

VALIDACIÓ DE MÈTODES ANALÍTICS QUALITATIUS

Itziar Ruisánchez, Esther Trullols i F. Xavier Rius.

Grup de Quimiometria i Qualimetria. Departament de Química Analítica i Química Orgànica. Universitat Rovira i Virgili. Pl. Imperial Tàrraco, 1, 43005 Tarragona.

Resum

És prou coneguda la necessitat de validar els mètodes d'anàlisi i, en aquest article particularment es parla de la validació dels mètodes d'anàlisi qualitativa. Si bé el concepte de validació no depèn exclusivament de que el mètode d'anàlisi en qüestió sigui quantitatiu o qualitatiu, a l'hora de validar els mètodes qualitativs els paràmetres de qualitat que els caracteritzen són diferents. D'aquesta manera, a més de la caracterització dels mètodes d'anàlisi qualitativa, es defineixen els paràmetres de qualitat principals com ara els falsos positius i negatius, sensibilitat, especificitat, regió d'incertesa, límit de tall o *cut-off*, etc. Finalment, es descriuen molt breument quatre opcions per tal d'establir aquests paràmetres.

Introducció als sistemes de mesura qualitativs

En els laboratoris d'anàlisi és cada vegada més habitual la introducció de sistemes de mesura de resposta ràpida que generen respostes de caire qualitatiu més que quantitatiu. La resposta qualitativa sol ser binària, de tipus SÍ/NO i pot respondre a diverses situacions: presència/absència d'un determinat analit en una mostra, presència/absència per sobre d'un nivell concret (normalment de concentració) que habitualment ve fixat per la legislació, el client o l'usuari final, etc.

Aquests sistemes s'anomenen habitualment sistemes d'*screening*. Considerant el punt de vista pràctic, el principal interès en el desenvolupament d'aquests sistemes radica en que s'utilitzen com a etapa prèvia de diferenciació de les mostres (figura 1). D'aquesta manera s'evita que totes les mostres siguin sotmeses a tot el procés de mesura químic. Únicament seguiran el procés d'anàlisi quantitatiu les mostres que donin una resposta 'SI' (positiu), les mostres en les que es detecti la

presència d'un analit o es detecti que està per sobre del nivell permès.

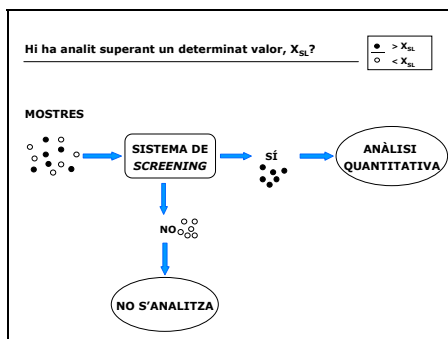


Figura 1. Sistemes d'*screening*.

Tipus de sistemes de mesura qualitatis.

Fer classificacions sempre resulta difícil donat que depèn molt dels criteris que es considerin. Si considerem el tipus de sistema utilitzat per a l'obtenció de la resposta (detecció) podem diferenciar dos grans grups: anàlisi qualitativa clàssica o sensorial i anàlisi qualitativa instrumental (Taula 1).

<p>► Detecció sensorial</p> <ul style="list-style-type: none"> □ Color: canvi, aparició, etc □ Olor □ ELISA <p>► Detecció instrumental</p> <ul style="list-style-type: none"> □ UV-Vis., Fluorescència, Potenciometria, etc. □ Sensors: químics, bio-químics, etc. □ ELISA
--

Taula 1. Tipus de sistemes de detecció utilitzats a l'*screening*.

En l'anàlisi qualitativa clàssica la detecció es fa sensorialment en base als sentits humans com ara la vista, l'olfacte, etc. El més utilitzat és la vista i la majoria de sistemes es basen en l'aparició o no d'una determinada coloració, que és la resultant d'una reacció química, bioquímica o immunològica.

En aquest grup de sistemes, la resposta binària 'SÍ/NO' s'obté de forma directa, sense cap tractament previ de les dades. Generalment, la presència d'un analit es determina mitjançant una comparació respecte d'un blanc (mostra sense l'analit), per exemple si apareix un determinat color hi ha analit, mentre que si no s'obté (color del blanc) aquest analit no és present a la mostra.

En aquest grup de sistemes qualitatis es troben continguts els anomenats '*tests kits*'. Aquests són dispositius comercials, dissenyats per a una aplicació concreta, que contenen tots els reactius necessaris i en alguns casos inclouen un sistema instrumental senzill, necessari per a l'obtenció de la resposta.

L'altre gran grup és el que hem anomenat anàlisi qualitativa

instrumental. En aquest cas, la detecció es realitza en base a una mesura instrumental (colorimetria, fluorescència, voltamperometria, etc.). La presència o no d'un determinat analit depèn del nivell en el que es vulgui detectar. Tant pot ser just al límit de detecció o a un nivell superior que generalment correspon al fixat per la legislació.

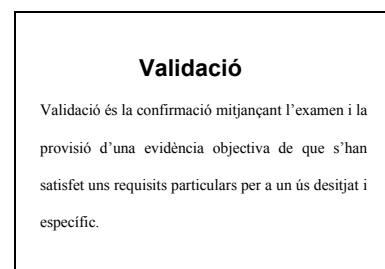
En aquest tipus de sistemes de mesura, la resposta instrumental que s'obté ha de transformar-se en resposta binària del tipus SÍ/NO i, per tant, implica un tractament de les dades. Primer s'ha d'establir la resposta instrumental, per exemple un valor d'absorbància per a la concentració corresponent al valor al que es volen separar les mostres (nivell de concentració establert per la legislació) i ha de comparar-se amb la resposta instrumental obtinguda per a cada mostra. Si és superior, hi ha analit i, per tant la resposta binària és SÍ, i si el valor és inferior, la resposta binària és NO.

Un altre grup de sistemes qualitius que podem considerar a més a més són els que utilitzen el mètode ELISA. Aquests es basen en una reacció immunològica i de la mateixa manera que els *test kits*, la

detecció pot ser tant sensorial com instrumental.

Validació dels sistemes de mesura qualitius

Com els mètodes de mesura quantitius, els mètodes qualitius també s'han de validar. La validació d'un mètode no depèn només de que sigui quantitiu o qualitiu, donat que validar consisteix en verificar i documentar la validesa, la seva adequació a uns requisits prèviament establerts. Aquesta és la idea que queda resumida en la definició proporcionada per la norma ISO 8402 (veure quadre 1) i que implica el concepte d'adequació a la finalitat o propòsit buscat o *'fit-for-purpose'*.



Quadre 1. Definició de validació segons la norma ISO 8402

Per tant, enfront de cada problemàtica concreta, la validació implica com a pas previ la definició dels requisits analítics, o paràmetres de qualitat (en anglès,

performance characteristics), que es necessiten i, un cop definits, la seva determinació.

La primera gran diferència que trobem entre l'anàlisi quantitativa i l'anàlisi qualitativa és la forma d'expressar el resultat (figura 2). És prou conegut que l'expressió d'un resultat quantitatiu es caracteritza per dos valors numèrics, el primer és l'estimació del valor vertader i el segon correspon a la incertesa que està associada a la dispersió del resultat o interval dins del qual hi ha una certa probabilitat de trobar el valor vertader.



Figura 2. Expressió del resultat analític.

El resultat d'una anàlisi qualitativa també es caracteritza per dos valors, el primer és binari 'SÍ/NO', i el segon és la probabilitat d'error associada a la decisió presa.

Paràmetres de qualitat

Els sistemes d'*screening* tenen certes connotacions especials que impliquen una acurada adaptació dels paràmetres de qualitat que estan ben definits i estudiats en els processos de mesura quantitatius, tant de tipus físic com químic. A la taula 2 es mostren els paràmetres més rellevants des del punt de vista matemàtico-estadístic, tant quantitatius com qualitius. Alguns dels paràmetres de tipus qualitatiu han estat definits recentment per l'AOAC 1995¹, però en d'altres casos encara s'està treballant en la seva definició.

Quantitativa	Binària/Cualitativa
▶ Sensibilitat, Especificitat	▶ Probabilitat de fals positiu i negatiu
▶ Selectivitat: interferents	▶ Sensibilitat, Especificitat
▶ Límit de detecció	▶ Selectivitat: interferents
▶ Rang i linealitat	▶ Límit de detecció
▶ Incertesa	▶ Límit de tall (cut-off)
▶ Exactitud: veracitat i precisió	▶ Incertesa o regió d'error
▶ Robustesa	▶ Robustesa
▶ ...	▶ ...

Taula 2. Paràmetres de qualitat.

És important destacar que alguns són específics de l'anàlisi qualitativa, com per exemple la proporció de falsos positius i negatius i el límit de tall o *cut-off*. D'altres, com l'especificitat, la sensibilitat i el límit de detecció, poden

tenir un significat lleugerament diferent tot i que mantenen igual nom que en el quantitatiu. Finalment, cal dir que la majoria dels paràmetres qualitatius, com són d'origen binari SÍ/NO, s'expressen en termes probabilístics.

De forma genèrica es podria dir que la validació de tot mètode d'anàlisi implica l'establiment dels paràmetres de qualitat considerats bàsics: traçabilitat, exactitud, representativitat, etc. Dels paràmetres de qualitat propis i característics de l'anàlisi qualitativa podem destacar els que fan referència o es relacionen amb nivells de concentració; d'aquesta manera podem diferenciar el límit de detecció, el límit de tall o *cut-off* i el límit legislatiu.

La situació habitual és aquella en la que la legislació indica el contingut màxim o mínim d'un analit concret en una mostra. Si prenem com a referència la presència d'un contaminant, la legislació fixa el contingut màxim. Per exemple, el contingut màxim d'Aflatoxina B₁ en fruita seca és de 2 ng/g.²

A partir d'ara, ens centrarem més en els mètodes qualitatius sensorials,

concretament en els *test kit* que avui dia s'utilitzen àmpliament. Un cas concret seria el que quan la mostra conté menys de 2 ng/g d'Aflatoxina B₁ la resposta obtinguda és negativa (apareix color), mentre que si la concentració és superior la resposta és positiva (no apareix color)³.

En aquest tipus de *test kits* l'ideal és que el límit real en el que el *test kit* diferencia les mostres, anomenat límit de tall o *cut-off*, coincideixi amb el límit legislatiu, però ara bé, el límit legislatiu és un límit teòric, i el de tall és un límit experimental (figura 3).

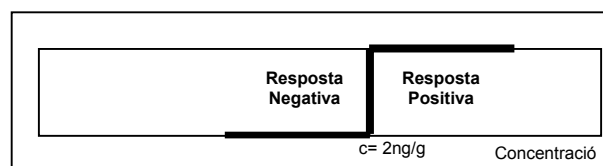


Figura 3. Resposta teòrica d'un *test kit*

Però, considerant el punt de vista experimental, s'ha de tenir present que si representem la resposta del *kit* en funció de la concentració es poden diferenciar 3 zones (figura 4): a) zona on s'obté sempre una resposta negativa (vertaders negatius), b) zona on per a una mateixa concentració es poden obtenir tant respostes positives com

negatives, correspon a la zona de falsos positius i negatius, i c) zona on s'obté sempre una resposta positiva (vertaders positius).

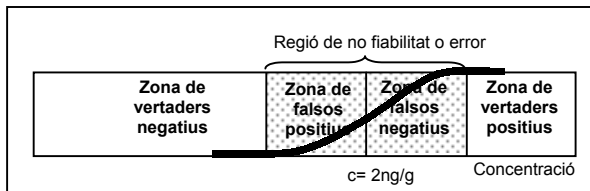


Figura 4. Resposta experimental d'un test kit

En l'exemple concret d'un contaminant com pot ser l'Aflatoxina B₁, es veu clar que interessa no donar cap fals negatiu (dir que no hi ha analit quan realment sí que n'hi ha), però d'altra banda, sí que podem acceptar falsos positius, donat que en cap cas implicaria un risc per a la salut perquè qualsevol mostra que doni una resposta positiva s'analitza posteriorment utilitzant el mètode confirmatori. En aquest cas concret, interessa que la zona b) quedi sota el límit legislatiu per a garantir l'absència de falsos negatius i a més que quedi el més propera possible al límit legislatiu, per a reduir el nombre de falsos positius, (figura 5).

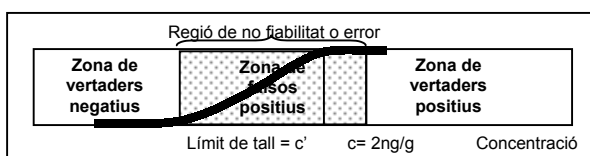


Fig. 5. Límit de tall en el cas d'un contaminant

Com es veu a les figures, al voltant del límit de tall es situa una zona o regió d'error o manca de fiabilitat (que en la literatura anglesa es denomina, *unreliability*). Aquesta zona correspon a l'interval de concentracions on s'obtenen els falsos positius i negatius, així queda definida per un valor superior i inferior de concentració d'analit en la mostra. Aquests dos tipus d'errors són dos paràmetres bàsics en la caracterització d'un sistema d'*screening* i es defineixen com⁴:

Falsos negatius. Corresponen a les mostres que contenen un o més analits superant el valor límit permès (límit legislatiu) i que en aplicar el test d'*screening* donen una resposta negativa. És a dir, com a resultat del test d'*screening* es conclou que la mostra no conté analit superant aquest nivell màxim establert quan realment sí que el conté.

Falsos positius. Corresponen a les mostres que realment no contenen analit superant el nivell màxim permès i que, tot i això, el test d'*screening* indica que

l'analit sí que supera aquest nivell de concentració.

Hi ha dos paràmetres que estan relacionats amb els límits inferior i superior de concentració que defineixen la zona de no fiabilitat i que són la sensibilitat i l'especificitat. Cal dir que, com ja s'ha esmentat anteriorment, ambdós paràmetres s'expressen com a probabilitats. La sensibilitat està relacionada amb el valor superior de la regió de no fiabilitat, mentre que l'especificitat està relacionada amb el valor inferior d'aquesta regió.

La **sensibilitat** es defineix com la capacitat o l'habilitat del sistema d'*screening* de detectar mostres positives quan realment són positives. De forma similar, l'**especificitat** es defineix com la capacitat o l'habilitat del sistema d'*screening* de detectar mostres negatives quan realment són negatives.

Finalment, d'igual manera que en els sistemes de mesura quantitativs, cal comprovar que el mètode que s'utilitza per a diferenciar les mostres tingui un límit de detecció inferior o igual al límit de tall.

Mètodes per a caracteritzar un sistema d'*screening*

S'han desenvolupat diversos mètodes que permeten caracteritzar un sistema d'*screening*. Cada opció té diferent grau d'adaptació a les diferents situacions i problemàtiques que poden presentar-se, atenent al tipus de mesura o resultat obtingut.

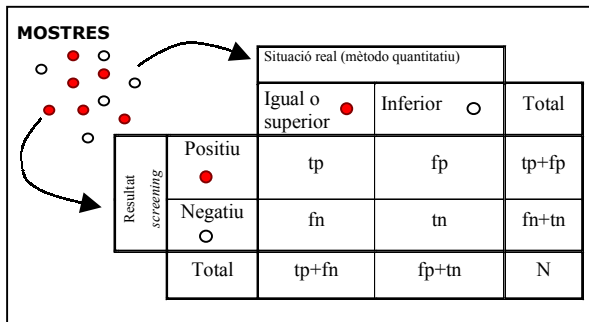
Dos d'aquests mètodes es fonamenten en el concepte de probabilitat i en aquest sentit associen una probabilitat al resultat de la mesura: les taules de contingència i el teorema de Bayes⁵. Un tercer mètode es basa en l'establiment de les corbes característiques de funcionament⁶, i finalment, es descriu un mètode basat en l'aplicació dels tests d'hipòtesis⁷. Aquest darrer és similar al que s'utilitza per a l'establiment dels paràmetres de qualitat en l'anàlisi quantitativa.

Tot seguit presentem unes pinzellades de cadascun d'aquests mètodes sense tractar-los a fons, però donant una idea general de quan seria més apropiat adoptar cadascun d'ells.

Taules de Contingència

Les més senzilles diferencien les mostres en dues categories, positiva/negativa segons el mètode d'*screening*, i construeixen una taula de comparació respecte al resultat obtingut mitjançant un mètode de referència o confirmatori (taula 3).

A partir de la taula, es calculen els quatre paràmetres bàsics, definits en la secció anterior: falsos positius, falsos negatius, sensibilitat i especificitat.



Taula 3. Taula de contingència amb 2 categories. Fp: falsos positius, fn: falsos negatius, tp: total positius, tn: total negatius, N: número total d'assaigs realitzats.

Un dels avantatges més importants és la fàcil aplicació a múltiples tipus de bioassaigs, camp en el que s'han fet la majoria de les aplicacions.^{8,9}

És interessant ressaltar que aquests paràmetres donen una mesura global de la capacitat del mètode d'*screening*. Això significa que es suposa que la mostra problema que s'ha d'examinar es comportarà estadísticament de forma similar a les ja analitzades, per tant, no es calcula una probabilitat d'error per a cada mostra en particular.

Un altre inconvenient destacable és que la capacitat de la taula depèn del número de mostres analitzades i el disseny experimental utilitzat per a elaborar aquesta taula i que, pel càlcul dels paràmetres de qualitat anomenats, totes les mostres han d'analitzar-se pels dos mètodes, el d'*screening* i el confirmatori.

Teorema de Bayes

El teorema de Bayes es basa en la teoria de probabilitats, per tant, considerant el punt de vista analític, la terminologia utilitzada és complexa i presenta la dificultat del càlcul de les diferents probabilitats intermèdies. En concret es calcula la probabilitat de donar com a resultat vàlid (positiu o negatiu) un resultat quant realment ho és (vàlid), $P(a/p)$. Aquesta probabilitat s'anomena condicional.

Cal destacar que obtenir una bona estimació de la incertesa o probabilitat d'error mitjançant aquest procediment implica un nombre elevat de mostres. Si es compara amb les taules de contingència, el teorema de Bayes sí que permet una estima individual (per a cada mostra analitzada) de la incertesa associada o probabilitat de donar un resultat erroni.

Corbes característiques

Aquest mètode consisteix en representar la probabilitat d'obtenir resultats positius a diferents nivells de concentració. Aquesta representació és sigmoïdal, i la posició i amplitud de la corba és característica de cada sistema d'*screening* (figura 6) ^{10,11}.

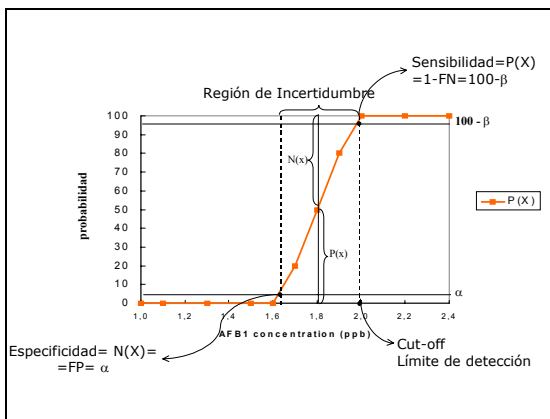


Figura 6. Corba característica de funció.

El principal problema d'aquest mètode, com també s'ha fet esment en el mètode de Bayes, és l'elevat nombre d'experiments que s'han de realitzar, donat que convé disposar d'un nombre prou gran d'anàlisis a diferents nivells de concentració.

Presenta l'avantatge, tal i com es veu a la figura 6, que a més dels quatre paràmetres bàsics (falsos positius, falsos negatius, sensibilitat i especificitat), s'obté molta informació addicional sobre les característiques del mètode d'*screening*: el límit de tall, el límit de detecció, la regió d'incertesa, etc.

Aplicació dels tests d'hipòtesis

En aquest cas, s'estableix el valor de resposta instrumental al nivell de concentració al que es pretén diferenciar les mostres (normalment amb un patró) i per a decidir si una mostra problema és positiva o negativa, es compara el valor de la resposta instrumental de la mostra problema amb el valor prèviament establert (del patró). Aquesta comparació es realitza mitjançant tests d'hipòtesis.⁷

Com a avantatge, es presenta la familiaritat des del punt de vista analític, per a treballar amb els tests d'hipòtesis. Si bé la probabilitat α d'error (probabilitat de cometre falsos positius) és majoritàriament coneguda i utilitzada en els laboratoris d'anàlisi, no és igual el cas de la probabilitat β (probabilitat de cometre falsos negatius).

Cal afegir, que és un mètode ràpid i senzill, de fàcil automatització i que

permet l'assignació d'una incertesa o probabilitat d'equivocar-nos utilitzant com a tècnica d'*screening* una tècnica instrumental. Però el principal inconvenient és que no és directament aplicable quan la tècnica d'*screening* utilitzada és un *kit* no instrumental o es basa en una observació visual no quantificable per part de l'analista.

Els autors agraeixen tots els comentaris relacionats amb els continguts d'aquest article. Poden dirigir-se, mitjançant correu electrònic, a l'adreça: quimio@quimica.urv.es. Poden trobar una versió en suport electrònic d'aquest article i informació suplementària a: <http://www.quimica.urv.es/quimio>

Bibliografia

-
- ¹ P. Feldsine, C. Abeyta, W. Andrews. AOAC International Method Committee. Guidelines for Validation of qualitative and quantitative microbiological official methods of analysis. Draft, November 2000. URL : <http://aoac.org/vmeth/MicrobiologyGuidelines112700.htm>.
 - ² Commission Regulation (EC) No 257/2002. Official Journal of the European Community, 12.2.2002, No. L 041, pp12-15.
 - ³ www.rhone-diagnostics.co.uk
 - ⁴ Massart et al. Data Handling in science and technology 20A. Handbook of Chemometrics and Qualimetrics: Part A. Ed. Elsevier Science, Amsterdam (1997).
 - ⁵ R.M. Mc Fall and T.A. Treat; Annu. Rev. Psychol; 50 (1999) 215-241

-
- ⁶ EURACHEM. Quantifying uncertainty in analytical measurement, Draft: EURACHEM Workshop, 2nd Edition, Helsinki, 1999.
- ⁷ A. Pulido, I. Ruisánchez, R. Boqué and F.X. Rius. *Analítica. Chim. Acta*; 455 (2002) 267-275.
- ⁸ B.J. Neil, E. Keeler and S. J. Adelstein; *The New England Journal of Medicine*; 293 (1975) 211-215.
- ⁹ N. E. Hawass; *The British Journal of Radiology*, 70 (1997) 360-366.
- ¹⁰ K. Ashley, R. Song and P. C. Schlecht,. *Journal of Hazardous Materials* 83 (2001) 29-39.
- ¹¹ K. Ashley, R. Song, P. C. Schlecht. *American Laboratory* 34 (12) (2002) 32-39.