc es
and actions

Clinical state,
setting, and
circumstances

Health care
resourc es

Decisione 2

Patient
preferenc es
and actions

Clinical state,
setting, and
circumstances

Health care
resourc es

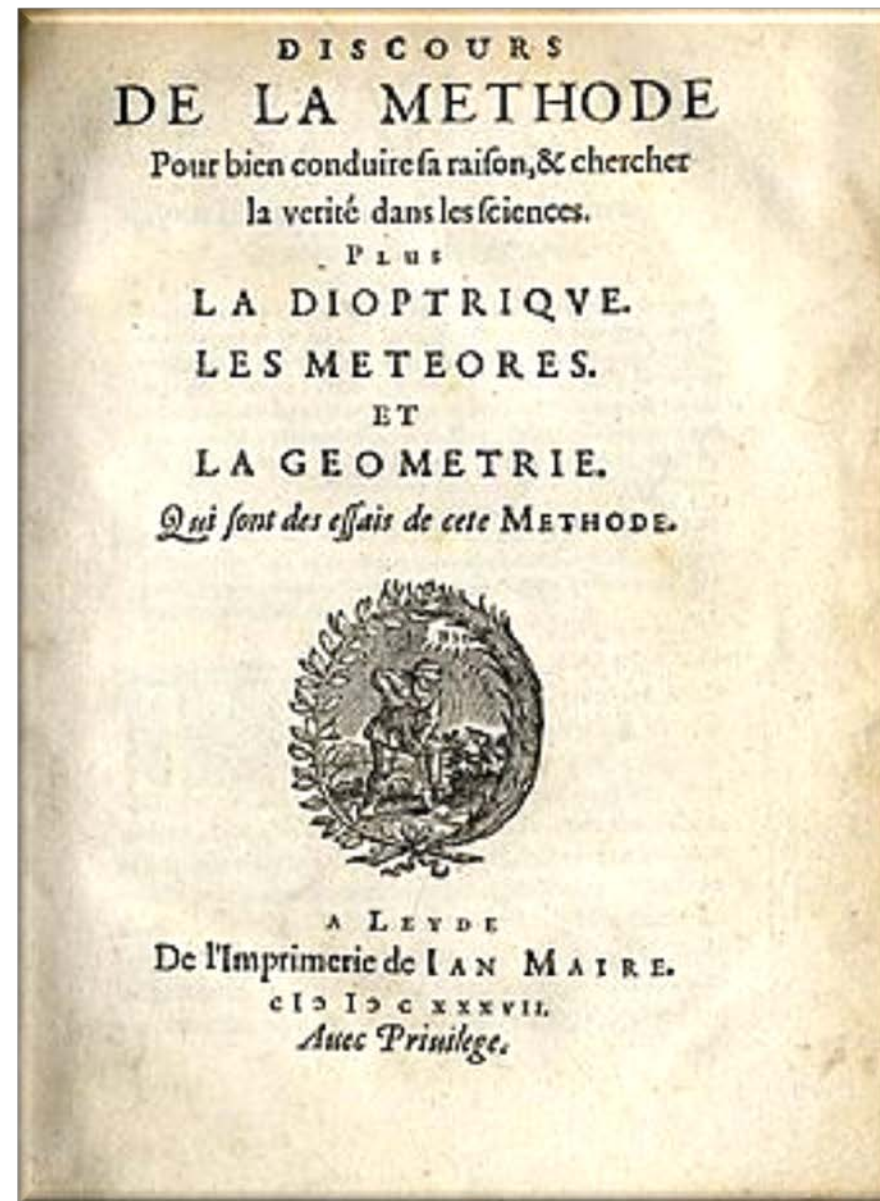
Decisione 3

Patient
preferenc es
and actions

Clinical state,
setting, and
circumstances

Health care
resourc es

L'uniformità è nel metodo, nel processo decisionale che seguiamo





RCT sull'efficacia del probiotico X (prezzo = € 20)
nel ridurre la durata della diarrea (N° di giorni di malattia)

Risultati: nel gruppo trattato con il probiotico X la durata della diarrea era inferiore del 20% (I.C. 5-40%) rispetto al gruppo controllo. Differenza statisticamente significativa

Noi ci armiamo delle nostre conoscenze di EBM, analizziamo criticamente il lavoro e scopriamo che la diarrea era proprio banale, durava mediamente 5 giorni e che la riduzione del 20% corrispondeva quindi ad 1 giorno in meno.

Decisione 1

Se siamo estensori di una linea guida sulla diarrea NON raccomandiamo la somministrazione del probiotico X,



se siamo **clinici** informiamo i nostri pazienti che non ci sono sostanziali benefici dall'assumere il probiotico X spendendo € 20.

...però...però...



Se il nostro paziente è un campione che sta partecipando alle Olimpiadi,



e le Olimpiadi sono il sogno della sua vita per il quale si sta preparando da anni, e può perdere la medaglia per una minima imperfezione, ed in quel momento anche mezza giornata di allenamento in meno può compromettere 4 anni di lavoro...



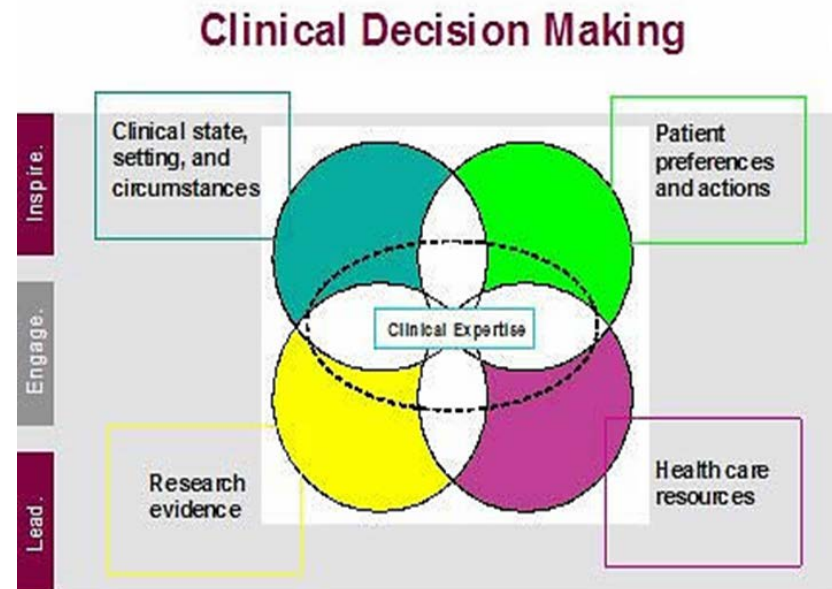
...risparmiare mezza giornata di malattia, anche banale, è per quello specifico paziente in quel momento, importantissimo.

Il nostro campione, in quel momento, sarebbe disposto a pagare anche € 200 o € 1.000 pur di guadagnare solo mezza giornata di allenamenti.



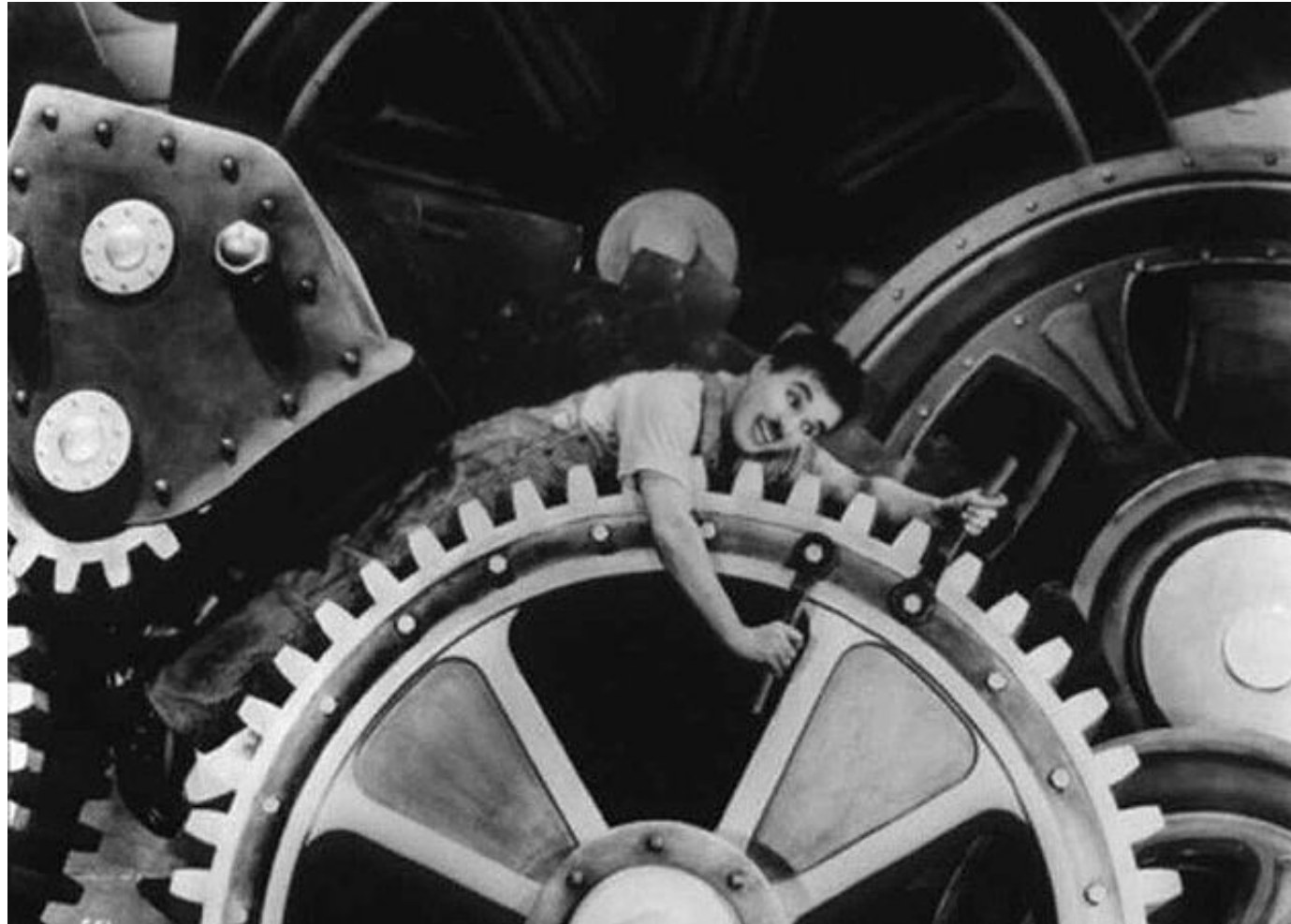
Decisione 2

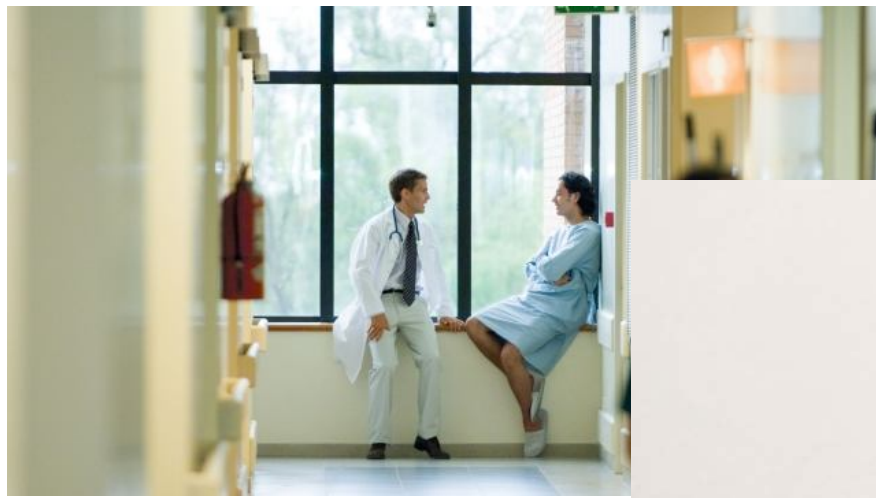
Ecco perchè il medico, proprio attenendosi ai principi dell'EBM, deve informare il paziente di questa possibilità, deve farlo secondo scienza e coscienza, conoscendo bene la reale efficacia e tutti gli eventuali eventi avversi.



...Metodo...Metodologia...

Troppo tecnicismo?



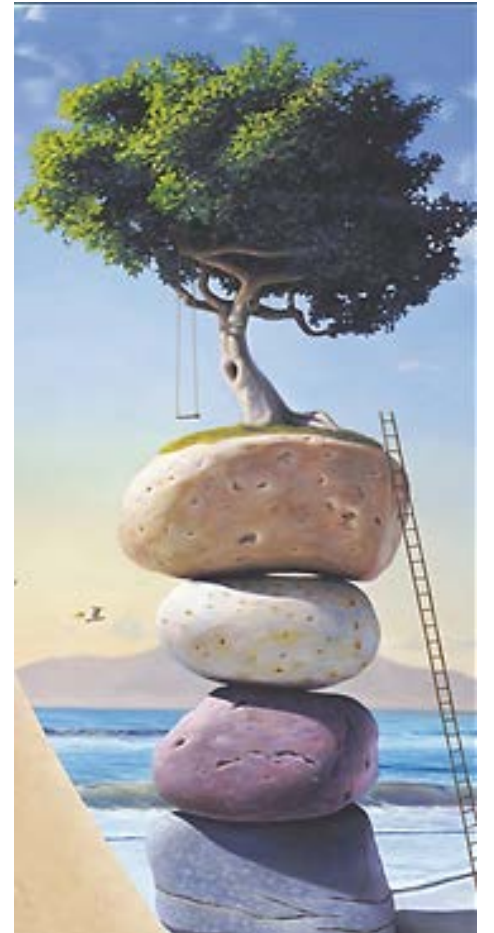


PRATICHE NARRATIVE E SALUTE

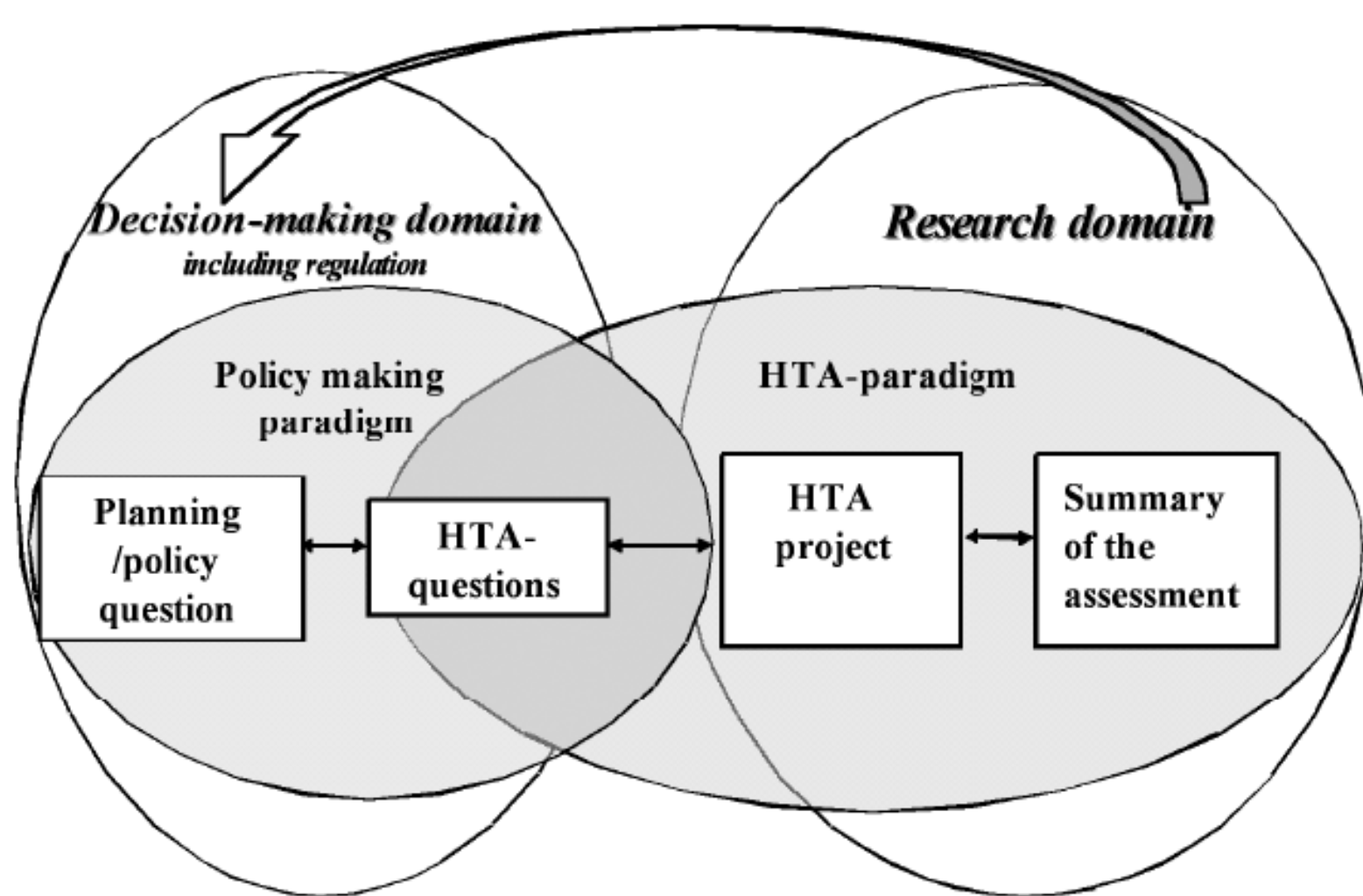
MEDICINA
NARRATIVA

CARING
NARRATIVES

STORYTELLING



I bioeticisti e l'Health Technology Assessment



L'HTA è la disciplina tesa a valutare le condizioni di realizzazione e gli esiti di scelte alternative.

Ha il fine di supportare razionalmente le decisioni di politica sanitaria



Dimensions of HTA

Gli approcci per la valutazione etica



Sulla base di un'indagine compiuta dal Network europeo di HTA (EUnetHTA), gli approcci più utilizzati nel panorama internazionale per realizzare le valutazioni etiche sono:

- Casuistry (Casuistica)
- Coherence analysis (Analisi di coerenza)
- Interactive, participatory HTA approach (Approccio interattivo-partecipativo)
- Principlism (Principialismo)
- Social shaping of technology (Modellamento sociale della tecnologia)
- Wide reflective equilibrium (Ampio equilibrio riflessivo)
- Approcci locali: (Approccio eclettico del FINOHTA; Analisi di valore del NKCHC; **Modello "triangolare"** dell'Istituto di Bioetica dell'UCSC)

Il processo di valutazione etico secondo l'EUnetHTA Project

Per ovviare al problema rappresentato dalla **eterogeneità** degli approcci e favorire lo **scambio di informazioni** tra i diversi paesi, l'EUnetHTA Project ha tentato di fornire una **base comune operativa** (senza la predilezione di nessuno tra gli approcci teorici individuati) al processo d'analisi etica, elaborando una sorta di **modello standard**



Criticità

- Necessità di **perfezionare la metodologia d'analisi etica** in modo da rendere più semplice lo scambio di informazioni e confrontabili i risultati
- Scarso **monitoraggio dell'effetto prodotto dalle valutazioni** (ma ciò riguarda tutto l'HTA!): oggi è difficile comprendere la "reale" influenza, che questi studi hanno sui decisori e sulle politiche

La valutazione di un nuovo test diagnostico

Ecografia polmonare

Esposito et al. *Italian Journal of Pediatrics* 2014, 40:37
<http://www.ijponline.net/content/40/1/37>



ITALIAN JOURNAL
OF PEDIATRICS

RESEARCH

Open Access

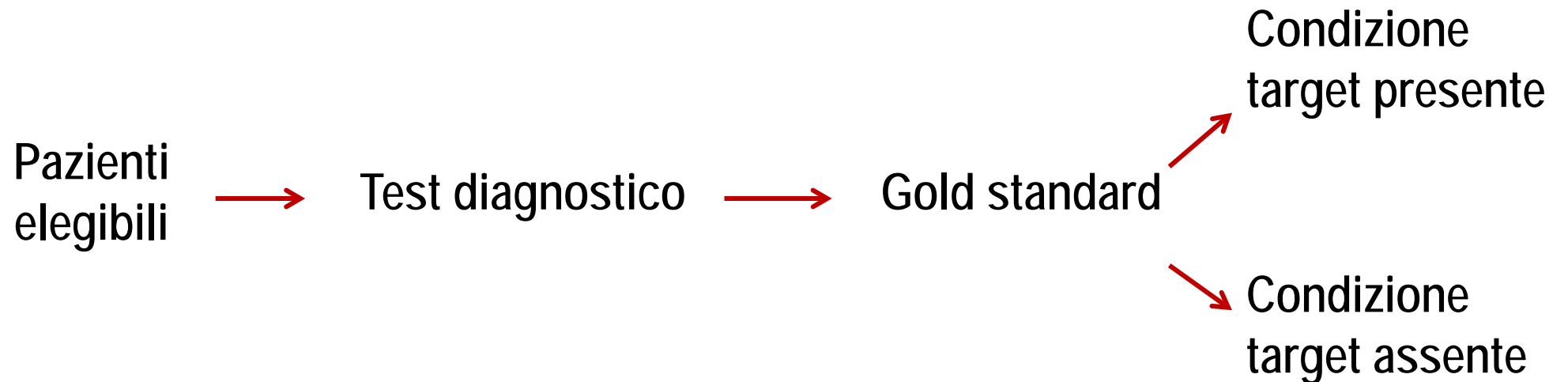
Performance of lung ultrasonography in children with community-acquired pneumonia

Susanna Esposito^{1*}, Simone Sferrazza Papa¹, Irene Borzani², Raffaella Pinzani¹, Caterina Giannitto², Dario Consonni³ and Nicola Principi¹

Test diagnostico

Determinare la capacità di un test diagnostico nel discriminare i soggetti che hanno o non hanno la patologia o la condizione *target*

Studio di accuratezza diagnostica, trasversale, *cross-sectional*



Validità di un test

Capacità di classificare correttamente sia gli individui malati che quelli sani, cioè capacità di generare risultati rispondenti al vero sia negli individui ammalati che in quelli sani.

La validità può essere calcolata facilmente qualora si conosca il vero stato degli individui che sono stati sottoposti al test.

		GOLDEN TEST (REALTA')	
		+	-
T+	VERO a	FALSO b	
T-	FALSO c	VERO d	

$$\text{Validità} = (a+d)/(a+b+c+d).$$

VALIDITA'
di un test
(o efficienza,
o accuratezza)

$$\frac{a + d}{a + b + c + d}$$

Il «problema Gold Standard»

La valutazione di un test diagnostico, oltre alle altre problematiche già note, è complicata, in alcune circostanze, dalla mancanza o dalla difficoltà di poter disporre di un Gold Standard.

1. Non è sempre possibile definire in modo netto degli standard di riferimento (es., nel caso dei disturbi psichiatrici)
2. Attenzione ad accettare in modo acritico gli standard di riferimento, anche quando si basano sulle interpretazioni esperte delle biopsie!
3. Pazienti negativi al test diagnostico >>> difficile sottoporli al GS di riferimento quando quest'ultimo è molto invasivo o comporta dei rischi (es. l'angiografia).

Concordanza fra due test

Quando si tratta di valutare la *performance* di un test, talvolta può essere necessario confrontarlo non con l'esito di golden test, bensì con un altro test, magari non eccellente ma di comune impiego nella pratica.

In questo caso, non si parla più di validità, ma di «concordanza».

La concordanza può riguardare

1. il grado di accordo che si osserva fra due test,
2. quello fra due (o più) operatori che interpretano l'esito di *uno stesso test* (es. radiografie, elettrocardiogramma, auscultazione cardiaca ecc.),
3. due letture effettuate da uno stesso operatore in tempi diversi.

Serve per stabilire se i criteri utilizzati per l'interpretazione del test siano efficienti e se la classificazione sia riproducibile.

		Test A	
		+	-
Test B	+	a	b
	-	c	d

CONCORDANZA
fra due test

$$\frac{a + d}{a + b + c + d}$$

Non tiene conto della quota di concordanza dovuta al caso >>> Calcolo dell'indice **Kappa di Cohen**.

- K < 0.2 = concordanza scarsa;
- K = 0.2 - 0.4 = concordanza modesta;
- K = 0.41 - 0.61 = concordanza moderata
- K = 0.61 - 0.80 = concordanza buona
- K > 0.80 = concordanza eccellente

Esiste un Gold Standard di riferimento?



Sì

Si calcola **SENSIBILITA'** e **SPECIFICITA'**
del test vs. il GS



No

Si calcola la misura K di **CONCORDANZA**
introdotta da Cohen

La validità di un test diagnostico: l'analisi dell'accuratezza

http://www.bris.ac.uk/quadas/ Bristol University | QUADAS... x

skip to content university home | study | research | global | contacting people | a-z index | news | help | use of cookies search

University of BRISTOL QUADAS

QUADAS-2

QUADAS home | Contacts | Training

A quality assessment tool for diagnostic accuracy studies

The development of QUADAS-2 was led by a team based at the [School of Social and Community Medicine](#) at the University of Bristol

[More about the current version of QUADAS](#)

QUADAS-2

- [QUADAS-2 tool](#)
- [Background document](#) (PDF, 396 kB)
- [Team members](#)

More about [QUADAS-2](#)»

History

History of the development of the QUADAS tool

More about [the history of QUADAS](#)»

Resources

Downloadable documents include:

- [Templates and Access database](#)
- [Publications](#)

See all [QUADAS resources](#)»

Training

There are no QUADAS training events planned.

More about [QUADAS training](#)»

COLLABORATING INSTITUTIONS

QUADAS home

Updated 21 June 2013 by the School of Social and Community Medicine | Feedback
University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK. Tel: +44 (0)117 928 7204

6:17 6/2013

Risk of bias and applicability judgments

The following table summarises QUADAS-2 and lists all signalling, risk of bias and applicability rating questions.

DOMAIN	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING
<u>Description</u>	<u>Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):</u>	<u>Describe the index test and how it was conducted and interpreted:</u>	<u>Describe the reference standard and how it was conducted and interpreted:</u>	<u>Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:</u>
<u>Signalling questions(yes/no/unclear)</u>	<u>Was a consecutive or random sample of patients enrolled?</u>	<u>Were the index test results interpreted without knowledge of the results of the reference standard?</u>	<u>Is the reference standard likely to correctly classify the target condition?</u>	<u>Was there an appropriate interval between index test(s) and reference standard?</u>
	<u>Was a case-control design avoided?</u>	<u>If a threshold was used, was it pre-specified?</u>	<u>Were the reference standard results interpreted without knowledge of the results of the index test?</u>	<u>Did all patients receive a reference standard?</u>
	<u>Did the study avoid inappropriate exclusions?</u>			<u>Did all patients receive the same reference standard?</u> <u>Were all patients included in the analysis?</u>

1.	Le caratteristiche della popolazione dei partecipanti rappresentano quelle dei soggetti che nella pratica saranno soggetti al test?	
----	---	--

1. Il test diagnostico deve essere stato valutato in un gruppo appropriato di pazienti, simili al tipo di popolazione assistita in cui l' esame può essere utilizzato.

2. La popolazione deve essere varia, con diversi livelli di gravità del disturbo (sia in forma lieve che grave, sia in uno stadio precoce che avanzato, sia pazienti trattati che non trattati).

Inoltre, particolare di non secondaria importanza,

3. il test *deve essere stato somministrato a pazienti "sicuramente" senza la malattia* da diagnosticare (o la complicanza) e a pazienti che hanno disturbi che spesso vengono confusi con quello in studio ma che in realtà non hanno la malattia in questione.

Stefano Miceli Sopo

DOMAIN	PATIENT SELECTION
<u>Description</u>	<u>Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):</u>
<u>Signalling questions(yes/no/unclear)</u>	<u>Was a consecutive or random sample of patients enrolled?</u>
	<u>Was a case-control design avoided?</u>
	<u>Did the study avoid inappropriate exclusions?</u>

All of the otherwise healthy children born at term aged between 1 month and 14 years admitted with fever (i.e. an axillary temperature of >38°C) and signs and symptoms consistent with CAP (i.e. cough, tachypnea, dyspnea or respiratory distress, and breathing with grunting or wheezing sounds with rales) and hospitalized in our pediatric ward were considered eligible for inclusion. The patients

- Reclutamento consecutivo
- Bambini tra 1 e 14 anni con segni e sintomi suggestivi di polmonite

Tutto ok?





Mancano pazienti asintomatici e con sintomatologia meno grave

Come faccio a testarlo per i veri negativi ed i falsi negativi?

DOMAIN	INDEX TEST
<u>Description</u>	<u>Describe the index test and how it was conducted and interpreted:</u>
<u>Signalling questions</u> (yes/no/ <u>unclear</u>)	<u>Were the index test results interpreted without knowledge of the results of the reference standard?</u>
	<u>If a threshold was used, was it pre-specified?</u>

on imaging session of normal models. The lung US was carried out using a MyLab™25 Gold (Esaote, Genoa, Italy) with a convex 2.5-6.6 MHz probe and a linear 7.5-12 MHz probe. The probes were placed perpendicularly, ob-

The physician performing lung US was always the same (SSP), was a resident in paediatrics and had limited experience in US prior to the study. He had no contact with the radiologist, and was blinded to the CR findings.

Il medico che interpreta il test indice è in cieco rispetto al referto della RX

Tutto ok?





DOMAIN	REFERENCE STANDARD
<u>Description</u>	<u>Describe the reference standard and how it was conducted and interpreted:</u>
<u>Signalling questions (yes/no/unclear)</u>	<u>Is the reference standard likely to correctly classify the target condition?</u>
	<u>Were the reference standard results interpreted without knowledge of the results of the index test?</u>

underwent CR, both postero-anterior and lateral views as recommended in recently published guidelines [6]. The radiographs were evaluated by an independent expert radiologist (IB), who classified the findings as alveolar pneumonia, non-alveolar (interstitial) pneumonia, mixed interstitial and alveolar pneumonia, or no pneumonia in accordance with the World Health Organisation (WHO) criteria for the standardised interpretation of pediatric chest radiographs for the diagnosis of pneumonia [18]. Immediately

Refertazione in base ai criteri del WHO
Il Gold standard per la polmonite è la TAC

after CR, in all of the patients lung US evaluation of the lungs was performed.

L'ecografia viene fatta dopo la RX

Tutto ok?





DOMAIN	FLOW AND TIMING
<u>Description</u>	Describe any patients who did not receive the <u>index test(s)</u> and/or <u>reference standard</u> or who were <u>excluded</u> from the <u>2x2 table</u> (refer to <u>flow diagram</u>): Describe the <u>time interval</u> and <u>any interventions</u> between <u>index test(s)</u> and <u>reference standard</u> :
<u>Signalling questions</u> (yes/no/ <u>unclear</u>)	<u>Was there an appropriate interval between index test(s) and reference standard?</u>
	<u>Did all patients receive a reference standard?</u>
	<u>Did all patients receive the same reference standard?</u>
	<u>Were all patients included in the analysis?</u>

graphs for the diagnosis of pneumonia [18]. Immediately after CR, in all of the patients lung US evaluation of the lungs was performed.

Forty-eight of the 103 enrolled children (56 males, 54.4%; mean age ± standard deviation [SD] 5.6 ± 4.6 years) had radiographically confirmed CAP. Table 1 shows the sensi-

L'articolo non cita perdite al follow-up

Tutto ok?





Definizione dello studio

This prospective observational study was carried out in Pediatric Highly Intensive Care Unit of the Department of Pediatric Intensive Care, University of Padua.

Analisi statistica

to published guidelines [5-7]. The concordance of US and CR in identifying specific types of CAP (i.e. alveolar, mixed, non-alveolar) and pleural effusion was evaluated using Cohen's weighted kappa (k) statistics [20]. The discrepancies between US and CR were assessed using marginal homogeneity Stuart-Maxwell statistics [21,22], which

Tutto ok?





Definizione dello studio

Gli studi di accuratezza diagnostica sono sempre trasversali o cross-sectional



Analisi statistica

La pubblicazione degli studi di accuratezza diagnostica

The *Standards for Reporting of Diagnostic Accuracy* (STARD)

Clinical Chemistry 49:1
1-6 (2003)

STARD Initiative

Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative

PATRICK M. BOSSUYT,^{1*} JOHANNES B. REITSMA,¹ DAVID E. BRUNS,^{2,3}
CONSTANTINE A. GATSONIS,⁴ PAUL P. GLASZIOU,⁵ LES M. IRWIG,⁶ JEROEN G. LIJMER,¹
DAVID MOHER,⁷ DRUMMOND RENNIE,^{8,9} and HENRICA C.W. DE VET,¹⁰ FOR THE STARD GROUP

STARD

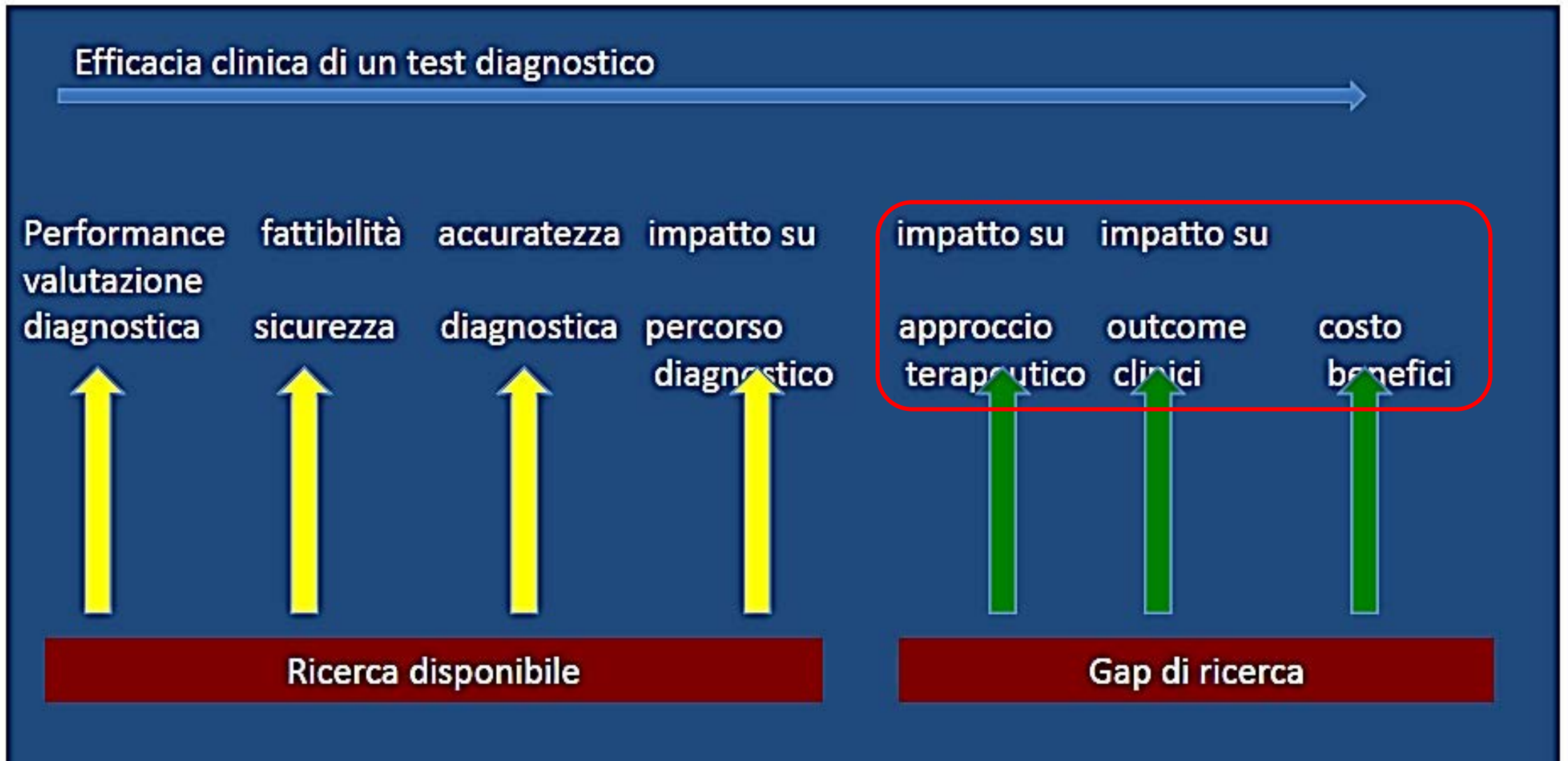
checklist per la pubblicazione degli studi di accuratezza diagnostica.

Prima versione ufficiale gennaio 2013

Sezione e Topic	Item n°	
TITOLO/ABSTRACT/ PAROLE CHIAVE	1	Identificare l'articolo come studio di accuratezza diagnostica (termini MeSH raccomandati "sensibilità e specificità")
INTRODUZIONE	2	Riportare i quesiti o lo scopo dello studio, come stimare l'accuratezza diagnostica , o confrontare l'accuratezza tra tests o tra gruppi di partecipanti
METODI		
Partecipanti	3	Descrivere la popolazione dello studio : criteri di inclusione e di esclusione, setting e luogo dove sono stati raccolti i dati
	4	Descrivere il reclutamento dei partecipanti : è stato basato sui sintomi presenti, sul risultato di precedenti tests o sul fatto che i partecipanti hanno ricevuto degli indici di test o dei riferimenti standard?
	5	Descrivere il campione dei partecipanti : la popolazione dello studio è stata definita secondo i criteri 3 e 4? Se no, specificare con quale ulteriore modalità i pazienti sono stati selezionati.
	6	Descrivere la raccolta dei dati : è stata pianificata prima di eseguire i test indici ed i riferimenti standard (studio prospettico) o dopo (studio retrospettivo)?
Metodi - test	7	Descrivere i riferimenti standard ed il loro razionale
	8	Descrivere i particolari tecnici di materiali e metodi usati , compresi come e quando le misurazioni sono state fatte e/o citare i riferimenti per i test indice e per le referenze standard
	9	Descrivere le definizioni ed il razionale per le unità, i cut-off e/o le categorie di risultati per i test indice ed i riferimenti standard
	10	Descrivere il numero, la formazione e l'esperienza delle persone che eseguono e leggono i test indice e i riferimenti standard
	11	Descrivere se chi legge i test indice ed i riferimenti standard siano stati o no in cieco rispetto ai risultati dell'altro teste descrivere qualunque altra informazione clinica fosse disponibile ai lettori
Metodi statistici	12	Descrivere i metodi per calcolare o confrontare le misure di accuratezza diagnostica ed i metodi statistici usati per quantificare le probabilità (es. 95% Intervallo di Confidenza)

RISULTATI		
Partecipanti	14	Riportare dove lo studio è stato fatto , incluse le date di inizio e fine del reclutamento
	15	Riportare le caratteristiche cliniche e demografiche della popolazione dello studio (es. età sesso, spettro dei sintomi presentati, comorbidità, terapie in atto, centri di reclutamento)
	16	Riportare il n° di partecipanti che hanno soddisfatto i criteri di inclusione , sottoposti o meno ai tests indice e/o non a quello di riferimento standard; descrivere perché i partecipanti non sono stati sottoposti ad entrambi i tests (è fortemente raccomandato un diagramma di flusso)
Risultato dei tests	17	Riportare l' intervallo di tempo dai tests indice ai riferimenti standard e qualunque trattamento somministrato tra i due
	18	Riportare la distribuzione di severità della patologia (definire i criteri) nei pazienti con la condizione target; le altre diagnosi nei pazienti senza la condizione target.
	19	Riportare la tabella con i risultati del test indice (compresi i risultati indeterminati e mancanti) incrociati con il risultato del riferimento standard ; per risultati continui, riportare la distribuzione dei risultati del test indice con quella dei risultati del riferimento standard.
	20	Riportare qualunque evento avverso durante l'esecuzione del test indice e del riferimento standard
Stime	21	Riportare le stime di accuratezza diagnostica e le misure di incertezza statistica (es. intervallo di confidenza 95%)
	22	Riportare come i risultati indeterminati, mancanti ed al di fuori del test indice sono stati gestiti
	23	Riportare le stime di variabilità dell'accuratezza diagnostica tra sottogruppi di partecipanti, lettori o centri, se fatto.
	24	Riportare le stime della riproducibilità del test , se fatto
DISCUSSIONE	25	Discutere l' applicabilità clinica dei risultati dello studio

L'analisi dell'appropriatezza



Grazie

