

# Dissenys d'estudi en recerca clinicoepidemiològica

## Part I: estudis descriptius

Cari Almazán, Mireia Espallargues, Glòria Oliva

Agència d'Avaluació de Tecnologia i Recerca Mèdiques

### OBJECTIUS FORMATIUS

1. Descriure els dissenys d'estudi clinicoepidemiològics més habituals en la investigació biomèdica.
2. Introduir els criteris bàsics per valorar la qualitat d'una investigació (precisió i validesa).
3. Descriure les principals característiques i aplicacions dels dissenys descriptius, les mesures de resultat més habituals i els seus avantatges i limitacions.
4. Proporcionar elements per saber quin és el disseny d'estudi més adequat a l'hora de respondre a una pregunta de recerca.

### Estructura de l'article

#### Introducció

#### Principals dissenys d'investigació

#### Criteris de qualitat d'una investigació

#### Tipus d'error

#### Factors de confusió

#### Estudis descriptius

#### Estudis descriptius transversals

#### Estudis de prevalença

#### Sèries de casos

#### Estudis d'associació encreuada

#### Estudis per avaluar proves diagnòstiques

#### Estudis de concordança

#### Estudis descriptius longitudinals

#### Estudis d'incidència

#### Bibliografia

### Introducció

La investigació és un procés sistemàtic en què s'utilitza el mètode científic per donar resposta a una pregunta d'investigació. A partir d'un coneixement sobre un tema, es formula una hipòtesi o objectiu, es recullen dades segons un disseny definit, aquestes s'analitzen, s'interpreten els resultats i s'obtenen unes conclusions que s'incorporaran al coneixement científic existent. Les preguntes a investigar poden ser diverses: des de definir i caracteritzar una malaltia, determinar la millor prova diagnòstica, descriure el patró d'ocurrència d'una malaltia, definir l'estratègia preventiva, els factors de risc d'una malaltia, establir el millor tractament, entre d'altres. Així, per respondre la pregunta d'investigació caldrà escollir el disseny més adequat.

En aquest article es descriurà, en primer lloc, la classificació dels dissenys més habituals dels estudis clínicopidemiològics, després es revisaran els principis bàsics per plantejar una investigació i finalment s'introduiran els dissenys d'estudi descriptius. En una segona part, es descriuran els dissenys de caràcter analític que es fan servir si es vol avaluar la presència d'una relació causal, ja sigui entre un factor de risc o l'efecte d'una intervenció i una malaltia.

### Principals dissenys d'investigació

El disseny d'un estudi d'investigació és el conjunt de procediments, mètodes i tècniques mitjançant els quals l'investigador porta a terme les etapes d'investigació (descripció i selecció de la mostra d'estudi, els mètodes i instruments de recollida de dades, l'anàlisi i interpretació dels resultats). Tots aquests aspectes han de quedar descrits en el protocol d'estudi que s'elaborarà abans d'iniciar-lo.

Els dissenys dels estudis es poden classificar segons els criteris següents:

#### 1) Finalitat de l'estudi i analítica o descriptiva.

Es considera analític l'estudi que avalua una presumpta relació causa-efecte, mentre que en l'estudi descriptiu les dades s'utilitzen per descriure un determinat fenomen.

#### 2) Control de l'assignació dels factors d'estudi: experimental o observacional.

En l'estudi experimental en què l'equip investigador assigna i controla el factor d'estudi (l'expo-

#### Correspondència:

Cari Almazán

Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATM)  
Travessera de les Corts, 131-159. 08028 Barcelona  
calmazan@olimpia.scs.es

Treball rebut: 08-03-2002. Treball acceptat: 12-04-2002.

Almazán C, Espallargues M, Oliva G.  
**Dissenys d'estudi en recerca clinicoepidemiològica.**  
**Part I: estudis descriptius.**  
*Pediatr Catalana* 2002; 62: 196-205.

sició, intervenció, tractament) de forma deliberada. Mentre que en un estudi observacional l'investigador es limita a observar, mesurar i analitzar el que ocorre de forma «natural» o «espontània», però no intervé en l'assignació i control dels fenòmens d'estudi.

### 3) Seqüència temporal: transversal o longitudinal (de seguiment)

La mesura del factor d'estudi, en els estudis transversals, correspon a un moment del temps. Mentre que en els estudis longitudinals hi ha una seqüència temporal entre les variables estudiades, això és, un període de seguiment.

### 4) Inici de l'estudi en relació a la cronologia dels fets: prospectiu o retrospectiu

Es considera un estudi prospectiu quan l'inici és anterior als fets estudiats i les dades es recullen a mesura que van succeint. En l'estudi retrospectiu es realitza posteriorment a l'ocurrència dels fets que s'estudien.

A partir de la combinació dels criteris anteriors es poden definir els dissenys d'estudi que es presenten a la Taula I. Cal tenir en compte, però, que hi ha altres classificacions diferents segons els autors. També és important remarcar que no hi ha un disseny d'estudi millor que un altre i, a més, una mateixa pregunta de recerca pot respondre's de diverses formes. Per tant, per a cada pregunta d'investigació s'ha de valorar quin disseny és més eficaç per obtenir una resposta vàlida i precisa.

TAULA I

#### Classificació dels tipus de disseny

| Estudis descriptius                                                                                                                                                                                                                                                                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Estudis descriptius transversals</b> <ul style="list-style-type: none"> <li>• Estudis de prevalença</li> <li>• Estudis d'associació encreuada</li> <li>• Sèries de casos transversals</li> <li>• Estudis d'avaluació de proves diagnòstiques</li> <li>• Estudis de concordança</li> </ul> |
| <b>Estudis descriptius longitudinals</b> <ul style="list-style-type: none"> <li>• Estudis d'incidència</li> <li>• Sèries de casos longitudinals</li> </ul>                                                                                                                                   |
| Estudis analítics                                                                                                                                                                                                                                                                            |
| <b>Estudis analítics experimentals</b> <ul style="list-style-type: none"> <li>• Assaigs clínics controlats aleatoritzats</li> </ul>                                                                                                                                                          |
| <b>Estudis analítics quasi-experimentals</b> <ul style="list-style-type: none"> <li>• Estudis abans-després</li> </ul>                                                                                                                                                                       |
| <b>Estudis analítics observacionals</b> <ul style="list-style-type: none"> <li>• Seqüència causa-efecte: estudis de cohorts (prospectius)</li> <li>• Seqüència efecte-causa: estudis de casos i controls (retrospectius)</li> </ul>                                                          |

## Criteris de qualitat d'una investigació

La planificació de la investigació tracta de donar resposta a la pregunta de recerca de forma vàlida i precisa. Aquestes propietats es poden veure modificades per diversos factors i poden influir en la qualitat de la investigació. Per tant, en dissenyar un estudi es tractarà d'evitar aquests factors: els errors (aleatori, sistemàtic) i els factors de confusió (Taula II).

TAULA II

#### Criteris de qualitat d'una investigació

| Factors que comprometen la precisió                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>• Error aleatori</li> </ul>                                                                                                                                                 |
| Factors que comprometen la validesa interna                                                                                                                                                                        |
| <ul style="list-style-type: none"> <li>• Error sistemàtic o biaixos <ul style="list-style-type: none"> <li>- Biaix de selecció</li> <li>- Biaix d'informació</li> </ul> </li> <li>• Factors de confusió</li> </ul> |

## Tipus d'error

Dos són els errors que s'han de minimitzar en qualsevol estudi: l'aleatori i el sistemàtic.

- L'*error aleatori* neix del fet que sovint es treballa amb mostres d'individus, i no amb tota la població. Procedeix, doncs, de la variabilitat inherent al mostreig. Així, l'error aleatori depèn de la grandària de la mostra: quan aquesta augmenta, l'error disminueix. Aquest error també es pot produir en mesurar les variables d'estudi; per això està molt relacionat amb el concepte de *precisió* o *fiabilitat*. Precisió (o fiabilitat) és el grau en què els resultats o mesures repetides d'un fenomen relativament estable cauen a prop unes de les altres. Una estimació o mesura serà més precisa com menor sigui l'error aleatori (Fig. 1).

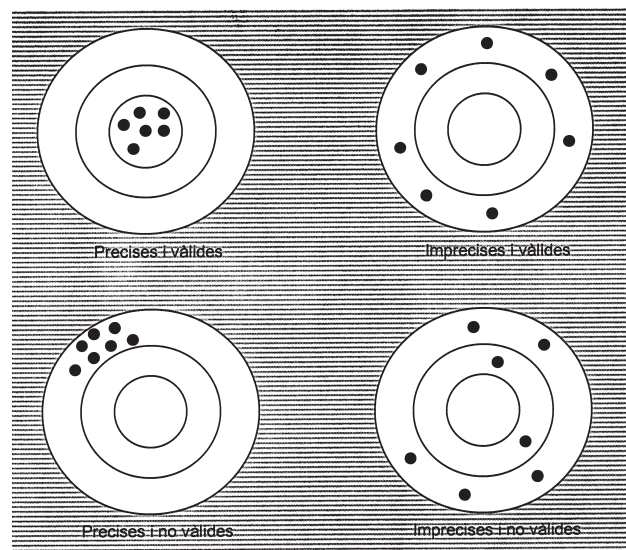


Fig. 1. Comparació entre validesa i precisió de les mesures

*L'error sistemàtic o biaix* apareix quan s'introdueix un error en el disseny de l'estudi, ja sigui en la selecció dels individus, en la informació recollida o en la seva anàlisi, de manera que es produeix una diferència sistemàtica entre els grups no atribuïble al factor que s'està estudiant. A diferència del que ocorre amb l'error aleatori, no s'atenua en augmentar la grandària de la mostra, i, una vegada introduït, és pràcticament impossible de controlar en la fase d'anàlisi. Aquest tipus d'error va molt lligat al concepte de *validesa*. La validesa (o exactitud) és el grau en què els resultats d'una observació es corresponen a la realitat del fenomen que s'està mesurant o avaluant. Aquest error condueix a una estimació incorrecta de l'efecte o del paràmetre que s'estudia (Fig. 1).

En general els biaixos es poden classificar en dos grups.

#### a) *Biaixos de selecció*

Aquest biaix està relacionat amb la forma en què els individus són identificats i seleccionats per a un estudi o amb la forma com són seguits en el temps. El biaix de selecció es pot donar en diverses circumstàncies, entre les quals destaquen:

1. Selecció dels diferents grups d'estudi. Per la forma com se seleccionen els grups (per diferents criteris de selecció entre els grups) aquests poden no ser comparables per a totes les variables pronòstiques (variables relacionades amb la malaltia o l'efecte). Aquesta comparabilitat, en general, és més difícil d'assegurar en els estudis observacionals. Així, en un cas extrem, per exemple, ens trobaríem enfront d'un biaix de selecció si tinguéssim nois en un grup de tractament i noies en un altre i el sexe fos una variable associada a la malaltia.
2. Pèrdues de seguiment. En estudis en què hi ha seguiment d'una cohort o grup d'individus, si la proporció d'individus que es perden en el seguiment és diferent entre els grups que s'han de comparar es pot sospitar la presència d'aquest biaix. L'error es produirà si la probabilitat de desenvolupar la malaltia entre els individus que es perden és diferent en cadascun dels grups que es comparen. El biaix relacionat amb la pèrdua d'individus durant el seguiment és característic dels estudis de caràcter prospectiu.
3. Supervivència selectiva. Aquest biaix és important en els estudis de prevalença i en els estudis de casos i controls en què s'inclouen casos prevalents, ja que aquests poden ser els supervivents de casos incidents i la supervivència pot estar relacionada amb l'exposició. A més, els casos prevalents canvien els seus hàbits en relació amb l'exposició. A més, els casos prevalents poden canviar els seus hàbits en relació amb l'exposició i en conseqüència també s'esbiaixa l'estimació de l'efecte.
4. Mostra no representativa de la població d'estudi. Aquesta situació pot ocórrer freqüentment en els

estudis descriptius transversals, encara que també pot presentar-se en els estudis analítics. S'introdueix en el procés de selecció de la mostra procedent de la població d'estudi (per exemple en un estudi en què es vulgui estimar la prevalença d'una malaltia en una població i es decideixi seleccionar una mostra a partir de les persones que acudeixen a centres sanitaris públics per qualsevol motiu. En aquest cas es podria obtenir una mesura esbiaixada de la prevalença, ja que les persones que consulten en aquests centres no són representatives de la població general).

5. Biaix de detecció. Aquest biaix es produeix quan el procediment utilitzat per identificar l'estat de la malaltia varia o es modifica segons la presència o absència d'exposició. Un exemple conegut d'aquest biaix és l'estudi del risc de tromboembolisme pulmonar i l'ús d'anticonceptius orals. En un estudi de casos i controls es va plantejar com una possible explicació d'aquesta associació el fet que alguns metges coneixien la hipòtesi de treball i per tant les dones que prenién anticonceptius orals tenien una major probabilitat de ser ingressades amb el diagnòstic de tromboembolisme pulmonar i a més eren sotmeses a un major nombre de proves per descartar la malaltia que no pas les dones del grup control.
6. Participació de voluntaris o efecte d'autoselecció. La participació de voluntaris en un estudi pot introduir un biaix d'aquestes característiques, atès que els voluntaris habitualment són diferents dels no voluntaris (autoselecció). Si la voluntat de participar està relacionada amb factors que poden influir sobre la relació que s'estudia, la selecció de voluntaris pot donar lloc a aquest biaix.

#### b) *Biaixos d'informació*

Aquest tipus de biaix es pot produir quan les mesures de les variables d'estudi són de mala qualitat o la informació sobre la malaltia es recull de forma sistemàticament diferent entre els grups d'estudi. Les causes més freqüents del biaix d'informació són:

- La utilització d'un instrument de mesura no adequat, o de definicions poc vàlides i/o precises (per exemple mesurar la tensió arterial amb un aparell no calibrat).
- Errors introduïts pels enquestadors o personal que recull la informació. Aquest biaix es refereix a qualsevol diferència sistemàtica en la forma de sol·licitar, recollir o interpretar la informació procedent dels participants en un estudi en funció del grup al qual pertanyen.
- Errors deguts als participants a l'estudi, com el «biaix de memòria» que es produeix quan els individus amb un determinat problema de salut recorden la seva història d'exposició de forma diferent que els que no tenen aquest problema.

La introducció d'un biaix d'informació suposa una mala classificació dels subjectes, respecte a l'exposició o a la intervenció, o bé respecte a la malaltia o variable de resposta.

Durant la fase de disseny és fonamental preveure els possibles biaixos de l'estudi, ja que un cop apareixen no es poden eliminar de l'anàlisi.

## Factors de confusió

Un altre tipus d'error procedeix de la presència de factors de confusió. El fenomen de confusió apareix quan l'associació observada entre el factor d'estudi i l'efecte (malaltia) pot ser totalment o parcialment explicada per una altra variable (factor de confusió), o al contrari, quan una associació real queda emmascarada per aquest factor. Així, si es vol mesurar la manca d'augment de pes durant l'embaràs (exposició) com a factor de risc del part preterme (malaltia/efecte) i es planteja la possibilitat que aquesta relació pugui estar afectada per l'atenció prenatal durant l'embaràs, s'han de complir els 3 criteris que es poden representar de la forma següent (Fig. 2):

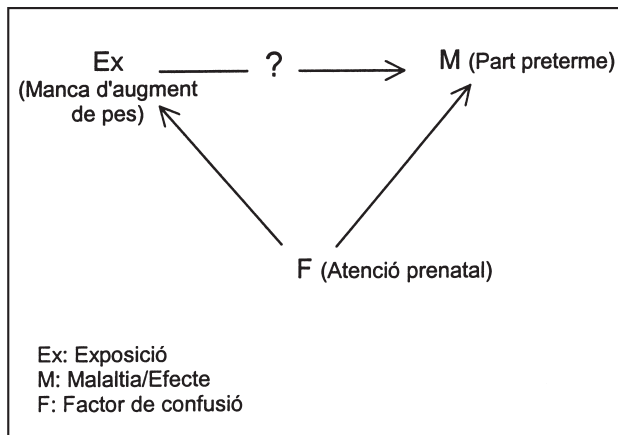


Fig. 2. Factors de confusió.

- El factor de confusió (F) ha de ser un «factor de risc» o estar molt correlacionat amb la malaltia (M). A més ha de produir l'efecte tot i que l'exposició (Ex) no estigui present.
- El factor de confusió (F) ha d'estar associat amb l'exposició (Ex) en la població d'estudi.
- El factor de confusió (F) no ha de ser un pas intermedi en la cadena causal (en l'associació entre el factor d'estudi i la malaltia).

La confusió es produeix quan alguna de les característiques que influeixen sobre els resultats (variables pronòstiques, a part del factor d'estudi) es distribueix de forma desigual entre els grups comparats. A diferència dels baixos, el fenomen de la confusió pot ser corregit en la fase d'anàlisi de les dades.

Els errors sistemàtics (biaixos) i els factors de confusió afecten la **validesa interna** de l'estudi, és a dir,

el grau en què els resultats són vàlids (lliures de biaixos o errors sistemàtics) per a la població que ha estat estudiada. La **validesa externa** es refereix al grau en què els resultats d'un estudi poden ser generalitzats a la població. Per tant, la primera condició perquè uns resultats puguin ser generalitzats és que tinguin validesa interna. A més, la generalització depèn de la consistència dels resultats entre diferents estudis i de mecanismes plausibles que expliquen raonablement un efecte, la qual cosa sovint implica judicis de valor per part dels propis investigadors.

## Estudis descriptius

Els estudis descriptius són observacionals i en general no es planteja una definició i contrast d'hipòtesis. Poden ser transversals si la recollida de dades es realitza en un únic moment en el temps o bé longitudinals si hi ha una seqüència temporal o seguiment dels individus estudiats.

### Estudis descriptius transversals

L'estudi transversal és un disseny bàsic en epidemiologia. Es caracteritza per l'absència de seguiment i dona una visió instantània del que ocorre en una població o una mostra seva en un moment donat. El procés que s'estudia (malaltia o un determinant de salut) es mesura una sola vegada.

Els estudis transversals es poden utilitzar per estimar la freqüència d'una malaltia en una determinada població en un moment donat, com són els estudis de prevalença. També es poden plantejar per descriure una sèrie de característiques seleccionades observades en un grup de malalts amb una determinada malaltia o subjectes que tenen una característica en comú: són el que s'anomena sèries de casos transversals. Un cas particular són els estudis d'associació encreuada o transversals analítics que són estudis amb un disseny transversal, però més enllà de voler descriure la població o mostra estudiada tenen una finalitat analítica, és a dir, s'utilitzen per investigar l'associació entre una determinada exposició i una malaltia, que es mesura de forma simultània en un moment donat.

Altres aplicacions més específiques del disseny descriptiu transversal són els estudis per avaluar proves diagnòstiques i els estudis de concordança. Aquests últims es realitzen per mesurar la fiabilitat d'un aparell de mesura o la concordança entre dos observadors en la mesura d'una variable o fenomen d'estudi.

### Estudis de prevalença

Els estudis de prevalença mesuren la freqüència d'una malaltia o problema de salut en un moment determinat o referit a un període definit de temps. Es poden



dur a terme a partir d'una mostra representativa de la població o de tota la població si és possible, en la qual es determinarà la proporció d'individus que tenen la malaltia o la característica d'estudi.

Aspectes rellevants a considerar en el disseny d'un estudi de prevalença són que la mostra d'estudi sigui representativa de la població i que la proporció de no respostes sigui el més baixa possible. En aquest estudi la definició de la població ha de ser precisa i quan es mesura la presència d'una malaltia s'ha de tenir en compte que es pot obtenir una major proporció de casos de llarga durada. En aquests casos, la mostra pot ser que no sigui representativa de tots els casos de malaltia que es presenten en una determinada població, sinó només d'aquells en què el procés sigui més llarg (biaix de supervivència selectiva). També en el disseny i en l'anàlisi de l'estudi s'ha de valorar la presència de no respostes. Així, si la prevalença de la malaltia és diferent en els que participen a l'estudi dels que no hi participen, tot i que hagin estat seleccionats (no respostes) es pot donar un biaix en l'estimació dels resultats. Com major sigui el percentatge de no respostes, major serà la probabilitat que hi hagi una diferència de prevalença observada en les respostes respecte a les no respostes. En plantejar l'estudi s'ha de definir l'estratègia per reduir el percentatge de no respostes, i també estudiar una mostra de les no respostes (o totes si està disponible la informació) per valorar la presència de diferències en ambdós grups.

La validesa dels resultats de l'estudi dependrà de la representativitat de la mostra, qualitat de les dades i un nombre reduït de no respostes.

La prevalença (puntual) és el quocient entre el nombre de casos d'una determinada malaltia (o factor de risc, característica, etc.) existent en una població en un moment donat i s'estima:

$$\text{Prevalença} = \frac{\text{Nombre de casos amb malaltia existents en el moment (t)}}{\text{Total de la població en el moment (t)}}$$

La prevalença depèn de la incidència i de la durada de la malaltia.

*Exemple:* Suposem que es vol conèixer la prevalença d'inici de la lactància materna exclusiva en una determinada població i en un moment donat (Suárez P, et al, 2000). Així, de tots els nadons nascuts als hospitals públics d'aquesta població i que tenien entre 0 i 8 mesos d'edat en el moment d'iniciar l'estudi es va seleccionar una mostra aleatòria de 453 nens. Les dades es van obtenir per entrevista telefònica a les mares mitjançant qüestionari i 233 mares van declarar que només feien alletament matern als seus fills. Per tant, la prevalença d'alletament matern exclusiu va ser del 51.4% (233/453).

## Sèries de casos transversals

Les sèries transversals consisteixen en la descripció d'una sèrie de característiques en un moment del temps en un grup de pacients que tenen una determinada malaltia o que estan rebent un mateix tractament o una altra característica d'interès. Per exemple, descriure les xifres actuals de colesterol en nens diabètics d'un centre de salut.

## Estudis d'associació encreuada

Són estudis que amb un disseny transversal (mesura simultània de l'exposició i de la malaltia en un moment donat o un període curt de temps) intenten investigar l'associació entre una determinada exposició i una malaltia. També s'anomenen estudis transversals analítics. La informació que es recull sobre l'exposició i la malaltia sovint fa referència al moment en què es fa l'estudi. En alguns casos i quan sigui etiològicament pertinent, es poden incloure preguntes sobre exposicions passades. La capacitat d'aquests estudis per establir una associació dependrà del grau en què l'exposició estigui relacionada amb l'etiologia de la malaltia. L'estudi de l'associació entre factors de risc i malaltia mitjançant aquest disseny està particularment indicat quan la malaltia és de llarga durada, en cas contrari, hi podria haver pocs casos amb la malaltia.

D'altra banda, la manca de seqüència temporal entre l'exposició i la malaltia sovint no permet establir si l'exposició és anterior o conseqüència de la malaltia. Per tant, aquest disseny permet estudiar associacions entre variables que posteriorment hauran de ser estudiades amb altres dissenys més adequats per demostrar la causalitat. A més, el valor d'una variable en un individu en el moment en què es realitza l'estudi pot ser diferent del que tenia en un període previ. Així, si es vol estudiar la dieta com a factor de risc d'una gastritis aguda, en el moment de fer l'estudi s'haurà de tenir en compte que aquesta pot haver estat modificada precisament per la presència de la malaltia. Aquests canvis en l'exposició són menys probables quan s'estudien característiques invariables relacionades amb factors genètics o amb el sexe.

En els estudis d'associació encreuada, com en el cas anterior dels estudis de prevalença, en l'etapa de disseny de l'estudi s'ha de tenir en compte la representativitat de la mostra i les no respostes.

*Exemple:* Suposem que es vol conèixer l'edat de contacte dels adolescents amb l'alcohol i la seva relació amb l'entorn familiar i social i els hàbits de vida (Paniagua H, et al, 2001). A partir d'una mostra de 2.178 adolescents de 12 a 16 anys representatius d'una determinada comunitat, es va estimar la prevalença del consum d'alcohol i la seva relació amb el consum d'altres drogues i diferents aspectes del seu entorn (oci i temps lliure, variables psicoafectives, alimentàries i relacionades amb l'aprenentatge). Es

va observar que l'entorn i els hàbits de vida (major d'edat, pare i mare que beuen, fumador actual, menor freqüència de fer esport i lectura, i major freqüència d'anar de copes, a la discoteca, de sortir amb els amics i d'escoltar música) es relacionaven estretament amb les conductes de risc (consum d'alcohol).

Els avantatges i limitacions dels estudis transversals es descriuen a la Taula III.

TAULA III

### Avantatges i limitacions dels estudis descriptius transversals

| Avantatges                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>– Permeten estudiar la prevalença de les malalties o problemes de salut i ajuden a la planificació del servei sanitari.</li> <li>– Es poden estudiar diverses malalties facilitant la generació d'hipòtesis causals que poden ser posteriorment provades amb estudis analítics.</li> <li>– Poden ser repetits en diferents moments en el temps per valorar l'evolució de diferents problemes de salut.</li> <li>– Se solen realitzar en mostres representatives de la població general, per la qual cosa tenen major validesa externa.</li> <li>– Són fàcils de dur a terme.</li> <li>– Es realitzen en curts períodes de temps, raó per la qual solen ser estudis amb un cost econòmic inferior a altres tipus de dissenys.</li> <li>– En determinades situacions poden ser l'únic disseny possible.</li> </ul> |
| Limitacions                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <ul style="list-style-type: none"> <li>– El fet que es mesuri simultàniament la presència o absència de factors de risc i la malaltia fa que no es pugui determinar si l'exposició és anterior o conseqüència de la malaltia, cosa que dificulta establir una associació causal.</li> <li>– No són útils per a l'estudi de malalties poc freqüents.</li> <li>– Possibilitat de sobreestimació de la prevalença en el cas de malalties cròniques per tenir una durada més llarga.</li> <li>– Possibilitat del biaix degut a les no respostes.</li> </ul>                                                                                                                                                                                                                                                                                                 |

### Estudis per avaluar proves diagnòstiques

L'objectiu dels estudis per avaluar una prova diagnòstica és determinar la capacitat de la prova per discriminar les persones que pateixen la malaltia i aquelles que no o bé per diferenciar diversos estats. És a dir, mesurar la seva validesa o capacitat diagnòstica. Aquests estudis consisteixen a aplicar la prova diagnòstica que s'està avaluant i una prova de referència o l'estàndard d'or (*"gold standard"*, criteri de referència o patró de referència) a una mostra de malalts i comparar els resultats d'ambdues proves.

Les característiques operacionals de la prova diagnòstica són la sensibilitat, l'especificitat i els valors predictius que quantifiquen la capacitat de la prova per classificar correcta o incorrectament una persona, segons la presència o absència d'una malaltia. Les definicions i el càlcul de les característiques operacionals es mostren a la Taula IV.

TAULA IV

### Càlcul i definició de les característiques operacionals de la validesa d'una prova diagnòstica

| Prova d'estudi (auscultació)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Resultats segons patró de referència |             | Total   |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|-------------|---------|
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Malaltia                             | No malaltia |         |
| Positiva                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | a<br>VP                              | b<br>FP     | a+b     |
| Negativa                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | c<br>FN                              | d<br>VN     | c+d     |
| Total                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | a+c                                  | b+d         | a+b+c+d |
| Veritables positius VP(a) = Resultat de la prova positiva en persones malaltes<br>Falsos positius FP(b) = Resultat de la prova positiva en persones no malaltes<br>Falsos negatius FN(c) = Resultat de la prova negativa en persones malaltes<br>Veritables negatius VN(d) = Resultat de la prova negativa en persones no malaltes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                      |             |         |
| <ul style="list-style-type: none"> <li>• <b>Sensibilitat</b> <math>a/(a+c)</math> o <math>VP/(VP+FN)</math>.<br/>Proporció d'individus amb la malaltia que tenen el resultat de la prova positiu.</li> <li>• <b>Especificitat</b> <math>d/(b+d)</math> o <math>VN/(VN+FP)</math>.<br/>Proporció d'individus sense malaltia amb el resultat de la prova negatiu.</li> <li>• <b>Valor predictiu positiu</b> <math>a/(a+b)</math> o <math>VP/(VP+FP)</math>.<br/>Probabilitat que un individu amb un resultat positiu tingui la malaltia.</li> <li>• <b>Valor predictiu negatiu</b> <math>d/(d+c)</math> o <math>VN/(FN+VN)</math>.<br/>Probabilitat que un individu amb resultat negatiu no tingui la malaltia.</li> <li>• <b>Raó de probabilitat positiva (RPP)</b> sensibilitat/(1-especificitat)</li> <li>• <b>Raó de probabilitat negativa (RPN)</b> (1-sensibilitat)/especificitat</li> </ul> |                                      |             |         |

Els valors predictius canvien segons la prevalença de la malaltia i són útils perquè inclouen informació sobre la prova i la població estudiada.

Una altra forma d'expressar els resultats de la prova és la raó de probabilitat o versemblança. Aquesta mesura informa de com es modificarà l'estimació de la probabilitat de presentació de la malaltia (o trastorn objectiu) que es vol diagnosticar a partir de la informació que s'obtingui amb la prova diagnòstica. Compara la probabilitat d'obtenir un determinat resultat en els individus que presenten la malaltia, amb la probabilitat d'obtenir el mateix resultat en els individus en els quals s'ha descartat la seva presència.

La raó de probabilitat té una aplicació rellevant a la pràctica clínica. Si un malalt prèviament a la realització de la prova diagnòstica té una probabilitat de tenir una malaltia estimada a partir de la informació aportada per la història clínica, l'exploració física i altres proves complementàries (o probabilitat pre-prova), la raó de probabilitat ens permetrà estimar com es pot modificar aquesta probabilitat prèvia amb el resultat de la prova diagnòstica (probabilitat post-prova). La utilització d'aquests valors en la pràctica clínica es facilita amb l'ús d'un nanograma (Fig. 3),

que permet un càlcul directe del canvi de les probabilitats de la malaltia, preprova i postprova, amb la informació de la prova diagnòstica. S'obté la probabilitat postprova (3a columna) fent una línia que uneixi el valor de la probabilitat preprova (1a columna) amb el valor de la raó de la probabilitat (2a columna) obtingut amb la prova diagnòstica, tal i com es descriu en l'exemple.

Quan els valors dels resultats de la prova diagnòstica segueixen una escala quantitativa (p.ex.: pressió

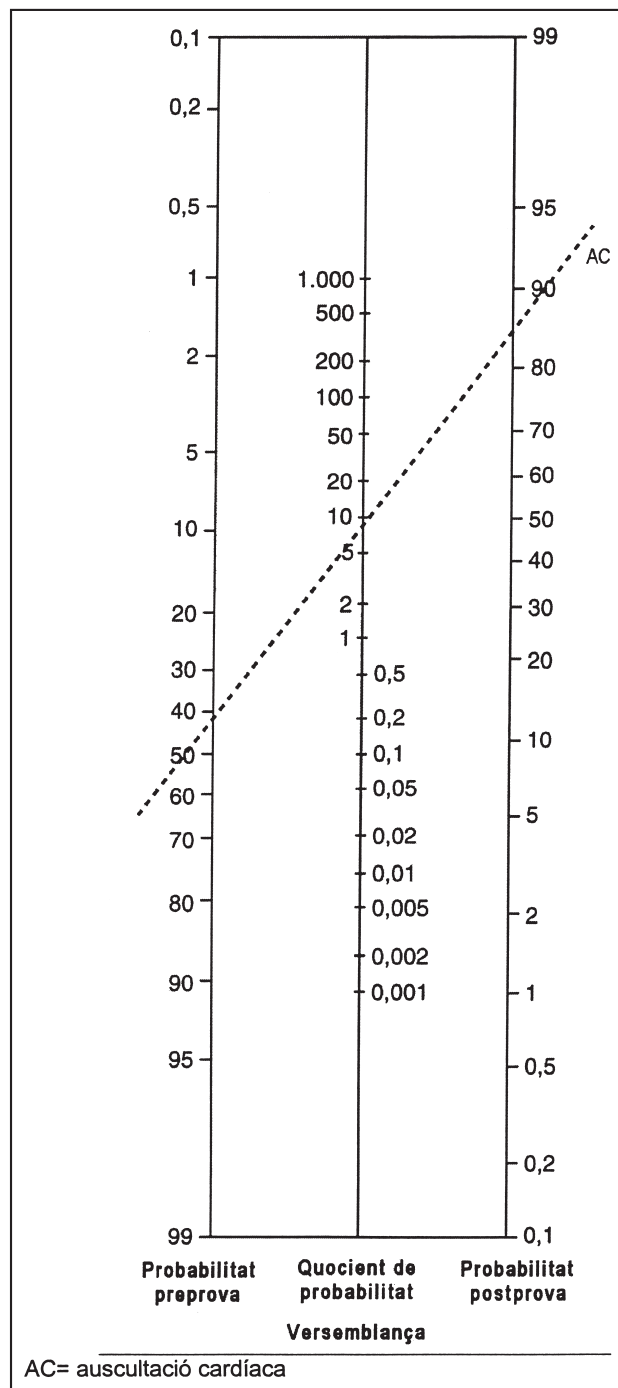


Fig. 3. Nanograma

arterial), la millor manera de conèixer de forma global la validesa d'una prova diagnòstica és mitjançant les corbes ROC (*receiver operating characteristics*). Es tracta d'un gràfic en què se situen els parells de valors de la sensibilitat, en l'eix d'ordenades, i del complementari de l'especificitat (1-especificitat), en l'eix d'abscisses, per a diferents categories o punts de tall dels resultats que s'hagin definit. La corba ROC facilita, doncs, l'elecció del millor punt de tall, és a dir, aquell que proporciona una major sensibilitat i especificitat a la prova (major àrea sota la corba). Els valors se situen entre 0 i 1.

**Exemple.** En l'estudi de Yubman et al 1991 en el qual es volia valorar la validesa diagnòstica de l'auscultació cardíaca per al diagnòstic de malformacions cardíques en nounats, es va seleccionar una mostra de 81 nounats consecutius amb la síndrome de Down als quals s'havia realitzat una ecocardiografia (prova de referència), també se'ls va fer una auscultació cardíaca registrant els valors. Els resultats d'ambdues proves es descriuen a la Taula V.

TAULA V

**Càlcul de les característiques operacionals de la validesa de l'auscultació cardíaca**

| Prova d'estudi (auscultació) | Resultats segons patró de referència (ecocardiografia) |                          | Total |
|------------------------------|--------------------------------------------------------|--------------------------|-------|
|                              | Malaltia (defecte cardíac)                             | No malaltia (No defecte) |       |
| Positiva                     | 18                                                     | 3                        | 21    |
| Negativa                     | 16                                                     | 44                       | 60    |
| Total                        | 34                                                     | 47                       | 81    |

- Sensibilitat = 53% (18/34)
- Especificitat = 94% (44/47)
- Valor predictiu positiu = 86% (18/21)
- Valor predictiu negatiu = 73% (44/60)
- Raó de probabilitat positiva = 9
- Raó de probabilitat negativa = 0,5

L'auscultació cardíaca va ser positiva en un 53% dels nadons amb malformació cardíaca (sensibilitat) i negativa en un 94% dels nadons sense malformació (especificitat).

En la mostra seleccionada, la prevalença de la malformació va ser d'un 42% (34/81) i la probabilitat de tenir una malformació cardíaca quan el resultat de l'auscultació era positiu va ser d'un 86% (valor predictiu positiu). La probabilitat de no tenir la malformació cardíaca quan el resultat de l'auscultació era negatiu va ser d'un 73% (valor predictiu negatiu).

Si es volguessin aplicar aquests resultats a un malalt en particular, amb unes característiques similars als malalts de la mostra i sense cap altra informació, i aquest tingués, per exemple, una probabilitat de

tenir una malformació cardíaca abans de l'auscultació del 42%, calculant la raó de probabilitat positiva i fent servir el nanograma, s'obtindrien els resultats següents: la probabilitat postprova de tenir una malformació cardíaca si l'auscultació és positiva, és d'un 85% aproximadament (vegeu línia del nanograma, Fig. 3).

La planificació d'un estudi per avaluar una prova diagnòstica ha de tenir en compte i valorar tres aspectes bàsics en el disseny: 1) la selecció dels malalts, 2) la selecció del criteri o prova de referència, i 3) l'aplicació de les proves als malalts o subjectes d'estudi.

**Selecció de malalts.** La mostra de malalts en els quals s'avalua una prova diagnòstica no pot diferir excessivament de la població a la qual s'aplica la prova en condicions de pràctica clínica habitual. A més s'haurà de seleccionar una mostra de malalts que inclogui diferents estadis de gravetat de la malaltia, i malalts amb altres diagnòstics que puguin induir a confusió (diagnòstic diferencial). Si la mostra està integrada per malalts molt greus, la prova donarà més resultats positius que si inclou malalts menys greus. El grup de no malalts permet determinar els falsos positius. Si aquest grup està integrat per individus molt sans el nombre de falsos positius pot ser inferior (major especificitat). Per tant, caldrà una representació de malalts amb l'espectre més ampli possible de la mostra.

Hi ha diferents alternatives per seleccionar la mostra d'estudi. Així, es pot seleccionar una mostra única, que representi els subjectes als quals s'aplicarà la prova per diagnosticar la malaltia. En aquest cas es va fer servir un disseny transversal. Una altra opció és la selecció de dos grups: malalts i no malalts. Aquest disseny és similar als estudis de casos i controls. En aquest cas, la prevalença de la malaltia molt probablement no representi la de la població i per tant dificultarà l'estimació dels valors predictius. Per tant, és important definir el context de pràctica clínica on es du a terme l'estudi, atès que els valors predictius d'una prova varien amb la prevalença de la malaltia.

**Selecció de la prova de referència.** La prova de referència és la prova que representa la veritat i amb la qual es compararan els resultats de la prova diagnòstica. D'aquesta manera, per avaluar una prova diagnòstica se seleccionarà la millor prova de referència disponible en el moment de l'estudi. Una bona prova de referència serà la que sempre és positiva en els malalts i sempre negativa en els no malalts. En el cas que la nova prova sigui millor que la prova de referència es podria arribar a concloure que aquesta última no és útil i ens trobaríem en presència del biaix del criteri de referència. També pot succeir que no existeixi una prova diagnòstica de referència; llavors es podria fer servir una combinació de proves com a estàndard de referència. En aquesta situació cal que la prova sotmesa a estudi no formi part del conjunt de proves anteriors: si no es pot incórrer en el biaix d'incorporació. Tots aquests biaixos poden

ocasionar una supra o infravaloració de la validesa diagnòstica de la prova.

**Aplicació de les proves.** Les dues proves (nova i de referència) s'han d'aplicar a tots els subjectes de l'estudi. Podria donar-se el cas que la prova de referència (per ser més invasiva o menys tolerada) només s'apliqués al grup de pacients que han donat positiu amb la nova prova. Aquesta situació podria donar lloc al biaix de verificació o de confirmació (*work-up bias*), ja que la confirmació diagnòstica només s'obté d'un grup de malalts, donant lloc a una valoració inadequada de les característiques operacionals de la nova prova.

La prova de referència s'ha d'aplicar de forma cega, és a dir, l'investigador no ha de tenir coneixement dels resultats de la prova que s'està avaluant. Així mateix l'aplicació i interpretació dels resultats de la prova que s'avalua han de ser cegues i independents de la prova de referència, de forma que els resultats que s'obtinguin no condueixin a verificacions posteriors en cas de no coincidir amb la prova de referència. D'aquesta forma s'evita el biaix de sospita diagnòstica.

### Estudis de concordança

Aquests estudis es duen a terme per avaluar la fiabilitat d'un instrument de mesura o la concordança entre dos observadors o un observador amb ell mateix, respecte d'una mesura. El terme fiabilitat fa referència al grau en què es poden reproduir els resultats obtinguts per un procediment de mesurament i és sinònim de reproductibilitat, precisió, estabilitat, consistència.

La mesura d'un fenomen està influïda per diverses fonts de variació, sovint difícils de valorar (individual de l'instrument i dels observadors). Així, quan s'avalua la fiabilitat d'una mesura s'han d'estudiar els aspectes següents:

- 1) Repetibilitat de la mesura: Suposa l'aplicació del mètode de mesura de la mateixa manera als mateixos subjectes en dos o més moments del temps (fiabilitat test-retest).
- 2) Concordança intraobservador. La seva avaluació implica que un mateix avaluador valori en dues o més ocasions els mateixos subjectes.
- 3) Concordança interobservador. La seva avaluació requereix que dos o més avaluadors valorin una mateixa prova en una mostra de subjectes.

Per avaluar la variabilitat anterior es fan servir diferents tècniques d'anàlisi estadística segons el tipus de variable. Així la concordança entre variables qualitatives es valora generalment mitjançant l'índex *Kappa* i en el cas de les variables quantitatives l'índex que resumeix el grau de concordança és el coeficient de correlació intraclass.



## Estudis descriptius longitudinals

Aquests estudis sovint es fan servir en clínica per descriure l'evolució temporal de determinades característiques en un grup de malalts que tenen en comú la malaltia, descriure la història natural d'una malaltia, el resultat d'un tractament o d'una intervenció, o estimar la incidència d'una malaltia en un grup d'individus. Els estudis més habituals amb aquest tipus de disseny són les sèries de casos longitudinals o el seguiment d'una cohort. En aquests dissenys tot i la seqüència temporal de les mesures de l'estudi no s'avalua la relació causa-efecte.

Dues característiques rellevants d'aquests estudis són per una banda que la informació sobre la malaltia o qualsevol fenomen que pugui aparèixer durant el temps de seguiment es poden relacionar amb l'evolució de la malaltia (canvis en la gravetat, tractaments previs, comorbiditats). D'altra banda, es pot fer una descripció de l'evolució d'una mostra de subjectes que tenen unes característiques comunes. Per tant, és important definir clarament els criteris de selecció dels subjectes d'estudis.

Els problemes més freqüents que poden donar una estimació dels resultats esbiaixada són el percentatge de no respostes i les pèrdues en el seguiment.

## Estudis d'incidència

Aquests estudis tenen com a objectiu estimar la incidència d'una malaltia o problema de salut en una població determinada. La incidència es defineix com el nombre de casos nous d'una malaltia que es desenvolupen en una població a risc durant un període de temps determinat (habitualment un any). Hi ha dos tipus de mesures d'incidència: la incidència acumulada i la densitat d'incidència:

**Incidència acumulada:** És la proporció de persones d'una població, inicialment lliures de la malaltia d'interès, que desenvoluparan la malaltia durant un període de temps determinat. Per tant, la incidència acumulada mesura la probabilitat o risc mitjà d'una persona a desenvolupar la malaltia a condició que no es desenvolupi una altra malaltia fatal que ocasioni la mort.

$$\text{Incidència acumulada} = \frac{\text{Nombre de nous casos en un període de temps}}{\text{Població lliure de la malaltia a l'inici del període}}$$

**Exemple:** Suposem que es vol conèixer la incidència d'al·lèrgia a la proteïna de llet de vaca en el primer any de vida. (Sanz J. et al. 2001). Es va realitzar un estudi descriptiu longitudinal prospectiu, en el qual es van estudiar 1.663 nadons durant el primer any de vida seleccionats de diferents centres de salut. Davant la sospita de reacció adversa a la llet de vaca es van enviar a l'hospital de referència per a l'estudi diagnòstic d'al·lèrgia. Es va confirmar el diagnòstic

d'al·lèrgia per IgE en 6 casos. La incidència acumulada d'al·lèrgia a la proteïna de la llet en el primer any de vida va ser del 0.36% (IA = 6/1.663). És a dir, el risc o la probabilitat d'un nadó de desenvolupar aquest tipus d'al·lèrgia en el primer any de vida va ser del 0.36%,

**Densitat d'incidència.** En general els períodes de seguiment de cada individu són diferents en durada i no es donen en el mateix moment en el temps. Per tenir en compte aquestes diferències relacionades amb el temps de seguiment es calcula la densitat d'incidència.

$$\text{Taxa d'incidència} = \frac{\text{Nombre de nous casos en un període de temps}}{\text{Suma dels períodes de risc de cadascun dels subjectes exposats durant el període de temps especificat}}$$

**Exemple:** A la figura 4 es representa un estudi que comença l'any 1995 i finalitza l'any 1998 amb 5 persones amb diferents períodes de seguiment. Dues d'aquestes persones es van incorporar més tard a la data d'inici de l'estudi (B i D) i dues van desenvolupar la malaltia (B i E). La incidència acumulada va ser de 2/5, que és el mateix que dir un risc del 40% de contreure la malaltia en 4 anys de seguiment. Però aquesta afirmació no té en compte que les persones han estat seguides durant períodes de temps diferents. En total sumen 14 persones any d'observació (p.ex. les persones B i E varen estar 3 anys a risc abans de desenvolupar la malaltia). La taxa d'incidència va ser de 2 casos/14 persones any, que és el mateix que 14,3/100 persones any de seguiment.

El concepte de persona-temps fa referència que 10 persones any pot procedir de l'observació de 10 persones durant un any cadascuna, o de 5 persones durant dos anys, o d'una persona en 10 anys.

|                                                                                     | Gener 1995 | Gener 1996 | Gener 1997 | Gener 1998 | Temps a risc |
|-------------------------------------------------------------------------------------|------------|------------|------------|------------|--------------|
| Persona A                                                                           | +-----     | -----      | --         |            | 2 anys       |
| Persona B                                                                           |            | +-----     | -----      | -----X     | 3 anys       |
| Persona C                                                                           | +-----     | -----      | -----      | -----      | 4 anys       |
| Persona D                                                                           |            |            | +-----     | -----      | 2 anys       |
| Persona E                                                                           | +-----     | -----      | -----      | X-----     | 3 anys       |
| Total d'anys que la cohort està en risc (2+3+4+2+3) = 14 anys                       |            |            |            |            |              |
| + = Inici del seguiment                                                             |            |            |            |            |              |
| -- = Període de seguiment (els subjectes estan en risc de desenvolupar la malaltia) |            |            |            |            |              |
| X = Diagnòstic de la malaltia                                                       |            |            |            |            |              |

Fig. 4. Càlcul del denominador persona-temps en la densitat d'incidència.

## Bibliografia

- 1** Argimón JM, Jiménez J. Métodos de investigación clínica y epidemiológica. Barcelona: Ed. Harcourt, 2000 (2a edició).
- 2** Baños JE, Brotons C, Farré N. Glosario de investigación clínica y epidemiológica. Barcelona: Ed. Doyma, S.A. 1998.
- 3** Fletcher RH, Fletcher WS, Wagner EH Fletcher. Epidemiología clínica: aspectos fundamentales. 2a ed. Barcelona: Masson, 1998.
- 4** Hulley SB, Cummings SR. Diseño de la investigación clínica: un enfoque epidemiológico. Barcelona. Ed. Doyma, S.A., 1993.
- 5** Kleinbaum DG, Kupper LI, Morgenstern H. Epidemiologic research. Principles and Quantitative methods. Van nostrand Reinhold, 1982.
- 6** Moyer VA, editor. Evidence based pediatrics and child health. London: BMJ Books, 2000.
- 7** Paniagua H, García S, Castellano G, Sarrallé R, Redondo C. Consumo de tabaco, alcohol y drogas no legales entre adolescentes y relación con los hábitos de vida y el entorno. *An Esp Pediatr* 2001; 55: 121-128.
- 8** Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Medicina basada en la evidencia. Como ejercer y enseñar la MBE. Churchill Livingstone España. Madrid 1997.
- 9** Sanz J, Martorell A, Michavila A, Nieto A y Grupo de Trabajo para el Estudio de la Alergia Alimentaria. Estudio de la incidencia de alergia mediada por IgE frente a la proteína de la leche de la vaca en el primer año de vida. *An Esp Pediatr* 2001; 54: 536-539.
- 10** Suárez P, Alonso JC, López AJ, Martín D, Martínez MM. Prevalença i durada de la lactància materna a Astúries. *Gac Sanit*; 15: 104-110.