

LLISTES GUIA PER A PUBLICACIONS

Llista STARD per a
estudis d'avaluació de
proves diagnòstiques

Barcelona, 28 d'octubre de 2008

Ítems de la llista STARD

1. Títol i resum - paraules clau
2. Introducció
- Mètodes - Participants
 3. Població d'estudi
 4. Reclutament
 5. Mostreig
 6. Obtenció de les dades

- Mètodes – Diagnòstics
 7. Prova de referència
 8. Especificacions tècniques
 9. Unitats i punts de tall
 10. Experiència professional
 11. Emmascarament
- Mètodes – Anàlisi
 12. Error aleatori
 13. Consistència
- Resultats
- Discussió

1 – Títol, resum i paraules clau

- Es recomana utilitzar *diagnostic accuracy* en el títol o el resum d'una publicació que compari
 - els resultats d'una o més proves d'estudi amb
 - els resultats d'una prova de referència
- Es recomana utilitzar “Sensitivity and specificity” com una paraula clau (és el descriptor del Medical Subject Headings)
- Justificació:
 - Facilita la recuperació de l'article

Exactitud (“diagnostic accuracy”)

		Resultat de la prova d'estudi		
		Present	Absent	
Resultat de la prova d'estudi	Positiva	VP	FP	VP+FP
	Negativa	FN	VN	FN+VN
		VP+FN	FP+VN	

$$\text{Sensibilitat} = \frac{VP}{VP + FN}$$

$$\text{Especificitat} = \frac{VN}{VN + FP}$$

Exactitud de la tira reactiva per al diagnòstic de la infecció urinària

Urocultiu

Tira reactiva	Positiu	Negatiu	
Positiva	86	433	519
Negativa	20	548	568
	106	981	1087

$$\text{Sens} = \frac{86}{106} = 81,13\% \quad \text{Esp} = \frac{548}{981} = 55,86\%$$

Multidetector computed tomography for acute pulmonary embolism

Stein PD, Fowler SE, Goodman LR, *et al*; PIOPED II Investigators.
N Engl J Med. 2006; 354(22): 2317-27.

BACKGROUND: The **accuracy** of multidetector computed tomographic angiography (CTA) for the diagnosis of acute pulmonary embolism has not been determined conclusively.

METHODS: The Prospective Investigation of Pulmonary Embolism Diagnosis II trial was a prospective, multicenter investigation of the **accuracy** of multidetector CTA alone and combined with venous-phase imaging (CTA-CTV) **for the diagnosis** of acute pulmonary embolism. We used a composite reference test to confirm or rule out the diagnosis of pulmonary embolism.

RESULTS: Among 824 patients with a reference diagnosis and a completed CT study, CTA was inconclusive in 51 because of poor image quality. Excluding such inconclusive studies, the sensitivity of CTA was 83 percent and the specificity was 96 percent. [...].

CONCLUSIONS: In patients with suspected pulmonary embolism, multidetector CTA-CTV has a higher diagnostic sensitivity than does CTA alone, with similar specificity [...].

Prospective study of the **diagnostic accuracy** of the Simplify D-dimer assay for pulmonary embolism in emergency department patients.

Kline JA, Runyon MS, Webb WB, Jones AE, Mitchell AM.

Chest. 2006; 129(6): 1417-23

OBJECTIVE: To determine if a d-dimer assay (Simplify D-dimer; Agen Biomedical; Brisbane, Australia) can reliably exclude pulmonary embolism (PE) by producing a posttest probability of PE < 1% in low-risk, symptomatic emergency department (ED) patients.

METHODS: Hemodynamically stable patients were evaluated for PE using a structured d-dimer-centered protocol; d-dimer testing was performed prior to imaging. Prior to testing, physicians completed an electronic data form that included their unstructured clinical estimate for the pretest probability of PE (< 15%, 15 to 40%, or > 40%) [...].

RESULTS: We enrolled 2,302 patients (mean age, 45 +/- 16 years [+/- SD]; 31% male); 108 patients received a diagnosis of PE (4.7%; 95% confidence interval [CI], 3.6 to 5.6%). The overall **sensitivity** and **specificity** of the d-dimer assay were 80.6% (95% CI, 71.8 to 87.5%) and 72.5% (95% CI, 70.6 to 74.4%), respectively. [...].

CONCLUSIONS: The Simplify D-dimer assay had moderate sensitivity and relatively high specificity for PE in low-risk ED patients. The combination of a physician's unstructured estimate of pretest probability of PE of < 15% and a negative d-dimer result produced a posttest probability of PE of 0.7% (95% CI, 0.3 to 1.4%).

KEY WORDS: d-dimer, diagnosis, pulmonary embolism, **sensitivity and specificity**, thromboembolism

2 – Introducció

- En la introducció d'un article cal
 - Descriure els **antecedents** i l'**estat actual** del tema,
 - Especificar els **dubtes** que persisteixen
 - Derivar-ne els **objectius** de l'estudi
- Les **preguntes** i **objectius** han d'ésser clars i explícits
- Exemples:
 - **Estimar** l'exactitud diagnòstica
 - **Comparar** l'exactitud entre proves
 - Comparar l'exactitud entre els grups inclosos.

CT venography and compression sonography are **diagnostically equivalent**: data from **PIOPED II**.

Goodman LR, Stein PD, Matta F, et al.

AJR Am J Roentgenol. 2007; 189: 1071-6

Although there are **some data suggesting diagnostic equivalence**, no large prospective database with uniform diagnostic criteria has been used to evaluate this issue.

This article evaluates **whether CTV and venous sonography are equivalent** in the diagnostic valuation of patients with suspected acute pulmonary embolism (PE) in a large, multiinstitutional prospective study.

Mètodes: participants

3. Població de l'estudi
4. Reclutament
5. Mostreig
6. Recollida de dades

3 - Població d'estudi

- Cal descriure els **critèris d'inclusió** i **d'exclusió** i l'**àmbit** i lloc on es van obtenir les dades
 - El **valor predictiu** d'una prova depèn del context clínic on s'aplica
 - L'exactitud diagnòstica fa referència als resultats d'una prova en **un context clínic** determinat
 - Per tant, cal conèixer la **població diana** de l'estudi

Càlcul del valor predictiu

		Malaltia		
		Present	Absent	
Resultat de la prova estudiada	Positiva	VP	FP	VP+FP
	Negativa	FN	VN	FN+VN
		VP+FN	FP+VN	

$$\text{Val pred +} = \frac{VP}{VP + FP}$$

$$\text{Val pred -} = \frac{VN}{FN + VN}$$

Valor predictiu positiu segons la prevalença de la malatia

Pacients pediàtrics simptomàtics

Tira reactiva	Urocultiu		
	+	-	
+	57	211	268
-	17	262	279
	74	473	547

Sens = 77,0% Esp = 55,4%
 VP + = 21,3% VP- = 93,9%
 Prevalença = 13,5%

Pacients adults

Tira reactiva	Urocultiu		
	+	-	
+	29	222	251
-	3	286	289
	32	508	540

Sens = 90,6% Esp = 56,3%
 VP + = 11,6% VP- = 99,0%
 Prevalença = 5,9%

4 - Reclutament dels participants

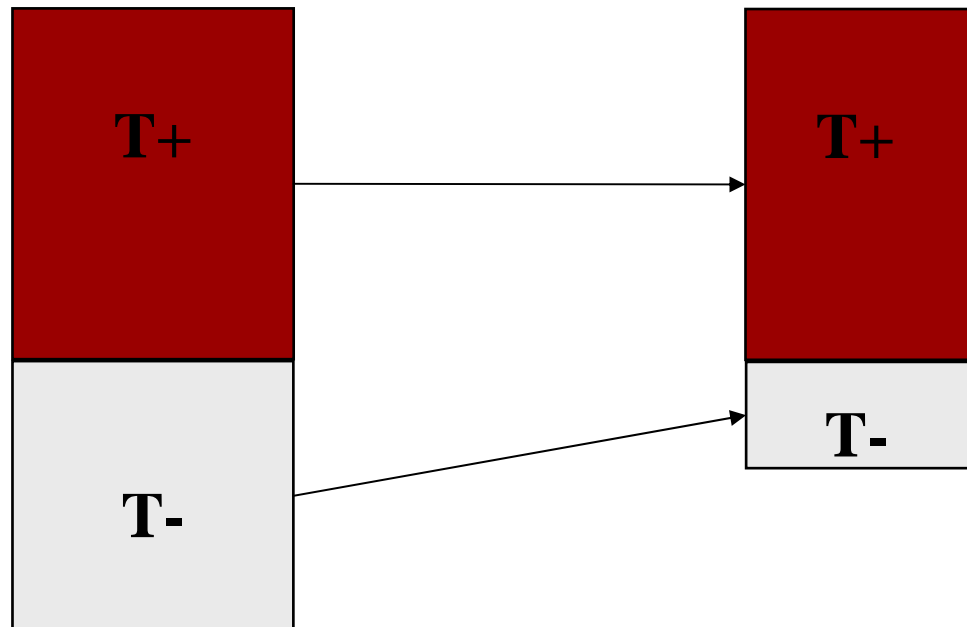
Cal descriure si el **reclutament** es va basar en:

- Els **síntomes** que presentaven
- Els **resultats** d'altres proves
- Haver rebut la **prova d'estudi** o de **referència**
 - En estudis en què els individus s'incloguin **segons el resultat** d'una prova de referència, es pot **supraestimar la sensibilitat** de la prova índex (biaix de verificació)
 - L'espectre clínic de la malaltia pot ser diferent

Biaix de verificació

Individus malalts a qui s'aplica la prova **índex**

Individus malalts a qui s'apliquen les proves **índex** i **de referència**



5 - Mostreig dels participants

- Cal especificar si:
 - La mostra era una **sèrie consecutiva** de pacients que complien els criteris de selecció
 - En cas contrari, com es va realitzar el **mostreig** (aleatori, etc...)
- L'estratègia de mostreig és important de cara a valorar si els resultats són **generalitzables**

Multidetector computed tomography for acute pulmonary embolism

Stein PD, Fowler SE, Goodman LR, *et al*; PIOPED II Investigators.
N Engl J Med. 2006; 354(22): 2317-27.

Study Population and Enrollment

All patients who were at least 18 years of age and had **clinically suspected acute pulmonary embolism** were seen on either an inpatient or outpatient basis at the eight participating clinical centers between September 2001 and July 2003. Patients who were referred for diagnostic imaging for **suspected pulmonary embolism** were identified for **recruitment** [...].

Patients were **recruited consecutively** during periods of staff availability, usually during the daytime on weekdays. **Exclusion criteria** are shown in Figure 1.

Prospective study of the **diagnostic accuracy** of the Simplify D-dimer assay for pulmonary embolism in emergency department patients.

Kline JA, Runyon MS, Webb WB, Jones AE, Mitchell AM.

Chest. 2006; 129(6): 1417-23

Research participants were ED patients of all ages who were selected by a board-certified emergency physician to undergo a structured **PE rule-out protocol** [...]

Typical **indications** for the PE rule-out protocol **included**

- **symptoms** of dyspnea, chest pain, or syncope, or
- physical **signs** such as a rapid pulse or low pulse oximetry reading that could not be explained by another disease process

6 - Obtenció de les dades

Cal descriure si es va planificar

- **abans** d'aplicar les proves d'estudi i de referència (estudi **concurrent** o prospectiu)
 - Permet auditories internes de l'obtenció de dades
 - Garantia de **qualitat** de les dades
- o bé **després** (estudi **històric** o retrospectiu)
 - Més **representatiu** de la pràctica habitual
 - Menor control sobre l'obtenció de les dades

Mètodes: diagnòstics

7. Prova de referència
8. Especificacions tècniques
9. Unitats i punts de tall
10. Experiència professional
11. Emmascarament

7 - Prova de referència

- La **prova de referència** s'ha de descriure i justificar
 - Estableix quins pacients tenen la malaltia i quins no.
- Quan per raons ètiques o pràctiques no hi ha una prova de referència que es pugui aplicar a tots els pacients, es sol utilitzar una **combinació de criteris**
 - Algunes patologies (malformacions congènites, etc...) es poden descartar **després d'un seguiment clínic** suficient

8 - Especificacions tècniques

- Cal descriure les **especificacions tècniques** dels materials i els mètodes implicats
 - **Quan** i **com** es van efectuar les mesures
 - Es poden aportar cites bibliogràfiques per a la prova de referència, o per a la prova estudiada.
- La descripció ha d'ésser detallada per a permetre:
 - Que altres investigadors puguin **replicar l'estudi**,
 - Que els lectors valorin la **factibilitat** d'aplicar la prova índex en el seu context clínic

9- Unitats i punts de tall

- Cal especificar la definició i el fonament de les **unitats**, **punts de tall** i **categories** de les proves índex i de referència
- Quan una variable contínua s'ha categoritzat, és important saber si els punts de tall es van definir **abans** d'obtenir les dades, o durant **l'anàlisi**
 - Si els punts de tall s'estableixen **després d'obtenir les dades**, se'n **facilita l'anàlisi** i la presentació, però és **menys probable** que els resultats es puguin **replicar** en un altre estudi

10 – Experiència dels professionals

- Cal descriure el **nombre**, **formació** i **experiència** de les persones que van aplicar i interpretar les proves índex i de referència
 - La variabilitat en l'**aplicació** o la **interpretació** de la prova índex o de referència pot afectar les mesures d'**exactitud diagnòstica**
 - Aquesta variabilitat és important en les proves d'imatge
- La informació sobre els professionals permet valorar **si en un altre àmbit** assistencial s'haurien pogut obtenir els **mateixos** resultats

11 - Emmascarament

- Cal especificar si **els professionals** que interpretaven les proves índex i de referència eren cecs (s'havien **emmascarat**) respecte al resultat de l'altra prova.
- L'exactitud diagnòstica quedarà **supraestimada** si:
 - la prova índex s'interpreta després de conèixer el resultat de la de referència (***test review bias***)
 - la prova de referència s'interpreta després de conèixer el resultat de la prova índex (***diagnostic review bias***).
- Cal especificar quina informació clínica es tenia en interpretar les proves

9- Unitats i punts de tall

- Cal definir i justificar les **unitats**, **punts de tall** i **categories** de les proves índex i de referència
 - Moltes proves diagnòstiques es basen en el valor d'una variable **contínua**
 - En alguns casos, la distribució de la variable en pacients amb la malaltia es **superposa** a la dels pacients sense la malaltia
 - Cal establir un **punt de tall** a partir del qual els individus es classificaran com sans o malalts

Sensibilitat i especificitat dels punts de tall

- Si la variable presenta valors més elevats en els individus amb la malaltia:
 - Com més baix és el punt de tall
 - Augmenta la **sensibilitat**
 - Disminueix l'especificitat
 - Com més alt és el punt de tall
 - Augmenta l'**especificitat**
 - Disminueix la sensibilitat

Definició de punts de tall

- Si els autors han definit **punts de tall**, han d'especificar si ho han fet abans de l'estudi o **durant** l'anàlisi de les dades
 - Si els punts de tall s'estableixen **després d'obtenir les dades**, és més probable que s'esculli un valor que optimitzi els resultats, però ho serà **menys** que els resultats es **repliquin** en un altre estudi

Corba ROC

- Una corba ROC és la representació gràfica de la **sensibilitat** i **especificitat** dels diferents punts de tall d'una prova diagnòstica basada en una variable contínua
- En **abscisses** es representa la **proporció de falsos positius** (1-especificitat) de cada punt i, en **ordenades**, la seva **sensibilitat**

Interpretació de la corba ROC

- La corba ROC informa de la **superposició** de la variable en **sans** i **malalts**
 - Com més s'acosta al vèrtex superior esquerre, millor els discrimina
 - Com més s'acosta a la diagonal de la gràfica, pitjor els discrimina
 - Com més elevada és l'**àrea sota la corba ROC**, millor els discrimina

Mètodes: Estadística

12. Error aleatori

13. Consistència

12 – Error aleatori

- Les estimacions d'exactitud diagnòstica depenen d'un **error aleatori** (deguda al mostreig), que serà **més important** com **més petita sigui la mostra**
- Cal descriure els mètodes estadístics utilitzats per a quantificar l'**error aleatori**
 - Intèrvals de confiança al 95%
- Els mètodes utilitzats per a **comparar l'exactitud diagnòstica**
 - Prova estadística de McNemar, etc..

Intervals de confiança al 95%

Sensibilitat

$$\text{sensibilitat} \pm 1,96 \cdot \sqrt{\frac{S \cdot (1 - S)}{VP + FN}}$$

Especificitat

$$\text{especificitat} \pm 1,96 \cdot \sqrt{\frac{Sp \cdot (1 - Sp)}{FP + VN}}$$

Quocient de probabilitats (“*likelihood ratio*”)

Quocient de probabilitats per a un resultat positiu

$$\frac{\text{Probabilitat d'un resultat positiu en malalts}}{\text{Probabilitat d'un resultat positiu en sans}} = \frac{S}{1 - E}$$

Quocient de probabilitats per a un resultat negatiu

$$\frac{\text{Probabilitat d'un resultat negatiu en malalts}}{\text{Probabilitat d'un resultat negatiu en sans}} = \frac{1 - S}{E}$$

Quocients de probabilitat de la tira reactiva per al diagnòstic de la infecció urinària

Tira reactiva	Urocultiu		Quocient de probabilitat
	Positiu	Negatiu	
Positiu	86	433	1,84
Negatiu	20	548	0,34
	106	981	

13 - Consistència

- Es recomana avaluar la **consistència** de les proves (reproduïbilitat)
- Cap prova (avaluada o de referència) **no és perfecta**
 - **Variabilitat** de l'observador, dels instruments, ...
 - La manca de consistència afecta l'exactitud de la prova
- Cal descriure el mètode de **càlcul de consistència** que s'hagi aplicat
 - Els estudis de concordança valoren el **grau d'acord** entre dues o més persones en fer un diagnòstic (índex Kappa)

Concordança

**Diagnòstic de
l'observador B**

**Diagnòstic de
l'observador A**

Positiu

Negatiu

Positiva

Concordant

No concordant

Negativa

No concordant

Concordant

	Positiu	Negatiu
Positiva	Concordant	No concordant
Negativa	No concordant	Concordant

Càlcul de l'índex Kappa

$$\kappa = \frac{\text{Concordança observada} - \text{Concordança esperada}}{100 - \text{Concordança esperada}}$$

User's Guides to the Medical Literature

III. How to Use an Article About a Diagnostic Test

A. Are the Results of the Study Valid?

- Criteris principals
 - La mostra de l'estudi és **representativa** de l'espectre de pacients als quals s'aplicarà la prova?
 - La prova avaluada es va comparar de forma **independent i a cegues** amb una tècnica de referència?
- Criteris secundaris
 - Els **resultats** de la prova avaluada **van influir** en la decisió d'aplicar la prova de referència?
 - L'aplicació de la prova avaluada **es va explicar amb prou detall** com per a poder replicar l'estudi?

JAMA 1994; 21: 389-391

Bibliografia

Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem. 2003; 49(1):7-18.

Pàgina web:

The STARD Initiative -- Towards Complete and Accurate Reporting of Studies on Diagnostic Accuracy. <http://www.stard-statement.org/>