



UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE CIENCIA E INGENIERÍA EN ALIMENTOS

CARRERA DE INGENIERÍA BIOQUÍMICA

Validación de la prueba de ELISA indirecto como herramienta de diagnóstico para el programa nacional de brucelosis bovina en el Ecuador

Trabajo de Titulación, modalidad Experiencia Práctica de Investigación y/o Intervención, presentado como requisito previo a la obtención del Título de Ingeniero Bioquímico otorgado por la Universidad Técnica de Ambato, a través de la Facultad de Ciencia e Ingeniería en Alimentos.

Autor: Edwin Enrique Martínez Barreno.

Tutor: Yunys Pérez Betancourt MSc.

Ambato – Ecuador

Septiembre-2017


APROBACIÓN POR EL TUTOR

Yunys Pérez Betancourt MSc.

CERTIFICA:

Que el presente trabajo de titulación ha sido prolijamente revisado. Por lo tanto, autorizo la presentación de este Trabajo de Titulación, modalidad Experiencia Práctica de Investigación y/o Intervención, el mismo que responde a las normas establecidas en el Reglamento de Títulos y Grados de la Facultad.

Ambato, 10 de julio de 2017



.....

Yunys Pérez Betancourt MSc.

C.I. 175647274-0

TUTOR

DECLARACIÓN DE AUTENTICIDAD

Yo, Edwin Enrique Martínez Barreno, manifiesto que los resultados obtenidos en el presente Proyecto de Investigación, previo a la obtención de título de ingeniero Bioquímico son absolutamente originales, auténticos y personales; a excepción de las citas y datos proporcionados por los registros de la empresa AGROCALIDAD.



Edwin Enrique Martínez Barreno

C.I. 180459427-1

AUTOR


APROBACIÓN DEL TRIBUNAL DE GRADO

Los suscritos Profesores Calificadores, aprueban el presente Trabajo de Titulación modalidad de Experiencias Prácticas de Investigación y/o Intervención, el mismo que ha sido elaborado de conformidad con las disposiciones emitidas por la Facultad de Ciencia e Ingeniería en Alimentos de la Universidad Técnica de Ambato.


Para constancia firman:



.....
Presidente del tribunal



.....
Dr. Carlos Alberto Rodríguez Meza
C.I. 180216650-2



.....
Lic. Mg. Danae Fernández Rivero
C.I. 175718120-9

Ambato, 28 de agosto de 2017

DERECHOS DEL AUTOR

Autorizo a la Universidad Técnica del Ambato, para que haga de este Proyecto de Intervención o parte de él, un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos en línea patrimoniales de mi Proyecto, con fines de difusión pública, además apruebo la reproducción de este proyecto dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor.



.....

Edwin Enrique Martínez Barreno

C.I. 180459427-1

AUTOR

DEDICATORIA

A Dios, por darme la vida.

A mimamá, por ser una persona que siempre me apoyo a pesar de todos los golpes
que nos dio la vida.

A mí esposa, por ser mi apoyo y acompañara en todo momento.

A mis hermanos Darwin. Judith, Liliana, Byron, Marcos y toda mi familia por haber
creído en mí.

AGRADECIMIENTOS

Agradezco de todo corazón a mi mamá, a mi esposa y a mis hermanos que supieron brindarme su apoyo incondicional en cada momento.

AGROCALIDAD por la oportunidad brindada para realizar la parte experimental del trabajo. Gracias especialmente a la Doctora Maritza Barrera, Doctor Patricio Sandoval y a la Licenciada Margoth Barrionuevo por compartirme sus conocimientos y ser mis guías en el tiempo que estuve con ellos.

A mi tutor Yunys Pérez Betancourt MSc., por su amistad, paciencia y colaboración en la realización de mi Tesis.

ÍNDICE GENERAL DE CONTENIDOS

PORTADA.....	I
APROBACIÓN POR EL TUTOR.....	II
DECLARACIÓN DE AUTENTICIDAD.....	III
APROBACIÓN DEL TRIBUNAL DE GRADO	IV
ÍNDICE GENERAL DE CONTENIDOS.....	VIII
ÍNDICE DE TABLAS	XII
ÍNDICE DE FIGURAS.....	XIII
RESUMEN.....	XIV
ABSTRACT.....	XV
INTRODUCCIÓN	1
CAPÍTULO I.....	3
EL PROBLEMA	3
1.1.- Tema.....	3
1.2.- Justificación	3
1.3.- Objetivos	4
1.3.1.- Objetivo General	4
1.3.2.- Objetivos Específicos	4
CAPÍTULO II	5
MARCO TEÓRICO.....	5
2.1.- Antecedentes investigativos.....	5
2.1.1.- Brucelosis	5
2.1.2.- Reservorios de <i>Brucella</i>	5
2.1.3.- Brucelosis bovina	6
2.1.4.- <i>Brucella abortus</i>	6
2.1.5 Periodo de incubación.....	7
2.1.6.- Formas de transmisión de la brucelosis en bovinos	7
2.1.7.- Patogénesis de <i>B. abortus</i>	7
2.1.8.- Respuesta inmunitaria	8
2.1.9.- Diagnóstico.....	8

2.1.10.- Validación de pruebas serológicos	10
2.1.11.- Criterios para el desarrollo y validación de una prueba	10
2.1.11.1.- Sensibilidad y Especificidad Analítica	10
2.1.11.2.- Sensibilidad y Especificidad Diagnóstica	11
2.1.11.3.- Valores predictivos	11
2.1.11.4.- Exactitud/Precisión	11
2.1.11.5.- Veracidad	11
2.1.11.6.- Repetibilidad/ Reproducibilidad	12
2.1.11.7.- Contaminación por arrastre (Efecto Carry-over)	12
2.1.5.8.- Incertidumbre de la medición	12
2.2.- Hipótesis	13
2.2.1.- Hipótesis nula	13
2.2.2.- Hipótesis alternativa	13
2.3.- Señalamiento de variables de la hipótesis	13
2.3.1.- Variables independientes	13
2.3.2.- Variables dependientes	13
CAPÍTULO III	13
MATERIAL Y MÉTODOS	14
3.1.- Materiales	14
3.1.1.- Material de Referencia Internacional Certificado (MIRC)	14
3.1.2.- Material de Referencia Nacional Certificado (MRNC)	14
3.1.3.- Reactivos	14
3.1.4.- Equipos	14
3.2.- Metodología	14
3.2.1.- ELISA indirecto	15
3.2.2.- Sensibilidad Analítica	16
3.2.3.- Especificidad Analítica	16

3.2.4.- Sensibilidad y Especificidad Diagnóstica	17
3.2.5.- Valores Predictivos Positivos y Negativos.....	17
3.2.6.- Exactitud.....	17
3.2.8.- Contaminación por arrastre	18
3.2.7.- Precisión (Repetibilidad/ Reproducibilidad).....	19
3.2.9.- Incertidumbre de la medición	19
CAPITULO IV	21
RESULTADOS Y DISCUSIÓN	21
4.1.- Análisis y discusión de los resultados	21
4.1.1.- Estimación de la sensibilidad analítica y punto de corte.....	21
4.1.2.- Estimación de la especificidad analítica.....	23
4.1.3.- Estimación de los valores de sensibilidad y la especificidad diagnóstica.....	24
4.1.4.- Estimación de los valores predictivos positivo y negativo.....	24
4.1.5.- Determinación de la exactitud de la prueba	25
4.1.6.- Análisis de la contaminación por arrastre	25
4.1.7.- Análisis de precisión.....	26
4.1.7.1.- Repetibilidad.....	26
4.1.7.2.- Reproducibilidad	27
4.1.8.- Análisis de la incertidumbre de la medición.	27
4.1.8.1.- Interpretación de la incertidumbre de la medición.	27
4.2.- Verificación de hipótesis	29
CAPÍTULO V	30
CONCLUSIONES Y RECOMENDACIONES.....	30
5.1.- Conclusiones.....	30
5.2.- Recomendaciones	30
BIBLIOGRAFÍA	32
ANEXOS	42
Anexo A. <i>Datos de Sensibilidad Analítica. Analista 1 y 2</i>	42

Anexo B. <i>Datos de sensibilidad y especificidad diagnóstica</i>	43
Anexo C. <i>Exactitud. Valores obtenidos</i>	45
Anexo D. <i>Cálculo del coeficiente Carry-over</i>	46
Anexo E. <i>Datos para la determinación de la incertidumbre</i>	47
Anexo F. <i>Certificado de análisis del suero control positivo</i>	48
Anexo G. <i>Certificado de análisis del suero control negativo</i>	49
Anexo H. <i>Certificado de análisis del Kit de brucelosis IDEXX</i>	50
Anexo I. <i>Declaración de Validación</i>	51
Anexo J. <i>Principios y métodos de validación de las pruebas de diagnóstico de las enfermedades infecciosas</i>	52
Anexo K. <i>Enfoques estadístico de la validación</i>	70
Anexo L. <i>Incertidumbre de la medición</i>	82

ÍNDICE DE TABLAS

Tabla 1. Reactivos del Kit de ELISA indirecto.....	14
Tabla 2. Valoración del coeficiente Kappa	18
Tabla 3. Objetivos de validación.....	19
Tabla 4. Especificidad analítica para la prueba de ELISA indirecto	23
Tabla 5. Estimaciones de la sensibilidad y la especificidad diagnósticas	24
Tabla 6. Determinación de la fuerza de concordancia y la exactitud por índice kappa	25
Tabla 7. Índice de contaminación por arrastre	25
Tabla 8. Criterios de estadísticos para la estimación de repetibilidad de la prueba de ELISA indirecto	26
Tabla 9. Análisis de varianza para la repetibilidad de la prueba de ELISA indirecto	26
Tabla 10. Prueba de t, análisis de medias entre el analista 1 y analista 2	27
Tabla 11. Interpretación de la incertidumbre	27

ÍNDICE DE FIGURAS

Figura 1. Esquema de “Tiro al Blanco” Esquema de “Tiro al Blanco”, utilizado para ejemplificar métodos de ensayo, los punto u orificios equivaldrían a los resultados analíticos y el círculo rojo al centro el rango en el cual se espera este el valor de referencia (o verdadero) (Daffau et al., 2010).....	12
Figura 2. Análisis Probit para cada analista. Cada dilución se repitió 3 veces. La sensibilidad analítica al 95% de confianza se indica con líneas discontinuas.	22
Figura 3. Análisis (ROC). Representa la sensibilidad (Se) y la especificidad (Sp) en función del punto de corte de la prueba. La línea vertical discontinua muestra el punto de corte al 95% de confianza.	22

RESUMEN

La brucelosis bovina es una de las principales enfermedades zoonóticas en el mundo causada por la bacteria *Brucella abortus*. Esta enfermedad ha causado daños en la salud pública y grandes pérdidas económicas. En el presente trabajo se validó una prueba de Elisa indirecto como una herramienta más de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador. Utilizando como material de referencia internacional 1 suero positivo y 1 suero negativo de IDvet y como material de referencia nacional 30 sueros negativos de predios ganaderos libres de brucelosis, certificados por AGROCALIDAD y 11 sueros positivos a brucelosis certificados por el Centro Internacional de Zoonosis (CIZ). Los datos de la prueba de ELISA se analizaron de acuerdo al Manual de la OIE y la Norma ISO/IEC 17025:2005. La especificidad y sensibilidad diagnóstica de la prueba fueron del 100% a un punto de corte del 121,13 (M/P%) corroborado por el análisis de características de funcionamiento del receptor (ROC) en la dilución 1/32 última dilución en la que se evidenció positividad. La prueba de ELISA demostró ser específica para *B. abortus*. Los valores predictivos tanto positivos como negativos fueron del 100%. Se estimó el 100% de precisión (repetibilidad y reproducibilidad), una exactitud del 97% con un índice Kappa de 0,927, no presentó contaminación por arrastre y la incertidumbre de la prueba fue del 4,8%. Los resultados demostraron que la prueba de ELISA indirecto es altamente sensible y específica, y se puede utilizar como herramienta para el diagnóstico de brucelosis bovina en el Ecuador.

Palabras clave: brucelosis bovina, validación, prueba de ELISA indirecto, sueros, AGROCALIDAD.

ABSTRACT

Bovine brucellosis is one of the major zoonotic diseases in the world caused by the bacterium *Brucella abortus*. This disease has caused damage to public health and great economic losses. In the present work, an indirect Elisa test was validated as a diagnostic tool for the National Bovine Brucellosis Program in Ecuador. Using as international reference material 1 positive serum and 1 negative serum from IDvet and as national reference material 30 negative sera from livestock breeding sites free of brucellosis, certified by AGROCALIDAD and 11 positive sera to brucellosis certified by the International Center for Zoonoses (CIZ). Data from the ELISA test were analyzed according to the OIE Manual and ISO / IEC 17025: 2005. The specificity and diagnostic sensitivity of the test were 100% at a cutoff point of 121.13 (M / P %) corroborated by the analysis of receiver operating characteristics (ROC) at the dilution 1/32 last dilution in the which showed positivity. The ELISA test proved to be specific for *B. abortus*. Both positive and negative predictive values were 100%. A 100% accuracy (repeatability and reproducibility) was estimated, an accuracy of 97% with a Kappa index of 0.927, no trawl contamination, and test uncertainty was 4.8%. The results demonstrated that the indirect ELISA test is highly sensitive and specific and can be used as a tool for the diagnosis of bovine brucellosis in Ecuador.

Keywords: bovine brucellosis, validation, indirect ELISA test, sera, AGROCALIDAD.

INTRODUCCIÓN

Brucelosis es el nombre general para una de las enfermedades zoonóticas más importantes del mundo causada por las bacterias del género *Brucella* (Ducrotoy et al., 2015), principalmente *Brucella abortus*, agente causante de infección en bovinos (OIE, 2016a). La infección bovina tiene una gran distribución y es la más importante en lo que se refiere en pérdidas económicas. Si bien la infección ha sido erradicada en varios países europeos, Japón, Israel, Canadá, Australia y Nueva Zelanda (Cold et al., 2017), su incidencia en otras partes del mundo ha aumentado debido al énfasis en el aumento de la producción animal, como es el caso de América Latina (U. Jain, Bisht, Sahzad, Pragati, & Dwivedi, 2013).

La forma más importante de transmisión de la enfermedad en el ganado es la ingestión de alimentos o agua contaminados con *B. abortus*, también ocurre mediante la transmisión venérea. En la hembras la principal característica clínica de la enfermedad es el aborto, placentas retenidas y reducción en la producción de leche, mientras que en los machos produce infección en las vesículas seminales y los testículos (Laboratorios, 2017). La transmisión de la brucelosis en seres humanos está fuertemente relacionada con el consumo y el manejo de animales infectados con *B. abortus*, provocando daños en diferentes órganos y articulaciones (Giambartolomei, Arriola Benitez, & Delpino, 2017).

La brucelosis no tiene características clínicas particulares que permitan un diagnóstico preciso, razón por la que se confía mucho en el diagnóstico de laboratorio tanto por métodos directos como indirectos. Los métodos directos incluyen técnicas mediante las cuales el organismo causante puede ser identificado por examen microscópico o por cultivo. Las cepas de campo de *B. abortus* sólo pueden diferenciarse de las cepas vacunales utilizando técnicas de cultivo y biotipificación. Las técnicas serológicas para el diagnóstico de la brucelosis son métodos de diagnóstico indirectos que no distinguen de las reacciones serológicas causadas por la cepa vacunal de *Brucella*, S19 (Departement Agriculture, Forestry and Fisheries, 2016).

Las pruebas serológicas son cruciales para el diagnóstico de la brucelosis por su bajo costo y la facilidad de análisis. La mayoría de los programas de control y erradicación

se basan en estos métodos las mismas que deben estar bajo regulaciones locales e internacionales (Geresu & Kassa, 2015).

CAPÍTULO I

EL PROBLEMA

1.1.- Tema

“VALIDACIÓN DE LA PRUEBA DE ELISA INDIRECTO COMO HERRAMIENTA DE DIAGNÓSTICO PARA EL PROGRAMA NACIONAL DE BRUCELOSIS BOVINA EN EL ECUADOR”

1.2.- Justificación

La brucelosis bovina es una enfermedad infecciosa zoonótica re-emergente con una alta distribución mundial, causada por las especies bacterianas del género *Brucella*, que son patógenas para una amplia variedad de animales y seres humanos. La única forma de controlar y erradicar esta zoonosis es mediante diferentes pruebas de diagnóstico, sacrificio de animales infectados y vacunación. Actualmente la brucelosis bovina es una de las principales enfermedades que afecta al ganado bovino en los países en vías de desarrollo en Latinoamérica. Según datos proporcionados por La Agencia Ecuatoriana de Aseguramiento de la Calidad del Agro (AGROCALIDAD) en el 2008 las pérdidas económicas por brucelosis bovina ascendieron a 5`436.908 dólares. Kirk et al. (2015) calculan que la incidencia mundial de brucelosis humana supera los 800.000 casos al año, de los cuales se estima que el 40% produce una infección crónica, por lo que es necesario el control y la erradicación de esta enfermedad. AGROCALIDAD como ente de control y erradicación de la brucelosis

bovina en el Ecuador requiere de la validación de métodos serológicos con alta sensibilidad y especificidad diagnóstica con los cuales se pueda analizar un gran número de muestras en poco tiempo. Con esto los ganaderos tendrán conocimientos de los animales infectados en su predio y podrán tomar medidas para minimizar las pérdidas económicas. La presente investigación se enfocó en la validación de la prueba de ELISA indirecto, según la Norma ISO/IEC 17025:2005 y el Manual de pruebas de diagnóstico y vacunas para animales terrestres, como una herramienta más de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador.

1.3.- Objetivos

1.3.1.- Objetivo General

Validar la prueba de ELISA indirecto como herramienta de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador.

1.3.2.- Objetivos Específicos

- Establecer los valores de especificidad y sensibilidad analítica de la prueba de ELISA indirecto.
- Determinar los valores de especificidad y sensibilidad diagnóstica de la prueba de ELISA indirecto.
- Determinar los parámetros de validación de la prueba ELISA indirecto.

CAPÍTULO II

MARCO TEÓRICO

2.1.-Antecedentes investigativos

2.1.1.- Brucelosis

Enfermedad causada por bacterias del género *Brucella* afectando animales domésticos, salvajes y la mayoría de bacterias son también patógenos para los seres humanos. En los animales, las bacterias suelen afectar a los órganos reproductivos causando el aborto siendo a menudo el único signo de la enfermedad. En los humanos es una enfermedad febril aguda persistente con una amplia variedad de síntomas. Es una verdadera zoonosis ya que prácticamente todas las infecciones humanas son adquiridas de los animales (Alton & Forsyth, 1996).

Además a la Brucelosis también se la conoce como: Fiebre de Malta (Wyatt, 2005), Fiebre Mediterránea (Roberts & Kemp, 2001), Fiebre Ondulante o Enfermedad de Bang (Schulze, 1959).

2.1.2.- Reservorios de *Brucella*

Algunos de los reservorios naturales de *Brucella* son: bovinos (*B. abortus*), caprinos (*B. melitensis*), ovinos (*B. melitensis* y *B. ovis*), cerdos (*B. suis*), caninos (*B. melitensis*, *B. abortus*, *B. canis* y *B. suis*). Enfermedad significativa en la fauna salvaje, afectando al cerdo salvaje, el bisonte, el alce y la liebre europea. La presencia de un reservorio

en la fauna salvaje complica la lucha por erradicar la enfermedad. También se han detectado en mamíferos marinos, así como pequeños roedores hasta camélidos (Villareal, 2016).

2.1.3.- Brucelosis bovina

Enfermedad producida por *Brucella abortus*, bacteria que también afecta a otras especies, entre ellas el bisonte, el búfalo y el alce. Algunas especies actúan como huéspedes de mantenimiento para este microorganismo. Las infecciones en los animales silvestres dificultan los esfuerzos en la erradicación en el ganado bovino (L. Jain, Rawat, Ramakrishnan, & Kumar, 2017).

La brucelosis bovina es una enfermedad zoonótica repartida por todo el mundo produciendo grandes pérdidas económicas, debido principalmente a la rebaja en la producción de leche, abortos y otras complicaciones de tipo reproductivo, así como el sacrificio de animales enfermos, y la restricción de la comercialización internacional de animales o sus productos. (Aznar, Samartino, Humblet, & Saegerman, 2014).

En el Ecuador, a pesar de que la brucelosis bovina es una de las principales enfermedades zoonóticas en el mundo, el número de casos de infección a personas es desconocido. Según el Ministerio de Salud Pública del Ecuador (MSP), existieron 111 casos en humanos entre 1990 y 2007, mientras que el Instituto Nacional de Estadísticas y Censos (INEC) se dio a conocer que 152 personas fueron hospitalizadas por Brucelosis entre 1998 y 2007. La relevancia de esta enfermedad está relacionada debido a la infección que se puede dar en humanos y sus implicaciones económicas, así por ejemplo, en los EE.UU se estimó que el tratamiento de un paciente con Brucelosis oscila entre 340 USD hasta los 4.095 USD, dependiendo de la gravedad (Villareal, 2016).

2.1.4.- *Brucella abortus*

Brucella abortus es un cocobacilo Gram-negativo, inmóvil, no forma esporas y carece de cápsula, miden aproximadamente 0,6 a 1,5 μm por 0,5-0,7 μm . La membrana celular externa se parece mucho a la de otros bacilos Gram-negativos con un componente lipopolisacárido dominante (LPS) y tres grupos principales de proteínas. El contenido en guanina-más-citosina del ADN es 55-58 moles/cm. No se ha

encontrado que ninguna especie de *Brucella* albergue plásmidos naturalmente aunque aceptan fácilmente plásmidos de amplio rango de hospedador (Alton & Forsyth, 1996).

El metabolismo de *B. abortus* es principalmente oxidativo y muestran poca acción sobre los carbohidratos en los medios convencionales. Son aerobios pero algunas especies requieren una atmósfera con CO₂ añadido (5-10 por ciento). La multiplicación es lenta a la temperatura óptima de 37 °C y se necesita un medio enriquecido para soportar un crecimiento adecuado. Las colonias de *B. abortus* se hacen visibles en medios sólidos adecuados en 2-3 días. Las colonias de cepas lisas son pequeñas, redondas y convexas, pero la disociación, con pérdida de las cadenas O del LPS, se produce fácilmente para formar variantes rugosas o mucoides (Alton & Forsyth, 1996; OIE, 2016a).

2.1.5.- Periodo de incubación

El período de incubación es a menudo difícil de determinar, pero suele ser de 2 a 4 semanas, pudiendo prolongarse hasta varios meses. Los síntomas de la enfermedad pueden aparecer de forma brusca (1-2 días) o gradual (1 o varias semana) (Alton & Forsyth, 1996; Villareal, 2016).

2.1.6.- Formas de transmisión de la brucelosis en bovinos

La forma más común de contraer la enfermedad ocurre por vía oral, al momento del parto las vacas tienden a lamer al feto o las descargas genitales producidas durante el parto. Los terneros la infección ocurre en el útero o al consumir leche de vacas infectadas. Una forma de transmisión indirecta puede ser por la interacción con animales salvajes, consumo de pastos o agua contaminadas (Villareal, 2016).

2.1.7.- Patogénesis de *B. abortus*

Los estudios en animales sugieren que la bacteria invasora es rápidamente fagocitada por leucocitos polimorfonucleares. *B. abortus* con frecuencia es capaz de sobrevivir y multiplicarse en estas células porque inhibe el sistema bactericida mieloperoxidasa-peróxido-haluro mediante la liberación de 5'-guanosina y adenina. Al principio de la infección, los macrófagos también son relativamente ineficaces al matar la bacteria intracelular. En la diseminación sistémica, no está claro si las bacterias se transportan dentro de los neutrófilos y macrófagos o en el torrente sanguíneo fuera de las células,

pero los organismos pueden diseminarse ampliamente del tejido linfoide regional apropiado al portal de entrada y pueden localizarse en ciertos órganos diana como los ganglios linfáticos, el bazo, el hígado, la médula ósea y (especialmente en los animales) los órganos reproductores. La presencia de meso-eritritol en los testículos y vesículas seminales de toros, carneros, cabras y jabalíes y en los productos de la concepción en rumiantes y cerdos preñados estimula una enorme multiplicación de brucelosis (Alton & Forsyth, 1996).

2.1.8.- Respuesta inmunitaria

Los mecanismos involucrados en la respuesta inmune contra la *B. abortus*, pueden ser divididos en dos grupos principales; el lipopolisacárido (LPS) que induce anticuerpos aglutinantes (IgM) (Cloekaert, Jacques, de Wergifosse, Dubray, & Limet, 1992) y anticuerpos no aglutinantes tipo IgG que pueden proteger al huésped contra la liberación de grandes cantidades de endotoxinas (Kurtz & Berman, 1986).

El LPS es el primer antígeno frente al cual aparecen anticuerpos de tipo IgM e IgG, después de la infección natural o de la vacunación con la cepa atenuada S19 (Saldarriaga, Ossa, & López, 2002), la misma que no se diferencia serológicamente entre animales infectados o vacunados (Li et al., 2017). Mientras que se ha demostrado que la inmunidad inducida por la infección con la sepa mutante atenuada RB51 se mide únicamente por la producción de células T. citotóxicas específicas, capaces de destruir los macrófagos infectados con *B. abortus* (Dorneles et al., 2015).

2.1.9.- Diagnóstico

El diagnóstico está basado en procedimientos serológicos (suero/leche) y bacteriológicos (Laboratorios, 2017). El estándar de oro para el diagnóstico de brucelosis es el aislamiento y la identificación de la bacteria causante, pero requiere un laboratorio de alta seguridad (instalaciones de contención biológica de nivel 3), personal altamente calificado, un tiempo de respuesta prolongado para los resultados y se considera un procedimiento peligroso. Por lo tanto, el diagnóstico de brucelosis se realiza principalmente mediante pruebas convencionales de detección serológica (Bharathan, Backhouse, Rawat, Naik, & Millar, 2016; Mirjalili, 2016). Mathew et al. (2015) informan sobre la importancia de aislar y caracterizar *Brucella*, ya que con los

métodos serológicos no es posible inferir con qué especie de *Brucella* está infectado el huésped.

Las pruebas moleculares como herramientas de diagnóstico están avanzando y pronto estará en el punto de reemplazar el aislamiento bacteriano real. El uso de la reacción en cadena de la polimerasa (PCR) para identificar el ADN de *Brucella* en el género, las especies e incluso los niveles de biovar se ha extendido para mejorar las pruebas de diagnóstico y se han desarrollado una diversidad de métodos. Las aplicaciones para los métodos de PCR van desde el diagnóstico de la enfermedad hasta la caracterización de aislamientos de campo con fines epidemiológicos, incluyendo estudios taxonómicos. Los ensayos basados en PCR también son útiles en pacientes crónicamente infectados donde el rendimiento de bacterias de hemocultivos es generalmente bajo. Es rápido, seguro y rentable, los únicos problemas reales son algunas incertidumbres con respecto a la especificidad. Además de los ensayos de PCR de uso común, se desarrolló un nuevo ensayo Multiplex-PCR que identificó específicamente *B. neotomae*, *B. pinnipedialis*, *B. ceti* y *B. microti*. Además, diferenció los biovares 1, 2, 4 de *B. abortus* de los biovares 3, 5, 6, 9, así como entre biozonas 1, biovares 3, 4 y biovares 2 y 5 de *B. suis* (Huber, Scholz, Lucero, & Busse, 2009).

En la actualidad la detección de anticuerpos contra (LPS) de *B. abortus* es la columna vertebral de la mayoría de los métodos serológicos para brucelosis. Sin embargo existe el riesgo de reacciones falsas positivas relacionadas con otros patógenos que tienen prominente reactividad cruzada con *B. abortus* (Hop et al., 2016).

En el Ecuador los laboratorios de diagnóstico animal de AGROCALIDAD realizan el diagnóstico de brucelosis bovina habitualmente mediante la prueba serológicas, Rosa de Bengala como prueba de screening, seguida de la prueba de ELISA competitivo para su confirmación (Morales, Espinosa, Torres, & Sandoval, 2009).

Tanto la prueba de ELISA indirecto como la prueba de ELISA competitivo y la prueba de Rosa de Bengala han sido aprobados por la Oficina Internacional de Epizootias (OIE) como pruebas prescritas para el comercio internacional. El ELISA indirecto puede utilizarse como prueba de screening o como prueba de confirmación. El rendimiento de la prueba del ELISA indirecto depende de la especificidad de la

inmunoglobulina del conjugado. Los conjugados con una amplia especificidad y el uso de diluciones de ensayo más bajas darán como resultado un ensayo de detección de mayor sensibilidad diagnóstica. Los conjugados con una especificidad estrecha y el uso de diluciones de ensayo más altas resultarán en un ensayo confirmatorio de mayor especificidad diagnóstica (Paweska et al., 2002).

2.1.10.- Validación de pruebas serológicas

La validación es un proceso que determina la aptitud de una prueba para un uso particular que ha sido debidamente desarrollada, optimizada y estandarizada incluyendo estimaciones de las características de rendimiento analítico y diagnóstico. Para que una prueba siga siendo válida es necesario evaluarla continuamente para garantizar que mantiene su idoneidad para el propósito deseado. Una prueba puede validarse para uno o varios propósitos deseados optimizando sus características de rendimiento para cada uno. Por ejemplo, fijando una alta sensibilidad diagnóstica (Dse) para una prueba de screening o por el contrario, fijando una alta especificidad diagnóstica (Dsp) para una prueba confirmativa (OIE, 2016e).

Según la Norma (ISO/IEC 17025, 2005), punto 5.4.5 Validación de métodos señala que «La validación es la confirmación, a través del examen y el aporte de evidencias objetivas, de que se cumplan los requisitos particulares para un uso específico previsto. La validación debe ser tan amplia como sea necesario para satisfacer las necesidades del tipo de aplicación o del campo de aplicación dados».

Huber como se citó en (Durán, 2016) afirma que los resultados de la validación del método pueden utilizarse para juzgar la calidad, confiabilidad y consistencia de los resultados analíticos.

2.1.11.- Criterios para el desarrollo y validación de una prueba

La sensibilidad, especificidad, exactitud, precisión, contaminación por arrastre y la incertidumbre son rasgos característicos de una prueba validada, los mismos que constituyen factores decisivos en los cuales se basa un juicio o decisión (OIE, 2016e).

2.1.11.1.- Sensibilidad y Especificidad Analítica

El límite de detección o sensibilidad analítica es la menor cantidad de analito en una muestra que se puede detectar mediante una prueba de detección directa en al menos

el 50% de las réplicas de cada dilución, en una serie de diluciones del analito en la matriz. La especificidad analítica es el grado en que la prueba distingue entre el analito en cuestión y otros componentes que pueden ser detectados en la prueba (OIE, 2016c).

2.1.11.2.- Sensibilidad y Especificidad Diagnóstica

La sensibilidad diagnóstica es la proporción de muestras de animales de referencia que se sabe que están infectados que dan un resultado positivo en una prueba. La especificidad diagnóstica es la proporción de muestras de animales de referencia que se sabe que no están infectados que dan un resultado negativo en una prueba. Los valores de sensibilidad y especificidad diagnóstica constituyen los parámetros más importantes que se establecen durante la validación de una prueba. Estas estimaciones son la base del cálculo de otros parámetros a partir de los cuales se realizan inferencias sobre los resultados de la prueba, por ejemplo, los valores predictivos de los resultados positivos y negativos de la prueba (OIE, 2016c).

2.1.11.3.- Valores predictivos

El valor predictivo positivo (PPV) es la probabilidad de que un animal que haya dado positivo en la prueba realmente sea positivo según el diagnóstico real. El valor predictivo negativo (NPV) es la probabilidad de que un animal que ha dado negativo en la prueba de hecho sea negativo según el diagnóstico real (OIE, 2016c).

2.1.11.4.- Exactitud/Precisión

La exactitud es el nivel de concordancia entre el valor obtenido en la prueba y el valor esperado para un estándar de referencia de actividad o título conocido. La precisión es el grado de dispersión de los resultados para una muestra ensayada repetidamente (OIE, 2016c).

2.1.11.5.- Veracidad

La veracidad se define como el grado de coincidencia entre el valor medio obtenido de una gran serie de resultados y un valor aceptado como referencia (Tafur, 2014).

Cuando se aplica a un método de ensayo, el término “exactitud” se refiere a una combinación de veracidad y precisión (Figura 1).

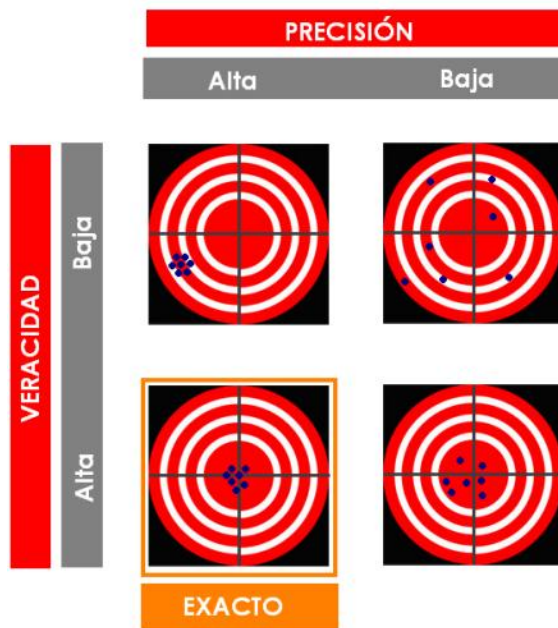


Figura 1. Esquema de “Tiro al Blanco”Esquema de “Tiro al Blanco”, utilizado para ejemplificar métodos de ensayo, los punto u orificios equivaldrían a los resultados analíticos y el círculo rojo al centro el rango en el cual se espera este el valor de referencia (o verdadero) (Daffau et al., 2010)

2.1.11.6.- Repetibilidad/ Reproducibilidad

La repetibilidad es el nivel de acuerdo entre las réplicas de una muestra, tanto dentro una aplicación como entre varias aplicaciones del mismo método de prueba en un mismo laboratorio. La reproducibilidad es la capacidad que una prueba tiene de proporcionar resultados consistentes cuando se aplica a las alícuotas de la misma muestra en diferentes laboratorios (OIE, 2016c).

2.1.11.7.- Contaminación por arrastre (Efecto Carry-over)

Ocurren cuando los residuos de un ensayo no desaparecen antes del inicio del ensayo siguiente. Esta situación constituye un problema importante porque el ensayo inicial puede haber alterado los resultados al ensayo posterior. Por tanto, si residuos del ensayo inicial están presentes cuando se da el segundo ensayo, el experimentador no será capaz de determinar si el resultado es debido al segundo tratamiento, a los residuos iniciales, o a la interacción de ambos (“Validez Interna,” 2005).

2.1.5.8.- Incertidumbre de la medición

La incertidumbre de una medición es el intervalo de medida dentro el cual se encuentra el valor verdadero, cuando se han tenido en cuenta todas las fuentes de error

(Committee, 1995). La estimación de la incertidumbre, denominada imprecisión de la medición, es un requisito para los laboratorios de análisis con resultados cuantitativos, que se basan en normas internacionales de calidad como la ISO/IEC 17025-2005(OIE, 2016d).

2.2.- Hipótesis

2.2.1.- Hipótesis nula

La prueba de ELISA indirecto si es válida como herramienta de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador.

2.2.2.- Hipótesis alternativa

La prueba de ELISA indirecto no es válida como herramienta de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador

2.3.- Señalamiento de variables de la hipótesis

2.3.1.- Variables independientes

La prueba de ELISA indirecto como herramienta de diagnóstico.

2.3.2.- Variables dependientes

Sensibilidad y especificidad analítica, sensibilidad y especificidad diagnóstica, valores predictivos, exactitud, contaminación por arrastre, precisión (repetibilidad, reproducibilidad), e incertidumbre del método.

CAPÍTULO III

MATERIAL Y MÉTODOS

3.1.- Materiales

3.1.1.- Material de Referencia Internacional Certificado (MIRC)

Material de referencia internacional ID.vet

1 suero positivo Anexo F

2 suero negativo Anexo G

3.1.2.- Material de Referencia Nacional Certificado (MRNC)

11 sueros positivos a brucelosis certificados y donados por el Centro Internacional de Zoonosis (CIZ).

30 sueros negativos de predios ganaderos libres de brucelosis de la provincia de Loja cantón Saraguro, certificados por AGROCALIDAD.

3.1.3.- Reactivos

Tabla 1. Reactivos del Kit de ELISA indirecto

1	Placa tapizada con LPS de <i>Brucella</i>	2	10
2	Control Positivo	1 x1,0ml	1 x 1,0ml
3	Control Negativo	1 x1,0ml	1 x 1,0ml
4 ^a	Conjugado Concentrado	1 x1,5ml	1 x 1,5ml
4b	Solución tampón de dilución n.º1	1 x120ml	1 x120ml
5	Solución tampón de dilución n.º2	1 x120ml	2 x 120 ml
A	Substrato TMB n.º13	1 x 60 ml	1 x 120 ml
B	Solución de frenado n.º3	1 x 60 ml	1 x 120 ml
C	Solución de lavado concentrada (20x)	1 x100ml	2 x 100 ml
Otros componentes: bolsa de plástico de cierre hermético reutilizable		1	1

3.1.4.- Equipos

Incubadora de microplaca VWR, Micropipetas, Lector de microplaca (espectrofotómetro) BioTek, Termo-agitador ALLSHENG

3.2.- Metodología

La metodología estaba basada en el Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres de la Organización Mundial de Sanidad Animal (OIE) 2016, el cual describe en diferentes capítulos el procedimiento de validación de pruebas serológicas Anexos J, K y L. Los análisis de sensibilidad y especificidad

analítica, repetibilidad y reproducibilidad fueron realizados por dos analistas, como analista 1 Edwin Martínez y analista 2 Jessica Valseca analista de AGROCALIDAD.

3.2.1.- ELISA indirecto

Las placas estaban tapizadas con un lipopolisacárido (LSP) de *Bucella abortus* en donde cualquier anticuerpo presente en la muestra específico frente a *Brucella abortus* formó un complejo LPS de antígeno-anticuerpo en la superficie del pocillo. Se dejó todos los reactivos y las muestras que adquirieran 18-26°C antes de usarlos temperatura óptima en la cual trabaja los reactivos evitando resultados erróneos. Los reactivos se mezclaron agitándolos suavemente para no provocar que sus componentes se destruyan por la fuerza del golpe. Se colocó 10 µl del control positivo y 10 µl del control negativo por duplicado en 190 µl de solución tampón de dilución n.º2 (buffer). 10 µl de las muestras a ser analizadas se diluyeron en 190 µl de solución tampón de dilución n.º2 sin repetición y se incubaron por 60 ± 5 min en los pocillos tiempo establecido por IDEXX. Se lavó 3 veces con 300 µl de solución de lavado 1x (se preparó a partir de la solución de lavado concentrada 20x del kit) para la eliminación de anticuerpos que no reaccionaron, se colocó 100 µl de una dilución 1:100 del conjugado (anti-anticuerpos ligados a una enzima) concentrado en solución tampón de dilución n.º1 los cuales reaccionaron con los anticuerpos específicos añadidos en el paso anterior y que se encentraron fijados al LSP. Se volvió a lavar para eliminar los anti-anticuerpos que no reaccionaron, se añadió a los pocillos el Sustrato TMB n.º13 sobre el que es capaz de actuar la enzima marcadora. En presencia del enzima peroxidasa, el sustrato se oxido generando una coloración azul, que se viro amarilla al añadir la solución de frenado n.º3 ácido sulfúrico concentrado. La densidad del color era proporcional a la concentración de anticuerpos específicos anti-*Brucella abortus* presentes en la muestra. El resultado se obtuvo comparando la densidad óptica (DO) de la muestra con la media del Control Positivo.

Cálculos

Controles

$$CN\bar{x} = [CN1A(450) + CN2A(450)]/2 \quad (ec.1)$$

$$CP\bar{x} = [CP1A(450) + CP2A(450)]/2 \quad (ec.2)$$

Dónde: $CN\bar{x}$ promedio de los controles negativo, $CP\bar{x}$ promedio de los controles positivos, CN densidad óptica de los controles negativos, CP densidad óptica de los controles positivos

Criterios de validación de los resultados de la prueba de ELISA indirecto

Si $CPx \geq 0,350 \bar{y}$ la relación $CPx/CNx \geq 3,00$ la prueba es válida.

Muestras

$$M/P \% = 100[(\text{Muestra A}(450) - CN\bar{x})/(\bar{CPx} - CN\bar{x})](ec.3)$$

Dónde: $M/P \%$ relación entre la densidad óptica de la muestra y la densidad óptica del control positivo, $\text{Muestra A}(450)$ densidad óptica de la muestra, $CN\bar{x}$ promedio de los controles negativo, $CP\bar{x}$ promedio.

Interpretación

Negativo	Dudoso	Positivo
$M/P \% \leq 110$	$110 < M/P \% < 120$	$M/P \% \geq 120$

3.2.2.- Sensibilidad Analítica

Para estimar la sensibilidad analítica o límite de detección, a través de la prueba ELISA indirecto se realizó 10 diluciones (1:2, 1:4, 1:8, 1:16, 1:32, 1:64, 1:128, 1:256, 1:512 hasta 1:1024) del suero positivo en el suero negativo (material de referencia internacional certificado). Se indicó el punto de corte mediante el análisis de la Característica Operativa del Recepto (ROC) en la última dilución que se evidenció la positividad, 3 repeticiones, 2 analistas, 3 días. Total de pruebas: 66

3.2.3.- Especificidad Analítica

Para estimar la especificidad analítica a través de la prueba ELISA indirecto y detectar la presencia de reacciones cruzadas con: Rinotraqueítis Infecciosa Bovina (IBR),

Diarrea Viral Bovina (DVB), leucosis, neosporosis bovina y lengua azul se utilizó los sueros controles positivos y negativos de los kits de ELISA de cada una de las diferentes enfermedades. Se indicó el M/P% en base a la (ec. 3) del kit de ELISA indirecto IDEXX para brucelosis, 3 repeticiones, 1 analista, 3 días. Total de pruebas: 60.

3.2.4.- Sensibilidad y Especificidad Diagnóstica

Para estimar la sensibilidad y especificidad diagnósticas, a través de la prueba ELISA indirecto, se analizó, un suero positivo (material de referencia internacional certificado), 11 sueros positivos y 30 sueros negativos (material de referencia nacional certificados) 1 repetición, 1 analista, 1 día. Total de pruebas: 42.

$$DSe = VP / (VP+FN)(ec. 4)$$

$$DSp = VN / (VN+FP)(ec. 5)$$

Dónde: (DSe) Sensibilidad diagnóstica, (DSp) Especificidad diagnóstica, (VP) Verdaderos positivos, (FP) Falsos positivos, (VN) Verdaderos negativos, (FN) Falsos negativos.

3.2.5.- Valores Predictivos Positivos y Negativos

Los valores predictivos positivos y negativos se calculó en base a los resultados obtenidos en el apartado 3.2.4.

$$PPV = (P * Dse) / (P * Dse + (1 - P) * (1 - Dse)) \quad (ec. 6)$$

$$NPV = ((1 - P) * DSp) / (P * (1 - Dse) + (1 - P) * DSp) \quad (ec. 7)$$

Dónde: (PPV) Valores predictivos positivos, (NPV) valores predictivos negativos (DSe) Sensibilidad diagnóstica, (DSp) Especificidad diagnóstica, (P) prevalencia.

3.2.6.- Exactitud

Para estimar la exactitud se realizó el análisis en el control negativo del kit, control positivo débil (dilución de punto de corte determinado mediante la sensibilidad analítica) y un suero positivo (material de referencia internacional certificado). 10 repeticiones, 1 analista, 1 día. Total de pruebas: 30. Se realizó los cálculos por INDICE DE KAPPA.

Tabla 2. Valoración del coeficiente Kappa

Coeficiente Kappa	Fuerza de concordancia
0,00	Pobre
0,01 - 0,20	Leve
0,21 - 0,40	Aceptable
0,41 - 0,60	Moderada
0,61 - 0,80	Considerable
0,81 - 1,00	Casi perfecta

(López & Fernández, 2001)

$$k = (Fc - F\infty)/(1 - F\infty) \quad (ec. 8)$$

Error estándar

$$Sek = (Fc - F\infty)/(1 - F\infty) \quad (ec. 9)$$

Intervalo de confianza 95%

$$IC = k \pm ZcSekZc = (Z\infty) 1.96(ec. 10)$$

$$\% \text{ exactitud} = (VP+VN)*100/(VP+FP+VN+FN) \quad (ec. 11)$$

Dónde: (VP) Verdaderos positivos, (FP) Falsos positivos, (VN) Verdaderos negativos, (FN) Falsos negativos.

3.2.8.- Contaminación por arrastre

Para estimar el arrastre, a través de la prueba ELISA Indirecto, se analizaron 11 sueros positivos y 24 sueros negativos (material de referencia nacional certificado). Se distribuyó en forma alternada de 2 sueros positivos y 3 sueros negativos. 1 repetición, 1 analista, 1 día. Total de pruebas 35. Se calculó a través del cálculo de Carry over.

$$q = (b1 - b3)/(a2 - b3) \quad (ec. 12)$$

Dónde: q: índice de arrastre, a2: control positivo 2, b1: control negativo 1, b3: control negativo 3.

3.2.7.- Precisión (Repetibilidad/ Reproducibilidad)

Para estimar la reproducibilidad y repetibilidad, a través de la prueba ELISA indirecto, se realizó 10 diluciones (1:2, 1:4, 1:8, 1:16, 1:32, 1:64, 1:128, 1:256, 1:512 hasta 1:1024) del suero positivo en el suero negativo (material de referencia internacional certificado). Se realizó los cálculos por ANOVA y t Student. 3 repeticiones, 2 analistas, 3 días. Total de pruebas: 66

3.2.9.- Incertidumbre de la medición

Para estimar la incertidumbre a través de la prueba ELISA indirecto, se realizó 20 repeticiones del control positivo débil (dilución de punto de corte determinada mediante la sensibilidad analítica). 1 repetición, 1 analista, 1 día. Total de pruebas: 20. Se calculó a través del cálculo de la Incertidumbre.

$$RSD (X) = SD (X) / (\bar{X}) \quad (ec. 13)$$

Dónde: (RSD) desviación estándar relativa o coeficiente de variación (CV), (SD) desviación estándar, (\bar{X}) promedio.

Incertidumbre ampliada

$$U (IC del 95\%) = 2 \times RSD(ec. 14)$$

Para evaluar el cumplimiento de la prueba de ELISA indirecto se definió los siguientes parámetros con sus respectivos objetivos de validación (Tabla 3)

Tabla 3. Objetivos de validación de la prueba de ELISA indirecto

Parámetros	Objetivos de validación
Sensibilidad analítica (limite detección)	Punto de corte
Especificidad analítica	92%
Sensibilidad diagnóstica	99%

Especificidad diagnóstica	98%
Valor predictivo positivo	$\geq 50\%$
Valor predictivo negativo	$\geq 50\%$
Exactitud	$> 95\%$
Arrastre	$< 1\%$
Repetibilidad	$\geq 92\%$
Reproducibilidad	$\geq 92\%$
Incertidumbre	$< 5\%$

CAPITULO IV

RESULTADOS Y DISCUSIÓN

4.1.- Análisis y discusión de los resultados

4.1.1.- Estimación de la sensibilidad analítica y punto de corte

La prueba de ELISA indirecto fue sensible hasta la dilución 1/32 ultima dilución en la que se evidencio positividad (Anexo A). La misma que fue justificada mediante un análisis Probit para cada analista (Figura 2) (Guthrie et al., 2013). La sensibilidad fue menor a la estimada por Vanzini et al. (1998) teniendo en cuenta que los trabajos fueron realizados con un Kit de ELISA indirecto de otra casa comercial. La sensibilidad de una prueba de ELISA va a depender del anticuerpo secundario utilizado (Saaid, 2014). Mantur et al. (2010) evaluaron la prueba de ELISA indirecto frente a pruebas convencionales demostrando que la prueba de ELISA indirecto es más sensible y recomiendan la prueba para un diagnóstico preciso en casos sospechosos de detección de casos de brucelosis.

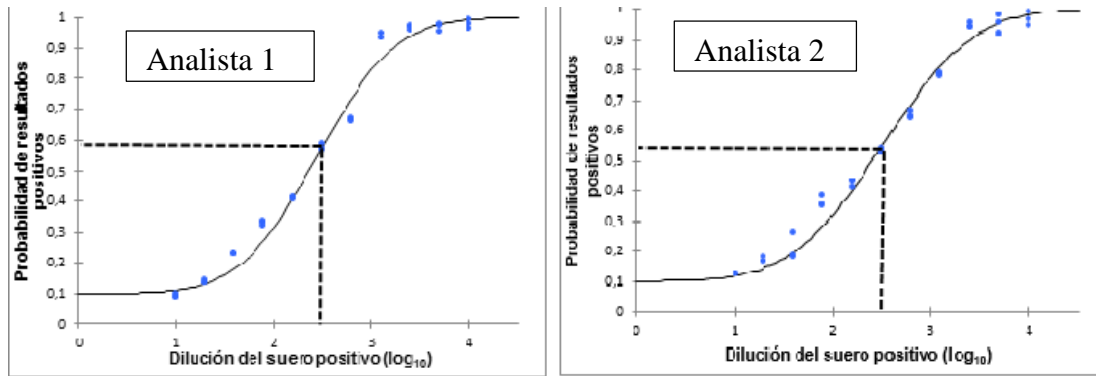


Figura 2. Análisis Probit para cada analista. Cada dilución se repitió 3 veces. La sensibilidad analítica al 95% de confianza se indica con líneas discontinuas.

El punto de corte de la prueba de ELISA indirecto se estimó mediante el análisis de la curva característica de funcionamiento del receptor (ROC) (Figura 3). El punto de corte estimado fue estadísticamente de 121,130% a valores máximos de sensibilidad y especificidad. Valor que no difiere estadísticamente de 120% punto de corte sugerido por los laboratorios IDEXX (Anexo C). Según Guthrie et al. (2013) la selección del punto de corte depende del uso pretendido de la prueba. Por ejemplo, puede ser aconsejable puntos de corte bajos cuando la prueba sea de pesquisa, mientras que cuando la prueba sea de confirmación es preferible puntos de corte altos (Greiner, Pfeiffer, & Smith, 2000). Gorsich, Bengis, Ezenwa, & Jolles (2015) propusieron un punto de corte diferente al sugerido para trabajar con otras especies de animales.

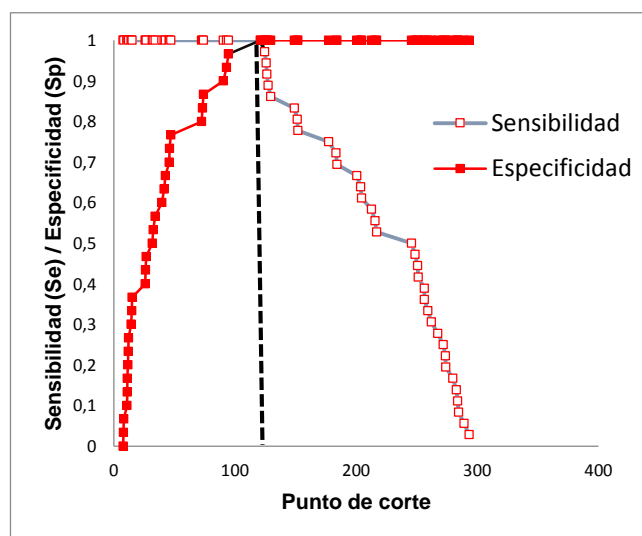


Figura 3. Análisis (ROC). Representa la sensibilidad (Se) y la especificidad (Sp) en función del punto de corte de la prueba. La línea vertical discontinua muestra el punto de corte al 95% de confianza.

4.1.2.- Estimación de la especificidad analítica

La especificidad analítica fue estimada en sueros controles del kit de ELISA de cada enfermedad (Tabla 4) que puede estimular reacción cruzada de los anticuerpos. La prueba fue específica para *B. abortus* al no mostrar positividad a ninguna de la enfermedades analizadas. El mayor riesgo de reacciones falsas positivas está relacionado con enterobacterias, especialmente con *Yersinia enterocolitica* que tiene la reactividad cruzada más prominente con el género *Brucella spp* (Hop et al., 2016). Sin embargo, estudios realizados en rinotraqueítis infecciosa bovina IBR (Kamaraj, Chinchkar, Rajendra, & Srinivasan, 2009), diarrea viral bovina DVB (Broaddus, Lamm, Kapil, Dawson, & Holyoak, 2009), neospora (Vemulapalli et al., 2007), leucosis (Oyejide, Adu, Makinde, & Ezch, 1987) y lengua azul (Waldvogel, Anderson, Phillips, & Osburn, 1992) muestran cómo interactúan cada una de ellas con *B. abortus*. Al Dahouk, Tomaso, Nöckler, Neubauer, & Frangoulidis(2003) describen las ventajas y desventajas de la utilización de pruebas ELISA al momento de trabajar con agentes infecciosos que pueden producir reacciones cruzados con *B. abortus*.

Tabla 4. Especificidad analítica para la prueba de ELISA indirecto

Controles de los Kits para cada enfermedad	Día 1		Día 2		Día 3		RESULTADOS
	M/P%		M/P%		M/P%		
	An1	An2	An1	An2	An1	An2	
IBR positivo	2,2	2,5	1,0	2,3	1,1	1,7	NEGATIVO
IBR negativo	-1,8	0,4	-0,7	0,5	-0,5	-0,3	NEGATIVO
DVB positivo	-0,9	0,1	0,3	0,5	0,1	-0,3	NEGATIVO
DVB negativo	0,3	0,9	0,8	1,0	-0,4	0,0	NEGATIVO
NEOSPORA positivo	-0,7	0,7	-0,8	0,3	0,1	-0,7	NEGATIVO
NEOSPORA negativo	-0,8	0,4	-0,6	-0,8	-0,2	-0,7	NEGATIVO
LEUCOSIS positivo	-1,1	0,6	-0,4	-0,4	-1,0	-0,6	NEGATIVO
LEUCOSIS negativo	-1,0	0,7	-0,4	-0,6	-0,9	-0,2	NEGATIVO
LENGUA AZUL positivo	-0,6	1,9	0,8	1,4	0,3	1,1	NEGATIVO
LENGUA AZUL negativo	5,9	6,3	5,3	7,4	2,8	8,4	NEGATIVO

Estudio realizado durante tres días por dos analistas. An1 analista 1, An2 analista 2, M/P% relación muestra control positivo. Se utilizaron los controles positivos y negativos de cada enfermedad. (IBR) rinotraqueítis infecciosa bovina, (DVB) diarrea viral bovina.

4.1.3.- Estimación de los valores de sensibilidad y la especificidad diagnóstica

Los valores de la sensibilidad y la especificidad diagnóstica fueron estimados a partir de los datos del Anexo By de las ec. 4 y 5 respectivamente (Tabla 5). Se estimó el 100% en cada uno de los parámetros, valores que no difieren de los sugerido por la OIE, al utilizar sueros certificados. Gorsich et al. (2015) demuestran que al trabajar con sueros no certificados los resultados de sensibilidad y especificidad diagnóstica son menores. Según Safronetz et al. (2017) una baja sensibilidad dará lugar a falsos negativos, de igual manera una baja especificidad dará lugar a falsos positivos.

Tabla 5. Estimaciones de la sensibilidad y la especificidad diagnósticas de la prueba de ELISA indirecto

Animales de referencia cuyo estado infeccioso es conocido			
	Positivas (12)	Negativas (30)	TOTAL
POSITIVOS	12	0	12
	VP	FP	
NEGATIVOS	0	30	30
	FN	VN	
TOTAL	12	30	42
	Sensibilidad de Diagnóstico (ec.4) 100%	Especificidad de Diagnóstico (ec.5) 100%	

Datos obtenidos partir de un conjunto de resultados de muestras analizadas procedentes de poblaciones cuyo estado infeccioso es conocido, material de referencia nacional. VP = Verdadero Positivo, FP = Falso Positivo, VN = Verdadero Negativo, FN = Falso Negativo.

4.1.4.- Estimación de los valores predictivos positivo y negativo

Los valores predictivos positivo y negativo se estimaron a partir de los resultados obtenidos de la sensibilidad y la especificidad diagnóstica (Tabla 5). Morales et al. (2009) establece el 2% de prevalencia de la enfermedad, dato requerido para el respectivo cálculo. Los valores predictivos positivo y negativo obtenidos mediante las ec. 6 y 7 demuestran que los sueros utilizados en los ensayos son verdaderos positivos y verdaderos negativos, al obtener el 100% en cada uno de los respectivos parámetros.

4.1.5.- Determinación de la exactitud de la prueba

Este parámetro fue evaluado a través de Kappa a un nivel de confianza del 95% para resultados binarios (Anexo C), como establece el Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres 2016.

Tabla 6. Determinación de la fuerza de concordancia y la exactitud por índice kappa

ÍNDICE KAPPA	FUERZA DE LA CONCORDANCIA	EXACTITUD
0,927	Casi perfecta	97%

El Índice Kappa fue utilizado para la evaluación de los resultados obtenidos en el control negativo del Kit, control positivo débil (dilución 1/32) y el material de referencia internacional positivo midiendo la fuerza de concordancia entre los valores obtenidos.

La fuerza de concordancia de la prueba fue casi perfecta al obtener un valor de Kappa igual a 0,927 (Tabla 6). Se puede afirmar que la exactitud calculada mediante la ec.11 de la prueba de ELISA indirecto es 97%.

4.1.6.- Análisis de la contaminación por arrastre

El índice de arrastre calculado (Anexo D) en cada uno de los test no fue mayor del 1% (Tabla 7). Se puede afirmar que no existió contaminación en ninguna de las etapas de la metodología ni en la utilización de equipos y materiales. Petrov, Quetel, & Taylor (2007) determinan que el manejo de la muestra y el ambiente son la mayor fuente de contaminación ya que contribuyen con ~75% del total.

Tabla 7. Índice de contaminación por arrastre

N° Test	Índice de arrastre %
test 1	-125,5
test 2	-404,8
test 3	-109,5
test 4	-292,7
test 5	0,37
test 6	-0,85
test 7	-2,5

El índice de arrastre se calculó mediante la ec. 12 clasificando como test a 2 positivos y 3 negativos.

4.1.7.- Análisis de precisión

4.1.7.1.- Repetibilidad

Para la estimación de la repetibilidad se calculó la desviación estándar (s) y el coeficiente de variación (%CVr) como se detalla en la (Tabla 8). Estos criterios estadísticos son aplicados a toda clase de mediciones con repeticiones (OIE, 2016b).

Tabla 8. Criterios de estadísticos para la estimación de repetibilidad de la prueba de ELISA indirecto

Muestras	Analista 1			Analista 2		
	\bar{X}	S	CVr%	\bar{X}	S	CVr%
PURO	279,12	5,15	1,84	288,98	4,38	1,52
1/2	262,96	4,16	1,58	276,22	5,72	2,07
1/4	248,46	2,51	1,01	254,59	2,85	1,12
1/8	202,94	2,07	1,02	215,05	2,17	1,01
1/16	150,89	1,54	1,02	181,61	3,72	2,05
1/32	127,54	1,86	1,45	124,02		2,15
1/64	73,25	0,77	1,05	92,72	2,15	2,32
1/128	46,26	0,79	1,70	41,26	1,50	3,64
1/256	26,35	0,42	1,61	32,79	1,13	3,43
1/512	14,82	0,43	2,90	11,92	0,15	1,27
1/1024	11,03	0,11	0,98	8,05	0,26	3,25

Los cálculos están realizados por diluciones y analistas, (X) media, (S) desviación estándar, (CVr%) coeficiente de variación (Anexo A).

Los valores del coeficiente de variación oscilaron entre 0,98% y 3,25% lo que indica una baja variación entre las diferentes repeticiones. Demostrando que el método es repetible, al no sobrepasar el 5% establecido por la Norma ISO/IEC 17025-2005. A través del análisis de varianza de un factor (ANOVA) se evaluó el cumplimiento de este parámetro por cada analista (Tabla 9).

Tabla 9. Análisis de varianza para la repetibilidad de la prueba de ELISA indirecto

	F	Probabilidad	Valor crítico para F
Analista 1	0,00086281	0,999137583	3,315829501
Analista 2	7,76E-5	0,999922307	3,315829501

La repetibilidad se evaluó por analista, mediante el estadístico F.

El análisis de varianza (ANOVA) realizado para cada analista, afirma que la prueba tiene una alta repetibilidad al 95% de confianza debido a que los F calculados son de menor valor al valor crítico de F 3,315.

4.1.7.2.- Reproducibilidad

Prueba t de student

La prueba fue utilizada para determinar si existe diferencia significativa entre los dos analistas (Tabla 10).

Tabla 10. Prueba de t, análisis de medias entre el analista 1 y analista 2

	Analista 1	Analista 2
Media	131,237089	138,836924
Varianza	10137,3237	11141,6903
Observaciones	33	33
Varianza agrupada	10639,507	
Diferencia hipotética de las medias	0	
Grados de libertad	64	
Estadístico t	-0,29928528	
P(T<=t) una cola	0,38284608	
Valor crítico de t (una cola)	1,66901303	
P(T<=t) dos colas	0,76569215	
Valor crítico de t (dos colas)	1,99772965	

Se utilizaron los datos de cada analista obtenidos durante los tres días de trabajo.

A un 95 % de confianza, no existe diferencia hipotética entre las medias y el estadístico t calculado es menor al valor crítico. Valores que permite afirmar que no existe diferencia significativa entre los dos analistas y que la prueba es reproducible.

4.1.8.- Análisis de la incertidumbre de la medición.

Se obtuvo una incertidumbre (Anexo E) de 4,8% mediante la (ec. 15). Cumpliendo con lo establecido en la Norma ISO/IEC 17025-2005 que no debe ser mayor al 5%. Requisito para los laboratorios de análisis que producen resultados cuantitativos y se basa en normas internacionales (OIE, 2016).

4.1.8.1.- Interpretación de la incertidumbre de la medición.

Tabla 11. Interpretación de la incertidumbre

Valores de M/P%	Resultado
$\leq 116,8$	Negativos
$>116,8$ y $< 125,93$	Sospechosos
$\geq 125,93$	Positivos

M/P% relación muestra control positivo

Finalmente, a través de la determinación y comprobación de los parámetros propuestos, se declaró validada la prueba de ELISA indirecto IDEXX (ANEXO D).

El kit de ELISA indirecto IDEXX fue sensible hasta una dilución 1/32 datos que concuerdan entre los dos analistas (Anexo A). Guthrie et al. (2013) sugieren abrir las diluciones para ser más exactos en la sensibilidad analítica. Se estimó 121,130% como valor de corte del kit de ELISA indirecto a condiciones del laboratorio de AGROCALIDAD, valor al que la sensibilidad y especificidad fueron del 100%. El valor de corte seleccionado debe tenerse en cuenta al comparar las pruebas de diagnóstico (Greiner et al., 2000). Por ejemplo, con el valor de corte sugerido para el ganado pero no en las mismas condiciones es 120%, el ELISA tuvo un valor de especificidad más bajo y un valor de sensibilidad más alto. Este resultado enfatiza la importancia de la optimización de las pruebas para cada población, especies y condiciones a las que se les aplica. La prueba de ELISA indirecto fue específica para anticuerpos producidos por *B. abortus* al no reaccionar con las diferentes enfermedades reportadas en la Tabla 4. Análisis realizado por dos analistas para la comparación y aseveración los resultados obtenidos. Hop et al. (2016) demuestras que anticuerpos producidos por *Yersinia enterocolitica* pueden generar reacciones cruzadas con anticuerpos producidos por el género *Brucella spp.* Estudios realizados por Gorsich et al. (2015) demuestran que el kit de ELISA indirecto IDEXX no tiende a reaccionar con anticuerpos producidos por la bacteria *Yersinia enterocolitica*. Los valores de sensibilidad y especificidad diagnósticas y los valores predictivos demostraron que la prueba de ELISA indirecto es capaz de reconocer sujetos enfermos y sanos de una población. Según Arruda et al. (2016) una prueba de diagnóstico entrega un resultado correcto al obtener valores altos en sensibilidad especificidad y en los valores predictivos. Los parámetros estadísticos demostraron que la prueba es confiable, reproducible, con una alta exactitud y la incertidumbre dentro de lo establecido como requiere la norma internacional ISO/IEC 17025-2005 y el Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres de la Organización Mundial de Sanidad Animal (OIE) 2016. Ensayos adicionales para la brucelosis, incluyendo el ensayo de polarización de fluorescencia (Gall et al., 2000),

ensayos de reacción en cadena de polimerasa (Bricker, 2002) y técnicas ELISA alternativas (Nielsen, 2002) han demostrado una mayor precisión en el diagnóstico de esta enfermedad en otras especies, los mismos que se deberían considerar para diagnósticos más acertados en ganado vacuno.

4.2.- Verificación de hipótesis

Mediante de las pruebas realizadas a lo largo de este trabajo, se confirma que la prueba de ELISA indirecto es válida como herramienta de diagnóstico para el Programa Nacional de Brucelosis Bovina en el Ecuador cumpliendo con lo establecido en el certificado de análisis del Kit de ELISA indirecto de los Laboratorios IDEXX (Anexo C), los criterios de calidad requeridos por el Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres de la Organización Mundial de Sanidad Animal (OIE) 2016 y la norma ISO/IEC 17025-2005.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

5.1.- Conclusiones

Se estableció que la prueba de Elisa indirecto es sensible hasta la dilución 1/32 y es 100% específica para anticuerpos producidos por la bacteria *Brucella abortus*.

Se determinó que los valores de sensibilidad y especificidad diagnóstica de la prueba de ELISA indirecto son del 100%. Valores que permiten decir que la prueba es capaz de dar como casos positivos sujetos enfermos y como casos negativos a sujetos sanos correctamente identificados.

Se determinó los parámetros de validación a un 95% de confianza cumpliendo con lo establecido en la norma ISO/IEC 17025, el Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres de la Organización Mundial de Sanidad Animal y el certificado de análisis del Kit de ELISA indirecto de IDEXX (Anexo C) permitiendo obtener la declaración de validación de la prueba (Anexo D).

5.2.- Recomendaciones

Los materiales deben estar debidamente almacenados en lugares en donde no de la luz del sol.

Para un diagnóstico más acertado se recomienda que la historia clínica del animal sea bien llevada desde su nacimiento hasta su sacrificio.

Los beneficios del ELISA indirecto son: relativamente barato, fácil de realizar en condiciones de campo, y da como resultado resultados cuantitativos de las pruebas.

Sin embargo, la elección de una prueba diagnóstica apropiada depende de su uso previsto.

BIBLIOGRAFÍA

- Al Dahouk, S., Tomaso, H., Nöckler, K., Neubauer, H., & Frangoulidis, D. (2003). Laboratory-based diagnosis of brucellosis--a review of the literature. Part II: serological tests for brucellosis. *Clinical Laboratory*, 49(11–12), 577–589.
- Alton, G. G., & Forsyth, J. R. L. (1996). Brucella. En S. Baron (Ed.), *Medical Microbiology* (4th ed.). Galveston (TX): University of Texas Medical Branch at Galveston. Recuperado a partir de <http://www.ncbi.nlm.nih.gov/books/NBK8572/>
- Arruda, M. M. de, Figueiredo, F. B., Marcelino, A. P., Barbosa, J. R., Werneck, G. L., Noronha, E. F., & Romero, G. A. S. (2016). Sensitivity and specificity of parallel or serial serological testing for detection of canine Leishmania infection. *Memórias do Instituto Oswaldo Cruz*, 111(3), 168–173. <https://doi.org/10.1590/0074-02760150364>
- Aznar, M. N., Samartino, L. E., Humblet, M.-F., & Saegerman, C. (2014). Bovine brucellosis in Argentina and bordering countries: update. *Transboundary and Emerging Diseases*, 61(2), 121–133. <https://doi.org/10.1111/tbed.12018>
- Bharathan, B., Backhouse, L., Rawat, D., Naik, S., & Millar, M. (2016). An unusual case of seronegative, 16S PCR positive Brucella infection. *JMM Case Reports*, 3(5). <https://doi.org/10.1099/jmmcr.0.005050>
- Bricker, B. J. (2002). PCR as a diagnostic tool for brucellosis. *Veterinary Microbiology*, 90(1), 435–446. [https://doi.org/10.1016/S0378-1135\(02\)00228-6](https://doi.org/10.1016/S0378-1135(02)00228-6)

- Broaddus, C. C., Lamm, C. G., Kapil, S., Dawson, L., & Holyoak, G. R. (2009). Bovine Viral Diarrhea Virus Abortion in Goats Housed with Persistently Infected Cattle. *Veterinary Pathology*, 46(1), 45–53. <https://doi.org/10.1354/vp.46-1-45>
- Cloeckaert, A., Jacques, I., de Wergifosse, P., Dubray, G., & Limet, J. N. (1992). Protection against *Brucella melitensis* or *Brucella abortus* in mice with immunoglobulin G (IgG), IgA, and IgM monoclonal antibodies specific for a common epitope shared by the *Brucella* A and M smooth lipopolysaccharides. *Infection and Immunity*, 60(1), 312–315.
- Cold, F., Health, E., Disease, P., Management, P., Conditions, S., & Problems, S. (2017). Brucellosis: Causes, Symptoms, and Treatment. Recuperado 13 de junio de 2017, a partir de <http://www.webmd.com/a-to-z-guides/brucellosis-symptoms-treatment>
- Committee, A. M. (1995). Uncertainty of measurement: implications of its use in analytical science. *Analyst*, 120(9), 2303–2308. <https://doi.org/10.1039/AN9952002303>
- Daffau, B., Rojas, F., Guerrero, I., Roa, L., Rodríguez, L., Soto, M., & Sandoval, S. (2010). *Validación de métodos y determinación de la incertidumbre de la medición: “Aspectos generales sobre la validación de métodos*. Santiago: Instituto de Salud Pública de Chile. Recuperado a partir de http://www.ispch.cl/sites/default/files/documento_tecnico/2010/12/Guia%20T%C3%A9cnica%20de%20validaci%C3%B3n%20de%20M%C3%A9todos%20y%20determinaci%C3%B3n%20de%20la%20incertidumbre%20de%20la%20medici%C3%B3n_1.pdf

- Departement Agriculture, Forestry and Fisheries, R. O. S. A. (2016). *BOVINE BRUCELLOSIS MANUAL*. Recuperado a partir de http://www.nda.agric.za/vetweb/pamphlets&Information/Policy/Brucellosis%20in%20Cattle%20Interim%20Manual%20for%20the%20Veterinarian%20%20&%20AHT%20-%20Sept2016_signed.pdf
- Dorneles, E. M. S., Lima, G. K., Teixeira-Carvalho, A., Ara?jo, M. S. S., Martins-Filho, O. A., Sriranganathan, N., ... Lage, A. P. (2015). Immune Response of Calves Vaccinated with *Brucella abortus* S19 or RB51 and Revaccinated with RB51. *PLOS ONE*, *10*(9), e0136696. <https://doi.org/10.1371/journal.pone.0136696>
- Ducrottoy, M. J., Ammary, K., Ait Lbacha, H., Zouagui, Z., Mick, V., Prevost, L., ... Benkirane, A. (2015). Narrative overview of animal and human brucellosis in Morocco: intensification of livestock production as a driver for emergence? *Infectious Diseases of Poverty*, *4*(1). <https://doi.org/10.1186/s40249-015-0086-5>
- Durán, F. (2016). *Validación de los Métodos de Ensayo de Punto de Inflamación en Cemento Asfáltico, Gasóleo y Jet Fuel*. Universidad Central del Ecuador, Quito-Ecuador. Recuperado a partir de <http://www.dspace.uce.edu.ec/bitstream/25000/7186/1/T-UCE-0017-0035-2016.pdf>
- Gall, D., Nielsen, K., Forbes, L., Davis, D., Elzer, P., Olsen, S., ... Joly, D. (2000). Validation of the fluorescence polarization assay and comparison to other serological assays for the detection of serum antibodies to *Brucella abortus* in

- bison. *Journal of Wildlife Diseases*, 36(3), 469–476.
<https://doi.org/10.7589/0090-3558-36.3.469>
- Geresu, M., & Kassa, G. (2015). A Review on Diagnostic Methods of Brucellosis. *Journal of Veterinary Science & Technology*, 07(03).
<https://doi.org/10.4172/2157-7579.1000323>
- Giambartolomei, G. H., Arriola Benitez, P. C., & Delpino, M. V. (2017). Brucella and Osteoarticular Cell Activation: Partners in Crime. *Frontiers in Microbiology*, 8. <https://doi.org/10.3389/fmicb.2017.00256>
- Gorsich, E. E., Bengis, R. G., Ezenwa, V. O., & Jolles, A. E. (2015). EVALUATION OF THE SENSITIVITY AND SPECIFICITY OF AN ENZYME-LINKED IMMUNOSORBENT ASSAY FOR DIAGNOSING BRUCELLOSIS IN AFRICAN BUFFALO (*SYNCERUS CAFFER*). *Journal of Wildlife Diseases*, 51(1), 9–18. <https://doi.org/10.7589/2013-12-334>
- Greiner, M., Pfeiffer, D., & Smith, R. . (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive Veterinary Medicine*, 45(1–2), 23–41. [https://doi.org/10.1016/S0167-5877\(00\)00115-X](https://doi.org/10.1016/S0167-5877(00)00115-X)
- Guthrie, A. J., MacLachlan, N. J., Joone, C., Lourens, C. W., Weyer, C. T., Quan, M., ... Gardner, I. A. (2013). Diagnostic accuracy of a duplex real-time reverse transcription quantitative PCR assay for detection of African horse sickness virus. *Journal of Virological Methods*, 189(1), 30–35.
<https://doi.org/10.1016/j.jviromet.2012.12.014>
- Hop, H. T., Arayan, L. T., Simborio, H. L., Reyes, A. W. B., Min, W., Lee, H. J., ... Kim, S. (2016). An evaluation of ELISA using recombinant Brucella abortus

- bacterioferritin (Bfr) for bovine brucellosis. *Comparative Immunology, Microbiology and Infectious Diseases*, 45, 16–19. <https://doi.org/10.1016/j.cimid.2016.01.006>
- Huber, B., Scholz, H. C., Lucero, N., & Busse, H.-J. (2009). Development of a PCR assay for typing and subtyping of *Brucella* species. *International Journal of Medical Microbiology: IJMM*, 299(8), 563–573. <https://doi.org/10.1016/j.ijmm.2009.05.002>
- ISO/IEC 17025. Requisitos generales para la competencia de los laboratorios de ensayo y de calibración (2005).
- Jain, L., Rawat, M., Ramakrishnan, S., & Kumar, B. (2017). Active immunization with *Brucella abortus* S19 phage lysate elicits serum IgG that protects guinea pigs against virulent *B. abortus* and protects mice by passive immunization. *Biologicals*, 45, 27–32. <https://doi.org/10.1016/j.biologicals.2016.10.006>
- Jain, U., Bisht, B., Sahzad, S., Pragati, P., & Dwivedi, K. (2013). Outbreak of brucellosis in buffaloes aborted in village Mahuan, district Mainpuri, UP, India- A case report. *Veterinary World*, 6(1), 51. <https://doi.org/10.5455/vetworld.2013.51-52>
- Kamaraj, G., Chinchkar, S. R., Rajendra, L., & Srinivasan, V. A. (2009). A combined vaccine against *Brucella abortus* and infectious bovine rhinotracheitis. *Indian Journal of Microbiology*, 49(2), 161–168. <https://doi.org/10.1007/s12088-009-0028-7>
- Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleeschauwer, B., ... Angulo, F. J. (2015). World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral

- Diseases, 2010: A Data Synthesis. *PLOS Medicine*, 12(12), e1001921.
<https://doi.org/10.1371/journal.pmed.1001921>
- Kurtz, R. S., & Berman, D. T. (1986). Influence of endotoxin-protein in immunoglobulin G isotype responses of mice to *Brucella abortus* lipopolysaccharide. *Infection and Immunity*, 54(3), 728–734.
- Laboratorios, I. (2017). *Brucella abortus* (B. abortus). Recuperado 13 de junio de 2017, a partir de <http://www.idexx.es/livestock-poultry/ruminant/b-abortus.html>
- Li, Z., Wang, S., Zhang, J., Yang, G., Yuan, B., Huang, J., ... Zhang, H. (2017). *Brucella abortus* 2308?NodV?NodW double-mutant is highly attenuated and confers protection against wild-type challenge in BALB/c mice. *Microbial Pathogenesis*, 106, 30–39. <https://doi.org/10.1016/j.micpath.2017.01.043>
- López, G., & Fernández, P. (2001). Medidas de concordancia: el índice de Kappa. Unidad de Epidemiología Clínica y Bioestadística. Recuperado a partir de <https://www.fisterra.com/mbe/investiga/kappa/kappa2.pdf>
- Mantur, B., Parande, A., Amarnath, S., Patil, G., Walvekar, R., Desai, A., ... Patil, S. (2010). ELISA versus Conventional Methods of Diagnosing Endemic Brucellosis. *The American Journal of Tropical Medicine and Hygiene*, 83(2), 314–318. <https://doi.org/10.4269/ajtmh.2010.09-0790>
- Mirjalili, A. (2016). Development of Indirect ELISA (IELISA) for Diagnosis of Bovine Brucellosis, Comparison of Three Different Labeled Detection Reagents. *MOJ Immunology*, 3(5).
<https://doi.org/10.15406/moji.2016.03.00104>
- Morales, R., Espinosa, R., Torres, H., & Sandoval, P. (2009). *Programa Nacional de Control de Brucelosis Bovina en el Ecuador*. Quito-Ecuador:

AGROCALIDAD. Recuperado a partir de http://www.agrocalidad.gob.ec/agrocalidad/images/pdfs/sanidadanimal/progr_ama_nacional_brucelosis_bovina.pdf

Nielsen, K. (2002). Diagnosis of brucellosis by serology. *Veterinary Microbiology*, 90(1-4), 447-459.

OIE. (2016a). Organización Mundial de Sanidad Animal. Brucelosis Bovina. En *Manual de pruebas de diagnóstico y vacunas para animales terrestres*. Recuperado a partir de http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.01.04_BRUCELLOSIS.pdf

OIE. (2016b). Organización Mundial de Sanidad Animal. Enfoques Estadísticos de la Validación. En *Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres*. Recuperado a partir de http://www.oie.int/fileadmin/Home/esp/Health_standards/tahm/3.6.5_STATISTICAL_VALIDATION.pdf

OIE. (2016c). Organización Mundial de Sanidad Animal. Glosario. En *Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres*. Recuperado a partir de http://www.oie.int/fileadmin/Home/esp/Health_standards/tahm/0.04_Glosario.pdf

OIE. (2016d). Organización Mundial de Sanidad Animal. Incertidumbre de la medición. En *Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres*. Recuperado a partir de

http://www.oie.int/fileadmin/Home/esp/Health_standards/aahm/current/GUIDELINE_3.6.4_MEASUREMENT_UNCERT.pdf

OIE. (2016e). Organización Mundial de Sanidad Animal. Principios y Métodos de Validación de las Pruebas de Diagnóstico de las Enfermedades Infecciosas. En *En Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres*. Recuperado a partir de http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/1.01.06_VALIDATION.pdf

Oyejide, A., Adu, F. D., Makinde, A. A., & Ezch, E. N. (1987). The Prevalence of antibodies of *Brucella abortus*, *Dermatophilus congolensis* and bovine leukaemia virus in Nigerian slaughter cattle. *Veterinary Quarterly*, 9(1), 83–85. <https://doi.org/10.1080/01652176.1987.9694080>

Paweska, J. T., Potts, A. D., Harris, H. J., Smith, S. J., Viljoen, G. J., Dungu, B., ... Prozesky, L. (2002). Validation of an indirect enzyme-linked immunosorbent assay for the detection of antibody against *Brucella abortus* in cattle sera using an automated ELISA workstation. *The Onderstepoort Journal of Veterinary Research*, 69(1), 61–77.

Petrov, I., Qu?tel, C. R., & Taylor, P. D. P. (2007). Investigation on sources of contamination, for an accurate analytical blank estimation, during measurements of the Fe content in seawater by isotope dilution inductively coupled plasma mass spectrometry. *J. Anal. At. Spectrom.*, 22(6), 608–615. <https://doi.org/10.1039/B700540G>

- Roberts, A., & Kemp, C. (2001). Brucellosis (Mediterranean fever, Gibraltar fever, Malta fever, Cyprus fever, undulant fever, typhomalarial fever). *Journal of the American Academy of Nurse Practitioners*, 13(3), 106–107.
- Safronetz, D., Sloan, A., Stein, D. R., Mendoza, E., Barairo, N., Ranadheera, C., ... Drebot, M. (2017). Evaluation of 5 Commercially Available Zika Virus Immunoassays. *Emerging Infectious Diseases*, 23(9). <https://doi.org/10.3201/eid2309.162043>
- Saldarriaga, O., Ossa, J., & Lopéz, R. (2002). Respuesta inmune y estrategias de evasión durante la infección con *Brucella* spp. (PDF Download Available). Recuperado 13 de junio de 2017, a partir de https://www.researchgate.net/publication/45161116_Respuesta_inmune_y_es_trategias_de_evasion_durante_la_infeccion_con_Brucella_spp
- Schulze, E. (1959). [Bang's disease, brucellosis]. *Deutsche Medizinische Wochenschrift (1946)*, 84(5), 194. <https://doi.org/10.1055/s-0028-1113570>
- Tafur, C. (2014). *Validación de un método para análisis de arsénico en agua potable por absorción atómica mediante generación de hidruros*. Ambato-Ecuador. Recuperado a partir de <http://repositorio.uta.edu.ec/bitstream/123456789/8459/1/BQ%2060.pdf>
- Vanzini, V. ., Aguirre, N., Lugaresi, C. ., de Echaide, S. ., de Canavesio, V. ., Guglielmono, A. ., ... Nielsen, K. (1998). Evaluation of an indirect ELISA for the diagnosis of bovine brucellosis in milk and serum samples in dairy cattle in Argentina. *Preventive Veterinary Medicine*, 36(3), 211–217. [https://doi.org/10.1016/S0167-5877\(98\)00080-4](https://doi.org/10.1016/S0167-5877(98)00080-4)

- Vemulapalli, R., Sanakkayala, N., Gulani, J., Schurig, G. G., Boyle, S. M., Lindsay, D. S., & Sriranganathan, N. (2007). Reduced cerebral infection of *Neospora caninum* in BALB/c mice vaccinated with recombinant *Brucella abortus* RB51 strains expressing *N. caninum* SRS2 and GRA7 proteins. *Veterinary Parasitology*, 148(3–4), 219–230.
<https://doi.org/10.1016/j.vetpar.2007.06.029>
- Villareal, V. (2016). Manual de Procedimientos para la Atención y Control de Brucelosis Bovina en el Ecuador. Recuperado a partir de <http://www.agrocalidad.gob.ec/wp-content/uploads/pdf/sanidad-animal/resolucion%200131%20rt%20-%20sa%20-%20manual%20de%20procedimientos%20para%20la%20atencion%20y%20control%20de%20brucelosis%20bovina.pdf>
- Waldvogel, A. S., Anderson, G. A., Phillips, D. L., & Osburn, B. I. (1992). Infection of bovine fetuses at 120 days' gestation with virulent and avirulent strains of bluetongue virus serotype 11. *Comparative Immunology, Microbiology and Infectious Diseases*, 15(1), 53–63.
- Wyatt, H. V. (2005). How Themistocles Zammit found Malta Fever (brucellosis) to be transmitted by the milk of goats. *Journal of the Royal Society of Medicine*, 98(10), 451–454.

ANEXOS

Anexo A. Datos de Sensibilidad Analítica. Analista 1 y 2

Analista	Dilución	Día 1		Día 2		Día 3	
		M/P%	Resultado	MP%	Resultado	MP%	Resultado
Analista 1	PURO	283,86	Positivo	279,84	Positivo	273,65	Positivo
	1/2.	262,04	Positivo	267,50	Positivo	259,32	Positivo
	1/4.	250,79	Positivo	248,76	Positivo	245,80	Positivo
	1/8.	204,53	Positivo	200,59	Positivo	203,68	Positivo
	1/16.	151,51	Positivo	149,13	Positivo	152,01	Positivo
	1/32.	125,78	Positivo	129,47	Positivo	127,36	Positivo
	1/64.	73,11	Negativo	74,07	Negativo	72,55	Negativo
	1/128.	47,15	Negativo	45,65	Negativo	45,97	Negativo
	1/256.	26,79	Negativo	26,28	Negativo	25,96	Negativo
	1/512.	15,17	Negativo	14,34	Negativo	14,93	Negativo
	1/1024.	11,12	Negativo	10,91	Negativo	11,05	Negativo
Analista 2	PURO	289,02	Positivo	284,57	Positivo	293,34	Positivo
	1/2.	271,99	Positivo	282,72	Positivo	273,94	Positivo
	1/4.	256,27	Positivo	256,19	Positivo	251,29	Positivo
	1/8.	216,86	Positivo	212,64	Positivo	215,65	Positivo
	1/16.	184,08	Positivo	177,32	Positivo	183,40	Positivo
	1/32.	121,13	Positivo	126,40	Positivo	124,52	Positivo
	1/64.	93,01	Negativo	90,43	Negativo	94,71	Negativo
	1/128.	42,54	Negativo	39,60	Negativo	41,63	Negativo
	1/256.	31,86	Negativo	34,04	Negativo	32,45	Negativo
	1/512.	12,07	Negativo	11,90	Negativo	11,76	Negativo
	1/1024.	8,04	Negativo	8,31	Negativo	7,79	Negativo

Anexo B. Datos de sensibilidad y especificidad diagnóstica

N°	PI	RESULTADOS
1	363,90	POSITIVO
2	351,22	POSITIVO
	1,11	NEGATIVO
3	5,16	NEGATIVO
	1,25	NEGATIVO
4	307,87	POSITIVO
5	285,85	POSITIVO
	2,23	NEGATIVO
6	0,28	NEGATIVO
	4,04	NEGATIVO
7	256,59	POSITIVO
8	292,40	POSITIVO
	6,97	NEGATIVO
9	1,53	NEGATIVO
	1,11	NEGATIVO
10	334,36	POSITIVO
11	310,52	POSITIVO
	2,79	NEGATIVO
12	1,81	NEGATIVO
	2,93	NEGATIVO
	206,13	POSITIVO
13	298,26	POSITIVO
	2,37	NEGATIVO
14	1,95	NEGATIVO
	-0,98	NEGATIVO
	281,11	POSITIVO
15	296,86	POSITIVO
	1,53	NEGATIVO
16	0,14	NEGATIVO
	1,95	NEGATIVO
17	2,65	NEGATIVO
18	5,57	NEGATIVO
	2,23	NEGATIVO
19	-2,79	NEGATIVO
	1,11	NEGATIVO
20	-0,56	NEGATIVO
21	0,70	NEGATIVO
22	0,42	NEGATIVO
23	6,55	NEGATIVO

Anexo B Continuación

24	9,34	NEGATIVO
25	0,84	NEGATIVO
26	-1,11	NEGATIVO
27	11,78	NEGATIVO
28	20,36	NEGATIVO
29	15,18	NEGATIVO
30	17,41	NEGATIVO
31	14,21	NEGATIVO
32	5,13	NEGATIVO
33	4,01	NEGATIVO
34	12,03	NEGATIVO
35	14,52	NEGATIVO
36	16,55	NEGATIVO
37	15,08	NEGATIVO
38	13,81	NEGATIVO
39	-0,15	NEGATIVO
40	8,02	NEGATIVO
41	6,35	NEGATIVO
42	14,11	NEGATIVO

Anexo C.Exactitud. Valores obtenidos

N°	MRC	M/P%	RESULTADO
1		1,3	NEGATIVO
2		1,4	NEGATIVO
3		1,7	NEGATIVO
4	CONTROL KIT NEGATIVO	0,3	NEGATIVO
5		-0,3	NEGATIVO
6		0,7	NEGATIVO
7		1,2	NEGATIVO
8		0,3	NEGATIVO
9		1,7	NEGATIVO
10		1,4	NEGATIVO
11		121,7	POSITIVO
12		123,4	POSITIVO
13		124,3	POSITIVO
14		119,3	NEGATIVO
15	MRC DIL. 1/32	120,2	POSITIVO
16		124,2	POSITIVO
17		122,3	POSITIVO
18		125,7	POSITIVO
19		121,3	POSITIVO
20		123,1	POSITIVO
21		287,9	POSITIVO
22		291,3	POSITIVO
23		291,6	POSITIVO
24		289,8	POSITIVO
25	MRCI POSITIVO	297,0	POSITIVO
26		290,0	POSITIVO
27		293,4	POSITIVO
28		299,6	POSITIVO
29		298,5	POSITIVO
30		297,6	POSITIVO

Anexo D. Cálculo del coeficiente Carry-over

test 1		test 2		test 3		test 4		test 5		test 6		test 7	
a1	363,9	a1	307,8 7	a1	256,5 9	a1	334,36	a1	206	a1	281,1 1	a1	2,65
a2	351,2 2	a2	285,8 5	a2	292,4	a2	310,52	a2	298	a2	296,8 6	a2	5,57
b1	1,11	b1	2,23	b1	6,97	b1	2,79	b1	2,37	b1	1,53	b1	2,23
b2	5,16	b2	0,28	b2	1,53	b2	1,81	b2	1,95	b2	0,13	b2	-2,79
b3	1,25	b3	4,04	b3	1,11	b3	2,93	b3	-1	b3	1,95	b3	1,11
q	- 1,255	q	- 4,048	q	- 1,095	q	-2,927	q	0,003	q	- 0,008	q	0,25
%	-	%	-	%	-	%	-292,7	%	0,37	%	-0,85	%	25,0
q	125,5	q	404,8	q	109,5	q		q		q		q	0

Anexo E. Datos para la determinación de la incertidumbre

Prueba	M/P%
1	119,70
2	123,44
3	124,31
4	119,54
5	120,23
6	124,18
7	122,33
8	125,74
9	121,33
10	123,08
11	126,92
12	127,49
13	123,22
14	122,51
15	119,93
16	118,93
17	124,03
18	124,17
19	123,18
20	124,23

Anexo F. Certificado de análisis del suero control positivo



January, Janvier 2016

Quality Control Data sheet & Instructions for Use
Fiche de Contrôle Qualité et Conditions d'Utilisation

Freeze-dried Brucellosis reference serum / Sérum de référence Brucellose lyophilisé

Product code / Code Produit: MRI-BRUJ

Batch / N° de lot: 004

Date of manufacture / Date de fabrication: 10/2011

Description:

Freeze-dried bovine serum (origin: Italy) containing anti-Brucella specific antibodies. Contains merthiolate (0.03%).

Sérum de bovin (origine: Italie) lyophilisé contenant des anticorps spécifiques anti-Brucella. Contient du merthiolate (0.03%).

Storage / Conservation:

- Store between +2°C and +26°C before reconstitution.
Conservation entre +2°C et +26°C avant reconstitution.
- Store at -20°C in aliquots after reconstitution up to 1 year.
Conservation à -20°C en aliquots après reconstitution pendant 1 an.
- Never thaw more than 3 times.
Ne jamais décongeler plus de 3 fois.
- May be stored undiluted at +4°C in liquid form up to 2 days.
Peut être stocké deux jours à +4°C sous forme liquide et non diluée.

Shelf life:

This product does not bear an exact expiry date because it is freeze-dried and very stable over time. As a general rule, it may be kept for up to 10 years at +2 – 26°C from its production date in its freeze-dried form.

Ce produit ne porte pas de date de péremption, étant donné qu'il est lyophilisé et par conséquent, très stable dans le temps. De règle générale, il peut être conservé jusqu'à 10 ans à partir de sa date de fabrication sous forme lyophilisée.

Quality Control / Contrôle de qualité:

This serum was found positive when tested by / Ce sérum a été trouvé positif lorsqu'il a été analysé par:

ID Screen® Brucellosis Serum Indirect Multi-species ELISA

Complement Fixation Test CFT / RFC

Rose Bengal RSA / ARL

Use / Utilisation:

GB

- 1- Reconstitute in 1 ml of distilled water.
- 2- Wait 1-2 min and mix well in order to homogenize the solution.
- 3- For ELISA: Dilute in a serum dilution buffer or a negative serum. Please note: The negative serum used may slightly influence the results. Always use the same dilution matrix (buffer or serum) when comparing results from different runs.
- 4- For Rose Bengal and Complement Fixation Test: Dilute in a negative serum. Please note: The negative serum used may slightly influence results. IDvet offers a freeze-dried negative serum (MRINEG-BRUJ) in which the dilutions may be performed.
- 5- Test the serum to the following dilutions depending on the method used and take the last positive or weak positive dilution as an internal standard.

FR

- 1- Reprendre dans 1 ml d'eau distillée.
- 2- Attendre 1-2 min et bien agiter pour tout homogénéiser.
- 3- ELISA: Diluer dans le tampon de dilution du sérum ou en sérum négatif. Nota: Le sérum négatif utilisé peut légèrement influencer les résultats. Toujours utiliser la même matrice de dilution (tampon ou sérum) pour la comparaison de résultats de différentes séries de tests.
- 4- Rose Bengale ou RFC: Diluer en sérum négatif. Nota: Le sérum négatif utilisé peut légèrement influencer les résultats. IDvet propose un sérum négatif lyophilisé (MRINEG-BRUJ).
- 5- Tester le sérum aux dilutions suivantes en fonction de l'application souhaitée et retenir comme standard interne la dernière dilution donnant un résultat positif ou faiblement positif.

Method / Méthode	Dilutions to test - Zone de dilution à tester
ELISA	From 1/6 to 1/32 - Du 1/2 au 1/32
RSA (Rose Bengal) / ARL	Pure to 1/16 - Pur au 1/16
Complement Fixation Test / RFC	Positive up to 1/32 - Trouvé positif jusqu'au 1/32

Quality Control Manager : S. Lescau
Responsable Contrôle Qualité
stephanie.lescau@id-vet.com

Director : Philippe Pourquier
Directeur
philippe.pourquier@id-vet.com

MRI-BRUJ
DOC 6432

IDvet, 310, rue Louis Pasteur – Gratelets - FRANCE
Tel + 33 (0)4 67 41 49 33 - Fax + 33 (0)4 67 45 36 95
www.id-vet.com - E-mail: info@id-vet.com

Anexo G. Certificado de análisis del suero control negativo



March, Mars 2016

Quality Control Data sheet & Instructions for Use
Fiche de Contrôle Qualité et Conditions d'Utilisation
Freeze-dried Brucellosis negative serum / Sérum négatif Brucellose lyophilisé

Product code / Code Produit: MRINEG-BRU

Batch / N° de lot: 002

Date of manufacture / Date de fabrication: 10/2012

Description:

Freeze-dried bovine serum free of anti-Brucella specific antibodies. Contains merthiolate (0.03%).

Sérum de bovin lyophilisé ne présentant aucun anticorps spécifique anti-Brucella. Contient du merthiolate (0.03%).

Storage / Conservation:

- Store between +2°C and +26°C before reconstitution.
Conservation entre +2°C et +26°C avant reconstitution.
- Store at -20°C in aliquots after reconstitution up to 1 year.
Conservation à -20°C en aliquots après reconstitution pendant 1 an.
- Never thaw more than 3 times.
Ne jamais décongeler plus de 3 fois.
- May be stored undiluted at +4°C in liquid form up to 2 days.
Peut être stocké deux jours à +4°C sous forme liquide et non diluée.

Shelf life:

This product does not bear an exact expiry date because it is freeze-dried and very stable over time. As a general rule, it may be kept for up to 10 years at +2 – 26°C from its production date in its freeze-dried form.

Ce produit ne porte pas de date de péremption, étant donné qu'il est lyophilisé et par conséquent, très stable dans le temps. De règle générale, il peut être conservé jusqu'à 10 ans à partir de sa date de fabrication sous forme lyophilisée.

Quality Control / Contrôle de qualité:

This serum was found negative when tested by / Ce sérum a été trouvé négatif lorsqu'il a été analysé par :

ID Screen® Brucellosis Serum Indirect Multi-species ELISA

Complement Fixation Test CFT / RFC

Rose Bengal RSA / EAT

Use / Utilisation:

GB

- 1- Reconstitute in 1 ml of distilled water.
- 2- Wait 1-2 min and mix well in order to homogenize the solution.

FR

- 1- Reprendre dans 1 ml d'eau distillée.
- 2- Attendre 1-2 min et bien agiter pour tout homogénéiser.

Quality Control Manager : S. Lesceu
Responsable Contrôle Qualité
stephanie.lesceu@id-vet.com

Director : Philippe Pourquier
Directeur
philippe.pourquier@id-vet.com

Anexo H. Certificado de análisis del Kit de brucelosis IDEXX



Livestock, Poultry and Dairy

CERTIFICATE OF ANALYSIS

Certificado de Análisis / Certificat de contrôle / Chargen-Prüfprotokoll

IDEXX Brucellosis Serum [Brucellosis Antibody Test Kit]

Product code :	P54130-10	Expiration date :	24-FEB-2016
Batch :	6008		
Tests per kit :	960		
Insert version :	06-04130-14	Manufacturing date :	FEB-2016

KIT COMPONENT	Lot number	KIT COMPONENT	Lot number
Plate	251-09	Dilution buffer	2-652
Positive control	2296	TMB Substrate	13-307
Negative control	2296	Stop solution	3-111
Conjugate	2-1274	Wash solution	1-186
Dilution buffer	1-676		

KIT CONTROLS

Mean Negative Control OD	0.054
Mean Positive Control OD	1.001

DETECTABILITY

Mean % S/P Serum Bru POS SE03* at 1/16	170%	Pass
Mean % S/P Serum Bru POS SE03* at 1/64	96%	Pass
Mean % S/P Serum Bru POS SE03* at 1/160	49%	Pass

REPEATABILITY

Low positive sample (n=252)	CV = 6%
-----------------------------	---------

SPECIFICITY

Bovine sera n=23	mean % S/P = 3%	Pass
------------------	-----------------	------

SENSITIVITY

IDEXX serum panel	1	Pass
-------------------	---	------

* French Sub-standard equivalent to the Brucellosis Strong Pos International standard OIEELISA55

This product was performance tested and has met all quality control specifications required for release.

This information is released by :

Title : Quality Assurance Manager I, Samsonoff Anne

Signature :

Manufacturer: IDEXX Montpellier SAS - 326, rue de la Galère - Parc Euromédecine - 34090 Montpellier - France

Tel : 33(0)4 98 23 24 25 - Fax : 33(0)4 67 04 20 25 - info-montpellier@idexx.com

EN 05-01-17 - Vers : C

Document printed on : 1-Mar-2016

Anexo I. Declaración de Validación

	DECLARACIÓN DE VALIDACIÓN		PGULAB-PURE
	LABORATORIO: SERIDLOGÍA		Rev. 3
			Hoja 1 de 1
I. MÉTODO DE ENSAYO			
INTERNO	PCE/SE/RS DIAGNÓSTICO DE BRUCELLOSIS MEDIANTE LA PRUEBA ELISA INDIRECTO		
REFERENCIA	Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres 2016. OIE		
II. RESULTADOS DE PRUEBAS DE VALIDACIÓN			
PARÁMETROS	OBJETIVO DE VALIDACIÓN	RESULTADO	CUMPLE / NO CUMPLE
LÍMITE DE DETECCIÓN	PUNTO DE CORTE	102	CUMPLE
ESPECIFICIDAD ANALÍTICA	92%	100%	CUMPLE
SENSIBILIDAD DIAGNÓSTICA	99%	100%	CUMPLE
ESPECIFICIDAD DIAGNÓSTICA	98%	100%	CUMPLE
VALOR PREDICTIVO POSITIVO	≥ 50%	100%	CUMPLE
VALOR PREDICTIVO NEGATIVO	≥ 50%	100%	CUMPLE
EXACTITUD			
a. EXACTITUD	≥ 99%	99%	CUMPLE
b. ARRASTRE	≤ 1%	0%	CUMPLE
PRECISION			
a. REPETIBILIDAD	≥ 92%	100%	CUMPLE
c. REPRODUCIBILIDAD	≥ 92%	100%	CUMPLE
LINEALIDAD			
INCERTIDUMBRE	≤ 5.0	5%	CUMPLE
OTROS:			
III. OBSERVACIONES Y COMENTARIOS A LOS RESULTADOS OBTENIDOS			
Sin comentarios			
IV. CONCLUSIONES SOBRE DECLARACIÓN DE VALIDACIÓN			
Una vez revisado los resultados de la validación y comparados con los objetivos establecidos, declaramos que el método de ensayo PCE/SE/RS DIAGNÓSTICO DE BRUCELLOSIS MEDIANTE LA PRUEBA ELISA INDIRECTO para una muestra de SUERO SANGUÍNEO en el Laboratorio de SERIDLOGÍA de AGROCALIDAD, se declara VALIDADO.			
REVISADO POR: RESPONSABLE TÉCNICO NOMBRE: FIRMA: FECHA: 21 DE MARZO DE 2017		APROBADO POR: DIRECTOR DE DIAGNÓSTICO ANIMAL NOMBRE: Dra. Mónica Barrios FIRMA: FECHA: 22 DE MARZO DE 2017	

Anexo J. Principios y métodos de validación de las pruebas de diagnóstico de las enfermedades infecciosas

CAPÍTULO 1.1.6.

PRINCIPIOS Y MÉTODOS DE VALIDACIÓN DE LAS PRUEBAS DE DIAGNÓSTICO DE LAS ENFERMEDADES INFECCIOSAS

INTRODUCCIÓN

La validación es un proceso que determina la idoneidad de una prueba¹, que se ha desarrollado, optimizado y estandarizado adecuadamente para un propósito concreto. Todas las pruebas de diagnóstico (pruebas de laboratorio y de campo) deben validarse para la especie en la que se utilizarán. La validación incluye estimaciones de las características de rendimiento analítico y diagnóstico de una prueba. En el contexto de este capítulo, una prueba que ha superado las tres primeras etapas de la validación (véase la Figura 1 abajo), incluida la caracterización del rendimiento, puede designarse como "válida para el objetivo inicial deseado(s)". Pero para que una prueba siga estando válida, es necesario realizar un cuidadoso seguimiento del rendimiento de la misma en las condiciones habituales de uso, a menudo controlando el comportamiento de los controles de la prueba a lo largo del tiempo. Esto garantiza que la prueba, según la validación original, mantenga siempre sus características de rendimiento. En el caso de que empiece a dar resultados que no concuerden con los datos de la validación original, la prueba podría no ajustarse al propósito/s deseado/s. Así, una prueba válida se evalúa continuamente para garantizar que mantiene su idoneidad para el propósito deseado, mediante una evaluación de los resultados de los controles de la prueba y mediante una evaluación continua durante el uso sistemático en la población de destino.

Las pruebas realizadas en individuos o en poblaciones tienen varios propósitos, tales como: documentar la ausencia de una determinada enfermedad en un país o región, evitar su propagación a través del comercio, contribuir a erradicar una infección de una zona o país, confirmar el diagnóstico de los casos clínicos, estimar la prevalencia de una infección para facilitar el análisis del riesgo, identificar a los animales infectados con vistas a implementar medidas de control, y clasificar los animales según su salud o estado inmunitario tras la vacunación. Una única prueba puede validarse para uno o varios propósitos deseados optimizando sus características de rendimiento para cada uno, como por ejemplo, fijando una alta sensibilidad diagnóstica (Dse), asociada a una baja especificidad diagnóstica (Dsp) para una prueba de cribado, o por el contrario, fijando una DSp alta asociada a una DSe más baja para una prueba confirmativa.

El siempre cambiante repertorio de nuevos y específicos reactivos de diagnóstico, junto con la existencia de muchos nuevos formatos y protocolos analíticos ha precipitado debates sobre cómo validar adecuadamente estas pruebas. Para orientar a los usuarios de pruebas en cuanto a la validación de pruebas más complejas, como las de detección de ácido nucleico, ya no basta con ofrecer simples ejemplos de pruebas serológicas, como el enzoinmunoanálisis. Con el fin de aportar coherencia al proceso de validación para todos los tipos de pruebas, este capítulo se centra en los criterios que deben cumplirse durante la realización y la validación de pruebas de cualquier tipo. La inclusión de la ejecución de la prueba como parte del proceso de validación de la misma puede parecer contra-intuitiva, pero en realidad, tres de los criterios de validación requeridos (definición de los objetivo/s buscado, optimización y estandarización) que deben evaluarse para llegar a la validación de una prueba incluyen pasos de su proceso de ejecución. Según esto, el proceso de ejecución de la prueba da lugar a un sistema de validación de la misma, y ambos contienen criterios de validación que deben cumplirse. Además, se aportan otras

1 "Prueba," "método analítico," y "test" son sinónimos a los efectos de este capítulo, y por lo tanto se utilizan indistintamente.

directrices en una serie de Recomendaciones de la OIE para la Validación de pruebas de diagnóstico² que se adaptan a varios tipos de prueba que en esencia son distintos (por ejemplo, la detección de ácido nucleico, anticuerpos o antígenos) y aportan más información sobre temas específicos relacionados con la validación de pruebas de diagnóstico. Para información específica relativa a especies de fauna salvaje, consúltese el capítulo 3.6.7 Principios y métodos para la validación de las pruebas de diagnóstico de las enfermedades infecciosas aplicables a la fauna salvaje. La información que se aporta en el capítulo, que es específica de especies de fauna salvaje, también podría ser de utilidad para la validación de pruebas realizadas en animales domésticos, por ejemplo, cuando se disponga de un escaso número de muestras.

CONSIDERACIONES PRELIMINARES EN EL DESARROLLO Y VALIDACIÓN DE PRUEBAS

Todos los laboratorios deben cumplir los requisitos establecidos en los Capítulos 1.1.1 (Manual Acuático) o 1.1.5 (Manual Terrestre) sobre Gestión de la calidad en los laboratorios de pruebas veterinarias. Con ello, se minimizará la influencia de factores que no dependen de la prueba en sí, como el instrumental, errores del operario, la elección del reactivo (químico o biológico) y la calibración, los recipientes y plataformas de reacción, la calidad del agua, el pH y la ionidad de tampones y diluyentes, las temperaturas y duraciones de incubación, y los errores en el rendimiento técnico de la prueba.

El primer paso en el desarrollo de una prueba es definir el objetivo de la prueba, ya que esto guiará todos los pasos subsiguientes del proceso de validación. Los criterios de validación de la prueba son los rasgos característicos de una prueba que constituyen factores decisivos, mediciones o estándares en los cuales se basa un juicio o decisión. Si tenemos en cuenta las variables que pueden afectar al rendimiento de una prueba, podremos apreciar más claramente los criterios que deben tenerse en cuenta para su validación. Estas variables pueden agruparse en categorías: (a) relativas a la muestra: si son individuales o agrupadas, la composición de la matriz, y las interacciones hospedador/organismo que afecten al análisis en cuestión cuantitativa o cualitativamente; (b) relativas al sistema analítico – que incluyen factores físicos, químicos, biológicos y técnicos que afectan a la capacidad de la prueba de detectar en la muestra un análisis específico; y (c) relativos a la interpretación del resultado de la prueba – es decir, la capacidad del sistema analítico de predecir de forma precisa el estado del individuo o de la población en relación con el propósito con el cual se aplica la prueba.

La elección, obtención, preparación, conservación y manejo de las muestras son variables cruciales en el diseño y desarrollo de una prueba para garantizar resultados analíticos válidos. Otras variables, como el transporte, la cadena de custodia, la rastreabilidad de las muestras y el sistema de gestión de la información del laboratorio también son fuentes clave de variación/errores que se convierten en especialmente importantes cuando la prueba se implementa para un análisis sistemático. La integridad de los resultados experimentales durante la ejecución y validación de una prueba será como mucho igual de buena que la calidad de las muestras que se utilicen. Antes de iniciar las tareas de validación de una prueba, es importante prever los factores que pueden afectar negativamente a la calidad de la muestra. Las muestras de referencia que se utilicen en la ejecución y validación de una prueba deberán encontrarse en la misma matriz que se va a utilizar en la prueba (por ejemplo, suero, tejido o sangre total) y deberán ser representativas de la especie que vaya a analizarse mediante dicha prueba. Los materiales de referencia deberán representar adecuadamente el intervalo de concentraciones de análisis a detectar mediante la prueba. En los capítulos 1.1.2 y 1.1.3 del Manual Terrestre se encuentra información sobre la obtención, preparación, conservación, manejo y transporte de las muestras.

La matriz (suero, heces, tejido, etc.) en la que puede residir el análisis buscado puede contener inhibidores endógenos o exógenos que impidan el funcionamiento de ciertas pruebas. Esto es especialmente importante en el caso de pruebas que dependen de enzimas, como la reacción en cadena de la polimerasa (PCR) o el ensayo inmunoenzimático (ELISA). Otros factores que también influyen en la concentración y en la composición de un análisis buscado (en concreto anticuerpos) en la muestra son en su mayor parte atribuibles al hospedador y pueden ser factores inherentes (como la edad, el sexo, la raza, el estado nutricional, la gestación o la capacidad de respuesta inmunitaria), o bien adquiridos (como la adquisición pasiva de anticuerpos o la inmunidad activa obtenida por la vacunación o la infección). Otros factores que no dependen del hospedador, como la contaminación o el deterioro de la muestra, también pueden llegar a afectar a la capacidad de la prueba de detectar el análisis buscado en la muestra. También es importante que los reactivos biológicos estén libres de agentes extraños que, de encontrarse presentes, pudieran conducir a resultados erróneos.

2 http://www.oie.int/fileadmin/Home/esp/Health_standards/tahm/3.6.6_INTRODUCTION.pdf

CRITERIOS DE DESARROLLO Y VALIDACIÓN DE PRUEBAS

El rendimiento de la prueba se ve afectado por muchos factores, empezando por la optimización de la prueba. Tras la optimización inicial para un propósito determinado, se comprobarán las características de rendimiento de la prueba. Es posible que la prueba requiera una optimización adicional o que se observe que es apta para un determinado propósito en base a los resultados del trabajo de validación.

Crterios para el Desarrollo y Validación de una Prueba

1. Definición de los propósito/s
2. Optimización
3. Estandarización
4. Repetibilidad
5. Sensibilidad analítica
6. Especificidad analítica
7. Umbrales
8. Sensibilidad diagnóstica
9. Especificidad diagnóstica
10. Reproducibilidad
11. Idoneidad para el/los propósito/s

A. FASE DE DESARROLLO DE LA PRUEBA

1. Definición de los propósitos deseados en una prueba

La Norma de la OIE sobre Gestión y Requisitos Técnicos para los Laboratorios que Realizan Pruebas para las Enfermedades Infecciosas (Organización Mundial de Sanidad Animal, 2008)³ establece que los métodos analíticos y los procedimientos relacionados deben ser adecuados para las aplicaciones diagnósticas específicas a fin de que los resultados de las pruebas tengan utilidad. En otras palabras, la prueba debe adecuarse al propósito deseado. La capacidad de evaluación cualitativa y cuantitativa de la capacidad de un resultado positivo o negativo de una prueba de predecir de forma precisa si un animal o población de animales están o no infectados o han estado o no expuestos es el criterio definitivo para la validación de dicha prueba. Esta capacidad depende de que una prueba, desarrollada mediante una cuidadosa optimización y estandarización (apartado A.2.3) ofrezca, mediante la acumulación de datos de validación, confianza en la capacidad de la prueba de rendir de acuerdo con el propósito pretendido. Con el fin de asegurar que los resultados de la prueba proporcionan útiles inferencias diagnósticas sobre animales o poblaciones de animales respecto al propósito deseado, el proceso de validación abarca documentación sobre el desarrollo inicial y la realización de la prueba, así como una evaluación continua de los programas de control y garantía de calidad. La Figura 1 muestra el proceso de validación de la prueba, desde el diseño hasta las fases de desarrollo y validación y la implementación, despliegue y mantenimiento del mismo.

El primer paso del desarrollo de una prueba consiste en escoger un tipo de prueba que sea adecuada y que tenga posibilidades de ser validada para un uso concreto (adecuada para el propósito deseado).

Los propósitos más frecuentes son los siguientes:

- 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).

³ Esta es una interpretación específica de los requisitos establecidos de forma más general en la norma internacional de calidad ISO/IEC 17025:2005 para laboratorios de análisis (2005). La Norma de la OIE establece, además, que para que un método analítico se considere apropiado debe estar validado adecuadamente, y que esta validación debe respetar los principios descritos en los capítulos sobre validación del Manual Terrestre y el Manual Acuático.

- 5) Estimar la prevalencia de la infección o exposición para facilitar el análisis de riesgos (encuestas, estatus sanitario del rebaño, medidas para el control de la enfermedad).
- 6) Determinar el estado inmunitario de animales determinados o de poblaciones (después de la vacunación).

Estos fines incluyen muchas aplicaciones más acotadas y específicas de las pruebas (para más detalles véanse las Recomendaciones de la OIE para la Validación de pruebas de diagnóstico (nota 2) para cada tipo de prueba). Tales aplicaciones específicas y sus fines concretos deben ser claramente definidos en el contexto de una prueba completamente validada.

Además de definir la prueba en cuanto al propósito deseado, debe definirse también en cuanto a la especie de destino, al/los agente/s patógeno/s o trastorno/s y al tipo de matriz en el que se encuentra la muestra.

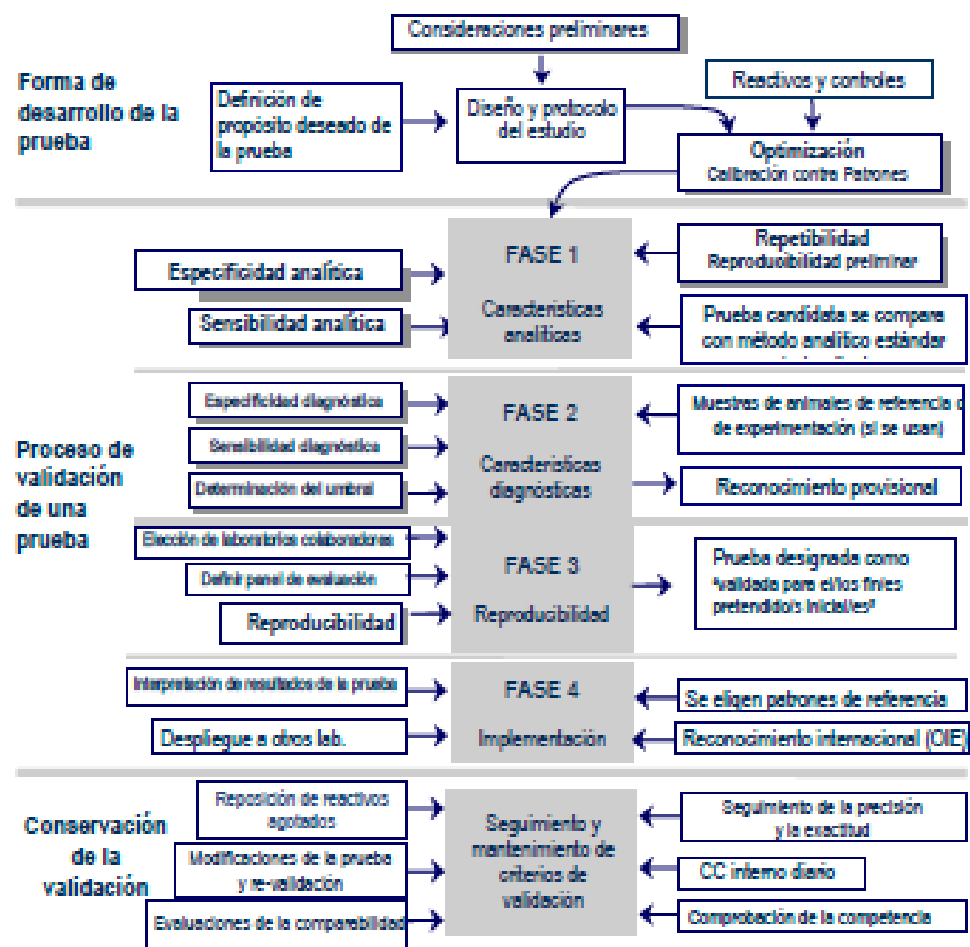


Figura 1. Las fases de desarrollo y validación de la prueba con los criterios de validación de la misma destacados en negrita en los recuadros sombreados.

2. Desarrollo de la prueba – los estudios experimentales

2.1. Diseño y demostración del método analítico

Para el diseño de todos los pasos de una nueva prueba todavía no validada, o de una prueba existente que se esté modificando, son necesarios un conocimiento previo, así como una buena reflexión y planificación. Se ofrece orientación las Directrices de Validación de la OIE (ref.), que cubren

las mejores prácticas para el desarrollo y validación de las pruebas destinadas a la detección de distintos analitos (como anticuerpos, antígenos o ácido nucleico).

El desarrollo de todas las pruebas depende de las muestras de referencia del analito, que representan el analito buscado, la matriz en la que se encuentra el mismo y la población a la que se destina la prueba. Las muestras de referencia pueden ser sueros, líquidos (Incluidos jugos de carne) o tejidos que contengan el analito de interés o una estructura genómica que concuerde con el analito buscado. Estos materiales de referencia se utilizan en experimentos llevados a cabo durante todo el proceso de desarrollo y también en la validación de la prueba.

2.2. Intervalo de funcionamiento de la prueba

El intervalo de funcionamiento de una prueba es el intervalo de concentraciones o títulos del analito en el cual el método aporta una exactitud y precisión idóneas. La exactitud es la cercanía de un valor analítico al valor esperado (real) (media o mediana) para un reactivo estándar de referencia de concentración o título conocidos. La precisión⁴ es el grado de dispersión (varianza, desviación estándar o coeficiente de variación) dentro de una serie de mediciones de la misma muestra analizada en unas condiciones especificadas. Durante el desarrollo de la prueba, se determinan los límites inferior y superior del intervalo de funcionamiento. Para determinar formalmente este intervalo, se selecciona una muestra de referencia positiva alta (lo ideal es que esta muestra se encuentre entre las tres muestras descritas en Apartado A.2.3). Se lleva a cabo una dilución seriada de esta muestra positiva alta hasta la extinción de respuesta de la prueba en una matriz sin analito que tenga la misma constitución que la matriz de muestra de los animales de la población a la que se destina la prueba. Los resultados se trazan en una "curva de respuesta", en la que la respuesta (por ejemplo, la densidad óptica, el ciclo umbral, los recuentos, etc.) es función de la concentración (cantidad) de analito. La curva establece el intervalo de funcionamiento de la prueba. Si se observa que el intervalo es inaceptable para el propósito deseado, tal vez sea necesaria una mayor optimización. La curva de calibración típica para la mayoría de pruebas tiene una forma sigmoidea. Los datos se transforman para aproximar una relación lineal entre la respuesta y la concentración empleando un algoritmo adecuado (Findlay y Dillard, 2007).

2.3. Estandarización y optimización

La optimización es el proceso mediante el cual la mayor parte de parámetros físicos, químicos y biológicos de una prueba se evalúan y se ajustan para garantizar que las características de rendimiento de la prueba se adapten mejor a la aplicación pretendida. Es útil seleccionar por lo menos tres muestras de referencia bien definidas, que contengan el analito abarcando resultados desde positivo alto a negativo (por ejemplo, positivo alto, positivo bajo y negativo). En teoría, las muestras deberían representar tanto a los animales que se consideran infectados como a los no infectados dentro de la población que finalmente va a ser objeto de la prueba. Sin embargo, no siempre es posible obtener estas muestras de referencia, sobre todo en el caso de las pruebas de detección de ácido nucleico y antígeno. La alternativa de preparar muestras de referencia a las que se hayan añadido agentes patógenos de cultivo o sueros positivos no es tan adecuada, puesto que estas muestras no representan exactamente la interacción entre matriz y agente que tiene lugar en condiciones naturales (véase también el Capítulo 3.6.6 Elección y uso de tipos y grupos de muestras de referencia). Cuando no existe otra alternativa, tal vez la única posibilidad sea añadir a la muestra una cantidad conocida del analito o agente derivado del cultivo, o bien diluir un suero positivo alto en suero negativo de la misma especie. En cualquier caso, es imprescindible que la matriz a la que se añade o en la que se diluye el analito sea idéntica o se parezca lo máximo posible a la de las muestras que finalmente se analizarán mediante la prueba. Lo ideal es que las muestras de referencia estén bien caracterizadas por una o, preferiblemente, al menos dos metodologías alternas. Estas muestras pueden utilizarse en experimentos para determinar si la prueba es capaz de distinguir diferentes cantidades del analito, distinguir los analitos buscados de otros que estén estrechamente relacionados con ellos, y optimizar las concentraciones de reactivos y perfeccionar el protocolo. En principio, para todos los tipos de prueba es muy conveniente preparar y almacenar una cantidad suficiente de cada muestra de referencia en alícuotas para su uso en cada ejecución de la prueba candidata, ya que se evalúa a lo largo de todo el proceso de desarrollo y validación. Cambiar las muestras de referencia durante el proceso de validación introduce una variable difícil de resolver que puede socavar gravemente la interpretación de los datos experimentales y, por tanto, la integridad del proceso de desarrollo y validación.

4 Las fuentes de variación que pueden afectar a la precisión de una prueba en el laboratorio son las siguientes: 1) en una misma ejecución de una única prueba, 2) entre ejecuciones consecutivas, 3a) entre ejecuciones de la prueba en distintos momentos del mismo día o en distintos días en condiciones similares, 3b) entre ejecuciones de la prueba en distintos días con distintos operarios, 4) entre laboratorios. En este capítulo, las categorías 1–3 son estimaciones de la repetibilidad, y la categoría 4 es sinónimo de reproducibilidad.

El laborioso proceso de optimización de una prueba es fundamental y crítico para lograr un rendimiento analítico fiable y predecible. La evaluación científica y la aplicación de las mejores prácticas científicas, como se indican en las Directrices de Validación de la OIE, se recomiendan para guiar la optimización de todos los elementos del desarrollo y validación de una prueba. El enfoque descrito proporciona una base firme para el desarrollo de una prueba fiable. A menudo se han desarrollado prototipos de pruebas utilizando los reactivos y el equipo que se tenían a mano en el laboratorio. Sin embargo, si la prueba ha sido diseñada para su ser utilizada con fines diagnósticos en múltiples laboratorios, la estandarización se vuelve crucial. Deben describirse detalladamente todas las formulaciones de sustancias químicas y tampones. Todos los reactivos deben ser definidos con respecto a la pureza y al grado (incluida el agua). Debe establecerse y documentarse cuáles son los intervalos de trabajo aceptables para parámetros como el pH, la molaridad, etc. De igual forma, para los productos biológicos también deben definirse los estándares de calidad, pureza, concentración y reactividad. Tanto en productos químicos como en productos biológicos también deben tenerse en cuenta los periodos de validez y las condiciones de almacenamiento. También deben establecerse los intervalos aceptables de los tiempos de reacción y las temperaturas. Debe describirse detalladamente el equipo esencial para optimizar el rendimiento de la prueba, incluyendo especificaciones de funcionamiento y de calibración. El control del proceso (calidad) debe formar parte de la optimización y debe tenerse en cuenta desde el principio, y no al final del desarrollo de la prueba, como suele ocurrir. Además de lo anterior, otros aspectos posteriores, como la obtención, manipulación e interpretación de los datos también pueden requerir estandarización y optimización. Finalmente, todos estos parámetros, una vez optimizados, deben describirse en detalle en el protocolo del método analítico.

Durante la optimización de la una prueba, es importante tomar nota de los pasos del procedimiento y de los parámetros de la prueba cuyo intervalo de valores en el que la prueba ofrece un rendimiento óptimo es estrecho, ya que son puntos críticos que en último término afectan a la fiabilidad de la prueba (véase el apartado A.2.7). En el caso de ciertos tipos de prueba, determinados pasos del procedimiento pueden influir más que otros en el rendimiento final de la prueba (véase el apartado B.5 abajo y el Capítulo 3.6.8 Comparabilidad entre pruebas tras realizar cambios en un método analítico validado para información adicional sobre cómo establecer la comparabilidad cuando se cambian reactivos o procesos).

En los siguientes apartados se presentan varias muestras de referencia de análisis y otros controles del proceso que se incluyen sistemáticamente en cualquier sistema analítico. Estos proporcionan las funciones críticas de seguimiento de la prueba que requieren atención especial durante la optimización del mismo. Además, para asegurar la estabilidad tienen que garantizarse la preparación y el almacenamiento adecuados de todos los reactivos biológicos y materiales de referencia (véase el capítulo 1.1.2).

2.4. Factores inhibidores en la matriz de la muestra

Cada matriz que vaya a utilizarse en una prueba debe emplearse en el proceso de validación. Algunas matrices de muestra incluyen factores inhibidores que interfieren con el rendimiento de determinados tipos de prueba. El suero, en concreto si está hemolizado, puede contener factores tóxicos para las células utilizadas en las pruebas de neutralización de virus, mientras que las sustancias endógenas que se encuentran en algunos tejidos y líquidos pueden interferir con las pruebas basadas en la fijación de ligandos y en enzimas, como el ELISA, o bien inhibirlas. Las muestras de heces, tejidos autolizados y semen tienden a contener más sustancias que interfieren y, por lo tanto, son más problemáticas para el rendimiento analítico que el suero, la sangre o los tejidos frescos.

2.5. Robustez

La robustez es la capacidad de una prueba de no resultar afectado por pequeñas variaciones en las situaciones en que se lleva a cabo el análisis, que podrían tener lugar en el curso de la prueba en un laboratorio determinado. La evaluación de la robustez debe comenzar durante las etapas de desarrollo y optimización de la prueba. Las variaciones intencionadas en los parámetros del método pueden abordarse en los experimentos una vez establecidas las condiciones óptimas de una prueba. Sin embargo, cuando se utilizan valoraciones multifactoriales de los reactivos para optimizar la prueba, pueden aparecer indicios de problemas en la robustez. Si pequeñas diferencias en las condiciones o en las concentraciones de reactivo causan una variabilidad inaceptable, lo más probable es que la prueba no sea robusta. Conocer cuanto antes esta situación constituye un punto crítico de la toma de decisiones para determinar si vale la pena seguir con la validación de la prueba, porque si esta no es robusta en un laboratorio en condiciones bastante ideales, es improbable que sea reproducible cuando se transfiera a otros laboratorios.

Los factores que con mayor probabilidad afectarán a la robustez de la prueba son el pH, la temperatura, el lote de los reactivos o la marca de las placas de microtitulación y factores relacionados

con la matriz acuosa u orgánica (Dejaegher y Vander Heyden, 2006). Una vez se ha terminado la optimización, la robustez de la prueba deviene parte de la evaluación de la repetibilidad.

2.6. Calibración de la prueba frente a reactivos estándar

2.6.1. Estándares de referencia a nivel internacional y nacional

Lo ideal es que los estándares de referencia a nivel internacional de la OIE y otros, que contienen una concentración o título conocidos del analito, sean los reactivos frente a los cuales todas las pruebas están estandarizadas (véase la Directriz nº 3 de la OIE: Estándares de Referencia a nivel Internacional de los Anticuerpos para las Pruebas de Anticuerpos⁵ y también el Capítulo 3.6.6). Estos estándares los preparan y distribuyen Laboratorios de Referencia de la OIE y otros laboratorios de referencia a nivel internacional. Los estándares de referencia a nivel nacional se calibran por comparación con un estándar de referencia a nivel internacional siempre que ello es posible, y los prepara y distribuye un laboratorio de referencia a nivel nacional. En ausencia de un estándar de referencia a nivel internacional, un estándar de referencia a nivel nacional se convierte en el estándar de comparación para la prueba candidata. Estos estándares se caracterizan con gran detalle mediante un exhaustivo análisis, y es preferible optar por métodos de caracterización, preparación y almacenamiento que hayan sido publicados en publicaciones revisadas por expertos.

2.6.2. Estándar interno

Un estándar de referencia interno en general debe calibrarse frente a un estándar internacional o Nacional. En ausencia de cualquiera de los calibradores y en la medida de lo posible, el estándar interno se caracteriza con gran detalle de igual forma que los estándares internacionales y nacionales (Capítulo 3.6.6). Este estándar interno local, por tanto, se convierte en el mejor estándar disponible, y se mantiene en volúmenes de alícuotas suficientes para un uso periódico como el estándar frente al cual se deberán calibrar los estándares de trabajo.

2.6.3. Estándar de trabajo

Se calibran uno o más estándares de trabajo, a menudo denominados controles del analito o del proceso, frente a un estándar internacional, nacional o interno, se preparan en grandes cantidades, en alícuotas, y se guardan para poder utilizarlos cada vez que se ejecute la prueba con fines de diagnóstico.

2.7. “Normalización” de los resultados de la prueba frente a un estándar de trabajo

Debido a la variación inherente en los resultados brutos de la prueba que se observan con frecuencia entre las ejecuciones de una misma prueba o entre laboratorios utilizando pruebas iguales o similares, es casi imposible comparar directamente datos (semi-)cuantitativos. Para mejorar visiblemente la comparabilidad de los resultados de la prueba, tanto intra como entre los laboratorios, se utiliza uno o más reactivos estándar de trabajo en cada ejecución de la prueba. A continuación, los valores brutos de la prueba para cada muestra problema pueden convertirse en unidades de actividad respecto al estándares de trabajo mediante un proceso denominado “normalización”. Los valores “normalizados” se pueden expresar de muchas maneras, como en porcentaje de un control positivo (por ejemplo, en un ELISA), o como la concentración o título estimados de un analito derivado de una curva estándar. Es una buena práctica el incluir estándares de trabajo en todas las ejecuciones de la prueba durante el desarrollo y validación de la prueba, ya que esto permite una “normalización” de los datos, lo cual proporciona un medio válido para la comparación directa de los resultados entre las ejecuciones de una prueba. Es indispensable controlar la variación (absoluta) de los estándares de normalización, puesto que de lo contrario la normalización puede introducir un sesgo. Véase los Capítulos 3.6.1 Desarrollo y optimización de las pruebas de detección de anticuerpos, 3.6.2 Desarrollo y optimización de las pruebas de detección de antígeno y 3.6.3 Desarrollo y optimización de las pruebas de detección de ácido nucleico, para más información.

2.8. Estudio provisional de la repetibilidad

La evaluación de la repetibilidad debe empezar durante las fases de desarrollo y optimización de la prueba. Conocer inicialmente esta situación resulta fundamental para determinar si vale o no la pena seguir adelante con la validación de la prueba.

⁵ http://www.oie.int/fileadmin/home/esp/Our_scientific_experts/docs/pdf/GUIDELINE_3_REF_STANDARD8_ESP.pdf

La repetibilidad se vuelve a verificar durante la Fase 1 de la validación de la prueba (apartado B.1.1). Cuando la prueba optimizada se ejecuta en un laboratorio de rutina o en condiciones de campo (Fase 4 de la validación de la prueba), se realiza un seguimiento continuo de la repetibilidad como parte de los procedimientos de control del proceso a lo largo de toda la prueba (véase el apartado B.5.1).

B. FASE DE VALIDACIÓN DE LA PRUEBA

La "validación" es un proceso que determina la idoneidad de una prueba que se ha desarrollado, optimizado y estandarizado adecuadamente para uno/s objetivo/s determinado/s. La validación incluye estimaciones de las características de rendimiento analítico y diagnóstico de una prueba. En el contexto de este documento, una prueba que ha superado las tres primeras etapas de la validación (Figura 1), incluida la caracterización del rendimiento, puede designarse como "validada para el/los propósito/s inicial/es deseado(s)".

1. Etapa 1 – Las características del rendimiento analítico

Lo ideal es que el diseño de los estudios descritos en los siguientes apartados se lleve a cabo con la ayuda de un estadístico y de un experto en enfermedades para asegurarse de que el tamaño de la muestra y el enfoque experimental sean válidos. Es posible diseñar experimentos que de manera eficiente proporcionen información sobre fuentes probables de variación en la precisión de la prueba intra e interlaboratorio (véase la nota 4 al pie, en el Apartado A.2.2, arriba), lo cual definirá las características de rendimiento de la prueba. La elección de los microorganismos, las cepas o los serotipos para evaluar la sensibilidad y la especificidad analíticas deberá reflejar el conocimiento actual y, por lo tanto, informar sobre cuál es el mejor diseño experimental posible para detectar análisis determinados.

1.1. Repetibilidad

La repetibilidad es el nivel de concordancia entre resultados de réplicas de una muestra tanto intra como entre ejecuciones del mismo método analítico en un laboratorio determinado. La repetibilidad se calcula mediante la evaluación de la variación en los resultados de varias réplicas. El número de réplicas debe determinarse preferiblemente consultando a un estadístico, y se sugiere un mínimo de tres muestras que contengan una actividad del analito que se sitúe en el intervalo de funcionamiento de la prueba. A continuación, de cada una de estas muestras se toman alícuotas y se depositan en un recipiente individual como réplicas idénticas de la muestra original que contienen las concentraciones originales de analito y matriz (véase el Capítulo 3.6.6). Después, cada réplica se procesa pasando por todos los pasos de la prueba, incluidos la creación de la dilución de trabajo, como si fuera una muestra problema procedente de la población objeto de la prueba. No resulta aceptable preparar una dilución final de trabajo de una muestra en un solo tubo a partir del cual se pipeteen alícuotas diluidas a vasos de reacción, ni crear réplicas de una extracción de ácido nucleico en lugar de extraer cada réplica antes de llevar a cabo la dilución en el interior de los vasos de reacción. Estas "muestras" no constituyen réplicas válidas para los estudios de repetibilidad. Se analiza la variación entre diferentes ejecuciones de la misma prueba utilizando las mismas muestras en múltiples ejecuciones llevadas a cabo por dos o más técnicos, preferiblemente en varios días. La variación en los resultados de las distintas réplicas se puede expresar como desviaciones estándar, coeficientes de variación (desviación estándar + media de réplicas), u otras opciones posibles (véase el Capítulo 3.6.4 *incertidumbre de la medición para las evaluaciones de la repetibilidad*).

1.2. Especificidad analítica

La especificidad analítica (ASp) es la capacidad de la prueba de distinguir entre el analito buscado (como un anticuerpo, microorganismo o secuencia genómica) y análisis no buscados, incluidos componentes de la matriz. La evaluación es cualitativa y la elección y procedencias de los tipos de muestra, microorganismos y secuencias escogidos para evaluar la ASp deben reflejar el propósito de la prueba y el tipo de prueba. En los Capítulos 3.6.1, 3.6.2 y 3.6.3 se ofrecen información sobre pruebas de anticuerpos, antígenos y ácidos nucleicos, respectivamente. La ASp se documenta durante la validación de la Fase 1, y se identifican reacciones cruzadas. La reactividad cruzada (ASp inferior al 100%) puede ser aceptable en función del uso propuesto de la prueba. La influencia de la reactividad cruzada se documenta en mayor medida durante la Fase 2 (establecimiento de la DSp) y se evalúa en el momento de implementar la Fase 4.

1.2.1. Selectividad

La selectividad es el grado en que un método puede cuantificar con exactitud el analito buscado en presencia de: 1) interferentes, tales como componentes de la matriz (por ejemplo, inhibidores de enzimas en la mezcla de reacción); 2) productos de degradación (por ejemplo,

factores tóxicos); 3) una unión no específica de los reactivos a una fase sólida (por ejemplo, conjugado de un ELISA adsorbido a un pocillo de la placa de microtitulación); 4) anticuerpos contra la vacunación, que pueden confundirse con anticuerpos contra la infección activa. Estos interferentes pueden causar falsas reducciones o incrementos en las respuestas de la prueba que afectarán negativamente a su especificidad analítica. Vessman, *et al* (2001) aportaron una útil visión general de la selectividad según se define para la química analítica, de la cual se dedujo una modificación aquí descrita para la aplicación a las pruebas diagnósticas.

1.2.2. Exclusividad

La exclusividad es la capacidad del método analítico de detectar un analito o secuencia genómica que es propia del microorganismo buscado, y excluye todos los demás microorganismos conocidos que pudieran dar una reacción cruzada. Esto también definiría una prueba confirmativa.

1.2.3. Inclusividad

La inclusividad es la capacidad de una prueba de detectar varias cepas o serovariedades de una especie, varias especies de un género o una agrupación similar de microorganismos o anticuerpos estrechamente emparentados. Caracteriza el ámbito de aplicación de una prueba de cribado.

1.3. Sensibilidad analítica

El límite de detección (LOD) es una medida de la sensibilidad analítica (ASe) de una prueba. El LOD es la cantidad estimada de analito en una matriz determinada que produciría un resultado positivo por lo menos durante parte del tiempo. Habitualmente, el LOD se estima añadiendo analito al matriz problema. La elección de/los analito/s (por ejemplo, especies o cepas) forma parte de la definición de la ASe y debe especificarse adecuadamente. Pueden diseñarse estas pruebas en función de la estimación precisa y exacta del nivel de probabilidad (por ejemplo, 50% o 100%), pero en determinadas circunstancias puede ser aceptable una estimación conservadora del LOD (por ejemplo, del 100%). Por ejemplo, en una titulación en la que se empleen diluciones decimales, todas las réplicas a todas las diluciones podrían mostrar una respuesta o bien del 100% o bien del 0%. En este caso existirían dos opciones. La última dilución que presente una respuesta del 100% puede aceptarse como estimación conservadora del límite inferior de detección. Una estimación más exacta puede conseguirse mediante una prueba de segunda fase empleando intervalos más estrechos en el esquema de dilución, centrándose en la región situada entre el 100% y el 0%. En el Capítulo 3.6.5 *Enfoques estadísticos de la validación* se muestran los métodos de evaluación estadística de los datos del LOD.

1.4. Exactitud analítica de las pruebas o procedimientos complementarios

Algunos métodos o procedimientos analíticos pueden ser calificados para su uso como herramientas analíticas en el laboratorio de diagnóstico. Estas suelen ser pruebas o procedimientos complementarios secundarios que se aplican a un analito que se ha detectado en una prueba primaria. El propósito de tales instrumentos de análisis es caracterizar en mayor grado el analito detectado en la prueba primaria. Algunos ejemplos de pruebas complementarias abarcan desde la neutralización vírica hasta la tipificación o la secuenciación molecular de un virus aislado.

Estas pruebas complementarias deben validarse en cuanto a las características de rendimiento analítico (Apartados A.2 a B.1.3, arriba), pero difieren de las pruebas diagnósticas en que no requieren validación relativa a las características de rendimiento diagnóstico (Apartados B.2 a B.4, abajo) si sus resultados no se utilizan para establecer un diagnóstico final respecto al propósito deseado. La exactitud analítica de estas herramientas se puede definir por comparación con un reactivo estándar de referencia, o por las características inherentes a la propia herramienta (como la titulación a punto final). En todos estos ejemplos, el analito en cuestión se caracteriza en mayor medida cuantitativa o cualitativamente mediante la herramienta analítica.

2. Etapa 2 – Rendimiento diagnóstico de la prueba

Las estimaciones de la DSe (proporción de muestras de animales de referencia que se sabe que están infectados y que dan positivo en una prueba) y de la DSp (la proporción de muestras de animales de referencia que se sabe que no están infectados y que dan negativo en una prueba) son los indicadores principales del rendimiento establecidos durante la validación de una prueba. Estas estimaciones son la base del cálculo de otros parámetros a partir de los cuales se realizan inferencias sobre los resultados de la prueba (por ejemplo, los valores predictivos de los resultados positivos y negativos de la prueba). Por consiguiente, es muy importante

que las estimaciones sobre la sensibilidad y la especificidad diagnósticas sean tan exactas como sea posible. Lo ideal es que deriven del análisis de un conjunto de muestras procedentes de animales de referencia, cuyos antecedentes y estado en cuanto a la enfermedad/infección en cuestión se conozcan y sean relevantes para el país o región en los cuales se va a utilizar la prueba. Una estimación del área bajo la curva característica operativa del receptor (ROC) es una herramienta útil para estimar la DSe y la DSp de una prueba diagnóstica cuantitativa porque evalúa la exactitud global teniendo en cuenta todos los posibles valores de la misma (Greiner et al, 2000; Zweig & Campbell, 1993). Este enfoque se describe en detalle en el Capítulo 3.6.5.

El número designado de muestras que se sepa que son positivas y muestras que se sepa que son negativas dependerá de cuáles sean los valores probables de DSe y de DSp de la prueba candidata y del nivel de confianza deseado para las estimaciones (Tabla 1 y Jacobson, 1998). En la Tabla 1, se indican dos conjuntos de las cantidades teóricas de muestras necesarias, permitiendo un error del 5% o del 2% en las estimaciones de la DSe o la Dsp. La comparación de un error del 5% frente a un error del 2% presenta una considerable reducción en el número de muestras necesarias. Para lograr una confianza alta (habitualmente, del 95%) en las estimaciones de DSe o DSp se requieren muchas muestras cuando se desea un margen de error pequeño en la estimación. Por ejemplo, al pasar de un error del 2% a un error del 5% para una DSe o DSp probables del 90% y un 95% de confianza se observa un aumento considerable (864 frente a 138) en el número de muestras necesarias. Las limitaciones logísticas y financieras podrían requerir que se evaluara un tamaño de muestra inferior al exigido estadísticamente, en cuyo caso el intervalo de confianza calculado para la DSe y la DSp indicarían menor confianza diagnóstica en los resultados. El tamaño de la muestra también puede resultar limitado por el hecho de que no se disponga de las poblaciones de referencia ni los estándares de referencia de la OIE (para más información, véase el Capítulo 3.6.5). Por lo tanto, inicialmente puede ser necesario utilizar una cantidad subóptima de muestras. Sin embargo, puede ser muy deseable potenciar la confianza y reducir el margen de error permitido en las estimaciones de la DSe y la DSp añadiendo más muestras (de estado equivalente respecto al conjunto original) a medida que se dispone de ellas.

Tabla 1. Cantidad teórica de muestras procedentes de animales cuyo estado de infección se sabe, necesaria para establecer las estimaciones de sensibilidad (DSe) y especificidad (DSp) diagnósticas en función del valor probable de DSe o DSp y del margen de error y el nivel de confianza deseados

Estimación de DSe o DSp	2% de error permitido en la estimación de la DSe y la DSp			5% de error permitido en la estimación de la DSe y la DSp		
	Confianza			Confianza		
	90%	95%	99%	90%	95%	99%
90%	610	864	1493	98	138	239
92%	468	707	1221	75	113	195
94%	382	542	935	61	87	150
95%	372	456	788	60	73	128
96%	280	369	637	42	59	102
97%	197	279	483	32	45	77
98%	133	188	325	21	30	52
99%	67	95	164	11	15	26

Los siguientes son ejemplos de poblaciones y metodologías de referencia que pueden ayudar a determinar las características de rendimiento de la prueba que está siendo validada.

2.1. Poblaciones de animales de referencia

Teóricamente, la selección de animales de referencia requiere que variables importantes del hospedador de la población estudiada estén representadas en los animales escogidos para ser infectados por el agente en cuestión o expuestos al mismo, o que nunca hayan sido infectados ni expuestos. Las variables a destacar son, aunque no exclusivamente, la especie, la edad, el sexo, la raza, el estado de la infección, el historial de vacunación y el historial de enfermedades relevantes en el rebaño (para más información, véase el Capítulo 3.6.6).

2.1.1. Muestras de referencia negativas

Puede ser difícil localizar muestras que se haya comprobado que son negativas, de animales que no hayan tenido posibilidad de infección ni exposición al agente en cuestión. A menudo es

posible obtener estas muestras de países o zonas en los que la enfermedad en cuestión se ha erradicado o nunca ha existido. Estas muestras son útiles siempre que la población objeto de la prueba sea suficientemente similar a la población de la que procede la muestra.

2.1.2. Muestras de referencia positivas

En general es problemático hallar cantidades suficientes de animales de referencia verdaderamente positivos, comprobados mediante el aislamiento del agente patógeno. Puede ser necesario recurrir a muestras de animales que se hayan identificado mediante otras pruebas de exactitud suficientemente alta, como las pruebas validadas de detección de ácido nucleico. La prueba candidata se aplica a estas muestras de referencia y los resultados (positivos y negativos) se clasifican de forma cruzada en una tabla 2x2. Esto se ha denominado "modelo de referencia" porque en él se considera que el estándar de referencia es perfecto. En la Tabla 2 del apartado B.2.5) se muestra un cálculo de la muestra.

2.2. Muestras de animales que no se sabe si están infectados

Cuando el denominado estándar de referencia es imperfecto, que es lo habitual en cualquier prueba de diagnóstico, las estimaciones de la DSe y de la DSp para la prueba candidata que se basen en este estándar no serán perfectas. Una forma de superar este problema es llevar a cabo un análisis de clases latentes de los resultados conjuntos de ambas pruebas asumiendo que ninguna prueba es perfecta.

Los modelos de clases latentes no parten de la suposición de una prueba de referencia perfecta sino que estiman la exactitud de la prueba candidata y del estándar de referencia con los resultados analíticos conjuntos (Branscum et al., 2005; Enøe et al., 2000; Georgiadis et al., 2003; Hul y Walter, 1980). Si se utiliza un análisis de clases latentes bayesiano, puede incorporarse al análisis la información de la que se disponga previamente sobre el rendimiento de la prueba de referencia y de la prueba candidata.

Dado que estos modelos estadísticos son complejos y requieren suposiciones críticas, debe solicitarse la ayuda de expertos en estadística para orientar el análisis y describir el muestreo de la/s poblaciones de destino y las características de otras pruebas incluidas en el análisis, y para elegir correctamente el modelo y los métodos de estimación en base a bibliografía revisada por expertos (más información en el Capítulo 3.6.5).

2.3. Animales de referencia infectados experimentalmente o vacunados

Las muestras obtenidas secuencialmente a partir de animales infectados experimentalmente o vacunados son útiles para determinar la cinética de las respuestas de anticuerpos o la presencia/ausencia de antígeno o microorganismos en muestras de dichos animales. Sin embargo, los resultados múltiples adquiridos de forma seriada antes y después de la exposición de animales individuales no son aceptables para establecer estimaciones de la DSe y la DSp porque se infringe el requisito estadístico de las observaciones independientes. Puede aceptarse el muestreo de animales de experimentación realizado en un solo momento (por ejemplo, una muestra tomada aleatoriamente de cada animal). No obstante, es importante tener en cuenta que en el caso de los métodos indirectos de detección del análisis, la exposición a microorganismos en condiciones experimentales o la vacunación pueden desencadenar respuestas de anticuerpos que podrían no ser cuantitativa y cualitativamente típicas de la infección natural en la población estudiada (Jacobson, 1998). La cepa del microorganismo, la dosis y la vía de administración a los animales de experimentación son ejemplos de variables que pueden inducir a error cuando se extrapolan las estimaciones de sensibilidad y especificidad diagnósticas a la población estudiada. Si debido a que es prácticamente imposible obtener muestras de referencia adecuadas de animales expuestos de forma natural se hace necesario utilizar muestras de animales de experimentación para los estudios de validación, las mediciones resultantes de DSe y de DSp deben considerarse estimaciones subóptimas de las verdaderas DSe y DSp.

2.4. Determinación de los umbrales (puntos de corte)

Para obtener estimaciones de la DSe y la DSp de la prueba candidata, que se mide en una escala continua, en primer lugar los resultados de la prueba deben agruparse en categorías: dos (positivos o negativos) o tres (positivos, intermedios [dudosos] o negativos). Esto se consigue insertando uno o dos puntos de corte (umbrales o límites de decisión) en la escala de resultados de la prueba. Los puntos de corte escogidos deben reflejar el objetivo pretendido de la prueba y su aplicación, y deben respaldar la DSe y la DSp requeridas para la prueba. Existen opciones y métodos descriptivos para determinar la mejor forma de expresar la DSe y la DSp (Branscum et al., 2005; Georgiadis et al., 2003;

Greiner et al., 19954; Greiner et al., 2000; Jacobson, 1998; Zweig y Campbell, 1993 y el Capítulo 3.6.5). Si se produce una superposición considerable en las distribuciones de los valores de las pruebas realizadas en animales que se sepa que están o no infectados, resulta imposible seleccionar un único punto de corte que permita clasificar adecuadamente a estos animales con relación a su estado de infección. En lugar de un único punto de corte, se pueden seleccionar dos puntos de corte que definan una DSe alta (por ejemplo, que incluya el 99% de los valores de los animales infectados), y una DSp alta (por ejemplo, que incluya el 99% de los valores de los animales no infectados).

La principal dificultad de establecer puntos de corte basados en las características del rendimiento diagnóstico es el no conocer el número necesario de muestras bien caracterizadas. Se explican alternativas en el Apartado B.2.6 sobre la aceptación provisional de una prueba durante la recogida de datos para mejorar las estimaciones de la DSe y de la DSp.

2.5. Cálculo de la DSe y de la DSp basado en los resultados del análisis de muestras de referencia

Un método habitual de determinación de las estimaciones de la DSe y la DSp es analizar las muestras de referencia con la prueba nueva, y tabular de forma cruzada los resultados categóricos de la prueba en una Tabla de 2 x 2. En un ejemplo hipotético, supongamos que el técnico que ejecutó la prueba eligió un tamaño de muestra suponiendo que los valores más probables de DSe y de DSp de la prueba nueva son del 97% (DSe) y del 99% (DSp), respectivamente, con una confianza deseada del 95% en ambas estimaciones. La cantidad de error permitido en las estimaciones se fijó en el 2%. La Tabla 1 indica que se necesitan 279 muestras de animales que se sepa que están infectados para evaluar la DSe, y que se necesitan 95 muestras que se sepa que son negativas para establecer la estimación de la DSp. A continuación, se analizaron las muestras mediante la prueba nueva. La Tabla 2 es un conjunto hipotético de resultados a partir de los cuales se han obtenido las estimaciones de la DSe y de la DSp.

Tabla 2. Estimaciones de la sensibilidad y la especificidad diagnósticas calculadas a partir de un conjunto hipotético de resultados de muestras analizadas procedentes de poblaciones que se sabe que están infectadas o de poblaciones que se sabe que no están infectadas.

		Cantidad de muestras de referencia necesarias*	
		Que se sabe que son positivas (279)	Que se sabe que son negativas (95)
Resultados de la prueba	Positivos	270	7
	Negativos	9	88
		VP	FP
		FN	VN
		Sensibilidad diagnóstica*	Especificidad diagnóstica*
		$VP/(VP + FN)$	$VN/(VN + FP)$
		98,8% (94,0 – 98,5%)**	92,8% (85,4 – 97,0%)**

*Basado en la Tabla 1 para una prueba con los siguientes parámetros:

- 1) Antes del análisis, DSe estimada del 97% y DSp estimada del 99%
- 2) 95% = confianza necesaria en las estimaciones de la DSe y la DSp
- 3) 2% = Error permitido en las estimaciones de la DSe y la DSp

VP y FP = Verdadero Positivo y Falso Positivo, respectivamente

VN y FN = Verdadero Negativo y Falso Negativo, respectivamente

**Intervalo de confianza binomial exacto del 95% para los valores calculados de DSe y DSp (véase el Capítulo 3.6.5 para información sobre límites de confianza)

En este ejemplo, las estimaciones de la DSe son las previstas, pero la DSp es muy inferior (92%) al valor previsto del 99%. Como consecuencia, la amplitud del intervalo de confianza para la DSp es muy superior a la esperada. Al volver a examinar la Tabla 1, se observa que son necesarias 707 muestras para lograr un margen de error de $\pm 2\%$ para una DSp del 92%, pero este aumento del tamaño de la muestra podría no ser factible (véase el Capítulo 3.6.5 para más información).

2.6. Reconocimiento de prueba provisional⁶

En determinadas situaciones no se puede o no se debe pasar a la Fase 2 del Proceso de Validación porque no se dispone de suficientes muestras adecuadas de la población estudiada y es difícil acceder a los animales (por ejemplo, debido a enfermedades infecciosas transfronterizas o a enfermedades de la fauna salvaje).

La experiencia ha mostrado que el principal obstáculo para pasar a la Fase 2 del Proceso de Validación es la cantidad de muestras específicas necesarias para calcular la DSe y la DS_p. La fórmula se conoce bien y existe tablas para determinar el número de muestras necesarias para estimar distintos niveles de DSe y DS_p, en función del margen de error y del nivel de confianza deseados en las estimaciones (Tabla 1 y Jacobson, 1998). En la fórmula se da por supuesto que está contemplada la infinidad de factores relacionados con el hospedador/microorganismo que pueden influir en el resultado de la prueba. Dado que este supuesto puede ser cuestionable, los tamaños de muestra estimados en el mejor de los casos son muy pequeños. En el caso de una enfermedad que no sea endémica ni esté diseminada, inicialmente podría ser imposible obtener el número de muestras requerido, pero con el tiempo, la acumulación de datos adicionales permitirá el ajuste del punto de corte (umbral) o, en el caso de que no se precise ajuste, una mejora de la confianza en las estimaciones.

El reconocimiento provisional define una prueba en cuya Fase 1 se han evaluado parámetros básicos (ASE, AS_p y repetibilidad) y, además, se ha realizado una estimación preliminar de la DS_p y la DSe basada en un conjunto pequeño de muestras bien caracterizadas que contengan el analito de interés y una estimación preliminar de la reproducibilidad. Esto supone haber terminado parte de la Fase 2. Pueden llevarse a cabo estimaciones preliminares de la reproducibilidad de la prueba candidata empleando el grupo escogido de muestras bien caracterizadas para mejorar el estado de aceptación provisional de la prueba. A continuación, el método analítico candidato se duplica en laboratorio de al menos dos centros distintos, y el conjunto de muestras se evalúa empleando la muestra candidata en cada uno de estos laboratorios, utilizando el mismo protocolo, los mismos reactivos que se han descrito en el protocolo y un equipo comparable. Es una versión reducida del estudio de reproducibilidad de la Fase 3 de la validación de la prueba. Siguiendo este procedimiento de reconocimiento provisional, el protocolo de la prueba no debe variar.

El reconocimiento provisional de una prueba por parte de autoridades estatales o nacionales significa que no se han evaluado las características de rendimiento diagnóstico de la prueba. Así, el laboratorio debe desarrollar y seguir un protocolo para añadir y evaluar muestras, a medida que se disponga de ellas, con el fin de cumplir este requisito. Teóricamente, este proceso debe limitarse a un período de tiempo concreto en el cual esta recopilación de muestras tendría por objetivo cumplir las Fases 2 y 3 del proceso de validación, y se reservaría a situaciones determinadas (emergencias, especies menores, situaciones en las que no se disponga de ninguna otra prueba, etc.).

3. Etapa 3 – Estimaciones de la reproducibilidad y del aumento de la repetibilidad

3.1. Reproducibilidad

La reproducibilidad es la capacidad de un método analítico de que sus resultados sean coherentes, según las estimaciones de la precisión, al aplicarlo a alícuotas de las mismas muestras analizadas en distintos laboratorios, preferiblemente situados en regiones o países distintos utilizando exactamente la misma prueba (protocolo, reactivos y controles). Para evaluar la reproducibilidad de una prueba, cada uno de al menos tres laboratorios debe analizar el mismo conjunto de un mínimo de 20 muestras, con idénticas alícuotas para cada laboratorio (véase el Capítulo 3.6.6). Este ejercicio también genera datos preliminares sobre efectos no aleatorios atribuibles a la utilización de la prueba en otros laboratorios. Además, las estimaciones de la repetibilidad intra-laboratorio aumentan por las réplicas que se utilizan en los estudios de reproducibilidad. Pueden estimarse las mediciones de la precisión de los datos tanto de reproducibilidad como de repetibilidad (en el Capítulo 3.6.4 se explica en mayor detalle este tema y su aplicación).

En el caso de las pruebas de campo, la reproducibilidad debe evaluarse en las condiciones en que pretenda utilizarse.

⁶ El reconocimiento de prueba provisional no implica la aceptación por parte de la OIE. No obstante, reconoce una decisión informada de las autoridades a nivel local, estatal, nacional o internacional de que aprueba de forma condicional una prueba validada parcialmente.

3.2. Designación de prueba validada

Al terminar la validación de Fase 3, suponiendo que las fases anteriores se hayan superado total y satisfactoriamente, la prueba puede considerarse "validada para el propósito inicial deseado". El que se mantenga esta designación depende de si se lleva a cabo un seguimiento continuo del rendimiento de la prueba, como se describe en el apartado 5.1.

4. Fase 4 – Implementación del programa

Cuando una prueba funciona bien, se obtiene información adicional y útil de su rendimiento respecto a las expectativas. Además, la prevalencia (real) del rasgo diagnóstico en la población en estudio es un factor importante que debe tenerse en cuenta como se describe abajo.

4.1. Idoneidad

Aunque este capítulo trata la validación y la idoneidad para un propósito determinado desde el punto de vista científico, también debe tenerse en cuenta que existen otros factores que podrían influir en la utilidad de una prueba para la aplicación prevista. Estos factores son no solo la idoneidad diagnóstica de la prueba, sino también su aceptabilidad por parte de las comunidades científica y reguladora, la aceptabilidad para el cliente y la viabilidad dados los recursos de laboratorio de los que se disponga. En algunas enfermedades, tal vez se disponga de varias pruebas que puedan utilizarse junto con programas de control y vigilancia de la enfermedad, en cuyo caso, la utilidad de la prueba tal vez tenga que valorarse evaluando incrementos graduales en los valores de la DSe, la DSp y predictivos de las dichas pruebas combinadas.

Cuando resulta imposible cumplir los requisitos de funcionamiento de una prueba, dicha prueba no resultará idónea para el propósito deseado. Estos requisitos pueden ser costes de funcionamiento, disponibilidad de equipo, nivel de sofisticación técnica y de capacidad de interpretación, disponibilidad de kits/reactivos, período de validez, requisitos de transporte, seguridad, bioseguridad, rendimiento de la muestra, tiempo necesario para la entrega de resultados, aspectos del control de calidad y de la garantía de calidad, y si es viable llevar a cabo la prueba en otros laboratorios. Los kits de las pruebas utilizados sobre el terreno son muy deseables en el sentido de que son muy fáciles de utilizar, pero dado que se usan fuera del entorno controlado de un laboratorio, deben añadirse precauciones para que sigan siendo adecuados para la finalidad en cuestión (Crowther et al., 2006).

4.2. Interpretación de los resultados de la prueba

Valores predictivos de los resultados de las pruebas: El valor predictivo positivo (PPV) es la probabilidad de que un animal que haya dado positivo en la prueba realmente sea positivo según el diagnóstico real. El valor predictivo negativo (NPV) es la probabilidad de que un animal que ha dado negativo en la prueba de hecho sea negativo según el diagnóstico real.

Los valores predictivos de los resultados de una prueba son una aplicación del teorema de Bayes y se calculan del siguiente modo:

$$PPV = \frac{P \times DSe}{P \times DSe + (1 - P) \times (1 - DSp)}$$

y

$$NPV = \frac{(1 - P) \times DSp}{P \times (1 - DSe) + (1 - P) \times DSp}$$

Donde:

- PPV = Valor predictivo de un resultado positivo en una prueba
- NPV = Valor predictivo de un resultado negativo en una prueba
- P = Prevalencia de infección
- DSe = Sensibilidad diagnóstica
- DSp = Especificidad diagnóstica

Contrariamente a lo que ocurre con la DSe y la DS_p, los valores predictivos están influidos por la prevalencia real del estado diagnóstico real de la población en estudio. En otras palabras, los valores predictivos no son características inherentes a una prueba diagnóstica determinada, sino una función de su DSe y su DS_p y de la prevalencia local de infección en una población definida en un momento determinado.

Los valores predictivos resultan de gran importancia para que los veterinarios de campo interpreten los resultados. Por ejemplo, un PPV de 0,9 significa que un animal que ha dado positivo en la prueba tiene un 90% de probabilidad de estar realmente infectado y un 10% de probabilidad de ser un falso positivo.

El valor predictivo de un resultado positivo también es muy importante para los servicios veterinarios responsables de la gestión de los programas de control o erradicación. En cuanto a la inversa del PPV (es decir, 1/PPV), da la información sobre cuánto dinero se gasta en el desvío de verdaderos y falsos positivos por cada animal realmente positivo que se detecta mediante las actividades de vigilancia. Dicho de otra forma, si el PPV es 0,67, significa que dos de cada tres animales positivos son realmente positivos y que el restante es un falso positivo. Dado que durante la aplicación de un programa de control, la prevalencia de infección está cambiando continuamente, el seguimiento del PPV es una forma de evaluar los costes del programa.

Además, durante la aplicación de un programa de control, normalmente es aconsejable cambiar la sensibilidad de las pruebas utilizadas; en base a la variación de la prevalencia de infección en la población estudiada y al objetivo del programa, el PPV puede utilizarse para aplicar cambios en la DSe y la DS_p en función de aspectos económicos. En otras palabras, cuando surge la necesidad de aplicar un cambio en la DSe y la DS_p de la prueba, pueden fijarse varios umbrales posibles a lo largo de la curva ROC de validación de la prueba, y pueden utilizarse los valores pertinentes de DSe y DS_p para cada punto de corte para evaluar el coste esperable del desvío de cada animal infectado.

Si el propósito es establecer evidencia de la ausencia de enfermedad, el NPV es la medición más importante. El NPV depende totalmente de la DSe.

4.3. Reconocimiento internacional

Tradicionalmente, la OIE ha reconocido pruebas a nivel internacional cuando están diseñadas como pruebas prescritas o alternativas para fines comerciales. Esto a menudo se ha basado en la evidencia de su utilidad a nivel nacional, regional o internacional. En el caso de los kits de diagnóstico comerciales que han pasado por el procedimiento de la OIE de validación y certificación de pruebas de diagnóstico, el paso final es incluir la prueba en el Registro de la OIE. Las pruebas incluidas en el Registro están certificadas como adecuadas para un propósito concreto si han superado las etapas de Validación 1, 2 y 3. El Registro tiene por objetivo proporcionar a los posibles usuarios de la prueba una fuente informada e imparcial de información sobre el kit y sus características de rendimiento para un propósito determinado. Este Registro puede consultarse en la página web de la OIE: <http://www.oie.int/es/nuestra-experiencia-ciencia/certificacion-de-pruebas-de-diagnostico/registro-de-pruebas-de-diagnostico/>.

4.4. Utilización de la prueba

La prueba definitiva de la utilidad de una prueba es/son su/s aplicación/es con éxito en otros laboratorios y su inclusión en programas nacionales, regionales y/o internacionales de control o vigilancia. Los laboratorios de referencia desempeñan un papel crucial en este proceso. En el avance natural de las mejoras diagnósticas y/o tecnológicas, nuevas pruebas se convertirán en el nuevo método de referencia frente al cual se compararán las demás. Como tales, poco a poco obtendrán el reconocimiento nacional, regional e internacional. Como estándares reconocidos, estas pruebas también se utilizarán para elaborar reactivos de referencia con fines de control de calidad, eficiencia y armonización. Esos reactivos de referencia también pueden convertirse en estándares internacionales.

Debe repetirse una evaluación de la reproducibilidad cuando la prueba se transfiere del laboratorio al campo, tanto para su uso en laboratorios locales como en aplicaciones a pie de explotación. Los cambios predecibles, como extremos de temperatura y niveles de experiencia del técnico, deben evaluarse como fuentes adicionales de variación en los resultados de la prueba que pueden afectar a las estimaciones de la reproducibilidad.

5. Seguimiento del rendimiento de la prueba tras la validación inicial

5.1. Seguimiento de la prueba

Para que una prueba siga considerándose validada, es necesario asegurarse de que la misma conserva las características de rendimiento tal como se definieron durante la validación (véase el apartado 6, abajo), lo cual puede determinarse mediante un programa de garantía de calidad que se caracterizará por un seguimiento cuidadoso del rendimiento diario de la prueba, principalmente con estimaciones de la precisión y la exactitud obtenidas al analizar los controles internos, así como de la tendencia a dar datos atípicos. Puede realizarse un seguimiento del rendimiento gráficamente marcando en diagramas de control los resultados obtenidos con los controles⁷. Deben estudiarse las desviaciones respecto al rendimiento esperado para poder aplicar medidas correctivas si es necesario. Este seguimiento aporta evidencias cruciales de que la prueba conserva su calificación de "validada" durante la fase de implementación de la prueba. La reproducibilidad se evalúa mediante programas de control de calidad externos, como la comprobación de la competencia. En caso de que la prueba empiece a dar resultados que no concuerden con los datos de la validación inicial, deberá clasificarse como no adecuada para el propósito deseado. Así, las pruebas validadas deben someterse a una evaluación continua para asegurarse de que siguen siendo adecuadas para el propósito en cuestión.

5.2. Modificaciones y mejoras – consideraciones a tener en cuenta para aplicar cambios en la prueba

Con el tiempo, es probable que sea necesario aplicar modificaciones en la prueba para abordar posibles cambios en el propósito de la prueba, en los análisis estudiados (es decir, modificaciones en la prueba para ajustar el rendimiento diagnóstico) o bien en las características técnicas, con el fin de mejorar la eficiencia o la relación coste-eficacia de la prueba. En caso de que se produzca un cambio en el propósito de la prueba, será obligatoria una validación revisada desde la Fase 2 en adelante.

Si la prueba va a aplicarse a otra región geográfica y/o población, es recomendable volver a validarla en las nuevas condiciones. Se sabe que los linajes o sublinajes de los agentes infecciosos de los animales varían en función de la región geográfica, lo cual implica tener que revalidar la prueba para cada población en estudio. Esto es especialmente cierto en el caso de los sistemas de detección de ácido nucleico (DAN) y es muy frecuente que se produzcan mutaciones puntuales en muchos agentes infecciosos (sobre todo, en virus de ARN). Las mutaciones, que pueden tener lugar en puntos del cebador o de la sonda, pueden afectar a la eficiencia de la prueba e incluso invalidar las características de rendimiento establecidas. También es aconsejable confirmar periódicamente la secuencia de acceso en las regiones genómicas escogidas de cepas nacionales o regionales de los agentes infecciosos. Esto es especialmente cierto para los puntos del cebador y de la sonda, para garantizar que permanecen estables y que la DSe y la DSp de la prueba no resultan comprometidas. Pueden surgir cuestiones similares con las pruebas basadas en la inmunología, tanto para detección de antígenos como de anticuerpos.

Puede producirse una situación similar con el surgimiento de nuevos subtipos de agentes patógenos existentes. En tales circunstancias, las pruebas DAN existentes tal vez tengan que modificarse.

5.2.1. Modificaciones técnicas y evaluaciones de la comparabilidad

Lo habitual es que las modificaciones técnicas realizadas en una prueba validada, como los cambios de instrumental o de los protocolos de extracción, así como la conversión de una prueba en un sistema semiautomático o totalmente automático utilizando la robótica, no exijan una completa revalidación de la prueba. En lugar de ello, se lleva a cabo una comparación de los métodos para determinar si las modificaciones relativamente pequeñas de la prueba afectan a las características de rendimiento documentadas con anterioridad. La comparabilidad puede establecerse ejecutando el procedimiento modificado y el original en paralelo, con el mismo conjunto de muestras en ambos, y realizando varias ejecuciones. El conjunto de muestras escogido para esta comparación debe representar el intervalo completo de funcionamiento de ambas pruebas. Si se determina que los resultados del procedimiento modificado y del método original validado son comparables en un experimento basado en un criterio pre-especificado, la prueba modificada sigue considerándose validada para el propósito deseado. Véase el Capítulo 3.5.8 por una descripción de experimentos que son adecuados para comprobar la comparabilidad y el Capítulo 3.5.6 por conjuntos de muestras de referencia.

⁷ Gráfico de control: Es una representación gráfica de datos obtenidos de la medición repetida de una o varias muestras control analizadas en distintas ejecuciones de la prueba a lo largo del tiempo.

5.2.2. Modificaciones biológicas y evaluaciones de la comparabilidad

Puede haber situaciones en que sea necesario y/o esté justificado realizar cambios en algunas de las sustancias biológicas utilizadas en la prueba. Estos cambios pueden realizarse en la muestra en sí (por ejemplo, una tejido distinto o incluso otra especie animal), en los reactivos (por ejemplo, la sustitución de un antígeno recombinante por un antígeno derivado de cultivo celular, o de un antígeno conjugado a anticuerpo por otro de especificidad inmunológica similar en un ELISA). La dificultad de cualquier modificación radica en el hecho de que debe determinarse si el cambio requiere una revalidación completa de la prueba tanto a nivel de laboratorio como de campo. Como mínimo, toda modificación requiere se evalúen los "requisitos analíticos" apropiados de la Fase 1. La decisión más difícil es la relativa al "rendimiento diagnóstico" de la Fase 2. Para facilitar dicha decisión, en primer lugar la prueba original (de referencia) debe compararse con la prueba modificada (candidata) en un ensayo controlado en el que se empleará un conjunto definido de muestras diagnósticas positivas y negativas. En el Capítulo 3.6.8 se ofrece una descripción de la evaluación de la comparabilidad. Si la evaluación de la comparabilidad no sugiere un cambio en el rendimiento diagnóstico, la prueba modificada se puede pasar a la fase de uso sistemático. Si, por el contrario, se observan diferencias en la DSp y la DSe, para adoptar la prueba modificada se requerirá otra Fase 2 o una validación sobre el terreno.

5.2.3. Reposición de reactivos agotados

Cuando un reactivo, como una muestra control o un estándar de trabajo, está a punto de agotarse, es fundamental preparar y analizar repetidamente un reactivo de repuesto antes de que aquellos se agoten. La futura muestra control debe incluirse en múltiples ejecuciones de la prueba paralelamente con el control original para establecer su relación de proporcionalidad. Siempre que sea posible, es importante cambiar cada vez solamente un reactivo para evitar agravar el problema al tener que evaluar más de una variable.

5.3. Mejora de la confianza en los criterios de validación

Dado que muchas de las variables del hospedador influyen en el rendimiento diagnóstico de las pruebas, es muy deseable aumentar con el tiempo el número de muestras de referencia o de muestras adecuadas para un análisis de clases latentes. El diseño de la obtención, recogida, transporte y entorno en el que se analicen nuevas muestras deben ser los mismos que los aplicados al estudio de validación inicial. Al aumentar el número de muestras mejora la precisión de las estimaciones globales de la DSe y la DSp, y puede permitir cálculos de las estimaciones de la DSe mediante factores como la edad, el estado de la enfermedad y la carga de microorganismos. Deben incluirse nuevos datos cada año en los dossiers de pruebas correspondientes.

5.4. Verificación de pruebas existentes (validación interna)

Si un laboratorio está planteándose utilizar un kit comercial validado o una prueba candidata en base a la bibliografía publicada sobre datos de validación, será preciso algún tipo de verificación para determinar si dicha prueba cumple con lo afirmado por el fabricante o el autor respecto a los criterios de validación de la Fase 1 y en el contexto de la aplicación pretendida. Ello puede requerir una verificación reducida tanto de la ASp como de la ASe empleando materiales de referencia de los que se disponga, tanto si se han adquirido fuera como dentro de la población en estudio. Una vez el laboratorio está seguro de que la prueba, desde el punto de vista analítico, está rindiendo como se ha descrito, para que pueda pasar a utilizarse de forma sistemática deberá plantearse el pasar antes por una Fase 2 reducida en el contexto de la aplicación pretendida y de la población en estudio.

BIBLIOGRAFÍA

BRANSCUM A.J., GARDNER L.A. JOHNSON W.O. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Prev. Vet. Med.*, **68**, 145–163.

GROWTHIER J.R., UNGER H. & VILJOEN G.J. (2006). Aspects of kit validation for tests used for the diagnosis and surveillance of livestock diseases: producer and end-user responsibilities. *Rev. sci. tech. Off. Int. Epiz.*, **26** (3), 913–935.

DEJAEGHER B. & VANDER HEYDEN Y. (2006). Robustness tests. *LCGC Europe*, **18** (7), online at <http://www.lcgeurope.com/lcgeurope/content/printContentPopup.jsp?id=357956>

- ENNE C., GEORGIADIS M.P. & JOHNSON W.O. (2000). Estimating the sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev. Vet. Med.*, **46**, 61–81.
- FINDLAY J.W.A. & DILLARD R.F. (2007). Appropriate calibration curve fitting in ligand binding assays. *AAPS J.*, **9** (2): E260-E267. (Also on-line as *AAPS Journal* [2007]; **9** [2], Article 29 [<http://www.aaps.org>]).
- GEORGIADIS M., JOHNSON, W., GARDNER I. & SINGH R. (2003). Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Appl. Statist.*, **62** (Part 1), 63–76.
- GRENER M., SOHR D. & GÖBEL P. (1995). A modified ROC analysis for the selection of cut-off values and the definition of intermediate results of serodiagnostic tests. *J. Immunol. Methods*, **185**,123–132.
- GRENER M., PFEFFER D. & SMITH R.D. (2000). Principles and practical application of the receiver operating characteristic (ROC) analysis for diagnostic tests. *Vet. Prev. Med.*, **46**, 23–41.
- HUI S.L. & WALTER S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics*, **38**, 167–171.
- JACOBSON R.H. (1998). Validation of serological assays for diagnosis of infectious diseases. *Rev. sci. tech. Off. Int. Epiz.*, **17**, 469–486.
- VESSMAN J., STEFAN R., VAN STADEN J., DANZER K., LINDNER W., BURNS D., FAJGELI A. & MULLER H. (2001). Selectivity in analytical chemistry. *Pure Appl. Chem.*, **73** (8), 1381–1386.
- WORLD ORGANISATION FOR ANIMAL HEALTH (OIE) (2008). OIE Standard for Management and Technical Requirements for Laboratories Conducting Tests for Infectious Diseases. *in: OIE Quality Standard and Guidelines for Veterinary Laboratories: Infectious Diseases*. OIE, Paris, France, 1–31.
- ZWIG M.H. & CAMPBELL G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.*, **38**, 561–577.



Anexo K. Enfoques estadístico de la validación

NB: Versión adoptada por la Asamblea Mundial de Delegados de la OIE en mayo de 2014

CAPÍTULO 3.6.5

ENFOQUES ESTADÍSTICOS DE LA VALIDACIÓN

INTRODUCCIÓN

Las Recomendaciones de la OIE para la Validación aportan información detallada y ejemplos que respaldan la norma de validación de la OIE publicada como Capítulo 1.1.6 del Manual Terrestre, o como Capítulo 1.1.2 del Manual Acuático. El Término "Norma de Validación de la OIE" de este Capítulo hace referencia a estos capítulos.

La elección de los métodos estadísticos para el análisis de los datos de validación de las pruebas obtenidos mediante estudios de laboratorio y de la evaluación de muestras de campo se basa en aspectos como el diseño experimental y la elección de las muestras (procedente, número de muestras, número de réplicas de las pruebas, etc.). Deben establecerse normas específicas en cuanto al "mejor enfoque" consultando a un estadístico, lo cual debe llevarse a cabo durante la fase de diseño, antes de que empiecen los estudios de validación.

Para abreviar, en este anexo se tienen en cuenta solo los enfoques de uso habitual para la validación de una prueba candidata, y por ello no incluyen todos los métodos estadísticos que podrían utilizarse en la práctica.

Se describen métodos que sirven para estimar la precisión de una prueba cuando se repite múltiples veces (repetibilidad y reproducibilidad), las características analíticas (sensibilidad y especificidad analíticas) y las características de diagnóstico (como la sensibilidad [DSe] y la especificidad diagnósticas [DSp], y el área bajo la curva de

Definiciones de las escalas de medición:

Binaria (dicotómica): Positivo o negativo, porque así es como se presenta el resultado, o positivo/negativo en un valor umbral (de corte) determinado cuando los resultados se miden en una escala ordinaria o continua.

Ordinal: Se mide en una escala con valores discretos en que valores más altos suelen indicar más análisis, por ejemplo, títulos de neutralización vírica en el suero.

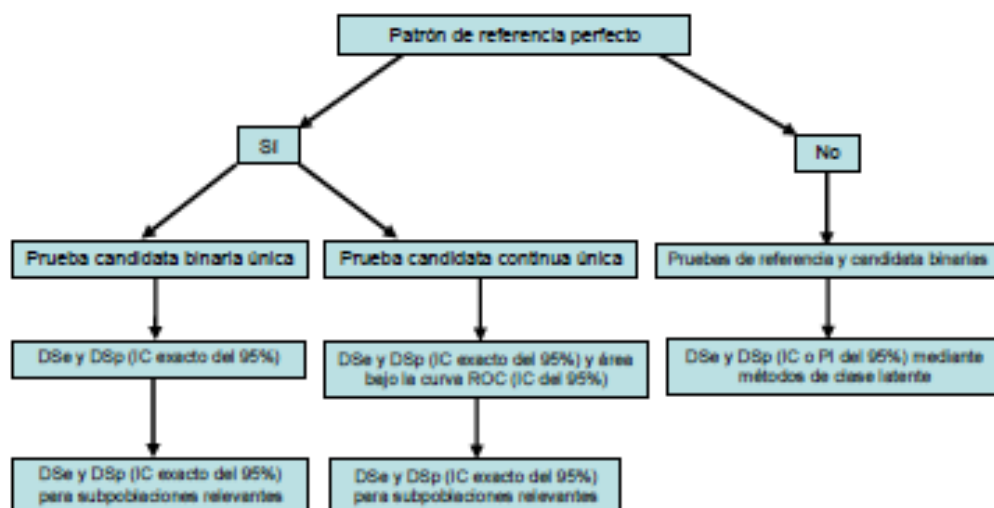
Continua: Teóricamente es posible un número infinito de valores de medición, en función del sistema de medición, por ejemplo, la densidad óptica o el porcentaje de valores positivos en una prueba de enzimoimmunoanálisis, o los valores de ciclo umbral en la reacción en cadena de la polimerasa en tiempo real.

la característica operativa del receptor de una prueba) que se utilizan para detectar un análisis en animales sueltos. Unos principios similares son aplicables cuando las pruebas se utilizan para detectar el mismo análisis en combinaciones de muestras generadas natural o artificialmente y procedentes de animales que viven en grupos (como rebaños o parvadas). En este caso, la unidad epidemiológica es el agregado y no animales sueltos.

Los métodos estadísticos difieren en función de si se evalúa una o varias pruebas, de sus escalas de medición (binaria, ordinal o continua), de si se utilizan muestras independientes o dependientes (pareadas) y de si hay un patrón de referencia perfecto para fines comparativos (Wilks, 2001). En las Figuras 1 y 2 se muestran diagramas de flujo que orientan acerca de la elección de los métodos estadísticos aplicables a la evaluación de las medidas de exactitud diagnóstica, como la sensibilidad y la especificidad.

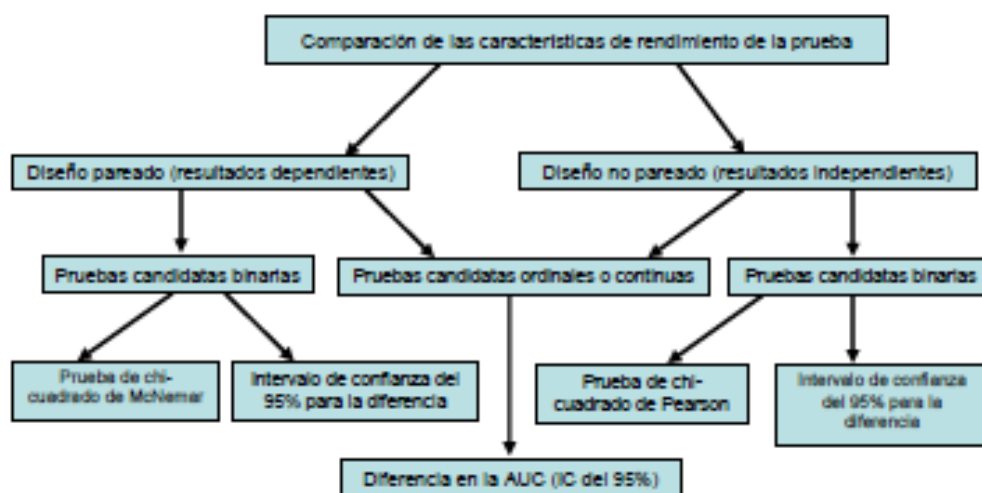
La idoneidad del diseño del estudio y del análisis estadístico es posible que no siempre se refleje en la calidad de la comunicación en las publicaciones científicas, de ahí que se aliente a los responsables del desarrollo y evaluación a seguir la lista de control de las normas para la publicación de informes sobre la exactitud diagnóstica (STARD, por las siglas de la expresión en inglés Standards for Reporting of Diagnostic Accuracy).

Para informarse acerca del análisis de los datos de Incertidumbre de la medición y para obtener datos relativos a los estudios de comparación de métodos, consúltese las Capítulos 3.6.4 y 3.6.8 de la OIE, respectivamente.



Abreviaciones: DSe = sensibilidad diagnóstica; DSp = especificidad diagnóstica; ROC = característica operativa del receptor; CI = intervalo de confianza; PI = intervalo de probabilidad.

Figura 1. Diagrama de flujo para métodos sugeridos de análisis estadístico cuando se evalúa una sola prueba candidata con y sin un patrón de referencia perfecto.



Abreviación: AUC = área bajo la curva de la característica operativa del receptor

Figura 2. Diagrama de flujo para métodos sugeridos de análisis estadístico cuando se evalúan las sensibilidades y especificidades, así como el AUC de múltiples pruebas con un patrón de referencia perfecto. Deben analizarse datos ordinales o continuos en su forma original y como resultados binarios respecto a los umbrales recomendados. Deben realizarse análisis tanto para la sensibilidad como para la especificidad cuando se disponga de estos datos.

1. Repetibilidad de la prueba en un mismo laboratorio

Para evaluar la repetibilidad intra-laboratorio de una prueba (a menudo denominada *precisión* cuando la medición se realiza en una escala continua), se analizan por triplicado al menos tres muestras que tengan concentraciones del analito situadas en el intervalo de funcionamiento de la prueba, y dicho análisis debe llevarlo a cabo un solo operario empleando un solo kit o lote de análisis. Habitualmente, estas ejecuciones se realizarán el mismo día, pero también es posible hacerlo en días distintos. Es aconsejable utilizar tres o cuatro réplicas de una muestra en lugar de dos porque se captura mejor la variabilidad inherente en los resultados analíticos de una misma ejecución. Sin embargo, por motivos de coste es posible que no en todos los tipos de prueba sea factible utilizar más de dos réplicas (por ejemplo, en la detección de ácido nucleico). Como se describe en la Norma de Validación de la OIE, la variación entre ejecuciones se puede evaluar realizando múltiples ejecuciones con la intervención de dos o más operarios a lo largo de varios días.

1.1. Resultados continuos

En el caso de los resultados dicotómicos, el enfoque más sencillo es estimar la desviación estándar (SD) de las réplicas de un grupo de muestras representativas del intervalo de funcionamiento de la prueba. Estos resultados inicialmente deben evaluarse en un diagrama de dispersión de la media de las réplicas representadas respecto a la SD. En el caso de las pruebas en que la SD es proporcional a la media, a menudo se calcula el coeficiente de variación (CV) intra-muestra. El CV a menudo se utiliza incluso cuando no existe proporcionalidad. En este caso, debe indicarse el CV de los niveles del analito en cuestión (por ejemplo, bajo, moderado y alto). Esto es necesario porque es habitual observar que el CV a menudo es mayor cuando la concentración del analito en cuestión es baja. En general, también debe calcularse una estimación de la incertidumbre de los valores del CV (por ejemplo, un intervalo de confianza [IC] del 95%). Cuando los valores del CV son bastante constantes a lo largo de todo el intervalo de valores de la prueba, esto se puede hacer utilizando los resultados de todas las muestras. Cuando el CV difiere en función de la concentración de analito, deben calcularse IC del 95% independientes para cada categoría de analito en función del número de muestras analizadas a cada nivel.

CV =	SD de las réplicas Media de las réplicas
donde:	CV = Coeficiente de variación SD = Desviación estándar

1.2. Resultados binarios

En general, cuando se dispone de resultados cuantitativos, son estos los que deben utilizarse para evaluar la precisión de la prueba, aunque puedan dicotomizarse para facilitar su comunicación. En el caso de las pruebas inherentemente binarias, que dan resultados positivos o negativos, puede utilizarse el parámetro estadístico kappa para cuantificar la coincidencia más allá del azar. El índice kappa oscila entre 0 (nada de coincidencia más allá del azar) y 1 (perfecta coincidencia más allá del azar) pero existen muchas más conjeturas acerca de cómo deben interpretarse los valores de kappa (Fleiss et al., 2003; Landis & Koch 1977). Se suele esperar una coincidencia mejor cuando los resultados de la prueba están muy alejados de los puntos de corte y, por lo tanto, deben evaluarse algunas muestras con valores intermedios/sospechosos para evitar determinaciones de la coincidencia excesivamente optimistas. Puede utilizarse una versión de kappa ponderada para resultados ordinales (por ejemplo, negativos, sospechosos y positivos) para reconocer que una discrepancia grande (por ejemplo, una diferencia de dos categorías) es más importante que una discrepancia más pequeña (por ejemplo, una discrepancia de una categoría). En las estimaciones de kappa, tanto ponderadas como no ponderadas, debe establecerse un IC del 95% (Fleiss et al., 2003).

Tabla 1. Ejemplos de los cálculos de Kappa para resultados binarios

Ejemplo 1: Cálculo de kappa basado en resultados de la prueba duplicados clasificados como positivos o negativos

Resultado de la prueba	Positivo	Negativo
Positivo	90	5
Negativo	10	95
	100	100

$$\text{Kappa} = 0,85 \text{ (IC del 95\% = 0,78 a 0,92)}$$

Ejemplo 2: Cálculo de kappa basado en resultados de la prueba duplicados clasificados en tres categorías (positivo, sospechoso o negativo)

Resultado de la prueba	Positivo	Sospechoso	Negativo
Positivo	80	10	10
Sospechoso	15	75	10
Negativo	5	15	80
	100	100	100

Kappa = 0,88 (IC del 95% = 0,81 a 0,75). Kappa ponderada = 0,70. (IC del 95% = 0,81 a 0,70)

2. Reproducibilidad de la prueba entre laboratorios

La precisión de la prueba variará en función de la implementación rutinaria, por ejemplo, según el operario, los lugares donde se realice, los lotes de kits de diagnóstico que se utilicen o el hecho de que se realice a lo largo de varios días. Lo más habitual es aplicar el término reproducibilidad a la evaluación de la precisión de la prueba en cuestión en varios laboratorios. Debe describirse qué factores se mantienen constantes, para interpretar los resultados en el contexto de la verdadera situación en que se desarrolla la prueba. Pueden realizarse estudios de reproducibilidad independientemente de los de repetibilidad o bien junto a estos, pero siempre con un diseño ciego. Como sugiere la Norma de Validación de la OIE, un mínimo de tres laboratorios deberán analizar al menos 20 muestras, de las cuales se enviarán idénticas alícuotas a cada laboratorio.

Los métodos estadísticos para el análisis de estudios de la reproducibilidad de la prueba entre laboratorios son similares a los que se utilizan para evaluar la repetibilidad dentro de un mismo laboratorio. No obstante, como parte de un estudio entre laboratorios, podría considerarse importante evaluar y ordenar la variabilidad en los resultados de la prueba derivada de múltiples fuentes (a menudo denominada clase). Por ejemplo, si se diseñara un estudio para evaluar una prueba en tres laboratorios, en cada uno de los cuales intervinieran dos técnicos altamente especializados y se analizaran las muestras por duplicado con dos lotes de kit, cada muestra se analizaría 24 veces. Los factores escogidos (laboratorio, técnico, lote de kit y réplica) pueden considerarse fijos o aleatorios en función de cómo se hayan escogido y de si son representativos de la población de destino. En este tipo de diseño del estudio, pueden estimarse los componentes de la varianza para cada clase (se ofrece un ejemplo en Dargatz et al., 2004), y el coeficiente de correlación intraclass (ICC) puede estimarse como el grado de similitud entre los resultados obtenidos con la muestra (Bartlett & Frost, 2008).

El coeficiente de correlación intra-clase constituye la similitud o correlación entre cualquier par de mediciones realizadas en la misma muestra. El ICC va de 0 a 1, y los valores cercanos a 1 indican un error muy bajo en la medición. Por el contrario, los valores cercanos a 0 indican una gran cantidad de error en la medición.

2.1. Cuando se realiza una modificación técnica en el método analítico

Una vez una técnica se ha validado para ser utilizada en un entorno de laboratorio controlado, puede plantearse su uso en un entorno muy distinto (como una aplicación a pie de explotación). Debido a que en estos casos se producen cambios más extremos, como las fluctuaciones drásticas de la temperatura que a menudo tienen lugar a pie de explotación, sería esperable que esta nueva aplicación de la prueba se comportara de un modo muy distinto a la original. De hecho, en lugar del error accidental de la medición a la medición en un mismo o entre distintos laboratorios, se prevé que el error de la medición de los valores de tal estudio probablemente se consideraría sistemático, lo cual sería aplicable si los valores proporcionan una sobreestimación o subestimación del valor real. En el caso del ejemplo de una prueba que se ejecuta con muestras divididas para poder ser analizadas tanto a pie de explotación como en un laboratorio, la media de las diferencias entre el valor a pie de explotación y el valor obtenido dentro de un mismo laboratorio (valor real) para una misma muestra debe indicarse con un IC del 95%. Si el IC del 95% excluye el cero, hay indicios de desviación sistemática de los resultados de la prueba cuando se utiliza a pie de explotación en comparación con lo que ocurre en el laboratorio. Cuando se produce tal desviación sistemática en los resultados de la prueba, los resultados de la prueba a pie de explotación no son comparables con los de la prueba validada para ser utilizada en el laboratorio. Para validar la prueba para su uso a pie de explotación, se somete a una "modificación técnica" que a continuación se evalúa en un estudio de comparación de métodos (véase el Capítulo 3.6.8) o bien se requiere una re-validación completa de la aplicación a pie de explotación.

Pueden utilizarse enfoques similares para evaluar cambios en el método dentro de un laboratorio, con el fin de determinar si la variación de los resultados es sistemática o accidental.

Ejemplo: los siguientes datos no publicados se obtuvieron comparando dos métodos de extracción (el antiguo y el nuevo, aplicados a las mismas muestras, divididas) a valores del ciclo umbral (Ct) para

una reacción en cadena de la polimerasa cuantitativa en tiempo real (qPCR) destinada a detectar la lengua azul. Los datos ($n=10$) representan las medias de los duplicados de la muestra.

Método antiguo: 25,6, 24,5, 21,3, 26,8, 25,2, 30,2, 31,2, 32,8, 31,8, 34,9

Método nuevo: 23,1, 21,0, 18,2, 25,2, 24,7, 28,6, 30,4, 32,2, 31,3, 34,7

La diferencia de las medias de ambos métodos (la del antiguo menos la del nuevo) fue de $-1,49$ (IC del 95% = $-2,33$ a $-0,64$) con una probabilidad bilateral de $p=0,003$. Ello indica un valor de Ct sistemáticamente inferior al utilizar el método nuevo de extracción. Puede emplearse un gráfico de Bland-Altman (Bland & Altman, 1999; Fig. 3) para observar gráficamente cómo la diferencia cambia en función del valor medio de los métodos antiguo y nuevo. En nuestro ejemplo, la diferencia parece disminuir a medida que aumentan los valores de Ct, pero el tamaño muestral es pequeño.

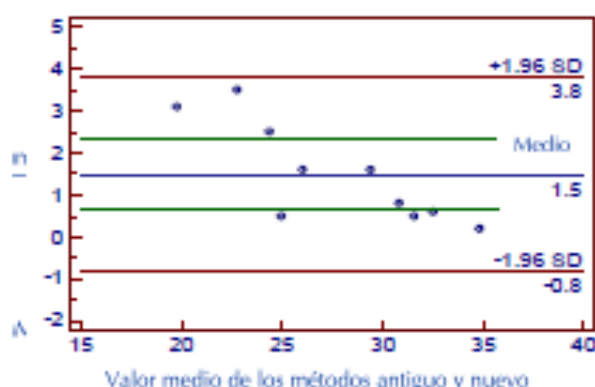


Figura 3. Gráfico de Bland-Altman de la diferencia media (eje de las y) en los valores de Ct como función del valor medio de los métodos antiguo y nuevo ($n=10$)

3. Sensibilidad analítica (ASe, sinónimo = límite de detección: LOD)

La sensibilidad analítica se puede estimar empleando un estudio de dilución hasta la extinción (DTE) en el que se generan diluciones seriadas de una cantidad cuantificada y conocida del analito en cuestión en la matriz de muestra apropiada. Esta cantidad cuantificada y conocida podría ser de un patrón interno o nacional/internacional de referencia o de una muestra de campo cuya concentración de analito se ha determinado. Pueden llevarse a cabo ejecuciones paralelas de un patrón de comparación pero no son fundamentales, a no ser que en el estudio se esté comparando un cambio pequeño en una prueba validada respecto a la prueba validada original. La estrategia del DTE puede utilizarse si el analito se mide cualitativa o cuantitativamente. En el segundo caso, el resultado de la prueba se vuelve a clasificar como positivo o negativo.

El enfoque del análisis de los datos de LOD depende del diseño experimental. Por ejemplo, supongamos que se realiza un estudio en el cual se añaden 10^8 unidades formadoras de colonia (UFC) de una bacteria a 10 g de heces para lograr una concentración de 10^7 UFC/g. A continuación, esta muestra se diluye en diluciones seriadas decimales hasta obtener 10^1 UFC/g. El estudio se repite tres veces. Si se detectara el analito en todas las réplicas de 10^3 UFC/g pero no se detectara en ninguna de las de 10^2 UFC/g, una estimación conservadora del LOD podría ser 10^3 UFC/g. Si se necesitara una estimación precisa del LOD, podría diseñarse un segundo estudio para determinar el LOD con mayor certidumbre empleando una serie de diluciones más finas, por ejemplo, a la mitad, que englobaran el intervalo entre el 100% de detección y el 0% de detección identificado en el primer estudio. El punto final del LOD a menudo se escoge que sea del 95%; en un estudio con 20 réplicas, esto corresponde a la dilución en que 19 réplicas daban positivo al analito. El aspecto importante es que la probabilidad escogida para el LOD (tanto si es del 95% como del 50% u otros valores) debe especificarse y utilizarse de forma constante para poder comparar los resultados de múltiples pruebas. El LOD puede estimarse utilizando empleando el enfoque no paramétrico de Spearman-Kärber, o mediante regresión logística o análisis probit. Cuanto mayor sea el número de réplicas de cada dilución, más precisa será la estimación del LOD.

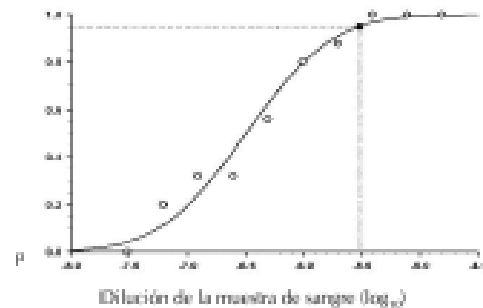


Figura 4. Límite estimado del 95% de detección del virus de la peste equina Africana en sangre de caballo (\log_{10}), que se representa mediante la línea de puntos

Ejemplo: Guthrie et al. (2013) generaron una serie de diluciones a la mitad con una sangre de caballo positiva para el virus de la peste equina africana (VPEA) (dilución a 10^{-2}), que cubrió el intervalo no lineal de la prueba. La extracción se repitió 25 veces y las muestras se analizaron mediante qPCR en tiempo real para el VPEA. Los resultados de la qPCR de los 15 niveles de dilución se utilizaron en un análisis probit para calcular el LOD al 95% (es decir, la concentración que daba un resultado positivo en la qPCR en tiempo real en un 95% de las réplicas (Burns & Valdivia, 2008)). El LOD al 95% se estimó que se encontraba en una dilución de $3,02 \times 10^{-8}$, como se muestra en la Figura 4, y que correspondía a un ciclo de cuantificación de 35,71 en la qPCR. No se indicó el IC para la estimación.

4. Especificidad analítica (ASp)

La especificidad analítica es la capacidad de la prueba de distinguir el analito en cuestión (por ejemplo, un anticuerpo, un microorganismo o una secuencia genómica) de analitos no buscados que podrían estar estrechamente relacionados con aquel (como se describe en la Norma de Validación de la OIE). Dado que la elección de los analitos relacionados (por ejemplo, microorganismos) es subjetiva y a menudo depende de los tipos y cantidades de muestras de los que se dispone, el resultado de la evaluación de la ASp debe indicarse cualitativamente, por ejemplo, como el porcentaje de agentes relacionados que generan reacción cruzada en la prueba, junto a un listado de los posibles agentes que podía generar reacción cruzada y que se evaluaron.

5. Rendimiento diagnóstico de la prueba

El rendimiento diagnóstico de una prueba se suele medir según la sensibilidad (DSe) o la especificidad (DSp) o combinación de la DSe y la DSp, como el porcentaje de probabilidad de obtener un resultado positivo o negativo. También pueden calcularse porcentajes de probabilidad para intervalos de resultados analíticos cuando es importante conservar la información sobre la magnitud del resultado analítico en lugar de utilizarlo en forma dicotomizada. Para más información sobre el uso y el cálculo de los porcentajes de probabilidad, consúltese Gardner & Greiner (2006) y Gardner et al. (2010). Este segundo artículo incluye un ejemplo relativo a la toxoplasmosis porcina con cálculos del IC mediante dos métodos.

La DSe y la DSp pueden estimarse cuando el método de referencia o comparativo es perfectamente sensible y específico o cuando el patrón de referencia es imperfecto. En general, la mayoría de patrones de referencia ante-mortem de uso habitual en los laboratorios de diagnóstico son imperfectos, motivo por el cual, si se va a considerar que los resultados del patrón de referencia son los verdaderos, a menudo es necesaria una necropsia con análisis de múltiples tejidos mediante pruebas auxiliares, como el cultivo y/o la histopatología. En la mayoría de estudios de validación de pruebas para enfermedades animales, esta última opción no es viable ni rentable excepto para un pequeño número de muestras.

La incertidumbre estadística relativa a los parámetros de rendimiento diagnóstico, como la DSe y la DSp, debe presentarse como intervalos de confianza (IC). Habitualmente, se utiliza un IC del 95% y su amplitud (precisión del valor estimado) depende en gran medida del tamaño de muestra utilizado para estimar el parámetro. Los IC exactos son preferibles a las aproximaciones aritméticas porque evitan límites superiores que superen el 100%.

5.1. DSe y DSp con un patrón de referencia perfecto

La prueba candidata puede dar resultados en escalas binarias (dicotómicas), ordinales (como un título) o continuas. En el caso de estas dos últimas escalas, para poder calcular la DSe y la DSp los resultados tienen que dicotimizarse, es decir debe

establecerse un valor de corte (umbral). Se recomiendan IC binomiales exactos del 95% para la DSe y la DSp (Greiner & Gardner, 2000) porque la aproximación normal podría no dar un IC apropiado cuando las estimaciones del parámetro se acercan a 1.

Ejemplo: enzoinmunoanálisis indirecto (HELISA)

		Número de animales			
		Se sabe que son positivos a anticuerpos (369)	Se sabe que son negativos a anticuerpos (198)		
Resultados de la prueba	Positivo	287	1		
	Negativo	82	197		
		TP	FP		
		FN	TN		
		Sensibilidad diagnóstica*		Especificidad diagnóstica*	
		TP/(TP + FN)		TN/(TN + FP)	
		77,8% (73,2 – 81,9%)*		88,6% (87,2 – 89,9%)*	

TP y FP = verdadero positivo y falso positivo, respectivamente

TN y FN = verdadero negativo y falso negativo, respectivamente

*límites de confianza binomiales exactos del 95% para la DSe y la DSp

Cuando el patrón de referencia no se aplica a todos los resultados de la prueba positivos y negativos (verificación parcial), deben realizarse estimaciones corregidas de la DSe y la DSp como se describe en Greiner & Gardner (2000) para justificar las distintas probabilidades de muestreo en los grupos positivos y negativos.

En el caso de las pruebas que dan resultados ordinales (como valores de título) o continuos (por ejemplo, la proporción del resultado de una muestra problema respecto al de la muestra control positiva en un ELISA), las estimaciones de la DSe y la DSp se pueden complementar con estimaciones del área bajo la curva de la característica operativa del receptor (ROC). El análisis de la ROC constituye un enfoque independiente del punto de corte para evaluar la exactitud global de una prueba cuyos resultados se miden como valores ordinales o continuos. El área bajo la curva ROC proporciona una única estimación numérica de la exactitud global que va de 0,5 (prueba inútil) a 1 (prueba perfecta). La principal justificación para un análisis de la ROC es que los valores de corte para interpretar la prueba pueden cambiar en función del propósito del análisis (por ejemplo, que la prueba pase de ser de cribado a ser confirmativa) y de la prevalencia de la infección, los costes de los errores analíticos y la disponibilidad de otras pruebas. Se presentan descripciones detalladas del análisis de la ROC en otras fuentes (Gardner & Greiner, 2006; Greiner et al., 2000; Zweig & Campbell, 1993). Cuando se comparan múltiples pruebas ordinales o continuas, debe calcularse la diferencia en el área bajo la curva con un IC del 95%. Los métodos para calcular estas diferencias varían en función de si las muestras son independientes o dependientes y se implementan en muchos programas estadísticos (Gardner & Greiner, 2006). En la Figuras 5 y 6 se ofrecen ejemplos de un diagrama de puntos con los resultados de un solo ELISA, y curvas ROC para dos ELISA.

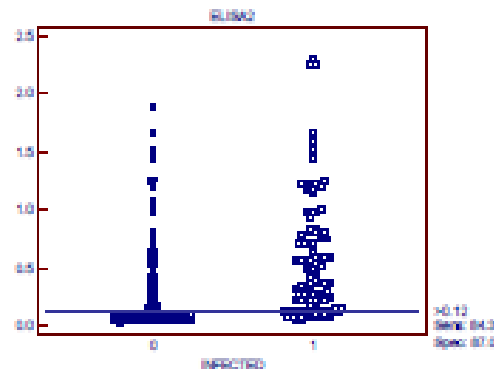


Figura 5. Diagrama de puntos de los resultados de ELISA en animales no infectados (Código = 0) e infectados (Código = 1).

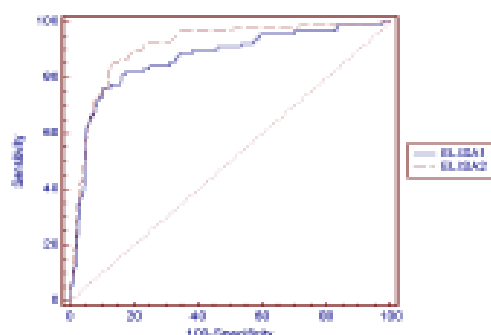


Figura 6. Curva de la característica operativa del receptor para dos ELISA.

5.2. Comparación de las estimaciones de DSe y DSp para dos pruebas con un patrón de referencia perfecto

A menudo, el investigador puede desear comparar los valores de DSe de distintas subpoblaciones de animales infectados, por ejemplo, clínicamente frente a subclínicamente infectados, o valores de DSp de distintas áreas geográficas. Dado que son muestras distintas, las comparaciones pueden realizarse estadísticamente mediante la χ^2 de Pearson, para asegurar la homogeneidad. Como alternativa, pueden calcularse un IC del 95% independiente y un IC del 95% para una diferencia en dos proporciones. Cuando la DSe (o la DSp) de dos pruebas se comparan con el mismo conjunto de muestras de animales infectados (o no infectados) en un diseño pareado, los resultados de la prueba dejan de ser independientes. Pueden utilizarse métodos estadísticos como la χ^2 de McNemar para analizar la hipótesis de sensibilidades equivalentes (especificidades) cuando la prueba se aplica a las mismas muestras.

Ejemplo: Se evaluaron cinco pruebas de detección de anticuerpos para el diagnóstico de la paratuberculosis bovina en vacas lecheras, y se trabajó en rebaños que se sabía que estaban infectados y otros que se sabía que no lo estaban, según los resultados de cultivos fecales y los antecedentes de los rebaños. Se generaron las siguientes tablas de datos en base a los datos originales antes de la posterior publicación en Collins et al. (2005). En la publicación, se eliminó un rebaño del análisis. El ejemplo se utiliza para mostrar la distribución tabular para el cálculo de la DSe y la DSp y la evaluación estadística.

		Infectado				No infectado			
		T_2^+	T_2^-			T_2^+	T_2^-		
T_1^+		124	74	198		3	27	30	
T_1^-		8	243	251		16	366	382	
		132	317	449		19	393	412	
		Sensibilidad de T1 = 198/449 = 44,1%				Especificidad de T1 = 382/412 = 92,7%			
		Sensibilidad de T2 = 132/449 = 29,4%				Especificidad de T2 = 393/412 = 95,4%			

Las sensibilidades difirieron significativamente ($p < 0,0001$), pero las especificidades no ($p = 0,126$), según una prueba de χ^2 de McNemar bilateral. También pueden calcularse las covarianzas de la sensibilidad y la especificidad (véanse los detalles en Gardner et al., 2000) para indicar si las pruebas son condicionalmente independientes o dependientes, dado el estado del rebaño respecto a la infección. En nuestro ejemplo, la covarianza de la sensibilidad (calculada mediante la tabla de la izquierda) fue de 0,147 ($p < 0,0001$ mediante la χ^2 de Pearson), lo cual indica una fuerte dependencia de las dos pruebas cuando se utilizan en animales infectados. La covarianza de la especificidad (calculada mediante la tabla para rebaños no infectados, a la derecha) fue de 0,004 ($p = 0,152$ mediante la χ^2 de Pearson), lo cual indica que no hay una dependencia significativa.

Gardner et al. (2010) presentan otro ejemplo basado en datos sobre la toxoplasmosis porcina.

5.3. DSe y DSp sin un patrón de referencia perfecto

Los avances en la metodología estadística, específicamente el desarrollo de modelos de clase latentes (a veces denominados "métodos no de referencia", permite a los investigadores liberarse de la restrictiva suposición de una prueba de referencia perfecta y estimar la exactitud de las pruebas candidatas y del patrón de referencia con los mismos datos (Enoe *et al.*, 2000; Hul & Walter, 1980).

Los modelos de clase latente (LC), tanto si se emplean métodos de probabilidad máxima como bayesianos, pueden utilizarse para estimar la DSe y la DSp cuando se dispone de resultados analíticos conjuntos de múltiples pruebas aplicadas a animales de varias poblaciones (por ejemplo, de distintos rebaños o de distintas zonas geográficas). No todos los modelos LC de estimación de la DSe y la DSp serán estadísticamente identificables en cuanto a posibilitar la inferencia. Un modelo es identificable si teóricamente es posible determinar el valor verdadero de parámetros modelo tras obtener un número infinito de observaciones del mismo. Básicamente, ello equivale a tener un único conjunto de valores de los parámetros de interés (DSe, DSp). Los enfoques bayesianos son especialmente adecuados para situaciones en que se dispone de información previa sobre la DSe y/o la DSp y cuando el problema de estimación no es identificable (Branscum *et al.*, 2005).

El modelo LC de una sola población más sencillo que es identificable es aquel en que se ejecutan tres pruebas condicionalmente independientes con las mismas muestras. La restricción de independencia de las tres pruebas puede ser difícil de lograr en la práctica a no ser que el análisis de interés difiera entre ellas. Así pues, un enfoque que se utiliza a menudo en sanidad animal es ejecutar dos pruebas con las mismas muestras de animales de dos poblaciones, porque es menos costoso y las suposiciones de independencia condicional pueden ser más razonables. El modelo de dos poblaciones y dos pruebas también requiere suponer una sensibilidad y especificidad constantes en la totalidad de las dos poblaciones, y prevalencias claramente diferenciadas. La suposición de sensibilidad constante puede ser difícil de verificar y es improbable que sea correcta si una población tiene animales afectados clínicamente y la otra tiene animales afectados subclínicamente, porque muchos estudios publicados han mostrado que la sensibilidad de una prueba es mayor en animales afectados clínicamente. Si una de las dos poblaciones se sabe que está libre del agente patógeno (la prevalencia es cero), mientras que la otra población se sabe que tiene una prevalencia distinta de cero, puede utilizarse la primera población para estimar la DSp, y ello facilitará la estimación de la DSe en la población infectada.

Las enfermedades de la lista de la OIE en que se ha estimado la DSe y la DSp con métodos bayesianos son la brucelosis ovina (Fraud *et al.*, 2012), la fiebre Q (Paul *et al.*, 2013), la tripanosomosis (Bronsvort *et al.*, 2010), la tuberculosis bovina (Clegg *et al.*, 2011), la fiebre aftosa (Bronsvort *et al.*, 2006), la peste equina africana (Guthrie *et al.*, 2011) Y el virus de la anemia infecciosa del salmón (Caraguel *et al.*, 2013).

El software WinBUGS¹ permite interpretar fácilmente los métodos de Monte Carlo vía cadena de Markov para la estimación bayesiana (Lunn *et al.*, 2000) y pueden realizarse análisis simples de probabilidad máxima empleando una interfaz basada en la red (Pouillot *et al.*, 2002). La información previa sobre los parámetros modelo que se emplee en los análisis bayesianos puede afectar a las estimaciones finales en función de la potencia relativa de las evidencias proporcionadas por los valores precedentes *priors* (nivel de incertidumbre previa) y por los datos (incertidumbre atribuible a tamaños de muestra finitos). Por lo tanto, las fuentes de información previa deben estar bien documentadas en análisis bayesianos y puede ser deseable repetir el análisis utilizando *priors* no informativos sobre todos los parámetros cuando el modelo sea identificable.

Es importante destacar que el análisis LC no puede corregir sesgos inherentes en estudios mal diseñados. Los métodos deben utilizarse cuidadosamente e incluir una evaluación exhaustiva de las suposiciones subyacentes (por ejemplo, dependencia condicional, sensibilidad y especificidad constante entre poblaciones, y prevalencias claramente diferenciadas), los efectos de utilizar las distribuciones previas seleccionadas sobre las inferencias posteriores, como se ha descrito en el párrafo previo, y la convergencia de las cadenas de Markov en un análisis bayesiano (Toft *et al.*, 2005).

Probabilidad máxima – algoritmo estadístico que estima los valores más probables para los parámetros de interés en base al valor de la función de probabilidad para los datos.
Métodos bayesianos – incorporan información o conocimientos previos relevantes sobre una o más pruebas además de la función de probabilidad para los datos. Con tamaños muestrales grandes, la probabilidad máxima y los métodos bayesianos dan inferencias similares.

1 Disponible en <http://www.mrc-biu.cam.ac.uk/bugs/winbugs/content.html>

Ejemplo: Guthrie et al. (2013) estimaron la DSe y la DSp de una PCR en tiempo real cuantitativa y del aislamiento vírico (VI) convencional para la detección del virus de la peste equina africana (PEA) en muestras de sangre total empleando un modelo de clase latente bayesiano para dos pruebas y dos poblaciones. Se analizaron mediante PCR y VI dos poblaciones de caballos purasangre de Sudáfrica (503 casos sospechosos de PEA, y 503 caballos sanos procedentes de la zona en que el virus de la PEA está controlado). En los 503 casos sospechosos, los resultados analíticos conjuntos fueron los siguientes: PCR+VI+ (n=156), PCR+VI- (n=184), PCR-VI+ (n=0), y PCR-VI- (n=163). Los 503 caballos sanos fueron PCR-VI-. Distintos modelos (Independencia condicional y dependencia condicional) se ajustaron para los datos y en algunos análisis también se incluyó una segunda población de caballos sanos.

Los modelos se procesaron en el programa WinBUGS 1.4.3 (Lunn et al., 2000), desechando las primeras 5.000 iteraciones, y las siguientes 50.000 iteraciones se utilizaron para inferencias posteriores (medianas e intervalos de probabilidad del 95% para la DSe y la DSp). Se evaluó la convergencia del modelo mediante inspección visual de los gráficos de trazo de valores iterados y ejecutando múltiples cadenas a partir de los valores iniciales dispersados. El modelo de independencia condicional ajustado con priors no informativos eta (1,1) sobre la DSe y la DSp de ambas pruebas dio resultados casi idénticos a los del modelo en que se utilizó un prior beta altamente informativo (9999,1) para la DSp de la VI. Las medias estimadas y los intervalos de probabilidad del 95% (a veces denominados intervalos creíbles), entre paréntesis, obtenidos con el modelo de independencia condicional con priors no informativos fueron los siguientes:

Sensibilidad de la PCR = 0,996 (0,977–0,999)

Especificidad de la PCR = 0,999 (0,993–1,0)

Sensibilidad de la VI = 0,458 (0,404–0,51)

Especificidad de la VI = 0,999 (0,998–1,0)

Los resultados indicaron una DSe dos veces más alta en la PCR que en la VI, y DSp comparables en ambas pruebas. En Guthrie et al. (2013) se ofrece una descripción completa del enfoque de modelado.

5.4. Comparación de las estimaciones de la IDSe y la IDSp de dos pruebas sin un patrón de referencia perfecto

Si se utiliza un enfoque bayesiano en el programa WinBUGS para analizar los datos analíticos conjuntos de múltiples poblaciones, la diferencia de sensibilidades (especificidades) se puede estimar fácilmente y la probabilidad de que la sensibilidad (especificidad) de una prueba supere a la de la otra puede estimarse con la función STEP.

Ejemplo: en el caso de los resultados de los datos de Guthrie et al. (2013) de la Sección 5.3, los intervalos de probabilidad (PI) del 95% para la DSe no se solaparon pero hubo un considerable solapamiento en el PI del 95% para la DSp. Los valores de probabilidad correspondientes obtenidos mediante la función STEP fueron de 1 y 0,24, respectivamente. Estos valores aportan la certeza de que las DSe difieren, pero la probabilidad de que las DSp difieran es pequeña (inferior a 0,5).

BIBLIOGRAFÍA

- BARTLETT J.W. & FROST C. (2008). Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstet. Gynecol.*, 31, 466–475.
- BLAND J.M. & ALTMAN D.G. (1999). Measuring agreement in method comparison studies. *Statist. Methods Med. Res.*, 8, 135–160.
- BOSSUYT P.M., RITSMA J.B., BRUNS D.E., GATSONIS C.A., GLASZIOU P.P., IRWIG L.M., LIMER J.G., MOHER D., RENNIE D. & H.C.M. de VET (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Clin. Chem.*, 48, 1–6.
- BRANSCUM A.J., GARDNER L.A. & JOHNSON W.O. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Prev. Vet. Med.*, 68, 145–163.

- BRONSVOORT B.M., TOFT N., BERGMANN I.E., SØRENSEN K.J., ANDERSON J., MALIRAT V., TANJA V.N., MORGAN K.L. (2006) Evaluation of three 3ABC ELISAs for foot-and-mouth disease non-structural antibodies using latent class analysis. *BMC Vet. Res.*, 2, 30.
- BRONSVOORT B.M., VON WISSMANN B., FÉVRE E.M., HANDEL I.G., PICOZZI K., & WELBURN S.C. (2010) No gold standard estimation of the sensitivity and specificity of two molecular diagnostic protocols for *Typanosoma brucei* spp. in Western Kenya. *PLoS One*; 5 (1), e8628.
- BURNS M & VALDIVIA H. (2008). Modelling the limit of detection in real-time quantitative PCR. *Eur. Food Res. Technol.*, 228, 1513–1524.
- CARAGUEL C., STRYHN H., GAGNÉ N., DOHO O. & HAMMILL L. (2012). Use of a third class in latent class modelling for the diagnostic evaluation of five infectious salmon anaemia virus detection tests. *Prev. Vet. Med.*, 104, 165–173.
- CLEGG T.A., DUGGAN A., WHELAN G., GORMLEY E., GOOD M., CLARKE J., TOFT N. & MORE S.J. (2011). Using latent class analysis to estimate the test characteristics of the γ -interferon test, the single intradermal comparative tuberculin test and a multiplex immunoassay under Irish conditions. *Vet. Microbiol.*, 161, 68–76.
- COLLINS M.T., WELLS S.J., PETRINI K.R., COLLINS J.E., SCHULTZ R.D., & WHITLOCK R.H. (2005). Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. *Clin. Diag. Lab. Immunol.*, 12, 685-692.
- DARGATZ D.A., BYRUM B.A., COLLINS M.T., GOYAL S.M., HIETALA S.K., JACOBSON R.H., KOPRAL C.A., MARTIN B.M., MCCLOSKEY B.J. & TEWARI D. (2004). A multilaboratory evaluation of a commercial enzyme-linked immunosorbent assay test for the detection of antibodies against *Mycobacterium avium* subsp. *paratuberculosis* in cattle. *J. Vet. Diagn. Invest.*, 18 (5), 509–514.
- ENGE C., GEORGIADIS M.P. & JOHNSON W.O. (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev. Vet. Med.*, 45, 61–81.
- FLEISS J.L., LEVIN B. & PAIK M.C. (2003). *Statistical Methods for Rates and Proportions*, Third Edition. John Wiley & Sons, New York, USA.
- GARDNER I.A., STRYHN H., LIND P., & COLLINS M.T. (2000). Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Prev. Vet. Med.*, 46, 107–122.
- GARDNER I.A. & GRENER M. (2006). Receiver-operating characteristic curves and likelihood ratios: Improvements over traditional methods for the evaluation and application of veterinary clinical pathology tests. *Vet. Clin. Pathol.*, 35, 8–17.
- GARDNER I.A., GRENER M. & DUBEY J.P. (2010). Statistical evaluation of test accuracy studies for *Toxoplasma gondii* in food animal intermediate hosts. *Zoonoses Public Health*, 67, 82–94.
- GRENER M. & GARDNER I.A. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. *Prev. Vet. Med.*, 46, 3–22.
- GRENER M., FREIFFER D. & SMITH R.D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev. Vet. Med.*, 46, 23–41.
- GUTHRIE A.J., MACLACHLAN N.J., JOONE C., LOURENS C.W., WEYER C.T., QUAN M., MONYI M.S. & GARDNER I.A. (2013). Diagnostic accuracy of a duplex real-time reverse transcription quantitative PCR assay for detection of African horse sickness virus. *J. Virol. Methods*, 188, 30–35.
- HUI S.L. & WALTER S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics*, 38, 167–171.
- LANDIS J.R. & KOCH G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- LUNN D.J., THOMAS A., BEST N. & SPIEGELHALTER D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statist. Comp.*, 10, 325–337.
- PAUL S., TOFT N., AGERHOLM J.S., CHRISTOFFERSEN A.B. & AGER J.F. (2013). Bayesian estimation of sensitivity and specificity of *Coxiella burnetii* antibody ELISAs in bovine blood and milk. *Prev. Vet. Med.*, 108, 258–263.

PRAUD A., CHAMPION J.L., CORDE Y., DRAPEAU A., MEYER L. & GARIN-BASTUJ B. (2012) Assessment of the diagnostic sensitivity and specificity of an indirect ELISA kit for the diagnosis of *Brucella ovis* infection in rams. *BMC Vet. Res.*, **8**, 68.

POUILLOT R., GERBER G. & GARDNER I.A. (2002). "TAGS", a program for the evaluation of test accuracy in the absence of a gold standard. *Prev. Vet. Med.*, **63**, 67–81.

TOFT N., JORGENSEN E. & HOUSSGAARD S. (2005). Diagnosing diagnostic tests: evaluating the assumptions underlying the estimated of sensitivity and specificity in the absence of a gold standard. *Prev. Vet. Med.*, **68**, 19–33.

WILKS C. (2001). Gold standards as fool's gold. *Aust. Vet. J.*, **79**, 115.

ZWEIG M.H. & CAMPBELL G. (1993). Receiver-operating characteristic (ROC) plots – a fundamental evaluation tool in clinical medicine. *Clin. Chem.*, **39**, 561–577.



Anexo L. Incertidumbre de la medición

INB: Versión adoptada por la Asamblea Mundial de Delegados de la OIE en mayo de 2014

CAPÍTULO 3.6.4.

INCERTIDUMBRE DE LA MEDICIÓN

INTRODUCCIÓN

Las Recomendaciones de la OIE para la Validación aportan información detallada y ejemplos que respaldan la norma de validación de la OIE publicada como Capítulo 1.1.6 del Manual Terrestre, o como Capítulo 1.1.2 del Manual Acuático. El término 'Norma de Validación de la OIE' de este Capítulo hace referencia a estos capítulos.

La estimación de la incertidumbre de la medición (MU), que a veces se denomina imprecisión de la medición, es un requisito para los laboratorios de análisis que se basa en normas internacionales de calidad como la ISO/IEC 17025-2005, Requisitos generales de competencia para laboratorios de análisis y calibración (ISO/IEC 17026). El proceso de medición para la detección de un analito en una muestra diagnóstica no es del todo reproducible y, por lo tanto, no existe un valor exacto que pueda asociarse al analito medido. Así pues, el resultado se expresa con mayor exactitud como una estimación junto a un nivel asociado de imprecisión. Esta imprecisión es la incertidumbre de la medición (MU). MU se limita al proceso de medición. No es cuestión de si la medición es adecuada e idónea para la aplicación prevista. No es una alternativa a la validación de la prueba, sino que se considera un componente de dicho proceso (véase la Norma de Validación de la OIE, Sección B.1.1).

1. La necesidad de determinar la MU

Para asegurar el cumplimiento de los requisitos de la ISO/IEC 17025-2005, los organismos nacionales de acreditación de laboratorios de diagnóstico exigen estimaciones de la MU en los métodos analíticos que producen resultados cuantitativos, como por ejemplo, densidades ópticas (OD), porcentaje de positividad o de inhibición (PP, PI), título, valores de ciclo umbral (Ct), etc. Ello incluye las pruebas cuyos resultados numéricos se calculan y a continuación se expresan como resultado positivo o negativo respecto a un valor de corte. Cuando se pretende estimar la MU de la serología y la RT-PCR, las mediciones estadísticas adecuadas son los valores objetivo medios ± 2 desviaciones estándar (SD), lo cual es aproximadamente equivalente a un intervalo de confianza (CI) del 95%, la desviación estándar relativa (RSD = SD /media de las réplicas) y el coeficiente de variación (CV = RSD x 100%). El concepto de MU no es aplicable a resultados estrictamente binarios (positivos o negativos).

1.1. Muestras útiles para determinar la MU

La repetibilidad es el nivel de concordancia entre los resultados de réplicas de una muestra tanto para una misma ejecución como entre distintas ejecuciones del mismo método analítico en un laboratorio determinado. Durante el desarrollo de la prueba, la repetibilidad se estima evaluando la variación de los resultados de réplicas independientes de un mínimo de tres (preferiblemente cinco) muestras que representen la actividad del analito dentro del intervalo de funcionamiento de la prueba (véase la Norma de validación de la OIE, Secciones A.2.5 y B.1.1). La variación entre resultados de una misma réplica se puede expresar como RSD, CV y/o intervalos de confianza del 95%. El rasgo significativo es que pueden utilizarse estudios de la repetibilidad para definir la precisión esperada de la prueba en la detección de un intervalo de concentraciones del analito.

El uso de controles internos de la calidad o del proceso que cubran un intervalo de resultados esperables se ha convertido en parte de las operaciones de control y garantía de la calidad de las instalaciones acreditadas (véase la Norma de Validación de la OIE, Secciones A.2.6 y B.5.1, y la el Capítulo 3.6.6, Sección 1.4). Estos resultados permiten realizar un seguimiento continuo respecto de los distintos aspectos de la repetibilidad, como la variación intra e inter-prueba, la variación de operario intra e inter-prueba y la variación de lote intra- e inter-prueba, lo cual, cuando se somete a un análisis estadístico, indica el nivel de robustez (precisión) de un procedimiento analítico. El seguimiento de los parámetros de control de calidad de la prueba en cuanto a la repetibilidad permite comprobar si la

prueba está o no rindiendo como se esperaba. Para que las muestras control permitan deducir correctamente la precisión de la prueba, deben tratarse exactamente del mismo modo que las muestras problema en cada ejecución, por ejemplo, incluyendo la preparación de la muestra, como por ejemplo los pasos de extracción o la dilución de muestras de suero en un ensayo inmunoenzimático (ELISA).

La variación entre resultados obtenidos con muestras control también puede utilizarse como una estimación de esas fuentes combinadas de incertidumbre y se denomina enfoque "de arriba abajo". Con este enfoque se tiene en cuenta que los componentes de precisión se manifestarán en la medición definitiva. Así pues, el seguimiento de la precisión de la medición a lo largo del tiempo mostrará de forma efectiva los efectos combinados de la imprecisión asociada a cada paso.

La imprecisión o incertidumbre del proceso de medición asociada a un resultado de la prueba se vuelve cada vez más importante a medida que el valor obtenido se acerca al valor de corte diagnóstico. Ello se debe a que el resultado se interpreta como positivo, negativo o inconcluyente respecto al umbral de la prueba (como se describirá en el siguiente ejemplo). En este contexto, las muestras débilmente positivas, como las que se utilizan en estudios de la repetibilidad o como el control débilmente positivo, son las más adecuadas para estimar la MU. Así pues, la MU, que es función de la precisión de la prueba, es más crítica en los puntos de decisión (es decir, en los valores umbral), que suelen estar cerca del límite inferior de detección de la prueba. En este Capítulo, se describe la aplicación de la MU respecto a los valores de corte (umbral), recomendados por los fabricantes del kit de análisis o bien determinados en el laboratorio de diagnóstico.

1.2. Ejemplo de cálculos de la MU en la serología con ELISA

En casi todas las pruebas de detección de anticuerpos es importante recordar que más de la mitad de las pruebas son mediciones de la actividad de anticuerpos respecto a un umbral contra el cual se aplica una interpretación dicotómica de positivo o negativo. Es algo importante porque contribuye a decidir cuándo es apropiado aplicar la MU. En serología, la incertidumbre a menudo es más relevante que el umbral que permite decidir si una determinación es positiva o negativa. Los resultados que entran en esta zona también se describen como intermedios, inconcluyentes, sospechosos o ambiguos (véase la Norma de Validación de la OIE, Sección B.2.4).

Se utiliza un grupo reducido de datos obtenidos con un ELISA de competición para la detección de anticuerpos contra el virus de la Influenza aviar como ejemplo de enfoque "de arriba abajo" para la serología. Se utilizó una muestra control débilmente positiva para calcular la MU en el nivel de corte.

1.2.1. Método de expresión de la MU

Dado que la incertidumbre debe estimarse en el nivel umbral, que no necesariamente es el nivel de reacción del suero control débilmente positivo, la desviación estándar relativa (RSD) o el coeficiente de variación (CV) (si se expresa como porcentaje) proporcionan una transformación conveniente:

$$RSD(X) = SD(X) / \bar{X}$$

Para simplificar la evaluación, el resultado transformado se considera el resultado final de la prueba, del cual se calcula la media por réplica (\bar{X}). En el caso de este ejemplo, que se un ELISA de competición, los resultados se "normalizan" (como se define en la Norma de Validación de la OIE, Sección A.2.7) respecto a un patrón de trabajo formando una proporción de todos los valores de densidad óptica (DO) respecto al resultado de OD de un control no reactivo (negativo) (OD_N). Esta proporción se resta a 1 para establecer el nivel de actividad de anticuerpos sobre una escala de correlación positiva; cuanto mayor es el nivel, mayor es el valor calculado. Este valor ajustado se expresa como porcentaje y se denomina porcentaje de inhibición o valor PI. Así pues, en el caso del suero control débilmente positivo (OD_L), la transformación para obtener los valores de porcentaje de la inhibición para el control débilmente positivo (PI_L) es la siguiente:

$$PI_L = 100 \times [1 - (OD_L / OD_N)]$$

La desviación estándar relativa resulta ser:

$$RSD(PI_L) = SD(PI_L) / (PI_L)$$

1.2.2. Ejemplo

A continuación se muestra un pequeño conjunto de datos para el ejemplo del ELISA de competición de detección de la IA. Teóricamente, en la aplicación de este método "de arriba abajo", se utilizaría un conjunto grande de datos, lo cual permitiría explicar los efectos que los cambios en el operario o en los componentes de la prueba (distintos de la mera variación en el suero control) tendrían en la precisión.

Prueba	PI (%)
1	56
2	56
3	61
4	64
5	51
6	49
7	59
8	70
9	55
10	42

PI medio = 56,3; Desv. Est. (SD) = 7,9; Pruebas (n) = 10; RSD = SD/Media: 0,14

1.2.3. Cálculo de la incertidumbre

En el caso del grupo pequeño de datos,

$$RSD (PI) = 7,9/56,3 = 0,14 \text{ (o como coeficiente de variación = 14\%)}$$

La Incertidumbre (U) ampliada es el parámetro estadístico que define el intervalo dentro del cual se considera que se encuentra el valor de la medición a un determinado nivel de confianza, normalmente del 95%. La incertidumbre se amplía multiplicando la RSD (PI) por un factor de 2; ello permite el cálculo de un intervalo de confianza de aproximadamente el 95% alrededor del valor umbral, suponiendo una distribución normal de los datos.

$$U \text{ (IC del 95\%)} = 2 \times RSD = 0,28$$

A continuación, esta estimación se puede aplicar al nivel umbral (en este caso, a un PI = 50%)

$$IC \text{ del 95\%} = 50 \pm (50 \times 0,28) = 50 \pm 14\%$$

1.2.4. Interpretación

Cualquier resultado positivo (PI > 50%) que sea inferior al 64% no es positivo si se exige una confianza del 95%. De forma similar, un resultado negativo (PI < 50%) que sea superior o igual al PI de 36 no es negativo si se exige un nivel de confianza del 95%. Esta zona de menor confianza se podría correlacionar con la "zona gris" o "zona inconcluyente/sospechosa" que a efectos de interpretación debe establecerse en todas las pruebas (Greiner et al., 1995).

2. Otras aplicaciones

En enfoque "de arriba abajo" debe ser ampliamente aplicable a varias pruebas diagnósticas, incluidas pruebas moleculares. Para el cálculo de pruebas mediante una serie clásica de diluciones a la mitad del control positivo, como la neutralización vírica, la fijación del complemento y la inhibición de la hemaglutinación, debe calcularse la media geométrica del título (es decir, la media y la DE del \log_2 de los títulos) del suero control positivo. A continuación, pueden aplicarse las desviaciones estándar relativas, basadas en estos valores logarítmicos, al título umbral (\log), y finalmente transformarse para que representen la incertidumbre en el valor umbral. No obstante, en todos los casos, con este enfoque se asume que la varianza respecto al control positivo utilizado para estimar la RSD es proporcionalmente similar en el punto de aplicación de la MU, por ejemplo, en el valor umbral. Si la RSD varía significativamente a lo largo de la escala de medición, el suero control positivo que se utiliza para estimar la MU en el valor umbral debe escogerse en función de que tenga un nivel de actividad cercano a dicho valor umbral. Las Australian Subcommittee for Animal Health Laboratory Standards (SCAHLIS)

ha producido ejemplos comprobados relativos a varias pruebas de diagnóstico, que pueden consultarse en la página web de SCAHLS a través del siguiente enlace:

http://www.scahls.org.au/guidelines/measurement_of_uncertainty

En el caso de las PCR en tiempo real cuantitativas (qPCR) pueden utilizarse réplicas de controles positivos con su respectivo ciclo umbral (Ct) para estimar la MU aplicando el enfoque "de arriba abajo".

Se han descrito otros enfoques y variaciones, como los que se aplican a las pruebas serológicas (Dimech et al., 2006; Goris et al., 2009; Toussaint et al., 2007). Existen otros trabajos y documentos normativos disponibles en el *National Pathology Accreditation Advisory Group and Life Science*. El documento esencial respecto a la MU es la Guía para la expresión de la incertidumbre en la medición (GUM, por las siglas en inglés de *Guide to the expression of uncertainty in measurement*), una Guía ISO/IEC (1995).

REFERENCES

General requirements for the competence of testing and calibration laboratories, AS ISO/IEC 17025-2005 2nd edition.

DIMECH W., FRANCIS B., KOX J. & ROBERTS G. (2006). Calculating uncertainty of measurement for serology assays by use of precision and bias. *Clin. Chem.*, 62, No. 3, 526–529.

GORIS N., VANDENBUSCHÉ F., HERR, C., VILLERS, J., VAN DER STEDE, Y. & DE CLERCK K. (2009). Validation of two real-time PCR methods for foot-and-mouth disease diagnosis: RNA-extraction, matrix effects, uncertainty of measurement and precision. *J. Virol. Methods*, 180, 157–162.

GRINER M., SOHR D. & GOESSEL P.A. (1995). Modified ROC analysis for the selection of cut-off values and the definition of intermediate results of serodiagnostic tests, *J. Immunol. Methods*, 186, 123–132.

TOUSSAINT J.F., ASSAM P., CAU B., DEKEYSER F., KNAPEN K., IMBERECHTS H., GORIS N., MOLENERGHS G., MINTENS K., & DE CLERCK K. (2007). Uncertainty of measurement for competitive and indirect ELISAs. *Rev. sci. tech. Off. Int. Epiz.*, 28 (3), 649–656.

Requirements for the estimation of measurement uncertainty, National pathology accreditation advisory group, (2007). [http://www.health.gov.au/Internet/main/publishing.nsf/Content/B1074B732F32282DCA257BF0001FA218/\\$File/dh_aeou.pdf](http://www.health.gov.au/Internet/main/publishing.nsf/Content/B1074B732F32282DCA257BF0001FA218/$File/dh_aeou.pdf) (accessed 14 October 2013).

American Association for Laboratory Accreditation. Policy on estimating measurement uncertainty for life science testing labs http://www.a2la.org/policies/A2LA_P103b.pdf (accessed 14 October 2013).

Guide to the expression of uncertainty in measurement (GUM), ISO/IEC Guide 98:1995.

Sub-committee for Animal Health Laboratory Standards, Australia and New Zealand. Examples for MU in Veterinary Diagnostic Testing: http://www.scahls.org.au/guidelines/measurement_of_uncertainty (accessed 8 July 2013).

