



Este material está protegido bajo la licencia de **Creative Commons 4.0. Reconocimiento- No Comercial - Sin Obra Derivada**

EPIDEMIOLOGÍA DIAGNÓSTICA

Los conceptos epidemiológicos relacionados con el diagnóstico suelen agruparse bajo el apartado de “Epidemiología serológica” pues muchos diagnósticos se basan en la detección de anticuerpos específicos en el suero sanguíneo. Sin embargo, hay otras opciones diagnósticas además del diagnóstico serológico, por lo que es más correcto hablar de “Epidemiología Diagnóstica”. Así, se puede recurrir a la detección de material genético a través de técnicas de biología molecular (PCR), la detección de portadores fecales de *Salmonella* en aves en programas oficiales de control de salmonelosis, o al recuento de células somáticas o bacteriológico en programas de calidad de la leche, entre otros ejemplos.

Una prueba diagnóstica puede definirse como “la utilización de uno o varios de nuestros sentidos durante el proceso de diagnóstico para la detección de signos o alteraciones de los tejidos” (Martin, 1997). El diagnóstico patognomónico (predicción absoluta de la enfermedad) resulta complejo, y costoso, no puede generalizarse a nivel poblacional y no se adapta a los estudios epidemiológicos. Debido al enfoque del diagnóstico poblacional (epidemiología secundaria, Martin, 1997), se necesitan pruebas masivas o “de cribado” (*screening*). Estas técnicas deben ser automatizables (ELISA, PCR en tiempo real, etc.), y su falta de predicción absoluta genera falsos diagnósticos. Así, un diagnóstico falso positivo (o negativo), es el que considera un individuo enfermo o infectado cuando en realidad no lo está.

MEDICIÓN DE ENFERMEDAD/INFECCIÓN

Uno de los objetivos de la investigación epidemiológica es medir la enfermedad para asociarla con posibles factores de riesgo. Esta medición de la enfermedad (o de la infección) genera una variable de naturaleza cualitativa que suele expresarse de forma dicotómica (binomial). Aunque algunas técnicas diagnósticas pueden generar variables numéricas (ej, densidades ópticas, recuentos celulares, ...), o cualitativas ordinales (0, +, ++, +++), sus resultados se transforman a una escala binomial y calificamos a los individuos como sanos/enfermos (no infectados/infectados). Este formato de variable binomial es el que mayoritariamente se utiliza en los estudios epidemiológicos, para asociar factores con enfermedades (tabla 2x2).

Debido a la existencia de errores diagnósticos (sesgo de medición) al usar técnicas que no son 100% perfectas, los resultados diagnósticos deben corregirse. Para corregir este sesgo de medición necesitamos validar la técnica, obteniendo sus parámetros de validez (sensibilidad, especificidad y valoresictivos). Especialmente importante es la estandarización de las técnicas diagnósticas para el diagnóstico de las enfermedades de declaración obligatoria (EDOs). La Organización Mundial de Sanidad Animal (OMSA), antiguamente conocida como la Organización Internacional de Epizootias (OIE) tiene un listado de técnicas estandarizadas de uso para el diagnóstico de las EDOs, y en esa estandarización se incluyen sus reactivos. La descripción de cada prueba la podemos encontrar en el “Manual de normas para pruebas de diagnóstico y vacunas” de la OMSA (2023).

Los objetivos principales del diagnóstico pueden ser confirmar la etiología de una enfermedad clínica (ámbito de la medicina individual más que la colectiva), detectar enfermos subclínicos, asintomáticos o portadores sanos. Se busca una doble finalidad: terapéutica, con el objetivo de eliminar las infecciones existentes (reducción de la prevalencia) y preventiva, pues al eliminar los animales infectados minimizamos el contagio al resto de animales sanos del colectivo (reducción de la incidencia).

El citado manual de la OMSA (2023), plantea diferentes finalidades del diagnóstico, entre ellas:

1) Contribuir a demostrar la ausencia de infección en una población definida (país/zona/compartimiento/rebaño) (prevalencia aparente del 0%):

- 1a) “Libre” con y/o sin vacunación,
- 1b) Restablecimiento de la ausencia después de los brotes

2) Certificar la ausencia de infección o presencia del agente causal en animales concretos o productos utilizados para el comercio o el transporte.

3) Contribuir a la erradicación de la enfermedad o eliminación de la infección en poblaciones definidas.

- 4) Confirmar el diagnóstico de casos clínicos o sospechosos (incluye la confirmación de un resultado positivo en una prueba de cribado).
- 5) Estimar la prevalencia de la infección o exposición para facilitar el análisis de riesgos (encuestas, situación sanitaria del rebaño respecto a la infección, medidas para el control de la enfermedad).
- 6) Determinar el estado inmunitario de animales determinados o de poblaciones (después de la vacunación).

Las técnicas de biología molecular, como la reacción en cadena de la polimerasa (PCR), han revolucionado el diagnóstico por sus altos parámetros de validez, reemplazando muchos de los métodos directos clásicos de detección de agentes patógenos. Entre sus diferentes modalidades, destaca la PCR en tiempo real, por su aceptable sensibilidad (próxima o igual a la PCR anidada tradicional) su menor riesgo de contaminación, permitir la realización de ensayos cuantitativos y su capacidad de automatización (OMSA, 2022).

FALSOS DIAGNÓSTICOS: INTERPRETACIÓN DE LA POSITIVIDAD/NEGATIVIDAD

Previo al diagnóstico laboratorial, se encadenan sucesos como la selección de las unidades de muestreo, toma de muestras, identificación, registro, conservación, transporte, almacenamiento, selección de la técnica y reactivos adecuados, procesamiento de los protocolos diagnósticos y generación y gestión de los resultados, entre otros. Cualquier eslabón de este proceso puede verse afectado por errores, por lo que la probabilidad de encontrar un falso diagnóstico estará siempre presente.

Al margen de lo anterior, las particularidades inmunológicas de las infecciones condicionan la veracidad del resultado. Así, la presencia de anticuerpos frente a un agente determinado (seropositividad) suele indicar la exposición previa a un agente que, en la mayoría de las ocasiones, comienza a partir de un período de una o dos semanas tras la exposición. En ocasiones, la ausencia de anticuerpos (seroconversión retardada, ausencia de seroconversión tras la exposición o pérdida de una seropositividad previa, entre otros) origina falsos negativos. A veces, los anticuerpos están presentes, pero no se detectan. En ocasiones (lentivirus), los anticuerpos pueden tardar 6 meses, o más, en aparecer tras la infección, y algunos animales previamente seropositivos pueden perder el título (seronegativizar) a pesar de estar permanentemente infectados (artritis encefalitis caprina o maedi visna ovino). De forma similar, un diagnóstico positivo no siempre significa que el individuo esté infectado, debido a la existencia de falsos diagnósticos positivos.

A continuación, se incluyen algunos ejemplos de falsos diagnósticos, pero las situaciones pueden ser muy variadas y le corresponde al profesional estar al corriente, en cada caso, de las particularidades etiológicas, epidemiológicas, inmunológicas y diagnósticas de las enfermedades objeto de estudio. Un porcentaje importante de falsos diagnósticos son achacables a los responsables de la toma, conservación y manipulación de las muestras hasta su traslado al laboratorio. Sin un adecuado control de calidad de esta primera fase, todo el posterior esfuerzo analítico del laboratorio acarreará errores.

Falsos positivos

Como se verá más adelante, a menor especificidad de la técnica, se detectará un mayor porcentaje de falsos positivos, y, de forma similar, una técnica muy sensible no produce mayor porcentaje de falsos positivos “*per se*” (este es un error muy común), sino porque una técnica que sea más sensible que específica generará un mayor porcentaje de falsos positivos, frente al porcentaje de falsos negativos producidos debido su menor especificidad.

Inmunidad maternal. Efecto a tener en cuenta en los animales en sus primeras semanas. Los anticuerpos maternos alcanzan los valores séricos más altos de 12 a 24 horas después del nacimiento, concentración que disminuirá debido al catabolismo de las proteínas. La velocidad catabólica se expresa en términos de vida media (tiempo que tarda en reducirse a la mitad la concentración de inmunoglobulinas).

Inmunización. La inmunización pasiva (sueroterapia) se usa solo ocasionalmente en las poblaciones animales, pero la inmunización activa (vacunación), muy extendida, puede generar diagnósticos falsos positivos. Las vacunas marcadas permiten inmunizar sin inducir falsos positivos debido a la vacunación (p. ej.: enfermedad de Aujeszky) y también son útiles las que protegen, pero cuyos anticuerpos dejan de ser detectables de forma precoz, como la que se utilizó en la campaña de erradicación de la brucelosis (Cepa Rev-1, dosis reducidas vía conjuntival).

Recuperación de una infección y presencia de anticuerpos. Supone el mejor ejemplo de que seropositividad e infección no deben utilizarse como sinónimos.

Inmunidad cruzada con anticuerpos frente a otros antígenos. Los anticuerpos de cerdos infectados por el virus de la diarrea vírica bovina reaccionan frente al virus de la peste porcina clásica (Pestivirus); algunos animales seropositivos a *Brucella* spp. son en realidad animales infectados con *Yersinia enterocolitica* serotipo 09, etc.

Reacciones inespecíficas debidas a inhibidores no específicos (inhibición de la hemaglutinación), de aglutininas inespecíficas (aglutinación) o de proteínas antigénicas fusionadas con glutatión S-transferasa (GST) en las técnicas ELISA, entre otros.

Técnicas diagnósticas inadecuadas, errores laborales, del equipamiento, de los reactivos, humanos, etc.

Falsos negativos

Similarmente a lo comentado para los falsos positivos, una menor sensibilidad de la técnica generará mayor porcentaje de falsos negativos, y viceversa. Igualmente, una técnica muy específica no produce mayor porcentaje de falsos negativos “*per se*”, sino porque al ser más específica que sensible, su menor sensibilidad generará mayor porcentaje de falsos negativos.

Incorrecta toma de muestras. Tras un aborto infeccioso, o tras un parto normal, las inmunoglobulinas se concentran en el calostro y los títulos de anticuerpos no se recuperan hasta pasada, al menos, una semana; también destacan los falsos negativos originados al someter al cultivo bacteriológico muestras de animales tratados con antibióticos, entre otros ejemplos.

Incorrecta manipulación o conservación de las muestras. La remisión al laboratorio de sueros hemolizados o contaminados es un punto crítico a controlar en las campañas de saneamiento.

Inmunotolerancia. Paradójicamente, animales que nacen infectados resultan seronegativos (p. ej.: animales persistentemente infectados –API– con el virus de la diarrea vírica bovina-BVD-).

Pérdida del título serológico. Algunos animales infectados y seropositivos pueden perder el título y pasar al estado de seronegativos (p. ej. lentivirus).

Utilización de técnicas diagnósticas inadecuadas: presencia de anticuerpos bloqueantes (el exceso de IgG₁ puede bloquear las IgG₂); errores humanos, laborales, del equipamiento, de los reactivos, etc.

PRESENTACIÓN DE LOS DATOS SEROLÓGICOS

Título serológico: Dilución más alta de suero que produce reacción al aplicar una técnica (menor cantidad de suero que permite detectar anticuerpos). Así, un título 1/32 significa que el suero sin diluir tiene 32 veces los anticuerpos necesarios para la reacción.

La vida media del título es el período que tardan los anticuerpos en disminuir a la mitad y varía con el tipo de infección. Los títulos de algunos anticuerpos perduran debido a una vida media elevada, una infección persistente o un contacto repetido con el agente. Una elevada persistencia de los anticuerpos no siempre está relacionada con la presencia del agente, sino que puede deberse a una elevada vida media, por lo que será necesario tener en consideración la cinética de los anticuerpos.

Por otra parte, una toma de muestras separadas dos semanas permite conocer si los títulos ascienden (respuesta reciente) se estabilizan o descienden (respuesta lejana).

Dilución de sueros. Los sueros suelen diluirse siguiendo series geométricas, la más habitual es la dilución doble:

$$1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128, \dots$$

A veces, para evitar los efectos indeseables del suero concentrado (reacciones inespecíficas) se realiza una dilución previa de 1/10, y a partir de ahí se continúa con diluciones dobles:

$$1/20, 1/40, 1/80, 1/160, 1/320, \dots$$

La distribución de los resultados serológicos mediante series geométricas impide la aplicación de técnicas estadísticas paramétricas sobre ellas, y deben usarse distribuciones lognormales. Para ello se realiza previamente una transformación logarítmica en base 2:

$$1/2^1, 1/2^2, 1/2^3, 1/2^4, 1/2^5$$

Codificación. Los títulos se codifican siguiendo la transformación logarítmica anterior. Así los títulos del ejemplo anterior se transformarían en:

$$1, 2, 3, 4, 5.$$

Títulos medios. Cuando tenemos varios títulos codificados podemos calcular su media aritmética. Ejemplo: 1/4, 1/16, 1/32, 1/8. Se transforman en 2, 4, 5, 3.

$$\text{Media aritmética} = \frac{2+4+5+3}{4} = 3,5 \text{ (título medio)}$$

El Título medio geométrico (TMG) es el antilogaritmo en base 2 de la media codificada. Hay que recordar que no se deben usar en este contexto resultados diagnósticos negativos (el logaritmo de 0 no se puede calcular). Para el ejemplo anterior sería

$$\text{TMG} = 2^{4,3} = 11,31$$

Cuando interpretamos los títulos de las poblaciones hay que considerar, tanto el porcentaje de animales seropositivos de la explotación, como el TMG de los animales seropositivos. Así, no es lo mismo, en dos explotaciones con un 20% de seropositivos, que el TMG sea de 20 o de 1280. La primera situación indicaría la persistencia de anticuerpos tras un episodio alejado en el tiempo, mientras que la segunda indica una reacción inmunológicamente activa (infección o reinfección reciente).

Dilución seriada simple, es la que habitualmente se realiza en campañas de saneamiento y supone una valoración débil. Lo ideal es hacer ensayos de dilución seriada múltiple, en los que cada dilución se valora al menos cinco veces y el punto límite que se considera es el que produce, al menos, resultados positivos en el 50% de los casos.

Perfiles serológicos

El estudio serológico de una explotación, permite determinar el momento idóneo en el que aplicar medidas preventivas (medicación, inmunoprofilaxis,...) con el fin de optimizar los parámetros productivos. Podemos distinguir tres tipos de perfiles serológicos:

Perfil serológico ocasional: sólo se dispone de un lote de muestras.

Perfil serológico pareado: se toman dos grupos de muestras de los mismos animales con un periodo de tiempo de un mes entre ambos muestreos. Los animales que tengan un título cuádruple en el segundo muestreo o que pasen de seronegativos a seropositivos se consideran infectados (demostración de la seroconversión).

Perfil serológico secuencial: constituye el método más útil. El muestreo se realiza de forma periódica sobre un número de animales elegidos, aleatoriamente o de forma predeterminada, en sus distintas fases de desarrollo o producción. Estos seroperfiles presentan la ventaja de permiten que con la misma muestra de suero se puedan estudiar varios agentes patógenos y nos dan información del momento en el que aparecen en la granja. Además, nos permite valorar la tasa de anticuerpos maternos presente en las crías, aumentando la efectividad de las vacunas al ponerlas en el momento más idóneo, y evitando fallos debido a la interferencia de la inmunidad pasiva (anticuerpos maternos), etc.

VALIDACIÓN DE LAS PRUEBAS DIAGNÓSTICAS

Para estandarizar una prueba (conocer su validez) hay que comparar sus resultados con otra que refleje la realidad. Los elementos que definen correctamente a una técnica diagnóstica son sus parámetros de validez: sensibilidad, especificidad y valores predictivos positivo y negativo. Cualquier estudio epidemiológico exige conocer la validez del método diagnóstico utilizado para poder corregir el sesgo de medición cometido. Para validar una prueba se comparan sus resultados con los resultados reales, para lo que habría que contrastarlo frente a un diagnóstico perfecto. Esto se puede conseguir a través de la “prueba de oro o *gold standard*”, o mediante una combinación de varias técnicas que nos aseguren un resultado “perfecto”. En una tabla de contingencia 2x2 se incluyen los resultados verdaderos y falsos, tanto para los positivos como para los negativos y a partir de ellos se obtienen los

parámetros de validez de la técnica. Al ser probabilidades, los resultados de su cálculo se expresan entre 0 y 1, si bien (al igual que ocurre con los indicadores de salud como prevalencias o incidencias), se suelen transformar en porcentajes.

		REALIDAD		
		POSITIVO	NEGATIVO	
TÉCNICA	POSITIVO	a	b	a + b
	NEGATIVO	c	d	c + d
		a + c	b + d	N = a + b + c + d

a = verdaderos positivos
b = falsos positivos
c = falsos negativos
d = verdaderos negativos

		Realidad	
		Positivo	Negativo
Técnica	Positivo	VP	FP
	Negativo	FN	VN

No hay que confundir el concepto de “verdadero positivo” (VP) o “verdadero negativo” (VN) (a o d, respectivamente) con los positivos o los negativos que hay en la realidad (a+c) o (b+d), respectivamente. La denominación de “verdadero positivo” o “verdadero negativo” nos puede inducir a confusión. Así, el significado epidemiológico de “verdadero positivo” corresponde con un resultado positivo obtenido por la técnica que, además, es

positivo en la realidad, es decir que cumple la condición de dar positivo en ambas situaciones. De forma similar debemos interpretar el concepto de “verdaderos negativos”.

PARAMETROS DE VALIDEZ INHERENTES A LA TÉCNICA

Sensibilidad de la técnica: proporción de verdaderos positivos que la técnica detecta del total de positivos que existen en la realidad:

$$S = a/(a+c)$$

$$S = VP/(VP+FN)$$

Para calcular la sensibilidad intervienen sólo los verdaderos positivos (VP) y los falsos negativos (VN). Una técnica será más sensible cuanto menos falsos negativos (FN) detecte. En el cálculo de la sensibilidad no intervienen los falsos positivos (FP).

Especificidad de la técnica: proporción de verdaderos negativos que la técnica detecta del total de negativos que existen en la realidad:

$$E = d/(b+d)$$

$$E = VN/(FP+VN)$$

Para calcular la especificidad intervienen los verdaderos negativos (d) y los falsos positivos (b). Una técnica será más específica cuanto menos falsos positivos detecte. En el cálculo de la especificidad no participan los falsos negativos.

La sensibilidad y la especificidad son parámetros inherentes a la técnica, por lo que no varían. Valoran la probabilidad que tiene una técnica de detectar los positivos o los negativos que existen en la realidad, respectivamente. Predicen, a priori, la eficacia que la técnica va a tener cuando se ejecute en muestras de diferentes poblaciones, con independencia de la prevalencia existente en cada población. Así, una sensibilidad del 98% supone que la técnica detectará como positivos el 98% de los positivos que existen en la realidad.

La mejor aplicación diagnóstica de la técnica vendrá condicionada por su S y E. Así, cuando necesitemos detectar el mayor número de positivos posibles habrá que usar técnicas más sensibles y por el contrario necesitaremos técnicas más específicas cuando nuestra orientación diagnóstica prioritaria sea la detección de individuos negativos.

PARAMETROS DE VALIDEZ RELACIONADOS CON EL RESULTADO

Valor predictivo del resultado positivo (VPP): proporción de verdaderos positivos del total de positivos detectados por la técnica:

$$VPP = a/(a+b)$$

$$VPP = VP/(VP+FP)$$

Valor predictivo del resultado negativo (VPN): proporción de verdaderos negativos del total de negativos detectados por la técnica:

$$VPN = d/(c+d) \quad VPN = VN/(FN+VN)$$

Los valores predictivos (VPs) valoran la probabilidad que tienen los resultados obtenidos por la técnica (positivos o negativos) de serlo realmente. Los VPs permiten interpretar, a posteriori, el resultado ya obtenido en una población concreta y por lo tanto con una prevalencia determinada (p. ej. un VPP del 88% supone que las muestras detectadas como positivas tienen una probabilidad de 0.88 de ser verdaderos positivos). Por ello, al contrario de lo que ocurre con la S y la E, los VPs se modifican con la prevalencia de las diferentes poblaciones.

El vídeo sobre validación de técnicas “Parámetros de validez de técnicas diagnósticas. Desarrollo mediante un ejemplo” <https://tv.um.es/video?id=82661&cod=a1> (Contreras, 2016) plantea un ejercicio de validación y se recomienda su visualización para preparar la sesión práctica de epidemiología diagnóstica.

OTROS INDICADORES DIAGNÓSTICOS

Además de los cuatro parámetros de validez anteriores, existen otros indicadores de la eficacia diagnóstica de las técnicas.

J de Youden: Este indicador no depende de la prevalencia, únicamente de la S y E, y mientras más se acerque su valor a “1”, mayor será la calidad del diagnóstico. Su limitación radica en que no indica cuál es la mejor orientación diagnóstica de la técnica.

$$J \text{ de Youden: } (Sensibilidad + Especificidad - 1)$$

Fiabilidad: valora el porcentaje de diagnósticos correctos (positivos o negativos) que se detectan con la técnica (porcentaje de muestras correctamente clasificadas). De forma similar a la J de Youden, aporta una información genérica sobre la orientación diagnóstica, por lo que es preferible usar los parámetros de validez (S y E). Se calcula sumando las proporciones de los verdaderos positivos y los verdaderos negativos detectados frente al total de las muestras estudiadas:

$$F = [a/(a+b+c+d)] + [d/(a+b+c+d)]$$

La gran limitación de estos dos parámetros consiste en que valoran conjuntamente la bondad de una prueba, sin saber si la técnica sería de elección para detectar animales enfermos (más sensible) o animales sanos (más específica) y en qué grado están equilibradas la sensibilidad y la especificidad, o si por el contrario alguno de estos dos parámetros de validez es muy elevado respecto al otro.

Al final del documento se incluyen algunos ejercicios relacionados con la validación de las técnicas. Los programas de epidemiología diagnóstica como *Win Epi* (De Blas, 2016) contienen aplicaciones (*Diagnóstico /Evaluación de prueba*) que nos facilitan estos cálculos y nos aportan los intervalos de confianza.

SENSIBILIDAD, ESPECIFICIDAD, PREVALENCIA APARENTE Y PREVALENCIA REAL

La prevalencia real cuantifica la proporción de enfermedad que existe en la población y corresponde a la relación entre los positivos que existen en la realidad (sean detectados o no por la técnica) y el total de muestras estudiadas.

$$\text{Prevalencia real} = (a+c) / (a+b+c+d)$$

Por el contrario, cuando desconocemos la realidad y valoramos la situación mediante una técnica imperfecta, sólo podemos conocer la prevalencia aparente o prevalencia detectada por la técnica, es decir la relación entre los positivos detectados por la técnica (ya sean verdaderos positivos o falsos positivos) y el total de muestras estudiadas.

$$\text{Prevalencia aparente} = (a+b) / (a+b+c+d)$$

La prevalencia aparente coincidirá con la real cuando la técnica sea perfecta y no detecte falsos diagnósticos ($b+c=0$), o cuando los falsos diagnósticos se equilibren, es decir el número de falsos positivos coincida con los falsos negativos ($b=c$). Sin embargo, en esta última situación, aunque ambas prevalencias coincidan, tendremos muestras incorrectamente clasificadas como positivas y negativas, respectivamente.

La prevalencia aparente contiene tanto los verdaderos positivos como los falsos positivos y el conocer la S y la E de una técnica permite calcular el sesgo cometido al calcular la prevalencia. Por ello, a partir de la prevalencia obtenida según la técnica (P_t , o prevalencia aparente) podemos conocer la prevalencia real (P_r)

		REALIDAD		
		POSITIVO	NEGATIVO	
TECNICA	POSITIVO	$S \times P_r$	$(1-E) \times (1-P_r)$	P_t
	NEGATIVO	$(1-S) \times P_r$	$E \times (1-P_r)$	$1-P_t$
		P_r	$1-P_r$	$N=1$

$$\text{Prevalencia aparente} = S \times P_r + (1-E) \times (1-P_r)$$

Prevalencia real =

$$P_r = \frac{P_t + E - 1}{S + E - 1}$$

Si, utilizando datos de la misma población, comparamos técnicas con igual sensibilidad, la prevalencia aparente será menor conforme aumenta la especificidad (a mayor especificidad menor número de falsos diagnósticos positivos). Similarmente, si comparamos técnicas con igual especificidad, la prevalencia aparente será mayor conforme aumenta la sensibilidad (a mayor sensibilidad mayor número de verdaderos diagnósticos positivos detectados). En la figura adjunta se incluye esta comparación efectuada con el programa *Win Episcopo* (Thrusfield et al., 2001).

Tres técnicas diferentes, datos de la misma población

Enfermos		No Enfermos		Total
Si	No	Si	No	
Test +	82	10	82	100
Test -	8	100	114	122
Total	90	110	200	

Enfermos		No Enfermos		Total
Si	No	Si	No	
Test +	82	4	88	100
Test -	8	100	114	122
Total	90	110	200	

Enfermos		No Enfermos		Total
Si	No	Si	No	
Test +	82	0	88	100
Test -	8	110	118	122
Total	90	110	200	

RESULTADOS		%	
Sensibilidad	91,111		
Especificidad	90,909		
Prevalencia Verdadera	45,000		
Prevalencia Aparente	46,000		
Valor Predictivo +	85,139		
Valor Predictivo -	92,593		
J de Youden	0,8202		

RESULTADOS		%	
Sensibilidad	91,111		
Especificidad	98,364		
Prevalencia Verdadera	45,000		
Prevalencia Aparente	43,000		
Valor Predictivo +	85,349		
Valor Predictivo -	92,902		
J de Youden	0,8747		

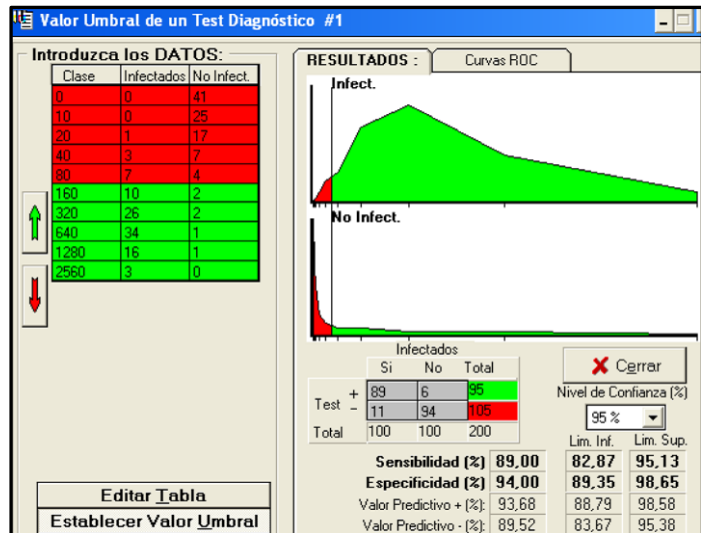
RESULTADOS		%	
Sensibilidad	91,111		
Especificidad	100,000		
Prevalencia Verdadera	45,000		
Prevalencia Aparente	41,000		
Valor Predictivo +	100,000		
Valor Predictivo -	93,228		
J de Youden	0,9111		

Misma sensibilidad
Especificidad en aumento
Prevalencia aparente descende

RELACIÓN ENTRE SENSIBILIDAD Y ESPECIFICIDAD DE LA TÉCNICA: VALOR UMBRAL

Una técnica se define por sus valores de sensibilidad y especificidad, y dichos parámetros son fijos (inherentes) a la técnica. Sin embargo, es posible modificarlos internamente, aumentando uno a expensas del otro debido a la existencia de una relación inversa entre sensibilidad y especificidad. Esta modificación se conoce como modificación del umbral o punto de corte (*cut off*). Así, si alteramos el punto de corte o umbral aumentando la sensibilidad de la técnica, disminuirá su especificidad. Por ello, al aumentar la sensibilidad aumenta la detección de positivos, tanto verdaderos positivos como falsos positivos, disminuyendo la especificidad. Esta modificación del punto de corte para clasificar los resultados tiene interés ya que una misma técnica puede utilizarse con más sensibilidad (y menor especificidad), o viceversa, dependiendo del objetivo del diagnóstico. Si nuestro interés fuera detectar el mayor número de verdaderos infectados podemos aumentar la sensibilidad a costa de la especificidad y, con ello disminuirán los falsos negativos, aunque aumentarán los falsos positivos. Por el contrario, cuando el objetivo sea detectar el mayor número posible de animales sanos podríamos aumentar la especificidad (en detrimento de la sensibilidad), disminuyendo los falsos positivos, si bien aumentarán los falsos negativos.

A veces es necesario determinar el valor umbral (valor límite para la positividad y negatividad) para un conjunto de resultados de una prueba. Una de las opciones para el cálculo de este valor umbral la presenta el programa *Win Episcopes*, que utiliza el área bajo la curva ROC (*Receiver Operating Characteristic*) y su intervalo de confianza mediante el método de Wilcoxon. Dado que el umbral no es nunca un valor, sino un límite entre dos valores se puede hablar de umbral de positividad o umbral de negatividad. En general el umbral del que se habla es el que permite detectar mejor los animales positivos (umbral de positividad). En la figura adjunta (*Win Episcopes*), el umbral se sitúa entre los títulos 1/80 y 1/160 por lo que el primero sería el umbral de negatividad y el segundo el de positividad. De forma habitual nos solemos referir como umbral al título a partir del cual se detectan los positivos (umbral de positividad), pero conviene precisarlo para evitar confusiones.



Selección de técnicas o umbrales diagnósticos

El umbral ideal será el que detecte la mayor proporción de muestras correctamente clasificadas (mejor fiabilidad y J de *Youden*), pero, en función del objetivo del diagnóstico, se puede desplazar el umbral para ganar sensibilidad o especificidad, según las necesidades, pero sin desviarse en exceso del mejor umbral previamente seleccionado, pues este es el que producirá los mejores resultados globales. En general, el parámetro más deseado en las técnicas que se utilizan en campañas de saneamiento (sobre todo en las primeras fases) es la sensibilidad, para detectar la mayor proporción de casos posible, aunque obtengan mayor proporción de falsos positivos por su menor especificidad. Esto es especialmente interesante en los programas de control de zoonosis, como los saneamientos ganaderos obligatorios. Con el avance de la campaña y la disminución de la prevalencia, los falsos positivos se estabilizarán, por lo que se buscarán técnicas más específicas.

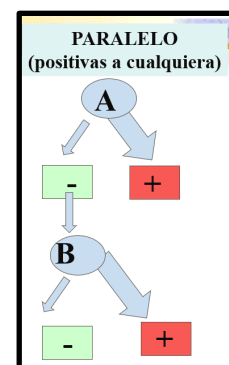
En determinadas situaciones se necesitan técnicas con mayor especificidad sobre todo cuando las consecuencias diagnósticas son radicales (sacrificio obligatorio), ante factores emocionales (mascotas) o rebaños o individuos de alto valor genético y/o económico. Para asegurarnos este objetivo se buscarán técnicas con mayor especificidad, aunque obtengan mayor proporción de falsos negativos por su menor sensibilidad. Otros elementos a considerar para la selección de la técnica pueden estar relacionados con sus aspectos económicos, técnicos y logísticos. Así, el coste por análisis, la capacidad de automatización, la sencillez para su realización, etc., deben ser considerados al diseñar los programas de diagnóstico de ámbito poblacional.

Aunque no existe un límite de los parámetros de validez para seleccionar una técnica, debemos tener en cuenta que ninguno de sus parámetros inherentes (sensibilidad o especificidad) debe aproximarse al 50% (si tiramos una moneda al aire para considerar un animal enfermo o sano nos aproximaríamos a esa probabilidad). Aunque lo ideal es utilizar técnicas diagnósticas con altos o muy altos parámetros de validez, incluso las técnicas con moderada validez pueden utilizarse, al combinar técnicas diagnósticas.

COMBINACIÓN DE DOS O MÁS PRUEBAS (ANÁLISIS MÚLTIPLES).

Permite aprovechar las ventajas y minimizar las limitaciones de las técnicas diagnósticas, incluso con bajos valores de alguno de sus parámetros de validez (S o E).

Análisis múltiples en paralelo. Realizamos dos o más pruebas y consideramos infectados los animales cuyas muestras han sido positivas a cualquiera de ellas, con lo que estamos potenciando la sensibilidad a costa de perder especificidad. Aunque ambas pruebas se pueden hacer de forma simultánea ("en paralelo"), lo ideal es hacerlas de forma secuencial lo que reduce costes, y a la segunda prueba sólo se someterían las muestras negativas a la primera.



Análisis múltiples en serie. Tras realizar la primera prueba, sólo las muestras positivas se procesan con la segunda (o siguientes). De esta forma consideramos positivas las muestras que han resultado positivas a ambas técnicas, dando como negativas todas las demás. Con este método potenciamos la especificidad en detrimento de la sensibilidad, lo que nos permite trabajar inicialmente con métodos de barrido más sensibles, pero menos específicos. Así, la campaña de saneamiento de brucelosis en pequeños rumiantes se ha realizado clásicamente combinando la técnica de aglutinación con rosa bengala (RB) (muy sensible pero poco específica) con la de fijación de complemento (FC) (muy específica). La primera es una técnica (“de cribado”), muy barata, sencilla de realizar y sensible, con la que se seleccionaban los “posibles positivos”, asumiendo la presencia de numerosos falsos positivos. Para confirmar los positivos reales, y evitar los falsos positivos, se volvían a analizar los sueros positivos a RB mediante la FC, que es más específica, aunque más costosa en tiempo y reactivos. De esta forma se realiza un cribado previo con todos los sueros a analizar.

Las imágenes siguientes (*Win Epi*), muestran el resultado de la combinación múltiple de dos pruebas. La

Prueba :	Prueba A	Prueba B
Sensibilidad :	61.6%	94.6%
Especificidad :	97.7%	69.3%
Prevalencia % :	22%	
Tamaño de población :	500	

combinación en serie reduce los falsos positivos aumentando la especificidad a costa de disminuir la sensibilidad. Por el contrario, la combinación en paralelo reduce los falsos negativos aumentando la sensibilidad a costa de disminuir la especificidad.

	Prueba A	Prueba B
Val. Pred. Positivo :	88.3%	46.5%
Val. Pred. Negativo :	90.0%	97.8%
Prev. aparente :	15.3%	44.8%
J de Youden :	59.3%	63.9%
Fiabilidad :	89.8%	74.8%
	Enf. San.	Enf. San.
Diag. +	68 9	104 120
Diag. -	42 381	6 270
	Serie	Paralelo
Sensibilidad :	58.3%	97.9%
Especificidad :	99.3%	67.7%
Val. Pred. Positivo :	95.9%	46.1%
Val. Pred. Negativo :	89.4%	99.1%
Prev. aparente :	13.4%	46.7%
J de Youden :	57.6%	65.6%
Fiabilidad :	90.2%	74.4%
	Enf. San.	Enf. San.
Diag. +	64 3	108 126
Diag. -	46 387	2 264

CONCORDANCIA ENTRE TÉCNICAS

A veces no es posible comparar las pruebas con la realidad al carecer de la prueba de oro (*gold standard*), o resultar económicamente inviable. La concordancia entre técnicas es útil cuando desconocemos los parámetros de validez de las técnicas, pero no tiene interés para técnicas validadas. Una de sus aplicaciones permite conocer el grado de similitud entre técnicas para sustituir la de uso habitual por otra cuya validez desconocemos. El grado en el que dos pruebas coinciden (excluyendo la coincidencia por azar) se valora mediante el coeficiente de *Kappa*. Hay diferentes enfoques para la interpretación del resultado. En la imagen adjunta se adjuntan los estratos de interpretación de los valores del coeficiente de *Kappa*, según Thrusfield, (2007). Así, se establece un 0.7 de coeficiente de *Kappa* como un nivel aceptable de concordancia entre técnicas y con ese valor podríamos sustituir una por otra y esperar un comportamiento similar. Por su parte, *Fleiss et al.* (2003) consideran que un *Kappa* igual o superior a 0,75 indica una excelente concordancia entre técnicas. Un elevado valor del coeficiente de *Kappa* no indica que las técnicas sean buenas, sino que sus resultados concuerdan. Siempre que dispongamos de datos de validación deberán utilizarse frente a los de concordancia, al darnos mucha más información sobre las ventajas y limitaciones de la técnica.

CONCORDANCIA ENTRE DOS O MÁS TÉCNICAS	
COEFICIENTE DE KAPPA	CONCORDANCIA
0,81 ⇒ 1	Casi perfecta
0,61 ⇒ 0,8	Adecuada
0,41 ⇒ 0,60	Moderada
0,21 ⇒ 0,40	Débil
0 ⇒ 0,2	Escasa

La aplicación de *Win Epi* “Diagnóstico: Concordancia entre pruebas” nos permite introducir datos en tablas de 2x2 hasta 6x6 y permite utilizar otro tipo de variables, además de las binomiales. La tabla de contingencia que se crea presenta los símbolos (-, +) en orden inverso a como se presentan en las tablas de contingencia para validación de técnicas (+, -). Esta aparente complicación puede utilizarse como un recurso nemotécnico para recordar que esta es una situación totalmente diferente de la validación de técnicas.

OTROS CONCEPTOS RELACIONADOS CON EL DIAGNÓSTICO

A veces se utilizan otros conceptos para indicar la bondad de una técnica. Por ejemplo, si se va a realizar un estudio epidemiológico sobre mastitis podemos enfocarlo a mastitis clínicas, subclínicas, ambas o a la etiología implicada. Por ello, podemos establecer un gradiente según la exactitud del diagnóstico similar al siguiente:

Mastitis clínica < Mastitis bacteriana < Mastitis por gérmenes gram negativos < Mastitis por Gram negativos no enterobacterias < Mastitis por *Pseudomonas spp.* < Mastitis por *Pseudomonas aeruginosa*

Según esta exactitud diagnóstica (sinónimo de precisión en este contexto) se podrán establecer unas u otras medidas de control frente al proceso mastítico. Así, ante un diagnóstico exacto de mastitis caprina por *Pseudomonas aeruginosa* el técnico debería considerar un reservorio hídrico en el establo, el uso inadecuado de agua de pozo sin clorar para la limpieza de la máquina de ordeño, una incorrecta desinfección de la misma, o la inadecuada manipulación de las cánulas intramamarias de tratamiento en el secado, entre otras. Por el contrario, un diagnóstico menos preciso (diagnóstico de mastitis bacteriana) no permitirá establecer estrategias tan específicas.

La exactitud y la precisión de una prueba se utilizan como sinónimos cuando hacen referencia al grado de adecuación de la medición a la realidad, es decir el grado de detalle. El diagnóstico de mastitis reflejado anteriormente tiene seis niveles diferentes y cada nivel inferior es más preciso o exacto que el anterior. La precisión tiene otra acepción, desde el punto de vista estadístico, y viene a reflejar el error o coherencia de una serie de mediciones (prevalencia esperada de $40\% \pm 5\%$), y en este contexto interesa que el porcentaje de precisión (error) sea lo más pequeño posible.

Por su parte, la fiabilidad o reproductibilidad de una técnica hacen referencia a su repetibilidad, es decir, el grado en el que se obtienen resultados similares al repetir la misma.

En algunos contextos se utiliza el término “sensibilidad” de una técnica para referirse a su capacidad de detectar una pequeña cantidad de sustrato problema (unidades formadoras de colonias, anticuerpos, toxinas, enzimas, hormonas, etc.). Para evitar estas confusiones debe referirse, en el caso anterior, al límite de detección de una técnica y no debe utilizarse el concepto de sensibilidad.



Con la excepción de la sensibilidad y la especificidad, los conceptos anteriores no son inherentes a las pruebas y pueden variar según los laboratorios o las personas que las lleven a cabo. Para controlar estas variaciones los laboratorios normalizados llevan a cabo con cierta frecuencia estudios intercolaborativos en los que diferentes laboratorios analizan las mismas muestras, estableciendo una relación de laboratorios según su grado de adecuación al resultado real.

EJEMPLO DE VALIDACIÓN

Una técnica se comparó con un Gold Standard. En total, la técnica detectó 37 positivos y 80 negativos. De los 36 positivos que había en la población, la técnica detectó 33.

		Realidad	
		Positivo	Negativo
Técnica a validar	Positivo	33	4
	Negativo	3	77

Prevalencia aparente= $37/117=0,3162=31,62\%$

Prevalencia real= $36/117=0,3077=30,77\%$

Sensibilidad= $33/36=0,9167=91,67\%$

Especificidad= $77/81=0,9506=95,06\%$

Valor Predictivo Positivo= $33/37=0,8919=89,19\%$

Valor Predictivo Negativo= $77/80=0,9625=96,25\%$

J Youden= $(0,9167 + 0,9506 - 1) = 0,8673=86,73\%$

Fiabilidad= $(33/117) + (77/117)=0,9402=94,02\%$

Sensibilidad :	91.7%	(82.6%, 100.7%)
Especificidad :	95.1%	(90.3%, 99.8%)
Valor Predictivo Positivo :	89.2%	(79.2%, 99.2%)
Valor Predictivo Negativo :	96.3%	(92.1%, 100.4%)
Prevalencia real :	30.8%	(22.4%, 39.1%)
Prevalencia aparente :	31.6%	(23.2%, 40.0%)
J de Youden :	86.7%	(76.5%, 96.9%)
Fiabilidad :	94.0%	(89.7%, 98.3%)

REFERENCIAS

- De Blas, I., 2006. *Win Epi (Working in Epidemiology)*. Accesible en red <http://www.winepi.net/>. Fecha de acceso 24 de febrero de 2025.
- De Blas, I., Ruiz-Zarzuela, I., Bayot, B., Ferreira, C. 2007. Capítulo 5: Pruebas diagnósticas. En: Manual de Epidemiología Veterinaria. Pag: 59-74. Universidad de Zaragoza.
- Contreras, A., 2016. Video tutorial “Parámetros de validez de técnicas diagnósticas. Desarrollo mediante un ejemplo (7’56’’). Televisión Universitaria. Universidad de Murcia. 30 jun 2016. <http://tv.um.es/video?id=82661&idioma=es>
- Contreras, A., Luengo, C., Sánchez, A., Miranda, R.E., Corrales, J.C., 2001. Capítulo 21: Epidemiología diagnóstica. Pág: 233-248. En “Epidemiología Veterinaria” Ed. Contreras, A., Sánchez, A., Corrales, J.C (Coordinadores). Diego Marín Librero. Murcia. España.
- Fleiss, J.L., Levin, B., and Paik, M.C., 2003. *Statistical Methods for Rates and Proportions*. Citado por Thrusfield, 2007.
- Martin S.W., Meek A.H., Willemberg P. 1997. Capítulo 3: Medida de la frecuencia de la enfermedad y de la producción. En: Epidemiología Veterinaria Principios y métodos. Editorial Acribia, S.A., Zaragoza pp. 55-86.
- OMSA (Organización Mundial de Sanidad Animal). 2023. Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres. <https://www.woah.org/es/que-hacemos/normas/codigos-y-manuales/acceso-en-linea-al-manual-terrestre/>. Fecha de acceso 24 de febrero de 2025.
- Thrusfield, M. 2007. Chapter 17. Diagnostic testing. In: Veterinary Epidemiology. 3rd Edition. Pages 305-330. Blackwell Science Ltd. Oxford, Reino Unido.
- Thrusfield, M., Ortega, C., de Blas, I., Noordhuizen, J.P., Frankena, K. 2001. *WIN EPISCOPE 2.0: improved epidemiological software for veterinary medicine. Veterinary Record*. 148(18):567-72. doi: 10.1136/vr.148.18.567.