

**MEMORIA DE ACTIVIDAD DEL  
PROGRAMA DE CONTROL DE CALIDAD DE LA  
SOCIEDAD ESPAÑOLA DE CITOLOGÍA (SEC)  
EJERCICIO 2022**



**q-PAP Morfológico & q-PAP Molecular**

---

**INDICE DE CONTENIDOS**

INTRODUCCIÓN	4
DEFINICIONES	7
DATOS GENERALES DEL PROVEEDOR Y DEL PROGRAMA	8
CARACTERÍSTICAS DEL PROGRAMA	9
CALENDARIO	9
DETERMINACIÓN DEL VALOR ASIGNADO Y SU INCERTIDUMBRE	10
ESTADÍSTICA EMPLEADA	10
ESTUDIOS DE HOMOGENEIDAD Y ESTABILIDAD	13
CONFIDENCIALIDAD Y PROTECCIÓN DE DATOS	13
CUADRO DE INDICADORES PARA EL PROGRAMA q-PAP DE CONTROL DE CALIDAD DE LA SOCIEDAD ESPAÑOLA DE CITOLOGÍA	14
DATOS ESPECÍFICOS DEL PROGRAMA 2022	15
DATOS ESPECÍFICOS DEL PROGRAMA Q-PAP MORFOLÓGICO	15
<i>Proceso de selección de muestras</i>	15
<i>Estudio de homogeneidad y estabilidad</i>	16
<i>Valor asignado y su incertidumbre</i>	16
<i>Parámetro</i>	19
<i>Materiales y transporte</i>	20
<i>Instrucciones</i>	20
Evaluación global del desempeño: Resultados	21
<i>Centros participantes, número de inscritos y tasas de respuesta.</i>	21
<i>Concordancias respecto al valor asignado y su incertidumbre</i>	21
<i>Discordancia según parámetro de riesgo para las distintas categorías diagnósticas</i>	24
<i>Valor de las respuestas respecto al valor asignado: Coeficiente Kappa de Cohen (k)</i>	26
<i>Valor de las respuestas entre observadores: Mediana robusta y Z Score kappa</i>	29
DATOS ESPECÍFICOS DEL PROGRAMA Q-PAP MOLECULAR	35
<i>Proceso de selección de las muestras</i>	35
<i>Estudio de homogeneidad y estabilidad</i>	35
<i>Valor asignado y su incertidumbre</i>	36
<i>Parámetro</i>	37
<i>Materiales y transporte</i>	37
<i>Instrucciones</i>	37
Evaluación global del desempeño: Resultados	38
<i>Centros participantes y tasas de respuesta.</i>	38

<i>Asignación muestras q-PAP Molecular 2022</i>	38
<i>Concordancia respecto al valor asignado y su incertidumbre</i>	39
<i>Valor de la determinación molecular de HPV de alto riesgo como prueba diagnóstica</i>	40
<i>Valor de las respuestas respecto al valor asignado: Coeficiente Kappa de Cohen (k)</i>	41
<i>Valor de las respuestas entre observadores: Mediana robusta y Z-Score Kappa.</i>	41
DISCUSIÓN	42
CONCLUSIONES	46
CUADRO DE INDICADORES PARA EL PROGRAMA q-PAP DE CONTROL DE CALIDAD DE LA SOCIEDAD ESPAÑOLA DE CITOLOGÍA. EJERCICIO 2022	46
BIBLIOGRAFÍA	47
INDICE DE TABLAS y GRÁFICAS	50
ANEXOS	52
ACTUACIONES A LLEVAR A CABO ANTE LA OBTENCIÓN DE RESULTADOS NO SATISFATORIOS. BENEFICIOS DE LA PARTICIPACIÓN EN UN PROGRAMA DE INTERCOMPARACIÓN.	53
CRITERIOS PARA LA EVALUACIÓN DE DISCREPANCIAS DIAGNÓSTICAS EN CITOLOGÍA GINECOLÓGICA	55

## INTRODUCCIÓN

La **Sociedad Española de Citología (SEC)** persigue la acreditación para la norma ISO/IEC 17043 actuando en el desarrollo de ensayos de aptitud respecto a las pruebas diagnósticas para la determinación de la patología tumoral y no tumoral del cérvix uterino en programas de cribado. El objetivo de este programa es organizar un ejercicio de intercomparación para los profesionales y los laboratorios que realicen esta práctica diagnóstica, además de servir de soporte para el cumplimiento de la norma UNE-EN-ISO 15189 que en su apartado 5.6.4 indica que *“el laboratorio debe participar en comparaciones entre laboratorios tales como las organizadas en el marco de programas de evaluación externa de la calidad.”*<sup>1</sup>

Como dice el documento NT-03 de Política de ENAC sobre intercomparaciones<sup>2</sup>, el objetivo fundamental de estos ejercicios es dotar a los laboratorios de una herramienta de control de la calidad que les permita comparar sus resultados con laboratorios similares, monitorizar sus resultados a lo largo del tiempo, detectar tendencias y considerar acciones correctivas cuando sea necesario. Esta política es complementada cuando también ENAC afirma, a tenor de lo recogido en la norma UNE-EN-ISO 17025<sup>3</sup>, que la herramienta incide en la capacidad de los laboratorios para la realización de un ensayo concreto obteniendo información externa con la que el laboratorio asegura, en la medida de lo posible, que *la validación de su procedimiento y su estrategia de control interno de calidad son suficientemente eficaces y, por tanto, pueden asegurar, con cierto grado de confianza, que no tienen sesgo en sus resultados de rutina.*

Por otro lado, la participación en estos programas incluye un potencial de mejora al obligar al laboratorio, ante resultados no satisfactorios, a *poner a prueba su capacidad para detectar la posible fuente del error, que podría ser desde la cualificación no adecuada del personal, la validación incompleta del procedimiento o un error puntual en el funcionamiento de un equipo.* A este respecto, el establecimiento de procedimientos normalizados de trabajo y los controles internos son la primera garantía de calidad de las medidas en un laboratorio. Sin embargo, la medida de que todo funciona bien está en la confirmación de una revisión externa que evalúe el desempeño de la actividad de una manera objetiva e imparcial. Esta evaluación externa es un poderoso argumento comercial ya que consolida la confianza de los clientes en cuanto a la aptitud que tiene el laboratorio para obtener resultados satisfactorios, o lo que es lo mismo, demostrar la competencia técnica frente a terceros.

---

<sup>1</sup> UNE-EN-ISO 15189

<sup>2</sup> Política de ENAC sobre intercomparaciones. NT-03 Rev. 7 abril 2021. Series 1,2 y 3.

<sup>3</sup> Guía sobre la participación en programas de intercomparaciones. G-ENAC-14 Rev.1 septiembre 2008.

Tal como decía William Osler<sup>4</sup> “*la medicina es la ciencia de la incertidumbre y el arte de la probabilidad*”, pero la incertidumbre, esa constante inherente en la toma de decisiones debe ser reconocida y presentada no como una limitación, sino todo lo contrario, como una fuerza acompañando al movimiento. Los programas de prevención de cáncer de cérvix han sido, y son, una poderosa herramienta en salud pública, por ello, cualquier aportación que contribuya a sostener la evidencia disponible debe ser entendida como otra forma de hacer.

La Sociedad Española de Citología hace suya esta responsabilidad y pone en marcha, con un gran esfuerzo, los recursos necesarios para su crear su Programa q-Pap de Control de la Calidad y con ello cumple, además, con el principio bioético de beneficencia, que conlleva acciones positivas para evitar daños a los demás y para beneficiarlos.

Para reforzar este argumento utilizamos a Tom L. Beauchamp y James F. Childress<sup>5</sup> que distinguen dos principios de beneficencia: la beneficencia positiva y la utilidad. Macario Alemany<sup>6</sup> dice respecto a la beneficencia positiva que muchos actos de beneficencia no son obligatorios, pero el principio de beneficencia impone una obligación (*prima facie*) de ayudar a promover sus intereses importantes y legítimos. Tom L. Beauchamp y James F. Childress sostienen que el *principio de beneficencia positiva apoya un conjunto de reglas morales de obligación más específicas*. Los ejemplos que citan de las reglas de beneficencia obligatoria son: 1) proteger y promover los derechos de los otros, 2) prevenir que suceda algún daño a otros, 3) suprimir las condiciones que puedan causar un daño a otros, 4) ayudar a las personas con discapacidades y 5) rescatar a las personas en peligro<sup>5</sup>.

Sin embargo, con este razonamiento se corre el riesgo de que prevalezca el principio de beneficencia sobre el respecto a las personas, o principio de autonomía, y con ello justificar el paternalismo. Macario Alemany sostiene que lo que puede justificar el paternalismo es la noción de “competencia”. La función de los juicios de competencia es distinguir entre las personas cuyas decisiones deben ser respetadas (competentes) y las personas cuyas decisiones pueden no ser respetadas (incompetentes). La competencia se define como la “capacidad para llevar a cabo una tarea” según la definición de Bernard Gert y Charles M. Culver<sup>7</sup>. Además, la capacidad para llevar a cabo una tarea es esencialmente graduable, unas personas pueden ser más competentes que otras para desempeñar la misma tarea, pero para que el concepto de competencia pueda cumplir su función de dividir a los individuos entre competentes e incompetentes es necesario fijar un nivel de

---

<sup>4</sup> Young, P, Finn BC, Bruetman JE, et al. El hombre y sus descripciones. Revista médica de Chile. Sociedad Médica de Santiago. 2012; 140(9): 1218-27

<sup>5</sup> Beauchamp, Tom L. y James F. Childress: Principles of Biomedical Ethics, ed. Oxford University Press, Nueva York, (ed 2001)

<sup>6</sup> Macario Alemany. El concepto y la justificación del paternalismo. Tesis de doctorado Departamento de Filosofía del Derecho. Universidad de Alicante 2005.

<sup>7</sup> Culver, Charles M. y Bernard Gert: Philosophy in Medicine (Conceptual and Ethical Issues in Medicine and Psychiatry), ed. Oxford University Press, Nueva York, 1982.

capacidad, con respecto a una tarea, por encima del cual los sujetos son competentes y por debajo del cual son incompetentes, Los estándares de competencia fijan el nivel de competencia exigible en relación con las tareas propuestas. En los contextos médicos, explican Tom L. Beauchamp y James F. Childress, se han propuesto diversos estándares de incompetencia que reflejan diferentes niveles de exigencia en relación con diversos tipos de capacidades. Los estándares menos exigentes se basan únicamente en la simple capacidad para formular una preferencia. Otros, con diversos grados de creciente exigencia, se basan en la capacidad para comprender la información y apreciar cómo ella se aplica a la propia situación. Finalmente, el estándar más exigente demanda la capacidad de alcanzar una decisión razonable<sup>8</sup>.

***Programa q-Pap de Control de Calidad de la Sociedad Española de Citología  
5 de abril de 2023***

---

<sup>8</sup> Beauchamp, Tom L. y James F. Childress: Principles of Biomedical Ethics (Fifth Edition), cit., p. 73. Sobre los estándares de competencia manejados en la práctica médica, véase: Roth, L. H., Meisel, A. y C. W. Lidz: “Tests of competency to Consent to Treatment”, en American Journal of Psychiatry, 134 (1977), pp. 279-284. Una visión general de este problema puede obtenerse en Simón Lorda, Pablo: “La evaluación de la capacidad de los pacientes para tomar decisiones y sus problemas”, en Feito Grande, Lydia: Estudios de Bioética, ed. Dykinson, Madrid, 1997, pp. 119-154.

## DEFINICIONES

- ***Incertidumbre:*** La falta de certeza un estado de conocimiento limitado donde es imposible describir exactamente el estado existente, un resultado futuro o más de un resultado posible. El Vocabulario Internacional de Metrología (VIM) define la incertidumbre de medida como un parámetro asociado al resultado de una medición, que caracteriza la dispersión de los valores que razonablemente podrían ser atribuidos al mensurando.
- ***Ítem:*** Muestra, material, instrumento, equipo, etc. sobre el que se lleva a cabo el ensayo/calibración objeto del ejercicio de inter comparación.
- ***Inter comparación:*** Proceso de organización, desarrollo y evaluación de ensayos del mismo ítem o ítems similares por varios laboratorios, de acuerdo con condiciones preestablecidas.
- ***Nivel de concordancia:*** acuerdo entre el valor asignado (valor de referencia o valor establecido por los expertos) y el valor emitido (respuesta de un participante).
- ***Parámetro:*** Una variable estadística es una característica que puede fluctuar y cuya variación es susceptible a adoptar diferentes valores, los cuales pueden medirse u observarse. Las variables adquieren valor cuando se relacionan con otras variables, es decir, si forman parte de una hipótesis o de una teoría. En este caso se las denomina constructos o construcciones hipotéticas. En estadística, un parámetro es un número que resume la gran cantidad de datos que pueden derivarse del estudio de una variable estadística. El cálculo de este número está bien definido, usualmente mediante una fórmula aritmética obtenida a partir de datos de la población de estudio.
- ***Proveedor:*** Organización que diseña y organiza el ejercicio de inter comparación.
- ***Rendimiento:*** Expresión de la evaluación de los participantes realizada por el proveedor a partir de los resultados emitidos por el laboratorio en un ejercicio de inter comparación.
- ***Valor asignado:*** Valor atribuido a una determinada propiedad de un ítem ensayado, en un ejercicio de inter comparación, con el objeto de calcular el rendimiento.
- ***Probabilidad de aleatoriedad:*** Probabilidad que un participante responda una cuestión al azar, independientemente de su acierto. La aleatoriedad se asocia a todo proceso cuyo resultado no es previsible más que en razón de la intervención del azar. El resultado de todo suceso aleatorio no puede determinarse en ningún caso antes de que este se produzca. El estudio de los fenómenos aleatorios queda dentro del ámbito de la teoría de la probabilidad. La palabra aleatorio se usa para expresar una aparente carencia de propósito, causa, u orden.

## DATOS GENERALES DEL PROVEEDOR Y DEL PROGRAMA

**Proveedor:** SOCIEDAD ESPAÑOLA DE CITOLOGÍA

**Domicilio social:** Calle Canarias, 38, 1º H, 28045 Madrid.

**Secretaria:** AZUL CONGRESOS

**Responsable:** Sra. Silvia Suarez Rivera

**Dirección:** Uría, 76 -3º Dcha 33003 OVIEDO

**Teléfono:** 984 051 604 – 984 051 671

**Email:** [secretaria@secitologia.org](mailto:secretaria@secitologia.org)

**Web:** [www.secitologia.org](http://www.secitologia.org)

### COMITÉ PARA EL PROGRAMA Q-PAP DE CONTROL DE CALIDAD DE LA SOCIEDAD ESPAÑOLA CITOLOGÍA

**Dra. Rosario Granados Carreño**

Presidenta del Programa de Control Calidad de la SEC.

**Dr. Francesc Alameda Quítllet**

Coordinador del Programa de Control de Calidad de la SEC.

**Dr. Antonio Martínez Lorente**

Responsable del Programa de Control de Calidad de la SEC.

**Dr. Francesc Tresserra Casas**

Vocal del Programa de Control Calidad de la SEC.

**Dra. Alejandra Caminoa-Lizarralde Aiza**

Vocal del Programa de Control Calidad de la SEC.

**Dra. Belén Lloveras Rubio**

Presidenta Saliente de la Sociedad Española de Citología SEC.

**Sr. Francesc Olivares Vegas**

Secretario Técnico del Programa de Control Calidad de la SEC.

### SECRETARÍA TÉCNICA DEL PROGRAMA Q-PAP DE CONTROL DE CALIDAD

**Secretario Técnico:** Sr. Francesc Olivares Vegas

**Dirección:** HOSPITAL DEL MAR. Servei de Patologia.

Passeig Marítim, 25-29.

08003 Barcelona

**Teléfono:** 932483394

**Email:** [calidad@secitologia.org](mailto:calidad@secitologia.org)

## CARACTERÍSTICAS DEL PROGRAMA

**Acreditación ISO/IEC 17043:2010: PROCESO DE IMPLEMENTACIÓN DE LA NORMA**

**Tipo de programa de aptitud: PROGRAMA INTERPRETATIVO CUALITATIVO**

**Subtipo: PROGRAMA DE EVALUACIÓN EXTERNA DE LA CALIDAD (EQA)**

**Evaluación del desempeño: FASE ANALÍTICA**

**Producto:** Programa de inter comparación para evaluar la eficacia del diagnóstico en el cribado del cáncer de cérvix uterino en muestras cervico-vaginales obtenidas con citología.

**Método de ensayo (q-Pap morfológico): interpretación diagnóstica de muestras mediante imágenes digitalizadas (ítems) distribuida de manera homogénea a todos los participantes.** Se pretende analizar la capacidad diagnóstica reproduciendo la práctica habitual, distanciándose de diagnósticos no usuales, como se presupone en un programa de cribado. Todas las determinaciones se encuentran dentro de las categorías diagnósticas recogidas en la clasificación Bethesda 2014 para el cribado de cáncer cervico uterino.

**Método de ensayo (q-Pap molecular): pool de casos concentrados y homogéneos (ítems) a partir de muestras citológicas obtenidas y preservadas en base líquida, con resultado previo conocido de HPV negativo y positivo alto riesgo incluyendo los genotipos de alto riesgo 16 y 18, y otros, mediante pruebas de detección del ADN y pruebas de detección de ARNm E6/E7.** Se pretende analizar la capacidad diagnóstica reproduciendo la práctica habitual, distanciándose de diagnósticos no usuales, (damos categoría a los genotipos con capacidad oncogénica más frecuentes en nuestro medio) como se presupone en un programa de control de calidad de cribado.

**Parámetro:** Nivel de concordancia de las distintas variables entre el valor asignado (valor de referencia) y el valor emitido (respuesta de un participante).

**Intervalo de trabajo:** 1 año natural (ajustado a los distintos programas de cribado para el cáncer cervical en muestras cervico-vaginales)

**Condiciones del ensayo:** aquellas que se realizan en práctica habitual

## CALENDARIO

Tabla 1. Calendario Programa Control de Calidad Diagnóstico de Citología Ginecológica: q-PAP 2022

Actividad/es	Ronda 1	Ronda 2
Inscripción de participantes	17-03-2022	
Apertura/Envío de muestras a participantes	28-04-2022	01-09-2022
Período de introducción/emisión de resultados por participantes	30-05-2022	15-10-2022
Fecha de emisión informe individual de participación	06-06-2022	22-10-2022
Fecha de emisión de informe global	30-11-2022	

## DETERMINACIÓN DEL VALOR ASIGNADO Y SU INCERTIDUMBRE

**Valor asignado:** Se determina según instrucciones de la norma ISO/IEC 17043:2010 en *Anexo B apartado B.2.1 (d)*: “Valores consensuados por participantes expertos –los expertos (que pueden, en algunos casos, ser laboratorios de referencia) deberán tener competencia demostrable en la determinación del o de los medrados a ensayar, utilizando métodos validados conocidos por ser altamente exactos y comprobables a los métodos utilizados habitualmente”.

**Nivel de incertidumbre:** se calcula atendiendo a las indicaciones de la ISO/IEC Guía 98-3:2008: “Además, en muchas aplicaciones industriales y comerciales, así como en las áreas de la salud y de la seguridad, a menudo es necesario proporcionar un intervalo en torno al resultado de la medición, en el que se espera encontrar la mayor parte de valores de la distribución que pueden ser razonablemente atribuidos a la magnitud objeto de la medición. Por tanto, el método ideal para evaluar y expresar la incertidumbre de medida debería ser capaz de proporcionar fácilmente un intervalo, en particular, aquel con la probabilidad o el nivel de confianza que corresponda de manera realista con lo requerido”.

## ESTADÍSTICA EMPLEADA

**Evaluación del rendimiento:** La evaluación del rendimiento se realiza mediante la relación entre dos aspectos diferentes. Por una parte, la diferencia entre los resultados ofrecidos por el laboratorio frente al valor asignado considerado como verdadero, y por otra un valor de referencia o diana de incertidumbre (habitualmente expresada como desviación estándar) y que utiliza el organizador para considerar que los resultados son adecuados. Dos son los estadísticos empleados:

### **Coefficiente Kappa de Cohen (k), (Índice/Coefficiente Kappa).**

Mide el grado de concordancia inter-observador y permite identificar la proporción de concordancias observadas sobre el total de observaciones, excluyendo las concordancias atribuibles al azar. El índice Kappa toma valores entre  $-\infty$  y  $+1$ , teniendo presente que los resultados más próximos a  $+1$ , disponen de un grado mayor de concordancia inter-observador, mientras que el valor Kappa más próximo o igual a  $0$ , refleja que la concordancia observada es más atribuible al azar.

Para la interpretación del coeficiente Kappa se suele posicionar su valor en una escala cualitativa que incluye cinco niveles de fuerza de concordancia y permite simplificar la comprensión del mismo (pobre, débil, moderada, buena, muy buena)

Tabla 2. Valoración estandarizada del índice Kappa

Valor Kappa	Grado Concordancia
<0,20	Pobre
0,21 – 0,40	Débil
0,41 – 0,60	Moderada
0,61 – 0,80	Buena
0,81 – 1,00	Muy Buena

**Utilidad:** compara el resultado de un participante con el resultado experto.

### Test de la Mediana Robusta y Z Score kappa

Se busca mediante medidas de tendencia central el posicionamiento de los datos y la identificación de valores de los extremos.

El test de la mediana robusta compara el resultado de un participante con respecto a la mediana del conjunto. El resultado se clasifica según se aleje más o menos de la mediana en función de la desviación típica.

Tabla 3. Valoración estandarizada del test de la Mediana Robusta

Mediana +3Sd	Excelente
Mediana +2Sd a Mediana +3Sd	Notable
Mediana +2Sd a Mediana -2Sd	Satisfactorio
Mediana -2Sd a Mediana -3Sd	Bajo
Mediana - 3Sd	Muy bajo

La puntuación Z Score kappa identifica la diferencia entre el resultado del participante y el valor de la mediana de los resultados emitidos por todos los participantes. Se mide en:

Tabla 4. Valoración estandarizada de Z Score kappa

Z= -1,99 a +1,99	Satisfactorio
Z= -2 a -2,99 o +2 a +2,99	Cuestionable
Z= <3,00 o +3,00	Insatisfactorio

(NPM Z= ± 3 y ±5)

Se define de acuerdo con la ecuación:

$$Z = (X - X_a) / \hat{f}_p$$

$X_a$  es el valor de la mediana de los resultados remitidos por los participantes

$\hat{f}_p$  es la desviación estándar diana o adecuada al fin pretendido

$X$  es la medida del participante

**Utilidad:** compara el resultado de un participante con los resultados del resto de participantes del ejercicio.

## *Evaluación de la validez de la determinación de HPV como prueba diagnóstica*

**Sensibilidad (S):** probabilidad de clasificar correctamente a los enfermos (la proporción de verdaderos positivos en una prueba diagnóstica).

**Especificidad (E):** probabilidad de clasificar correctamente a los sanos (la proporción de verdaderos negativos en una prueba diagnóstica).

**Prevalencia o probabilidad preprueba (Ppre):** prevalencia de enfermedad en la población de estudio, esto es, que la probabilidad que un positivo esté enfermo depende de lo frecuente o rara que sea la enfermedad en la población de la que procede.

**Valor Predictivo positivo (VPP):** valor de que el sujeto esté enfermo una vez que el resultado de la prueba es positivo.

**Valor Predictivo negativo (VPN):** valor de que el sujeto esté sano una vez que el resultado de la prueba es negativo.

**Cociente de probabilidades positivo (CPP):** cuánto más probable es tener un positivo en un enfermo que en un sano.

$$CPP = S / (1 - E)$$

**Cociente de probabilidades negativo (CPN):** cuánto más probable es tener un test negativo en un enfermo que en un sano.

$$CPN = (1 - S) / E$$

Tabla 5. Clasificación de la potencia de una prueba diagnóstica según los valores de los cocientes de probabilidades

CPP	CPN	Valor de la prueba
>10	<0,1	Prueba muy potente
5-10	0,1-0,2	Prueba poco potente
2-5	0,2-0,5	Prueba de aportación dudosa
1-2	0,5-1	Prueba sin utilidad diagnóstica

**Odd ratio:** medida de asociación entre dos variables (correlación bivariada) que indica la fortaleza de su relación.

**Utilidad:** sensibilidad y especificidad son propiedades intrínsecas a una prueba diagnóstica que indican su capacidad para clasificar correctamente enfermos y sanos. Por su parte, los valores predictivos nos informan de la probabilidad de que un positivo esté enfermo o de que un negativo esté sano, aunque su valor depende de la prevalencia de la enfermedad en la población en que se calcula.

Los cocientes de probabilidades nos ayudan a calcular la probabilidad real de padecer la enfermedad conociendo la probabilidad previa y el resultado de la prueba.

## **ESTUDIOS DE HOMOGENEIDAD Y ESTABILIDAD**

Se determina según instrucciones de la norma ISO/IEC 17043:2010 identificadas en los apartados:

*4.4.3.4 Se debe demostrar que los ítems del ensayo de aptitud son suficientemente estables para asegurarse de que no sufrirán cambios significativos a lo largo de la realización del ensayo de aptitud, incluyendo las condiciones de almacenamiento y transporte.*

*4.4.3.5 Cuando se retienen para su posterior uso los ítems de ensayo de aptitud de rondas anteriores, el proveedor de ensayos de aptitud debe confirmar, antes de la distribución, los valores de la propiedad a determinar en el programa de ensayos de aptitud.*

En función del número total de muestras se seleccionan para el estudio de homogeneidad y estabilidad una muestra del mismo volumen y características.

Se establece como norma la temporalidad desde la recepción de la muestra hasta la comunicación del resultado, correspondiendo a un periodo de dos meses. De este modo, se toma una muestra al inicio, una a mitad y otra al final, garantizando que el resultado sea el mismo en todos los periodos medidos.

## **CONFIDENCIALIDAD Y PROTECCIÓN DE DATOS**

Toda la información relativa a los participantes inscritos en el Programa de Control de Calidad de Citología Ginecológica de la SEC será tratada con total confidencialidad, incluyendo la evaluación obtenida y no se revelará a terceras partes salvo autorización explícita del participante para un fin concreto.

En los diferentes informes, la identificación de los participantes permanecerá anónima, a través del código de identificación. Este código varía anualmente y es proporcionado por el proveedor del ejercicio, que garantizará su trazabilidad y confidencialidad.

La Sociedad Española de Citología (SEC) se compromete a mantener durante todo el programa de inter comparación los requisitos de la norma UNE-EN ISO/IEC 17043:2010. Del mismo modo, la SEC se compromete a respetar los principios éticos básicos en relación a las muestras biológicas utilizadas, siempre provenientes de imágenes de diagnósticos ya emitidos e incorporados en las historias clínicas de las pacientes, así como las normas establecidas por la Ley Orgánica 3/2018, del 5 de diciembre, de Protección de Datos Personales y de Garantía de los Derechos Digitales, que adapta la legislación española al Reglamento General de Protección de Datos de la Unión Europea (RGPD), en relación a los participantes e instituciones que forman parte del programa.

**CUADRO DE INDICADORES PARA EL PROGRAMA q-PAP DE CONTROL DE CALIDAD DE LA SOCIEDAD ESPAÑOLA DE CITOLOGÍA**

Tabla 6: Distribución cuadro de indicadores q-PAP según medida y definición

<b>Indicador</b>	<b>Medida</b>	<b>Definición</b>
<b>Nº de Centros</b>	nº entero	Institución que agrupa a los inscritos en el programa.
<b>Nº de Inscritos</b>	nº entero	Participantes inscritos sometidos a evaluación.
<b>Nº de Ítems</b>	nº entero	Muestra sobre el que se lleva a cabo el ensayo objeto del ejercicio de inter comparación
<b>Grado Incertidumbre pre prueba</b>	%	Parámetro asociado al resultado de una medición, que caracteriza la dispersión de los valores que razonablemente podrían ser atribuidos al mensurando. Se mide antes de ofertar la prueba mediante un ensayo.
<b>Grado Incertidumbre post prueba</b>	%	Ídem. Se mide con los resultados de los participantes al finalizar el ejercicio.
<b>Probabilidad de aleatoriedad</b>	%	Probabilidad que un participante responda una cuestión al azar, independientemente de su acierto.
<b>Sensibilidad</b>	%	Determina la validez de una prueba midiendo la probabilidad de que la prueba dé positiva si la enfermedad está presente.
<b>Especificidad</b>	%	Determina la validez de una prueba midiendo la probabilidad de que la prueba dé negativa si la enfermedad está ausente.
<b>Valor predictivo positivo</b>	%	Probabilidad de tener la enfermedad si la prueba ha dado positiva.
<b>Valor predictivo negativo</b>	%	Probabilidad de no tener la enfermedad si la prueba ha dado negativa.
<b>Nivel de Concordancia</b>	%	Acuerdo entre el valor asignado (valor de referencia) y el valor emitido (respuesta de un participante).
<b>Nivel de Discordancia</b>	%	Descuerdo entre el valor asignado (valor de referencia) y el valor emitido (respuesta de un participante).
<b>Rango Kappa</b>	Intervalo $(-\infty, 1)$	Concordancia inter observador y permite identificar la proporción de concordancias observadas sobre el total de observaciones, excluyendo las concordancias atribuibles al azar.
<b>Valor Kappa medio</b>	Valor continuo en el rango de Kappa	Valor medio de rango kappa de la totalidad de los participantes.
<b>Mediana Robusta</b>	% Categórico: (Excelente, Notable, Satisfactorio, Bajo, Muy bajo)	Compara el resultado de un participante con respecto a la mediana del conjunto. El resultado se mide en unidades $\pm$ alejadas de la desviación típica.
<b>Z Score</b>	% Categórico: (Satisfactorio, cuestionable, insatisfactorio)	Identifica la diferencia entre el resultado del participante y el valor asignado (mediana de participantes).

## DATOS ESPECÍFICOS DEL PROGRAMA 2022

### DATOS ESPECÍFICOS DEL PROGRAMA Q-PAP MORFOLÓGICO

#### *Proceso de selección de muestras*

Se seleccionaron 20 muestras de citología cervico-vaginal, anonimizadas, obtenidas en medio líquido correspondientes a los centros colaboradores, que para el ejercicio del año 2022 fueron los Servicios de Anatomía Patológica del Hospital de Barbastro, de la Fundación Instituto Valenciano de Oncología y del Hospital del Mar de Barcelona, todos ellos acreditados con la norma UNE-EN-ISO 15189 para los alcances de Citología Ginecológica.

El diagnóstico se realizó, en su día, dentro de la práctica habitual y de acuerdo a los criterios de la clasificación Bethesda 2014 para el cribado de cáncer de cérvix, atendiendo a las categorías:

- Muestra insatisfactoria para evaluación
- Negativo para lesión intraepitelial o malignidad
- ASC-US: Células escamosas atípicas de significado incierto
- L-SIL: lesión intraepitelial escamosa de bajo grado.
- ASC-H: Células escamosas atípicas que no excluyen una lesión de alto grado
- H-SIL: Lesión intraepitelial escamosa de alto grado. (Incluye displasia moderada, displasia severa y carcinoma escamoso in situ)
  - Carcinoma escamoso (CAESC)
  - ACG: Células glandulares atípicas (endocervicales, endometriales, glandulares NOS)
  - AIS: Adenocarcinoma in situ
  - Adenocarcinoma (ADC)

Cada centro colaborador elaboró una tabla en la que se indica el código del caso, la edad, antecedentes, resultado del test de HPV si se tiene, y el diagnóstico citológico emitido así como el resultado de la biopsia y la conización posterior si es que se realizaron.

Siempre se preparará una cantidad de muestras superior a las requeridas en el ejercicio del programa, a fin de atender posibles contingencias.

Las tablas con los diagnósticos emitidos de ambos Centros Colaboradores se enviaron, junto con las muestras de los casos, al Coordinador del Programa. Se escanearon los casos y se reenviaron las tablas correspondientes, sin la categoría *diagnóstico citológico* a un segundo experto de manera que este, efectuó el diagnóstico sobre las muestras ya escaneadas.

Finalizado el ejercicio, cada centro habrá diagnosticado las muestras recibidas del otro experto, lo que permite llevar a cabo la fase de comparación. En este punto, el Coordinador del Programa recibió los diagnósticos y seleccionó solo aquellos que cumplían **los criterios de inclusión**:

1. Concordancia 100% diagnóstico morfológico citológico entre los expertos
2. Correspondencia entre los diagnósticos morfológicos (citología/biopsia) citológico y tisular cuando existe biopsia.

### Estudio de homogeneidad y estabilidad

Al tratarse de muestras humanas digitalizadas y anonimizadas, el control no requiere estudios de homogeneidad y estabilidad.

### Valor asignado y su incertidumbre

Tras la selección del total de casos concordantes, el Comité para el Programa de Control de Calidad de la SEC, selecciona para el ejercicio 2022 los siguientes casos (Tabla 7)<sup>9</sup>:

Tabla 7: Distribución casos ejercicio 2022 según historia clínica, identificación caso-ronda y resultado

DESCRIPCIÓN CASO	IDENTIFICACIÓN CASO Y RONDA	RESULTADO EMITIDO
CASO 1: Mujer de 36 años. ASCUS 6 meses antes. HPV16 positivo.	2022CG_R1/Caso1	ASCUS
CASO 2: Mujer de 42 años, sin antecedentes. HPV16 positivo.	2022CG_R1/Caso2	HSIL
CASO 3: Mujer de 29 años sin antecedentes. HPV negativo.	2022CG_R1/Caso3	LSIL
CASO 4: Mujer de 26 años. ASCUS y LSIL en 2019 (HPV positivo para alto riesgo)	2022CG_R1/Caso4	LSIL
CASO 5: Mujer DE 64 años. VAIN II/III hace 4 años. HPV no realizado.	2022CG_R1/Caso5	NEGATIVO
CASO 6: Mujer de 43 años. Antecedentes de LSIL. HPV positivo no especificado	2022CG_R1/Caso6	LSIL
CASO 7: Mujer de 50 años. Antecedentes ginecológicos sin interés. Carcinoma de mama hace 4 años. HPV no realizado.	2022CG_R1/Caso7	NEGATIVO
CASO 8: Mujer de 21 años, sin antecedentes HPV no realizado.	2022CG_R1/Caso8	LSIL
CASO 1: Mujer de 29 años. Sin antecedentes. HPV no realizado.	2022CG_R2/Caso1	NEGATIVO
CASO 2: Mujer de 27 años. Antecedentes de HSIL (CINII) en 2016 por biopsia y controles citológicos posteriores con diagnóstico de LSIL. HPV 18 positivo.	2022CG_R2/Caso2	ASCUS
CASO 3: Mujer de 31 años. Sin antecedentes. HPV no 16 no 18 positivo.	2022CG_R2/Caso3	LSIL
CASO 4: Mujer de 34 años. Sin antecedentes. HPV no 16 no 18 positivo.	2022CG_R2/Caso4	HSIL
CASO 5: Mujer de 38 años. Cribado. HPV positivo.	2022CG_R2/Caso5	ASCUS
CASO 6: Mujer de 25 años. Cribado. HPV positivo.	2022CG_R2/Caso6	LSIL
CASO 7: Mujer de 31 años. Cribado. HPV positivo.	2022CG_R2/Caso7	HSIL
CASO 8: Mujer de 40 años. Cribado. HPV positivo.	2022CG_R2/Caso8	LSIL

Aunque se trata de un “Programa de Interpretación Cualitativo”, la SEC propone objetivar el resultado de la evaluación atribuyendo a los diagnósticos coincidentes la categoría de 1 y a los no coincidentes la categoría de 0. De esta manera, hacemos una distribución de las coincidencias por cada uno de los participantes y por cada uno de los casos. El objetivo es conocer el valor de la dispersión de los resultados obtenidos por el conjunto de los participantes como indicador de calidad.

<sup>9</sup> Acta reunión del día martes 8 de febrero de 2022 Comité para el Programa de Control de Calidad de la SEC

Para hacer el cálculo de la incertidumbre del valor asignado, utilizamos los criterios de G-ENAC-14 Rev.1 sep 2008: *En aquellos casos en que la incertidumbre del valor asignado obtenida por consenso  $u_x$ , en el que participan  $n$  laboratorios y obtienen una dispersión de resultados que pueden describirse mediante la Desviación Estándar  $\hat{\sigma}_{ejercicio}$ , podría aproximarse al valor del error estándar definido como:*

$$u_x = \hat{\sigma}_{ejercicio} / \sqrt{n}$$

Debemos tener en cuenta que esta estimación de la incertidumbre del valor asignado es una aproximación y tiene utilidad exclusivamente para evaluar la calidad de un ejercicio de intercomparación.

A modo de ejemplo, incluimos a continuación un modelo simulado con un motor de cálculo en el que se estiman 8 ítems y 10 participantes con respuestas concordantes y no concordantes. Sobre ellos se calcula el promedio, la desviación estándar, IC al 95%, la incertidumbre, el límite de control superior e inferior (Tabla 8).

Tabla 8: Tabla simulación calculado grado concordancia, desviación estándar, IC95%, la incertidumbre, el límite de control superior e inferior.

CASO	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	% ACIERTOS	(XI-M) <sup>2</sup>
1	1	1	1	1	1	0	0	1	0	0	60%	0,0189
2	1	1	1	1	1	1	1	1	1	1	100%	0,0689
3	1	1	0	1	1	0	0	1	1	0	60%	0,0189
4	1	1	1	1	1	1	1	1	1	1	100%	0,0689
5	1	1	1	1	0	1	0	1	0	0	60%	0,0189
6	1	1	1	1	1	1	1	1	1	1	100%	0,0689
7	0	0	0	1	0	0	0	0	1	1	30%	0,1914
8	1	1	1	1	0	1	1	1	1	0	80%	0,0039

**P. Participante**

**LCS.** Límite control superior:  $\bar{x}_m + Z_{\alpha/2} * \delta$

**LCI.** Límite control inferior:  $\bar{x}_m - Z_{\alpha/2} * \delta$

Promedio	73,75%
Desv. Est	25,60%
Z <sub>α/2</sub> (95%)	1,96
Incertidumbre	25,60%
LCS	100%
LCI	23,57%

En base a ello realizaremos dos cálculos que establecerán el nivel de incertidumbre:

- **Pre-prueba:** 10 participantes escogidos al azar de los Centros Colaboradores realizaran el ejercicio en cada una de las rondas sin conocer el destino final de sus resultados.
- **Post-prueba:** recoge la variabilidad real obtenida por todos los participantes en cada una de las rondas en el ejercicio de ese año.

## Parámetro

### Nivel de concordancia entre el valor asignado (valor de referencia) y el valor emitido (respuesta de un participante).

En el caso de Q-PAP morfológico se gradúa la no concordancia en MENOR, INTERMEDIA, MAYOR y GRAVE entre las distintas categorías diagnósticas. Dicha categorización se hace en relación a la importancia clínica<sup>10,11,12</sup> que puede tener el error en el proceso asistencial del diagnóstico (anexo I) y que queda resumido en la *Tabla 5*:

Tabla 9: Distribución del valor de referencia según grado de las distintas categorías diagnósticas

	INSATISFACTORIA	NEGATIVO	ASCUS	LSIL	ASCH	HSIL	CAESC	ACG	AIS	ADC
INSATISFACTORIA	OK	INTER	MAYOR	MAYOR	GRAVE	GRAVE	GRAVE	GRAVE	GRAVE	GRAVE
NEGATIVO	INTER	OK	MAYOR	MAYOR	GRAVE	GRAVE	GRAVE	MAYOR	GRAVE	GRAVE
ASCUS	MAYOR	MAYOR	OK	MENOR	INTER	INTER	INTER	MENOR	INTER	INTER
LSIL	MAYOR	MAYOR	MENOR	OK	INTER	INTER	INTER	INTER	INTER	INTER
ASCH	GRAVE	GRAVE	INTER	INTER	OK	MENOR	MENOR	INTER	MENOR	MENOR
HSIL	GRAVE	GRAVE	INTER	INTER	MENOR	OK	MENOR	INTER	MENOR	MENOR
CARCINOMA ESC(CAESC)	GRAVE	GRAVE	INTER	INTER	MENOR	MENOR	OK	INTER	MENOR	MENOR
ACG	GRAVE	MAYOR	MENOR	INTER	INTER	INTER	INTER	OK	INTER	INTER
AIS	GRAVE	GRAVE	INTER	INTER	MENOR	MENOR	MENOR	INTER	OK	MENOR
ADENOCARCINOMA(ADC)	GRAVE	GRAVE	INTER	INTER	MENOR	MENOR	MENOR	INTER	MENOR	OK

<sup>10</sup> Alameda F, Bernet L, Cano R, et al. Control de calidad de la citología ginecológica. Programa de Calidad de la Sociedad Española de Citología. Resultados de la primera ronda. RevEsp Patol. 2017;50:154-160

<sup>11</sup> Alameda F, Asó S, Catalina I, et al. Control de calidad de la citología ginecológica. Programa de Calidad de la Sociedad Española de Citología. Resultados de la segunda ronda. RevEsp Patol. 2018;51:71-76

<sup>12</sup> Torné Bladé A, del Pino Saladrigues M, Cusidó Gimferrer M, et al. Guía de cribado del cáncer de cuello de útero en España, 2014. RevEsp Patol. 2014;47 (Supl 1):1-43

## *Materiales y transporte*

Este apartado queda definido por la naturaleza del Método de ensayo: ***Interpretación diagnóstica de muestras mediante imágenes digitalizadas (ítems) distribuidas de manera homogénea a todos los participantes.*** Las imágenes son publicadas en el visor VENTANA Vector 1.1 software alojado en el Hospital del Mar (<https://hospitaldelmar.ventana-vector.com/>) y en donde al participante se le da un código de acceso para realizar su ejercicio.

## *Instrucciones*

La Sociedad Española de Citología en calidad de proveedor del programa de control de calidad Q-PAP Morfológico tiene registrada y accesible a todos los participantes y usuarios, a través de su página web (<http://calidad.secitologia.org/>), las instrucciones para la participación en el control de citología ginecológica (<http://calidad.secitologia.org/instrucciones-de-participacion-control-de-citologia-ginecologica/>), que registran de forma precisa y adecuada el objeto del programa, material diagnóstico, accesibilidad a las muestras, especificaciones de la realización y envío de resultados y calendario detallado del mismo.

Del mismo modo y previo al inicio de cada ronda del programa, desde la secretaria técnica se comunica mediante correo electrónico personalizado a cada participante inscrito la ubicación y acceso a las instrucciones de participación en el control de citología ginecológica.

## *Evaluación global del desempeño: Resultados*

### *Centros participantes, número de inscritos y tasas de respuesta.*

El número total de participantes inscritos fue de 42, pertenecientes a 12 centros o instituciones (Tabla 6):

Tabla 10: Distribución inscripciones según número de instituciones.  
Programa Control de Calidad q-Pap Morfológico. Ejercicio 2022

<b>Atrys Health S.A</b>
<b>Cerba Internacional</b>
<b>Fundación Instituto Valenciano de Oncología</b>
<b>Hospital Clínico Universitario Lozano Blesa</b>
<b>Hospital de Barbastro</b>
<b>Hospital Universitari Dexeus</b>
<b>Hospital Universitari Vall d'Hebron</b>
<b>Hospital Universitario de Getafe</b>
<b>Hospital Universitario Vinalopó</b>
<b>Hospital Universitario Donostia</b>
<b>Hospital Universitario Príncipe de Asturias</b>
<b>Parc Salut Mar de Barcelona (Hospital del Mar)</b>

42 fueron el total de participantes inscritos en el programa 2022

- La tasa de respuesta para la primera ronda fue del 100% (42/42)
- La tasa de respuesta para la segunda ronda fue del 90% (38/42)

### *Concordancias respecto al valor asignado y su incertidumbre*

- **Pre-prueba:** 10 participantes escogidos al azar de los Centros Colaboradores realizarán el ejercicio sin conocer el destino final de sus resultados. El cálculo de la incertidumbre para la primera ronda fue de 4,42% y para la segunda ronda de 29,57%, siendo el grado de incertidumbre promedio para el total del ejercicio 16,99%.
- **Post-prueba:** recoge la variabilidad real obtenida por todos los participantes en el ejercicio de ese año.

Para abordar este apartado, debemos considerar que en cada ronda se evalúan los niveles de concordancia en las respuestas de un caso respecto al valor asignado (1/0). De esta manera podemos

ver que, para el cálculo de la desviación estándar, el promedio de concordancia para la Ronda 1 es de 79,46% y si consideramos que la población se comporta como una distribución normalizada, los cálculos serán:  $79,846\% \pm 0,70\%$  sitúa los límites en 80,16 % y 78,76% (Tabla 11). El mismo cálculo hacemos para la Ronda 2:  $82,89\% \pm 1,30\%$  con una horquilla de valores entre 84,19% y 81,59% (Tabla 12).

Tabla 11. Distribución concordancia Ronda1.Programa Q-Pap Morfológico. Ejercicio 2022

Caso	Valor referencia	% Concordancia
1	ASCUS	52,3%
2	HSIL	88,1%
3	LSIL	100%
4	LSIL	61,9%
5	Negativo	92,8%
6	LSIL	95,2%
7	Negativo	47,6%
8	LSIL	97,6%

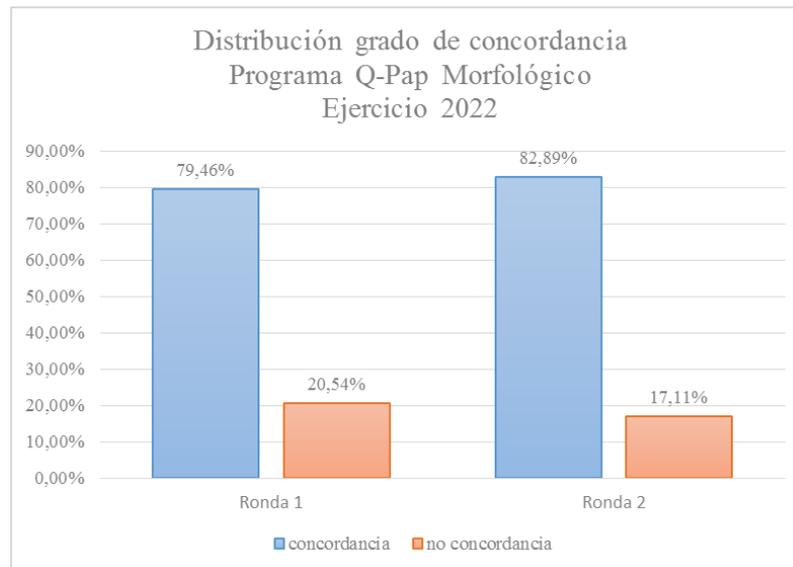
<b>Promedio</b>	79,46%
<b>Desv. Est</b>	0,70%
<b>Z<math>\alpha</math>/2 (95%)</b>	1,96
<b>Incertidumbre</b>	0,70%
<b>LCS</b>	80,16%
<b>LCI</b>	78,76%

Tabla 12. Distribución concordancia Ronda 2. Programa Q-Pap Morfológico. Ejercicio 2022

Caso	Valor referencia	% Concordancia
1	Negativo	97,3%
2	ASCUS	84,2%
3	LSIL	94,7%
4	HSIL	81,5%
5	ASCUS	71,0%
6	LSIL	100%
7	HSIL	73,6%
8	LSIL	60,5%

<b>Promedio</b>	82,89%
<b>Desv. Est</b>	1,30%
<b>Z<math>\alpha</math>/2 (95%)</b>	1,96
<b>Incertidumbre</b>	1.30%
<b>LCS</b>	84,19%
<b>LCI</b>	81,59%



Gráfica 1: Distribución grado concordancia y no concordancia Rondas 1 y 2

Varias lecturas debemos hacer ante estos resultados.

En primer lugar, el ejercicio nos dice que hay un nivel de concordancia en diagnósticos entre el valor asignado por los expertos y los participantes del 79,46% y 82,89% para la totalidad de los casos en cada una de las rondas. La desviación estándar, y con ello la incertidumbre, definida como *un parámetro asociado al resultado de una medición, que caracteriza la dispersión de los valores que razonablemente podrían ser atribuidos al mensurando* se sitúa en un 0,70% para la ronda 1 y un 1,30% para la ronda 2. De estos resultados hay que destacar que han mejorado la incertidumbre “preprueba” que fue del 4,42% para la primera ronda y del 29,7% para la segunda, lo que nos hace considerar a la población de inscritos más homogénea que la que realizó el ensayo previo.

En segundo lugar, se aporta, tal como hicimos el ejercicio pasado 2021, como parámetro, la probabilidad de aleatoriedad, definida como *Probabilidad que un participante responda una cuestión al azar, independientemente de su acierto*. Los resultados para cada una de las rondas han sido: 32,40% en la primera ronda y 27,10% en la segunda lo que nos indica que en 2,5 y 2,1 respuestas en la primera y segunda ronda probablemente fueron hechas de manera aleatoria.

En tercer lugar, los límites de control superior e inferior (LCS/LCI) también nos aportan información. En la primera ronda los casos 1(ASCUS/58,3%), 4 (LSIL/61,9%) y 7 (NEGATIVO /47,6%) están fuera del límite de control inferior (78,76%) y en la segunda son los casos 5 (ASCUS/71%), 7 (HSIL/73,6%) Y 8 (LSIL/60,5%) con un LCI del 81,59%. Con estos resultados podemos considerar que una parte importante de los inscritos se alejan de lo que es la distribución normal del grupo para el total del ejercicio. Pero cabe, tal vez, otra forma de aproximarnos a estos resultados y es considerar la totalidad de los casos (16), el peso de los distintos diagnósticos y las concordancias medias con los valores de los expertos. Así vemos que hay un grado de concordancia en el diagnóstico superior al 70% en todos los casos. El programa de este año tenía como objetivo fundamental presentar variables citológicas de lesión de bajo grado LSIL (7/16) teniendo una presencia del 43,7% en todo el ejercicio y podemos decir que la identificación de esta lesión ha sido identificada por todos los inscritos en el 87,1% de los casos. (Tabla 11).

Tabla 13. Distribución concordancia según valor de referencia del total de casos en ambas rondas  
 Programa q-Pap Morfológico. Ejercicio 2022

NEGATIVO (3/16)	ASCUS (3/16)	ASCH (0/16)	LSIL (7/16)	HSIL (3/16)	
18,7	18,7	0	43,7	18,7	% peso
79,2	71,1	0	87,1	81,0	% concordancia

### *Discordancia según parámetro de riesgo para las distintas categorías diagnósticas*

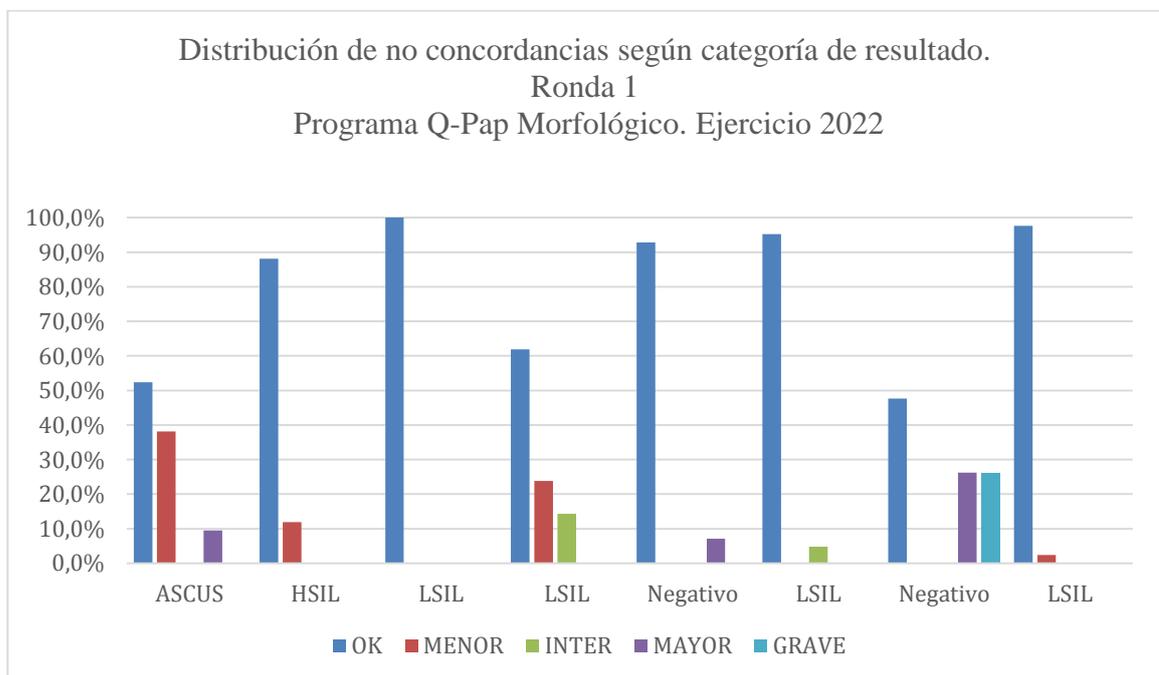
Otra manera de evaluar los resultados tiene que ver con la categorización del riesgo según la no concordancia. Un error diagnóstico mayor o grave, con repercusiones para la paciente y posible pérdida de oportunidad, supuso en la primera ronda el 49,1% (26% y 15,9%) y en la segunda un 16% (14% y 2%). Las discrepancias sin repercusiones y con posibilidades de diagnosticar correctamente las alteraciones estructurales en la siguiente toma fueron del 57,8% (46,3% y 11,5%) en la ronda 1 y 84% (60% y 24%) en la ronda 2. Podemos concluir que el modo fallo fue mayor en la primera ronda (69 frente a 50) y también la gravedad del fallo (49,1% frente a 16%) (Tablas 14 y 15)

 Tabla 14. Distribución de no concordancias según categoría del resultado en la Ronda 1.  
 Programa Q-Pap Morfológico. Ejercicio 2022

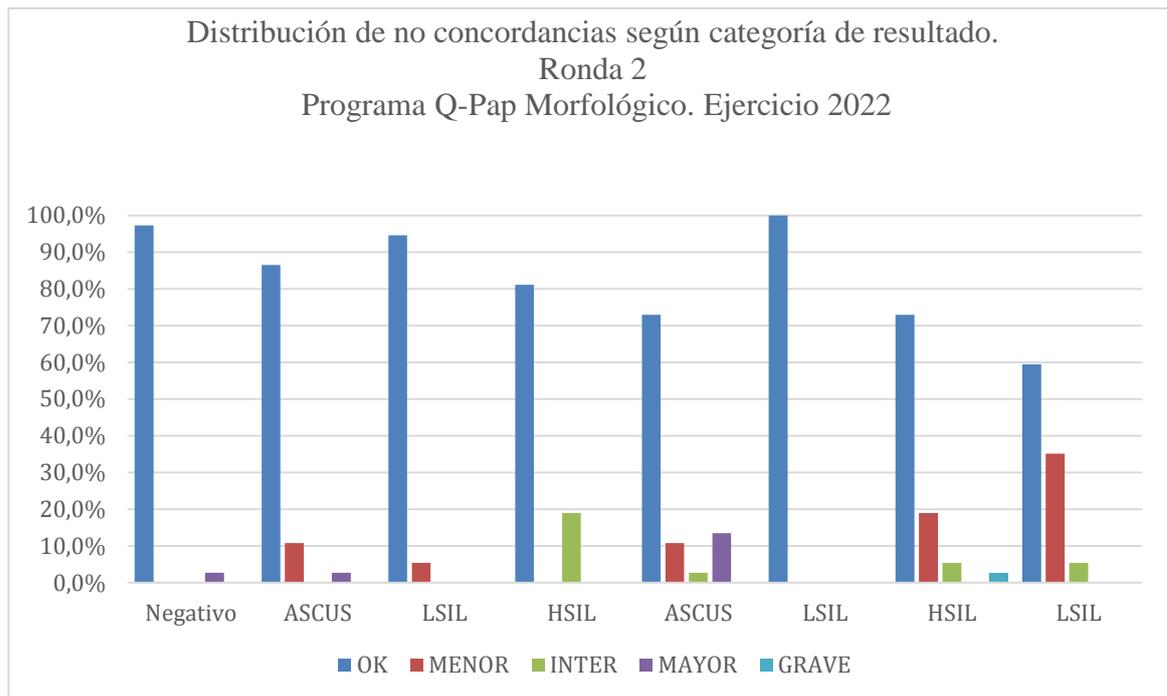
Nº Caso	Valor de Referencia	Categoría Resultado (n/%)				Total
		Menor	Intermedio	Mayor	Grave	
<b>Caso 1</b>	ASCUS	16 (38,1%)	-	4 (9,5%)	-	20
<b>Caso 2</b>	HSIL	5 (11,9%)	-	-	-	5
<b>Caso 3</b>	LSIL	-	-	-	-	-
<b>Caso 4</b>	LSIL	10 (23,8%)	6 (14,3%)	-	-	16
<b>Caso 5</b>	Negativo	-	-	3(7,1%)	-	3
<b>Caso 6</b>	LSIL	-	2 (4,8%)	-	-	2
<b>Caso 7</b>	Negativo	-	-	11 (26,2%)	11 (26,2%)	22
<b>Caso 8</b>	LSIL	1 (2,4%)	-	-	-	1
<b>Total</b>		32 (46,3%)	8 (11,5%)	18(26%)	11 (15,9%)	69

Tabla 15. Distribución de no concordancias según categoría del resultado en la Ronda 2.  
Programa Q-Pap Morfológico. Ejercicio 2022

N° Caso	Valor de Referencia	Categoría Resultado (n/%)				Total
		Menor	Intermedio	Mayor	Grave	
Caso 1	Negativo	.	-	1(2,7%)	-	1
Caso 2	ASCUS	4(10,8%)	-	1(2,7%)	-	5
Caso 3	LSIL	2(5,4%)	-	-	-	2
Caso 4	HSIL	-	7(19,9%)	-	-	7
Caso 5	ASCUS	4(10,8%)	1(2,7%)	5(13,5%)	-	10
Caso 6	LSIL	-	-	-	-	-
Caso 7	HSIL	7(18,9%)	2(5,4%)	-	1(2,7%)	10
Caso 8	LSIL	13(35,1%)	2(5,4%)	-	-	15
<b>Total</b>		30 (60%)	12(24%)	7(14%)	1 (2%)	50



Gráfica 2: Distribución de no concordancias según categoría de resultado. Ronda 1  
Programa Q-Pap Morfológico. Ejercicio 2022



Gráfica 3: Distribución de no concordancias según categoría de resultado. Ronda 2  
Programa Q-Pap Morfológico. Ejercicio 2022

### *Valor de las respuestas respecto al valor asignado: Coeficiente Kappa de Cohen (k)*

Tal como dijimos anteriormente, el Coeficiente kappa de Cohen es una medida estadística que ajusta el efecto del azar sobre elementos cualitativos (variables categóricas). Este coeficiente permite identificar la proporción de concordancias observadas, sobre el total de observaciones, excluyendo las concordancias atribuibles al azar. El objetivo es evaluar el rendimiento midiendo el grado de concordancia entre dos observadores: experto VRS inscrito, excluyendo las concordancias por azar.

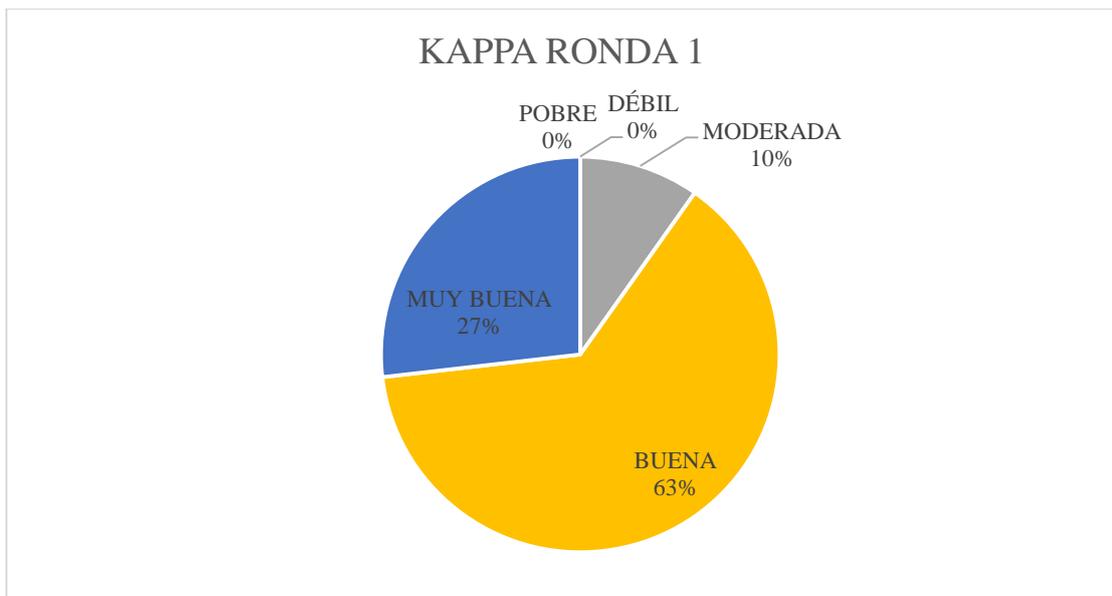
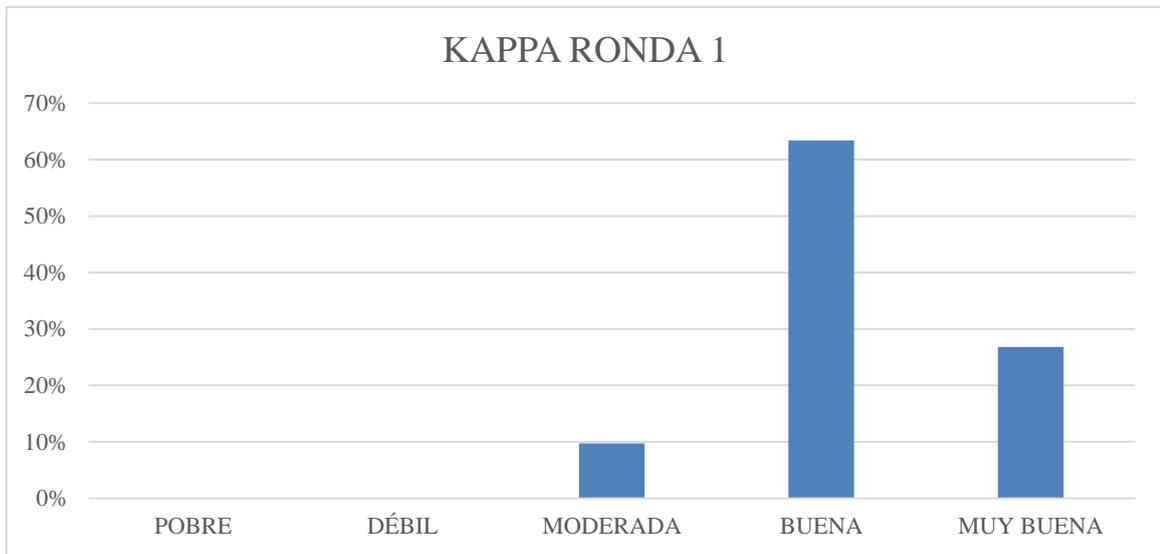
Tabla 16. Valoración estandarizada del índice Kappa

Valor Kappa	Grado Concordancia
<0,20	Pobre
0,21 – 0,40	Débil
0,41 – 0,60	Moderada
0,61 – 0,80	Buena
0,81 – 1,00	Muy Buena

El índice Kappa toma valores entre  $-\infty$  y  $+1$  (el límite inferior depende de ciertos elementos que se usan en su cálculo), teniendo presente que los resultados más próximos a  $+1$  disponen de un grado mayor de concordancia inter-observador. El valor Kappa más próximo o igual a 0 refleja que la concordancia observada es más atribuible al azar

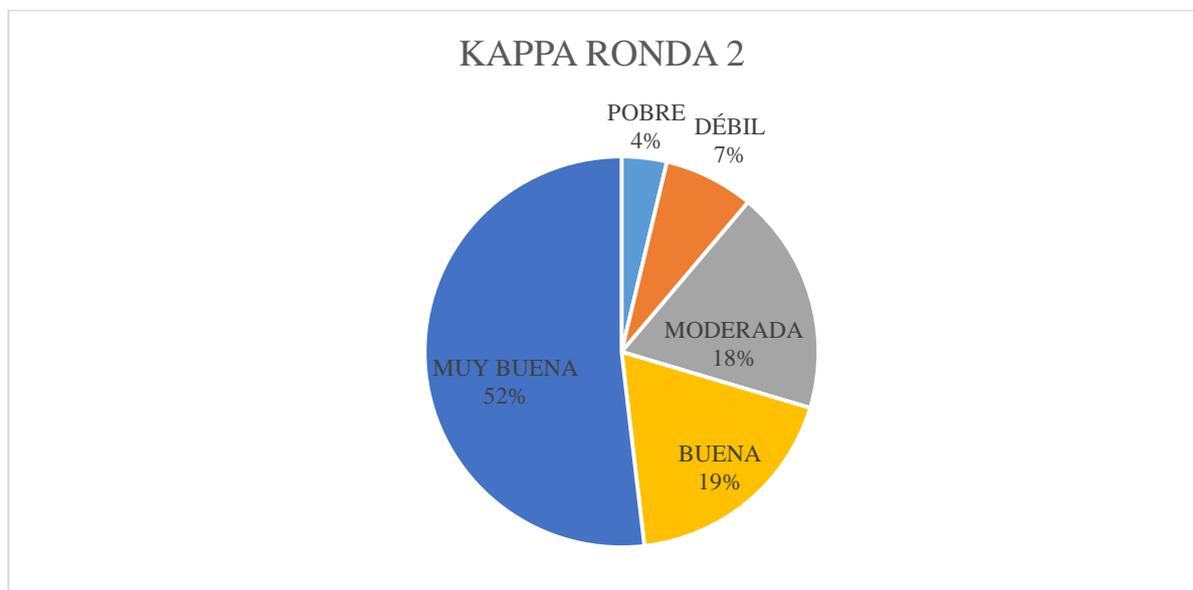
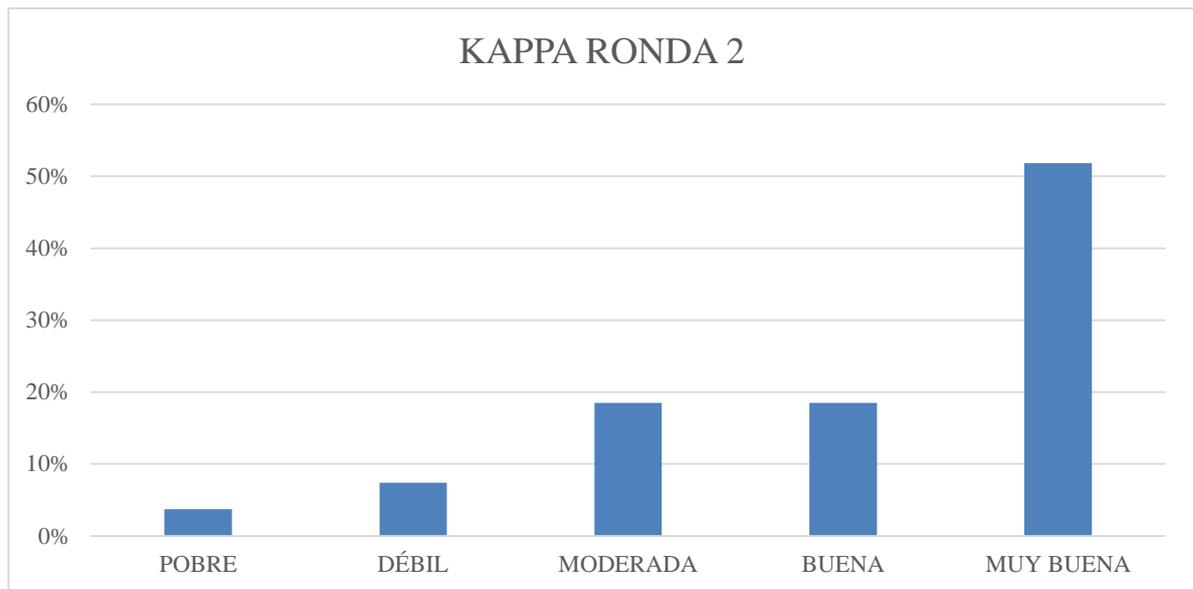
De este modo y según la categoría del coeficiente Kappa, para el total de participantes que finalizaron el ejercicio en la primera ronda, supuso que un 10% de los inscritos mantienen una

concordancia moderada, un 63% la mantienen buena y muy buena la mantienen el 27% de los participantes.



Gráfica 4: Distribución del valor kappa Ronda 1  
Programa Q-Pap Morfológico. Ejercicio 2022

En la segunda ronda, la concordancia pobre, débil o moderada asciende al 29% siendo el resto buena (19%) o muy buena (52%).



Gráfica 5: Distribución del valor kappa Ronda 2 Programa Q-Pap Morfológico. Ejercicio 2022

Los niveles de concordancia mayoritarios (buena y muy buena) en el total del ejercicio de este año son para la primera ronda del 90% y del 71% para la segunda ronda, que ponen de manifiesto el acierto diagnóstico en el conjunto de los centros inscritos en el programa y marca una tendencia en alza respecto a ejercicios de años anteriores.

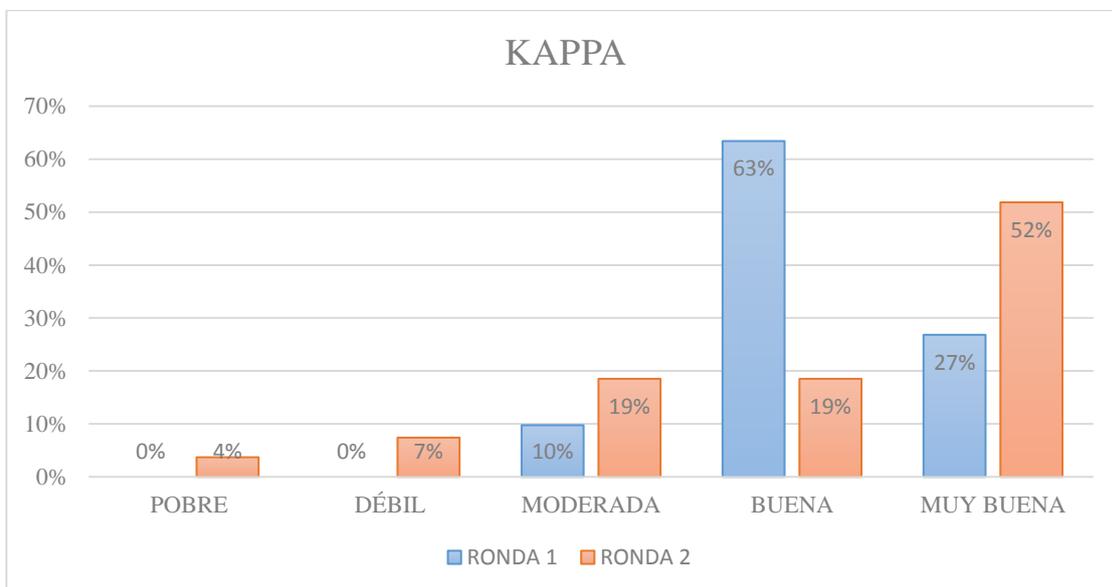
En este orden de cosas, debemos comparar los rangos kappa obtenidos por los participantes en cada una de las rondas, su valor medio y la calificación del grado de concordancia para cada ronda y para el conjunto del ejercicio de este año. Así vemos que el valor kappa medio de 2022 mantiene un grado de concordancia buena para el total de los inscritos, destacando hasta nueve participantes con una concordancia en todos los casos y que han servido para subir el rango de kappa medio.

Tabla 17. Valoración rango kappa, valor kappa medio y grado de concordancia.  
Rondas 1 y 2 Programa q-Pap Morfológico. Ejercicio 2022

	Rango Kappa	Valor Kappa medio	Grado Concordancia
<b>Ronda 1</b>	0,45-1	0,6978	Buena

	Rango Kappa	Valor Kappa medio	Grado Concordancia
<b>Ronda 2</b>	0,31-1	0,7666	Buena

	Rango Kappa	Valor Kappa medio	Grado Concordancia
<b>Ejercicio 2022</b>	0,31-1	0,7322	Buena



Gráfica 6: Distribución grado concordancia inter-observador para la Ronda 1 y 2, según valoración estandarizada del índice Kappa

### Valor de las respuestas entre observadores: Mediana robusta y Z Score kappa

El valor de estas medidas viene a determinar la tendencia central de los valores y establecer su posición respecto a valores en los extremos. Los métodos robustos pueden ser de utilidad para la realización de inferencias sin tener que «depurar» los datos extremos, ya que están diseñados para realizar inferencias sobre el modelo, reduciendo la posible influencia que pudiera tener la presencia de datos anómalos<sup>13</sup>. La principal de sus ventajas es que no se afectan por la existencia de datos anómalos. Un problema conocido en el cálculo de medidas de posición o centralización es aquel en el que una distribución de datos no sigue una ley normal; en esta circunstancia la media no es un buen estimador del promedio de los datos<sup>14</sup>.

<sup>13</sup> García A. Métodos avanzados de estadística aplicada. Madrid: UNED, 2001

<sup>14</sup> Ramalle-Gómara E, et al. Utilización de métodos robustos en la estadística inferencial. Aten Primaria 2003;32(3):177-82

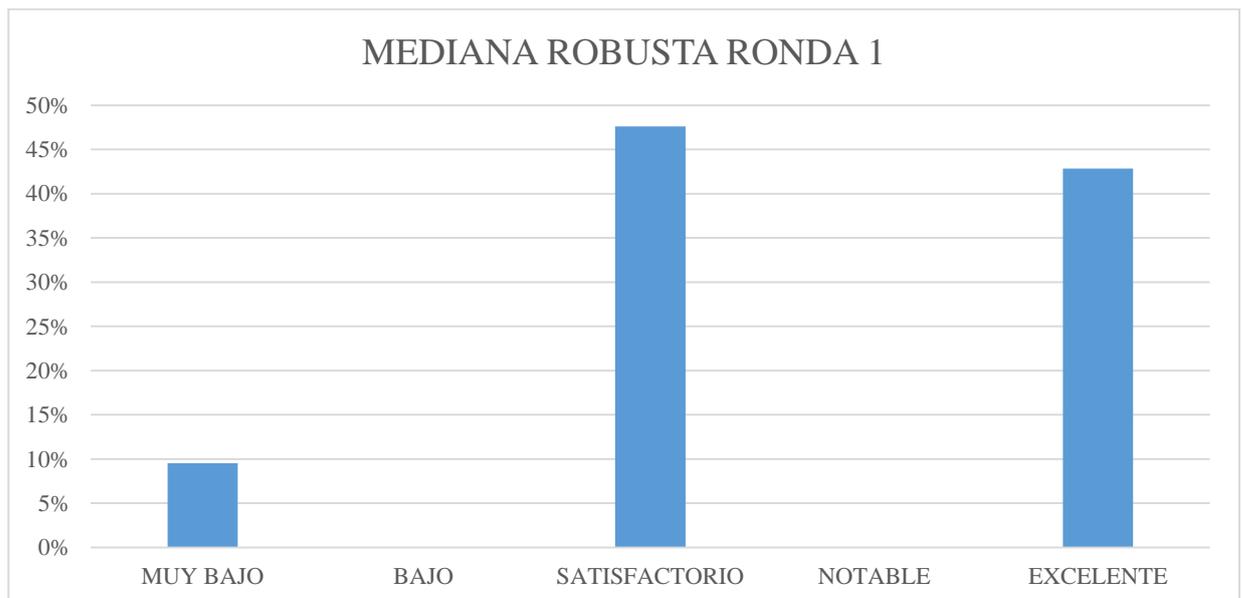
Así, la mediana robusta identifica una distribución en cada una de las rondas del ejercicio 2022 que quedan representadas en la Tabla 18 y en los Gráficos 8 y 9 y cuyos valores de referencia son:

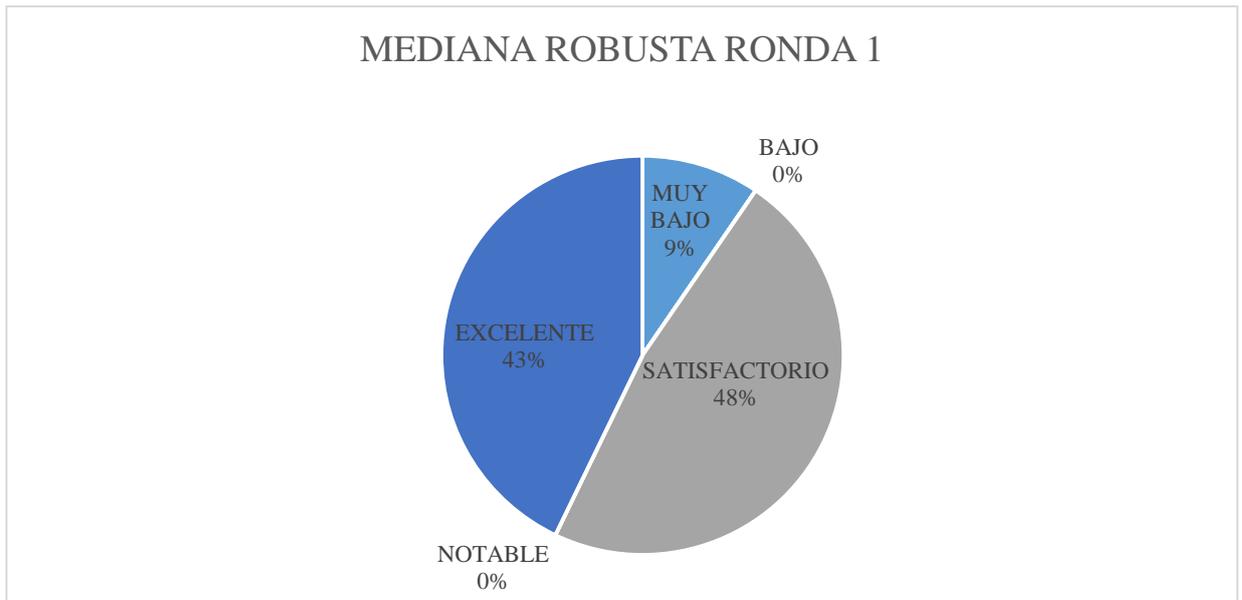
< mediana -3sd: **muy bajo**  
 (mediana-3sd, mediana -2sd): **bajo**  
 (mediana-2sd, mediana +2sd): **satisfactorio**  
 (mediana+2sd, mediana +3sd): **notable**  
 >mediana +3sd): **excelente**

Tabla 18: Distribución mediana robusta por frecuencia y porcentajes.  
Ronda 1 Programa q-Pap Morfológico. Ejercicio 2022

#### MEDIANA ROBUSTA RONDA 1

	Frecuencia	%	
Válidos	Muy bajo	4	9
	Bajo	-	-
	Satisfactorio	20	48
	Notable	-	-
	Excelente	18	43
	<b>Total</b>	<b>42</b>	<b>100,0</b>



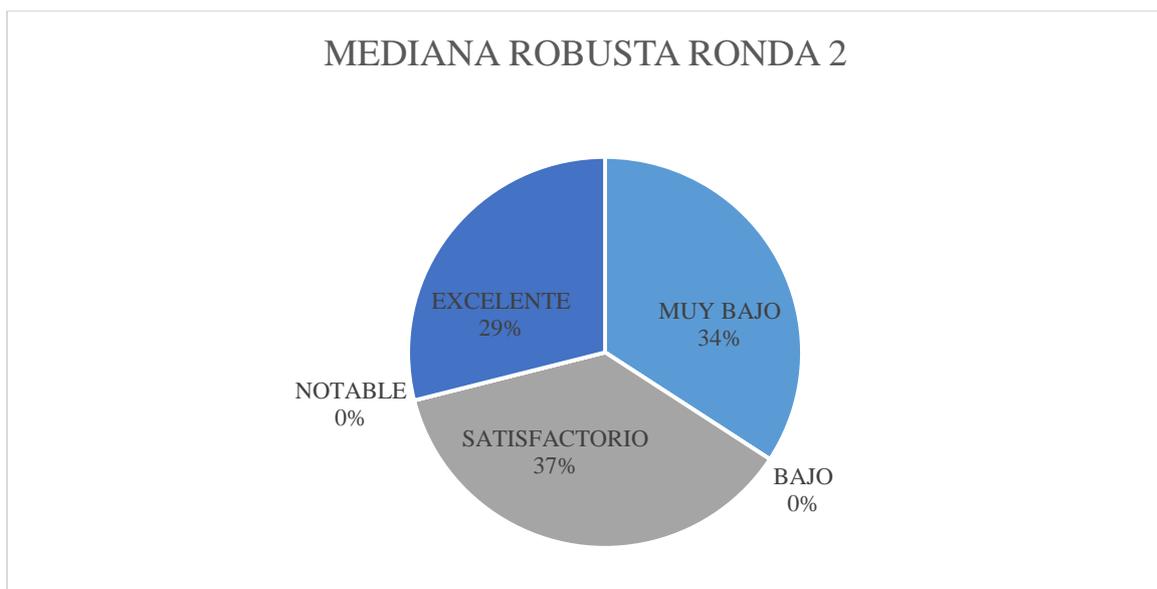
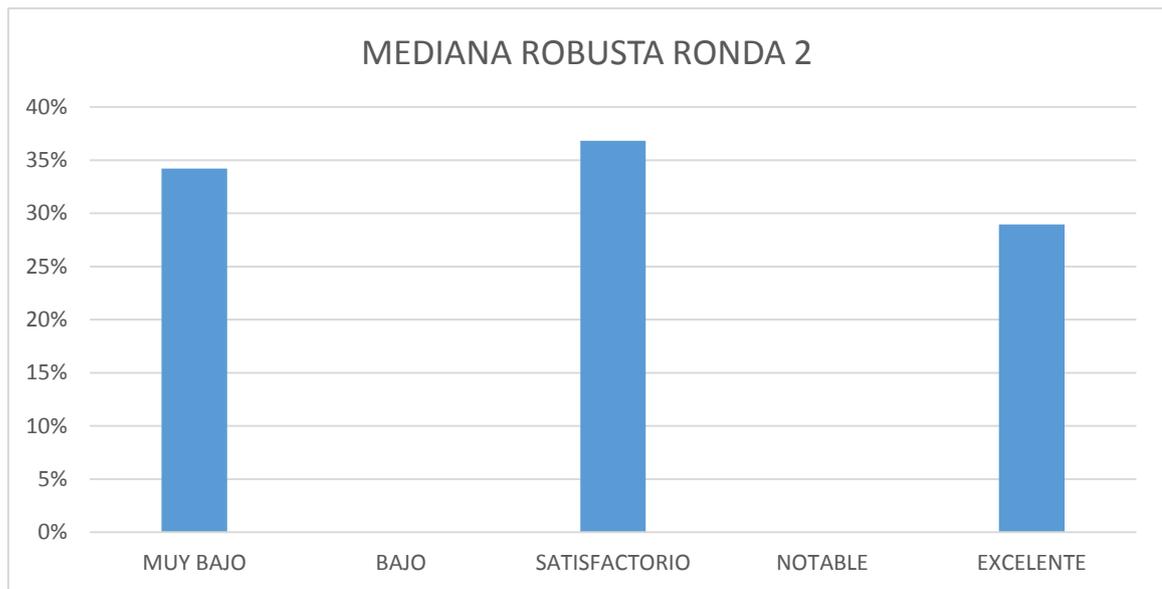


Gráfica 7: Distribución de los valores de la Mediana Robusta Ronda 1 Programa Q-Pap Morfológico. Ejercicio 2022

Tabla 19: Distribución mediana robusta por frecuencia y porcentajes. Ronda 2 Programa q-Pap Morfológico. Ejercicio 2022

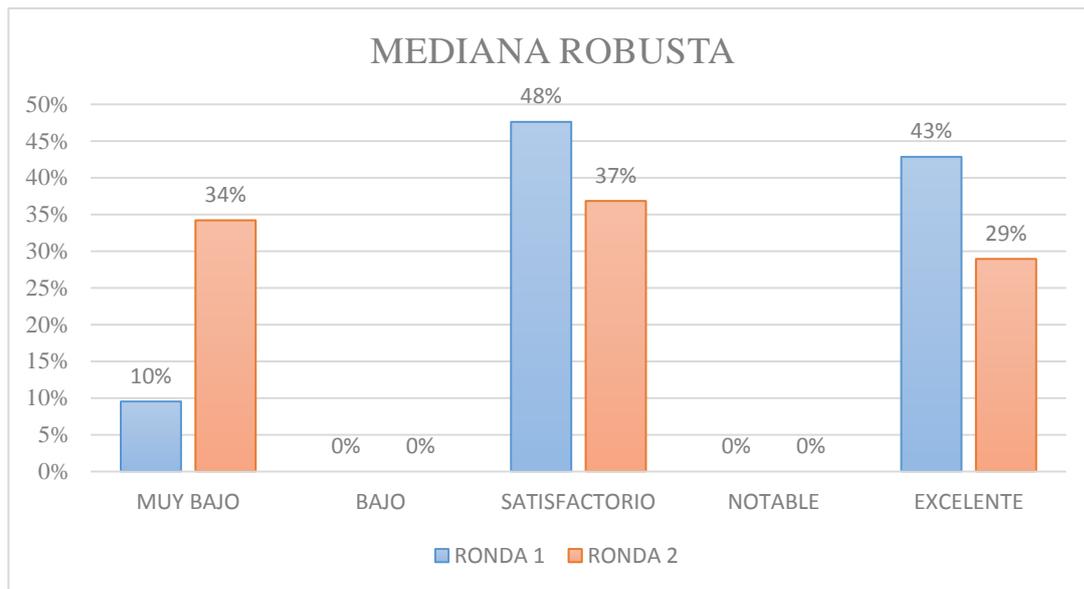
**MEDIANA ROBUSTA RONDA 2**

		<b>Frecuencia</b>	<b>%</b>
Válidos	Muy bajo	13	34
	Bajo	-	-
	Satisfactorio	14	37
	Notable	-	-
	Excelente	11	29
	<b>Total</b>	<b>38</b>	<b>100,0</b>



Gráfica 8: Distribución de los valores de la Mediana Robusta Ronda 2 Programa Q-Pap Morfológico. Ejercicio 2022

Como puede verse, tan solo hay tres valores extremos en la ronda 1 y representan un 6% del total de los participantes. El comportamiento del total de resultados nos confirma una distribución normal sin apenas valores extremos que distorsionen al conjunto y con la mayoría de respuestas en rango de excelente que puede servir de indicador para valorar la eficacia del diagnóstico en el cribado del cáncer de cérvix uterino en muestras cervico-vaginales obtenidas con citología, objetivo último de nuestro Programa de Intercomparación.



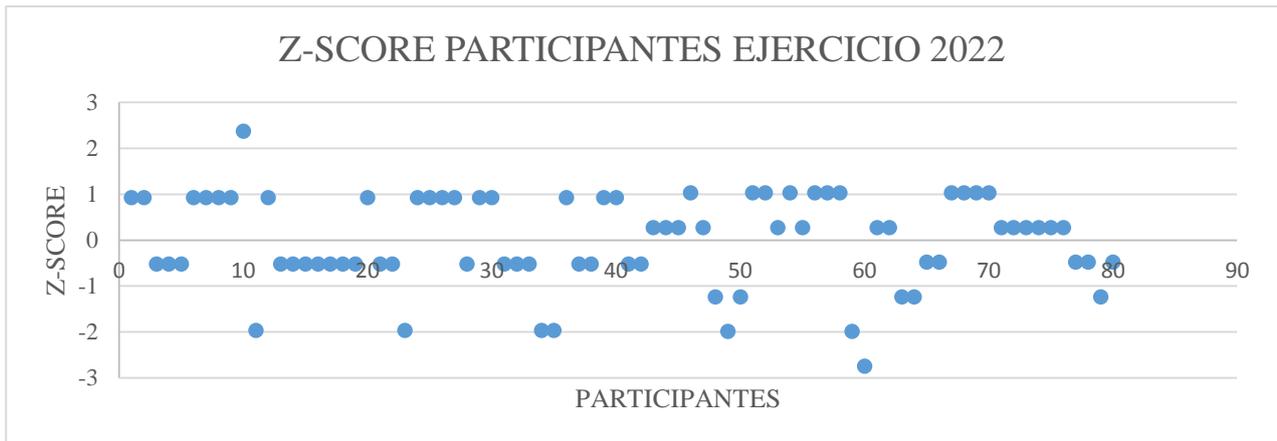
Gráfica 9: Distribución de los valores de la Mediana Robusta Total Ejercicio 2022 Programa Q-Pap Morfológico

La última medida de tendencia central es ZScore que nos aporta otro dato relevante. En estadística, la puntuación estándar es el número de desviaciones estándar por las que el valor de una puntuación bruta (es decir, un valor observado) está por encima o por debajo del valor medio de lo que se observa o mide. Las puntuaciones brutas por encima de la media tienen puntuaciones estándar positivas, mientras que las que están por debajo de la media tienen puntuaciones estándar negativas.

Este proceso de convertir un puntaje bruto en un puntaje estándar se denomina estandarización o normalización. Calcular una puntuación z requiere conocer la media y la desviación estándar de la población completa a la que pertenece un punto de datos. La distribución agrupada de ZScore por categorías queda representada en la tabla 19 y en la Gráfica 10:

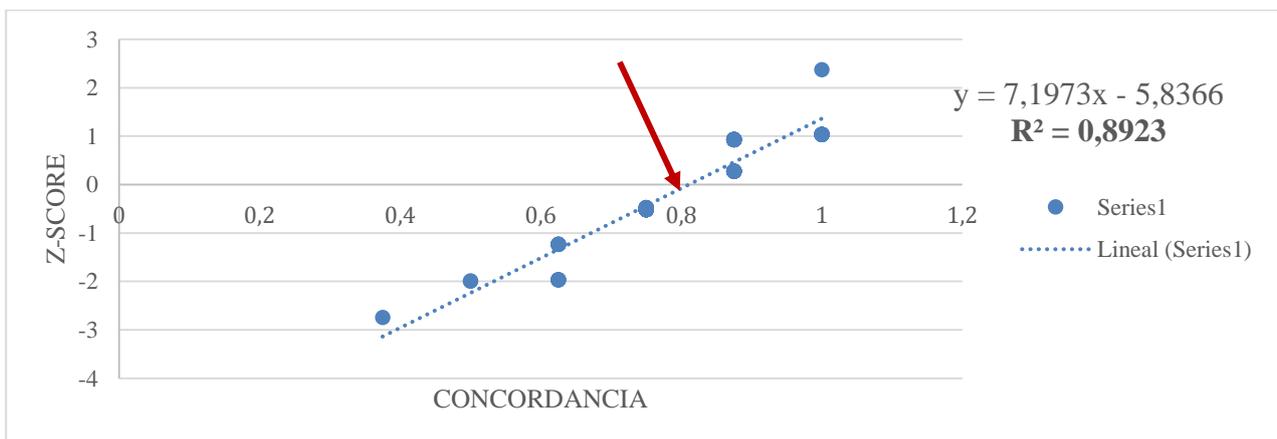
Tabla 19: Distribución Z Score kappa por categorías Rondas 1 y 2 Programa q-Pap Morfológico. Ejercicio 2022

Z-SCORE<0	Z-SCORE>=0
NO Satisf	Satisfac
37	43
46,25%	53,75%



Gráfica 10: Distribución de los valores ZScore atendiendo a resultados de los participantes.  
Programa Q-Pap Morfológico.Ejercicio 2022

Pero estos datos deben mostrarnos su validez a la hora de ser interpretados. El punto de referencia de aceptabilidad viene mostrado por el corte de la recta de regresión lineal con el eje de ordenadas. El modelo de regresión lineal proyecta una  $R^2 = 0,8923$ , lo que valida la correlación entre modelo y población (válido  $>0,7$ ) siendo el punto de corte del 81% (ver flecha Gráfica 11) que hace escorar la muestra a la derecha: valores  $>81\%$  de aciertos se consideran como satisfactorios y los inferiores a estos como no satisfactorios. La serie de este ejercicio 2022 debe ser calificada como **NOTABLEMENTE SATISFACTORIA** (Gráfica11)



Gráfica 11: Dispersión lineal ZScore atendiendo a resultados de concordancia de los participantes.  
Programa Q-Pap Morfológico.Ejercicio 2022

## **DATOS ESPECÍFICOS DEL PROGRAMA Q-PAP MOLECULAR**

### **Proceso de selección de las muestras**

El Servicio de Anatomía Patológica de Hospital del Mar (Laboratorio de Biología Molecular), actúa en el ejercicio 2022 como centro colaborador y experto. Se encuentra acreditado por UNE-EN ISO 15189 para el alcance: *Detección del Virus del Papiloma Humano para genotipos de alto riesgo*. Se encargó de elaborar cuatro *pooles* de casos, a partir de muestras obtenidas en medio líquido, con resultado previo conocido de HPV positivo alto riesgo incluyendo los genotipos 16, 18, o negativos.

La primera alícuota de cada una de las rondas que constituyen el programa, fue remitida por el Laboratorio de Biología Molecular del Hospital del Mar al Servicio de Anatomía Patológica del Hospital de Barbastro, centro que actúa como experto y se encuentra, también, acreditado por UNE-EN ISO 15189 para el alcance: *Test de VPH alto riesgo - Detección molecular y genotipado (HPV 16-18)*. Los resultados de ambos servicios expertos deben ser 100% concordantes.

Para el Programa de Control de Calidad de 2022 se incluyen un total de 8 muestras de citología ginecológica en medio líquido, repartidas en cada una de las rondas: 4 muestras por ronda. Las muestras están contenidas en viales de 4 µl con medio ThinPrep.

La cantidad de muestras (alícuotas) a preparar, por pool y para cada ronda, viene determinada por la siguiente ecuación:

#### **N+H+E+R**

**N:** número de participantes en la ronda

**H:** número de muestras para homogeneidad ( $H=3$ )

**E:** número de muestras para estabilidad ( $E=2$ )

**R:** número de muestras de reserva (*10% respecto del total de participantes*)

### **Estudio de homogeneidad y estabilidad**

Se determina según instrucciones de la norma ISO/IEC 17043:2010 identificadas en los apartados:

*4.4.3.4 Se debe demostrar que los ítems del ensayo de aptitud son suficientemente estables para asegurarse de que no sufrirán cambios significativos a lo largo de la realización del ensayo de aptitud, incluyendo las condiciones de almacenamiento y transporte.*

*4.4.3.5 Cuando se retienen para su posterior uso los ítems de ensayo de aptitud de rondas anteriores, el proveedor de ensayos de aptitud debe confirmar, antes de la distribución, los valores de la propiedad a determinar en el programa de ensayos de aptitud.*

En función del número total de muestras se seleccionan para el estudio de homogeneidad y estabilidad una muestra del mismo volumen y características.

Se establece como norma la temporalidad desde la recepción de la muestra hasta la comunicación del resultado, correspondiendo a un periodo de dos meses. De este modo, se toma

una muestra al inicio, una a mitad y otra al final, garantizando que el resultado sea el mismo en todos los periodos medidos.

Todos los resultados son canalizados y registrados por la Secretaría Técnica del Programa.

### Valor asignado y su incertidumbre

**Valor asignado:** Se determina según instrucciones de la norma ISO/IEC 17043:2010 en Anexo B apartado B.2.1 (d):

*“valores consensuados por participantes expertos –los expertos (que pueden, en algunos casos, ser laboratorios de referencia) deberán tener competencia demostrable en la determinación del o de los medidos a ensayar, utilizando métodos validados conocidos por ser altamente exactos y comprobables a los métodos utilizados habitualmente”.*

El Comité para el Programa de Control de Calidad de la SEC, selecciona para el ejercicio 2022 los siguientes casos (Tabla 20):

Tabla 21: Distribución casos-ronda seleccionados y valor de referencia  
Programa q-Pap Molecular. Ejercicio 2022

IDENTIFICACIÓN CASO Y RONDA	VALOR DE REFERENCIA
2022HPV_R1/Caso1	Negativo
2022HPV_R1/Caso2	Positivo para HPV alto riesgo
2022HPV_R1/Caso3	Positivo para HPV alto riesgo
2022HPV_R1/Caso4	Negativo
2022HPV_R2/Caso1	Negativo
2022HPV_R2/Caso2	Positivo para HPV alto riesgo
2022HPV_R2/Caso3	Negativo
2022HPV_R2/Caso4	Positivo para HPV alto riesgo

**Nivel de incertidumbre:** se calcula atendiendo a las indicaciones de la ISO/IEC Guía 98-3:2008:

*“Además, en muchas aplicaciones industriales y comerciales, así como en las áreas de la salud y de la seguridad, a menudo es necesario proporcionar un intervalo en torno al resultado de la medición, en el que se espera encontrar la mayor parte de valores de la distribución que pueden ser razonablemente atribuidos a la magnitud objeto de la medición. Por tanto, el método ideal para evaluar y expresar la incertidumbre de medida debería ser capaz de proporcionar fácilmente un intervalo, en particular, aquel con la probabilidad o el nivel de confianza que corresponda de manera realista con lo requerido”.*

Valoraremos la incertidumbre con los resultados pos prueba (Tabla 24)

### *Parámetro*

**Nivel de concordancia entre el valor asignado (valor de referencia) y el valor emitido (respuesta de un participante).**

Los participantes deben expresar sus diagnósticos con dos categorías:

- ✓ POSITIVO GENOTIPOS 16,18 y OTROS DE ALTO RIESGO
- ✓ NEGATIVO GENOTIPOS 16,18 y OTROS DE ALTO RIESGO

En el apartado *Observaciones*, el participante puede incluir comentarios adicionales que se considere adecuado indicar.

Una vez cerrada la ronda, los participantes podrán acceder de nuevo a los casos, por si desean revisar los casos discordantes.

Los resultados enviados fuera del plazo máximo no son evaluados.

### *Materiales y transporte*

Los materiales de ensayo se envían a los laboratorios participantes en la fecha de comienzo establecida en el programa. El envío se realiza por correo en condiciones de embalaje, almacenamiento e identificación de acuerdo a los **requerimientos de IATA sin necesidad de adoptar condiciones especiales.**

### *Instrucciones*

La Sociedad Española de Citología en calidad de proveedor del programa de control de calidad Q-PAP Molecular tiene registrada y accesible a todos los participantes y usuarios, a través de su página web (<http://calidad.secitologia.org/>), las instrucciones para la participación en el citado control (<http://calidad.secitologia.org/instrucciones-de-participacion-control-de-virus-papiloma-humano/>), que registra de forma precisa y adecuada el objeto del programa, material diagnóstico, envío de muestras, especificaciones de la realización y envío de resultados y calendario detallado del mismo.

Del mismo modo y previo al inicio de cada ronda del programa, desde la secretaría técnica se adjunta al envío de las muestras a cada participante inscrito, en documento impreso, las instrucciones para la participación en el control de calidad de HPV (Documento Anexo).

## Evaluación global del desempeño: Resultados

### Centros participantes y tasas de respuesta.

Tabla 22: Distribución inscripciones según número de instituciones

<b>Hospital Universitari Dexeus</b>
<b>Hospital Universitario Donostia</b>
<b>Hospital Universitario de Getafe</b>
<b>Parc Salut Mar de Barcelona</b>

El número total de participantes fue de cuatro centros o instituciones. De ellos, se identificó una tasa de respuesta del 100% de los participantes cumplimentando todos los casos.

### Asignación muestras q-PAP Molecular 2022

Tabla 23: Asignación muestras participantes inscritos q-Pap Molecular. Ejercicio 2022

IDENTIFICACIÓN CASO Y RONDA	DISTRIBUCIÓN	DESTINO
2022HPV_Ronda 1/Caso_061	Estudio Homogeneidad y Estabilidad	Hospital del Mar
2022HPV_Ronda 1/Caso_062	ID 202	Centro Participante
2022HPV_Ronda 1/Caso_063	ID 222	Centro Participante
2022HPV_Ronda 1/Caso_064	Estudio Homogeneidad	Hospital del Mar
2022HPV_Ronda 1/Caso_065	ID 248	Centro Participante
2022HPV_Ronda 1/Caso_066	ID 213	Centro Participante
2022HPV_Ronda 1/Caso_067	Reserva	Centro Participante
2022HPV_Ronda 1/Caso_068	Reserva	Centro Participante
2022HPV_Ronda 1/Caso_069	Estudio Homogeneidad y Estabilidad	Hospital del Mar
2022HPV_Ronda 2/Caso_070	Estudio Homogeneidad y Estabilidad	Hospital del Mar
2022HPV_Ronda 2/Caso_071	ID 202	Centro Participante
2022HPV_Ronda 2/Caso_072	ID 222	Centro Participante
2022HPV_Ronda 2/Caso_073	Estudio Homogeneidad	Hospital del Mar
2022HPV_Ronda 2/Caso_074	ID 248	Centro Participante
2022HPV_Ronda 2/Caso_075	ID 213	Centro Participante
2022HPV_Ronda 2/Caso_076	Reserva	Centro Participante
2022HPV_Ronda 2/Caso_077	Reserva	Centro Participante
2022HPV_Ronda 2/Caso_078	Estudio Homogeneidad y Estabilidad	Hospital del Mar

### Concordancia respecto al valor asignado y su incertidumbre

La concordancia de respuestas tanto en la primera, como en la segunda ronda fue del 100%, respecto a todos los valores de referencia y de todos los centros inscritos.

Tabla 24: Relación de concordancia de los casos evaluados en la Ronda 1 Programa q-Pap Molecular. Ejercicio 2022

Nº Caso	Valor de referencia	Concordancia	
		n/4	%
<b>Caso 1</b>	Negativo	4	100
<b>Caso 2</b>	Positivo para HPV alto riesgo	4	100
<b>Caso 3</b>	Positivo para HPV alto riesgo	4	100
<b>Caso 4</b>	Negativo	4	100
<b>Media</b>			100%

Tabla 25: Relación de concordancia de los casos evaluados en la Ronda 2, Programa q-Pap Molecular. Ejercicio 2022

Nº Caso	Valor de referencia	Concordancia	
		n/4	%
<b>Caso 1</b>	Negativo	4	100
<b>Caso 2</b>	Positivo para HPV alto riesgo	4	100
<b>Caso 3</b>	Negativo	4	100
<b>Caso 4</b>	Positivo para HPV alto riesgo	4	100
<b>Media</b>			100%

Tabla 26. Valores Ronda 1 y Ronda 2 Programa Q-PAP Molecular

Caso	Valor referencia	% Concordancia		
1	negativo alto riesgo	100%	<b>Promedio</b>	100%
2	positivo alto riesgo	100%	<b>Desv. Est</b>	0%
3	positivo alto riesgo	100%	<b>Za/2 (95%)</b>	1,96
4	negativo alto riesgo	100%	<b>Incertidumbre</b>	0%
5	negativo alto riesgo	100%	<b>LCS</b>	100%
6	positivo alto riesgo	100%	<b>LCI</b>	100%
7	negativo alto riesgo	100%		
8	positivo alto riesgo	100%		

### Valor de la determinación molecular de HPV de alto riesgo como prueba diagnóstica

Tabla 27. Resultados valoración determinación molecular HPV como prueba diagnóstica.  
Programa q-Pap Molecular Ejercicio 2022

Caso	Diagnóstico	Alto Riesgo +	%	Alto Riesgo -	%	Total	%
1	negativo	0	0	4	100	4	100
2	positivo	4	100	0	0	4	100
3	positivo	4	100	4	0	4	100
4	negativo	0	0	4	100	4	100
5	negativo	0	0	4	100	4	100
6	positivo	4	100	0	0	4	100
7	negativo	0	0	4	100	4	100
8	positivo	4	100	0	0	4	100

Determinación Molecular HPV alto riesgo	Positiva	Negativa	
Positiva	16 VP(a)	0 FP(b)	16 a+b
Negativa	0 FN(c)	16 VN(d)	16 c+d
	16 a+c	16 b+d	32 a+b+c+d

(a) Verdaderos positivos (VP): enfermos con la prueba positiva = 16

(b) Falsos positivos (FP): no enfermos con la prueba positiva = 0

(c) Falsos negativos (FN): enfermos con la prueba negativa = 0

(d) Verdaderos negativos (VN): no enfermos con prueba negativa = 16

a+c: casos con patrón de referencia positivo (enfermos) = 16

b+d: casos con patrón de referencia negativo (no enfermos) = 16

a+b: casos con prueba diagnóstica positiva = 16

c+d: casos con prueba diagnóstica negativa = 16

Sensibilidad (S) =  $a/(a+c) = 16/16 = 1$

Especificidad (E) =  $d/(b+d) = 16/16 = 1$

Falsos negativos =  $1-S = 0\%$

Falsos positivos =  $1-E = 0\%$

Valor Predictivo Positivo (VPP) =  $a/(a+b) = 1$

Valor Predictivo Negativo (VPN) =  $d/(c+d) = 1$

Cociente de probabilidades positivo (CPP) =  $S/(1-E) = \infty$  (no evaluable)

Cociente de probabilidades negativo (CPN) =  $(1-S)/E = 1-1/1 = 0$  (no evaluable)

Prevalencia o probabilidad pre prueba (Ppre) =  $(a+c)/(a+b+c+d) = 16/32 = 0,5$

Odds pre prueba =  $(Ppre)/(1-Ppre) = 0,5/1-0,5 = 1$

Odds post prueba = CPP x Odds pre prueba = no evaluable

Probabilidad post prueba (Ppost) =  $(\text{Odds post prueba} / (\text{odds post prueba} + 1)) = \text{no evaluable}$

*Valor de las respuestas respecto al valor asignado: Coeficiente Kappa de Cohen ( $k$ )*

No procede.

*Valor de las respuestas entre observadores: Mediana robusta y Z-Score Kappa.*

No procede.

## DISCUSIÓN

El programa de inter comparación para evaluar la eficacia del diagnóstico en el cribado del cáncer de cérvix uterino en muestras cervico-vaginales obtenidas con citología o *Programa Q-Pap de Control de Calidad de la Sociedad Española Citología*, puesto en marcha en 2019 presenta, en estos momentos, sus primeros datos significativos. Hay que tener presente que en su diseño se consideró un gran objetivo: *alcanzar un diagnóstico certero con un nivel tolerable de incertidumbre*, y la consecución de este objetivo debe incluir dos variables imprescindibles y relacionadas.

La primera variable, es el hecho en sí del diagnóstico. El diagnóstico médico convive, en su proceso dinámico, con la **toma de decisiones en presencia de incertidumbre**, e incierto es toda aquella situación que se deriva de la falta de certeza.

En lo que respecta a la citología, y su valor para vehiculizar un diagnóstico, la incertidumbre tiene un gran espectro donde actuar, y particularmente en la citología ginecológica por las diversas alteraciones estructurales secundarias a la infección por el virus del papiloma humano. En este punto concreto, la discusión debe plantearse sobre el valor diagnóstico de los resultados obtenidos en la determinación molecular del HPV frente a la alteración morfológica, y los distintos grados de incertidumbre que ambas determinaciones tienen.

La base Cochrane ofrece una revisión sobre todos los estudios relevantes hasta 2015<sup>15</sup> que nos permite documentar los niveles de evidencia de estas pruebas. Cuarenta estudios compararon la prueba de HPV con la prueba de Papanicolaou en más de 140.000 mujeres de entre 20 y 70 años que acudieron a consulta para un cribado cervical de rutina. Los estudios examinaron qué prueba puede detectar cambios cervicales precancerosos con CIN2+ y CIN3+. La calidad de evidencia de estos estudios fue calificada de moderada a alta, teniendo en cuenta que los estudios de Europa fueron más exactos que los de Asia o América Central o del Sur.

Una prueba diagnóstica perfecta es la que identifica correctamente a todos los pacientes positivos y a todos los negativos. La revisión citada determinó que por cada 1000 mujeres sometidas a cribado, alrededor de 20 tendrían cambios precancerosos. La prueba de HPV identificará a 16 de estas mujeres (aunque omitirá a 4) y la prueba de Papanicolaou identificará a 12 (pero omitirá a 8 mujeres). Por cada 1000 mujeres sometidas a cribado, habrá 980 que no tendrán cambios precancerosos. La prueba de HPV identificará correctamente a 879 mujeres (aunque a 101 se les dirá incorrectamente que presentan una lesión). La prueba de Papanicolaou identificará correctamente a 951 mujeres (aunque a 29 se les dirá incorrectamente que presentan lesión).

Las conclusiones de los autores de la revisión Cochrane son que, aunque las pruebas de HPV tienen menor probabilidad de omitir casos CIN2+ y CIN3+, estas pruebas dan lugar a más derivaciones innecesarias. Sin embargo, una prueba para HPV negativa es más tranquilizadora que

---

<sup>15</sup>Koliopoulos G, Nyaga VN, Santesso N, et al. Prueba del virus del papiloma humano (HPV) en comparación con prueba de Papanicolaou (Pap) para el cribado del cáncer de cuello de útero. En: Biblioteca Cochrane Plus, 2017. Oxford: Update software Ltd. Disponible en: <http://www.update-software.com>

una citología negativa, debido a que la citología tiene mayores posibilidades de presentar un falso negativo. Según esta fuente, se necesita evidencia de estudios longitudinales prospectivos para establecer las implicaciones clínicas relativas de estas pruebas, lo que abre un espectro amplio de investigación.

La segunda variable, tiene que ver con las **competencias para el desempeño**. La Organización Internacional del Trabajo (OIT) ha definido tres grandes grupos de competencias: aprender a pensar, aprender a hacer y aprender a ser, y se refiere al concepto de “competencia profesional” como la idoneidad para realizar una tarea o desempeñar un puesto de trabajo eficazmente por poseer las condiciones requeridas para ello. Tras esta definición de la OIT se plantean, desde nuestro punto de vista, al menos, dos cuestiones: ¿la competencia la otorga el puesto? y ¿el trabajador construye su competencia o es transferida desde la organización?

Para encontrar respuestas debemos, en primer lugar, diferenciar entre capacidad y competencia. La primera hace referencia a una atribución oficial otorgada: las instituciones académicas confieren la capacitación para el ejercicio de una profesión. La segunda, sin embargo, se refiere a las habilidades logradas. En definitiva, ¿qué debemos entender por competencia profesional? Utilizamos la definición que Núñez Martínez hace en su magnífico trabajo de revisión: “...conjunto de capacidades, habilidades y conocimientos que permite a las personas enfrentar de manera exitosa los dinámicos cambios que se producen en sus campos de acción y esferas de actuación laboral a tenor con los avances de la ciencia y la tecnología...”<sup>16</sup>.

En definitiva, si seguimos las palabras de Gallart<sup>17</sup>, podríamos decir que las competencias no son patrimonio del puesto de trabajo, sino que son atributos del trabajador, lo que nos introduce directamente en el concepto de *acreditación de la competencia*. Tradicionalmente, y en líneas generales, se entiende la acreditación de la competencia como la certificación de la superación de pruebas siempre en un tiempo, que suele ser pasado. Aunque más bien debemos aspirar a acreditar la calidad de la tarea de los profesionales sanitarios como “*un mecanismo para asegurar a la población atendida, que mantienen sus competencias y dar así garantía de su correcta actuación*”<sup>18</sup>, tal como dice Núñez Martínez.

Es necesario pues superar la confusión atemporal que se establece entre formación continuada y acreditación de la competencia. La diferencia entre ambos probablemente reside en que la acreditación de la competencia requiere de una estrategia, como sistema, que opere desde la determinación de necesidades hasta la evaluación y corrección. Roger Kaufman define este sistema como “...*proceso lógico mediante el cual se identifican necesidades, se seleccionan problemas, se determinan los requisitos para la solución de los mismos, se escogen soluciones entre las*

---

<sup>16</sup> Núñez Martínez MC, Llerena Bermudez FM. Competencias: modelos de diagnóstico y evaluación. Revista electrónica de las ciencias médicas en Cienfuegos. Medisur 2009;7(6).

<sup>17</sup> Gallart M, Jacinto C. Competencias laborales. Revista de educación Técnico y Profesional (Madrid) 2007;2(4):15-23

<sup>18</sup> Cf. Núñez Martínez. Competencias: modelos de diagnóstico y evaluación.

*alternativas, se obtienen y aplican métodos y medios, se evalúan resultados y se efectúan las revisiones que requiera todo o parte del sistema, de modo que se eliminen las carencias”<sup>19</sup>.*

Analizar necesidades de superación es un proceso complejo que implica explorar, conocer, identificar, categorizar, diagnosticar, priorizar y evaluar partiendo de los objetivos y recursos existentes. Una completa evaluación de necesidades, según Antonio Noguez, en su obra *Sistematización de la enseñanza, determinación de necesidades educativas*<sup>20</sup>, debe considerar los siguientes aspectos:

- Recoger datos reales
- No considerar a la evaluación de necesidades como un hecho único, a realizar una sola vez.
- Las discrepancias deben identificarse con los fines de la institución y no con sus medios, para no confundir la determinación de necesidades como discrepancias.
- Incluir, hasta donde sea posible, a todos los participantes del proceso.

Para hablar de la metodología seguiremos a Felipe Lara Rosario<sup>21</sup>. Propone un enfoque sistémico que sea acompañado de un proceso de determinación de necesidades sobre bases científicas para lo cual hay que crear una metodología y plantea que *“la justificación de esta metodología reside en la necesidad de superar el conocimiento vulgo o del sentido común y sustituirlo por un conocimiento rigurosamente científico de la realidad”*.

De los distintos modelos consultados. Inductivo (tipo I), Deductivo (tipo D), Tradicional (tipo T) creemos que el que mejor se ajusta a nuestra manera de trabajo es aquel en el que se involucran los participantes y se enfoca fundamentalmente a la evaluación.

Es el modelo CIPP de Stufflebean: Contexto, Insumo, Proceso y Producto nombrando el proceso como un programa de intervención. Evaluar es emitir juicios de valor que deben partir de la definición de los objetivos. Llegados a este punto debemos aportar la experiencia acumulada en INACEPS (Instituto para la Acreditación y Evaluación de las Prácticas Sanitarias de la Comunidad Valenciana) en donde la metodología utilizada estaba en consonancia a la propuesta por Michel Scriven<sup>22</sup> que también toma como criterios de valor las necesidades de los usuarios/consumidores. Establece tres fases para su proceso de evaluación que son:

- *Desarrollo de criterios/méritos justificables para que el programa funcione y pueda categorizarse como bueno.*
- *Fijar estándares/normas para cada criterio.*

---

<sup>19</sup>Kaufman,R. Vegara,A. Determinación de las necesidades de donde vienen los objetivos. Revista Tecnología Educativa (Buenos Aires) 2004;9(2):259

<sup>20</sup>Noguez, A. Sistematización de la enseñanza, determinación de necesidades educativas. México DF; Ed. Instituto Latinoamericano de comunicación Educativa; 1992.

<sup>21</sup> Lara Rosario, F. Metodología para la evaluación de sistemas: Un enfoque prospectivo. México DF: Prentice;1990

<sup>22</sup>Scriven M. Modelo de Evaluación. En: Castellanos Quintero S. Apuntes bibliográficos para la evaluación de programas educativos y educación comparada. Riobamba: CDES; 1997

- *Evaluar recogiendo información sobre el comportamiento de esos criterios, para determinar si se alcanzan o no los estándares de funcionamiento prefijados*<sup>23</sup>.

Llegados a este punto nos vemos con la obligación de asumir el compromiso de aportar al proceso de capacitación una evaluación de los resultados, o del impacto que mida los efectos producidos en la calidad del servicio y el grado de satisfacción tanto del personal involucrado como del usuario que lo recibe. Esta valoración, creemos del mismo modo que lo hace Núñez Martínez, “...permitirá reiniciar el ciclo que evolucionará en espiral ascendente, ante las crecientes demandas siempre insatisfechas”<sup>24</sup>.

Y en este momento es cuando aportamos parte de los resultados obtenidos en la secuencia de años 2020, 2021 y 2022 (leer los resultados en este orden) del programa de Control de Calidad de la SEC :

#### Q-Pap morfológico:

- El nivel de concordancia entre el valor asignado y las respuestas de los participantes en la secuencia de los tres ejercicios ha sido de 59,5%, 71,53% y 81,18%
- El nivel de discordancia: 40,5%, 28,47%, 18,82%
- Valor kappa medio: 0,44 (moderado), 0,64 (moderado) y 0,732 (bueno)
- Z-score kappa: 71% (2021) y 81% (2022)

#### Q-Pap molecular

- Sensibilidad: 93,3%, 100%, 100%
- Especificidad: 88,2%, 100%, 100%
- Valor predictivo positivo: 87,5%, 100%, 100%
- Valor predictivo negativo: 93,7%, 100%, 100%

A lo largo de estos años el programa ha realizado el ejercicio de intercomparación a 150 inscritos de 20 centros diferentes de España y Portugal y en una primera lectura estos resultados dan valor intrínseco a este ejercicio de intercomparación porque el nivel de competencia de los participantes ha aumentado y con ello la mejora en el diagnóstico con un nivel de incertidumbre tolerable.

Tal vez, en este momento podemos decir que habría valido la pena acometer este proyecto del programa de intercomparación para evaluar la eficacia del diagnóstico en el cribado del cáncer de cérvix uterino en muestras cervico-vaginales obtenidas con citología o *Programa Q-Pap de Control de Calidad de la Sociedad Española Citología* iniciado en 2019 como una herramienta puesta al servicio de la sociedad en la mejora de la práctica médica.

---

<sup>23</sup> Martínez Lorente A, Aranda López I, Escandón Álvarez J, et al. Manual de Autoevaluación (EFQM) Cómo otorgar crédito a nuestra calidad. *Grupo de trabajo de la Conselleria de Sanitat para la elaboració del Manual para la Acreditación de los Servicios de Anatomía Patológica. Dirección General de Ordenación, Evaluación, Investigación, Calidad y Atención al Paciente de la Conselleria de Sanitat. Generalitat Valenciana (INACEPS 2007-2011) XXI Congreso Nacional de la SEAP. Cádiz 2013.*

<sup>24</sup> Cf. Núñez Martínez. Competencias: modelos de diagnóstico y evaluación

## CONCLUSIONES

### CUADRO DE INDICADORES PARA EL PROGRAMA q-PAP DE CONTROL DE CALIDAD DE LA SOCIEDAD ESPAÑOLA DE CITOLOGÍA. EJERCICIO 2022

Tabla 28. Cuadro de indicadores ejercicio 2022 de q-PAP Morfológico y q-PAP Molecular

Indicador	q-PAP Morfológico	q-PAP Molecular
Nº de Centros	12	4
Nº de Inscritos	42	4
Nº de Ítems	16	8
Grado Incertidumbre pre prueba	16,99%	No determinado
Grado Incertidumbre post prueba	1%	-
Sensibilidad	No determinado	100%
Especificidad	No determinado	100%
Valor predictivo positivo	No determinado	100%
Valor predictivo negativo	No determinado	100%
Nivel de Concordancia	81,18%	100%
Nivel de Discordancia	18,82%	0
Probabilidad de aleatoriedad	29,75%	0
Rango Kappa	0,31-1 (intervalo)	No procede
Valor Kappa medio	0,732 (grado de concordancia buena)	No procede
Mediana Robusta	75% (ronda 1) 87,5% (ronda 2)	No procede
Z Score kappa	81% notablemente satisfactorio	No procede

#### Indicadores para la evaluación de prueba diagnóstica

Grado Incertidumbre pre prueba
Grado Incertidumbre post prueba
Sensibilidad
Especificidad
Valor predictivo positivo
Valor predictivo negativo

#### Indicadores para la evaluación del nivel de desempeño

Nivel de Concordancia
Nivel de Discordancia
Probabilidad de aleatoriedad
Rango Kappa
Valor Kappa medio
Mediana Robusta
Z Score kappa

## BIBLIOGRAFÍA

ALAMEDA F, ASO S, CATALINA I, et al. Control de calidad de la citología ginecológica. Programa de Calidad de la Sociedad Española de Citología. Resultados de la segunda ronda. Rev. Esp. Patol. 2018; 51:71-76

ALAMEDA F, BERNET L, CANO R, et al. Control de calidad de la citología ginecológica. Programa de Calidad de la Sociedad Española de Citología. Resultados de la primera ronda. Rev. Esp. Patol. 2017; 50:154-160

ALEMANY GARCIA, Macario. El concepto y la justificación del paternalismo. Tesis de doctorado. Departamento de Filosofía del Derecho. Universidad de Alicante 2005.

BEAUCHAMP, TOM L. Y JAMES F. CHILDRESS: Principles of Biomedical Ethics, ed. Oxford University Press, Nueva York, (ed 2001)

CULVER, CHARLES M. Y BERNARD GERT: Philosophy in Medicine (Conceptual and Ethical Issues in Medicine and Psychiatry), ed. Oxford University Press, Nueva York, 1982

ENAC. Política de ENAC sobre Intercomparaciones. NT-03. Re. 7 abril 2021. Series 1,2 y 3. Disponible en:  
<https://www.enac.es/documents/7020/d3b14f16-55ed-424f-b12c-674764c1d8e8>

ENAC. Guía sobre la participación en programas de intercomparaciones. G-ENAC-14. Rev. 1 Septiembre de 2008. Series 1, 2, 3 y 10. Disponible en:  
<https://www.enac.es/documents/7020/55d344fa-35e1-4320-9cba-307e0f3704a1>

FERLAY J, COLOMBET M, SOERJOMATARAM I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int. J. Cancer. 2019;144(8):1941-1953

GALLART M, JACINTO C. Competencias laborales. Revista de educación Técnico y Profesional (Madrid) 2007;2(4):15-23

GARCÍA A. Métodos avanzados de estadística aplicada. Madrid: UNED, 2001

ISO 15189:2012. Laboratorios clínicos. Requisitos particulares para la Calidad y la competencia

KAUFMAN, R. VEGARA, A. Determinación de las necesidades de donde vienen los objetivos. Revista Tecnológica Educativa (Buenos Aires) 2004;9(2):259

KOLIOPOULOS G, NYAGA VN, SANTESSO N, et al. Prueba del virus del papiloma humano (HPV) en comparación con prueba de Papanicolaou (Pap) para el cribado del cáncer de cuello de útero. En: Biblioteca Cochrane Plus, 2017. Oxford: Update software Ltd. Disponible en: <http://www.update-software.com>

LARA ROSARIO, F. Metodología para la evaluación de sistemas: Un enfoque prospectivo. México DF: Prentice;1990

<sup>1</sup> Culver, Charles M. y Bernard Gert: Philosophy in Medicine (Conceptual and Ethical Issues in Medicine and Psychiatry), ed. Oxford University Press, Nueva York, 1982

LÓPEZ DE ARGUMEDO GONZÁLEZ DE DURANA M, BAYÓN YUSTA JC, MATEOS DEL PINO M. Impacto de la implantación de un programa de cribado poblacional de cáncer de cérvix, siguiendo las recomendaciones europeas (prueba/intervalo) en relación a la situación actual. Informes de Evaluación de Tecnologías Sanitarias: OSTEBA. Ministerio de Sanidad, Servicios Sociales e Igualdad. Servicio de Evaluación de Tecnologías Sanitarias del País Vasco; 2016. Disponible en:

[http://www.ogasun.ejgv.euskadi.eus/r51-catpub/es/k75aWebPublicacionesWar/k75aObtenerPublicacionDigitalServlet?R01HNoPortal=true&N\\_LIBR=051788&N\\_EDIC=0001&C\\_IDIOM=es&FORMATO=.pdf](http://www.ogasun.ejgv.euskadi.eus/r51-catpub/es/k75aWebPublicacionesWar/k75aObtenerPublicacionDigitalServlet?R01HNoPortal=true&N_LIBR=051788&N_EDIC=0001&C_IDIOM=es&FORMATO=.pdf)

MARTÍNEZ LORENTE A, ARANDA LÓPEZ I, ESCANDÓN ÁLVAREZ J, et al. Manual de Autoevaluación (EFQM) Cómo otorgar crédito a nuestra calidad. *Grupo de trabajo de la Conselleria de Sanitat para la elaboració del Manual para la Acreditació de los Servicios de Anatomía Patológica. Direcció General de Ordenació, Evaluació, Investigació, Calidad y Atención al Paciente de la Conselleria de Sanitat. Generalitat Valenciana (INACEPS 2007-2011) XXI Congreso Nacional de la SEAP. Cádiz 2013.*

NOGUEZ, A. Sistematización de la enseñanza, determinación de necesidades educativas. México DF; Ed. Instituto Latinoamericano de comunicación Educativa; 1992.

NÚÑEZ MARTÍNEZ MC, LLERENA BERMUDEZ FM. Competencias: modelos de diagnóstico y evaluación. Revista electrónica de las ciencias médicas en Cienfuegos. Medisur 2009;7(6).

ORDEN SCB/480/2019, de 26 de abril, por la que se modifican los anexos I, III y VI del Real Decreto 1030/2006, de 15 de septiembre, por el que se establece la cartera de servicios comunes del Sistema Nacional de Salud y el procedimiento para su actualización. Boletín Oficial del Estado (BOE) Núm. 101 Sec. I. Pág.: 43018-43028. Disponible en: <https://www.boe.es/boe/dias/2019/04/27/pdfs/BOE-A-2019-6277.pdf>

PATRONES DE MORTALIDAD EN ESPAÑA 2016. Información y estadísticas sanitarias 2019. Ministerio de Sanidad, Servicios Sociales e Igualdad. Disponible en: [https://www.mscbs.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/mortalidad/docs/Patrones\\_Mortalidad\\_2016.pdf](https://www.mscbs.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/mortalidad/docs/Patrones_Mortalidad_2016.pdf)

PUBLIC HEALTH. Propuesta de Recomendación del Consejo sobre cribado de cáncer., Bruselas: CE; 2003. Disponible en: <http://ec.europa.eu/health/>

RAMALLE-GÓMARA E, et al. Utilización de métodos robustos en la estadística inferencial. Aten Primaria 2003;32(3):177-82

SCRIVEN M. Modelo de Evaluación. En: Castellanos Quintero S. Apuntes bibliográficos para la evaluación de programas educativos y educación comparada. Riobamba: CDES; 1997

TORNÉ BLADÉ A, DEL PINO SALADRIGUES M, CUSIDÓ GIMFERRER M, et al. Guía de cribado del cáncer de cuello de útero en España, 2014. Rev EspPatol. 2014;47 (Supl 1):1-43

WORLD HEALTH ORGANIZATION (OMS), Preventing chronic diseases: a vital investment. WHO global report. Ginebra, Suiza: WHO, 2005:1107. Disponible en: [https://apps.who.int/iris/bitstream/handle/10665/43314/9241563001\\_eng.pdf;jsessionid=2DF61BE6F1F18A3149B75F489A706D90?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/43314/9241563001_eng.pdf;jsessionid=2DF61BE6F1F18A3149B75F489A706D90?sequence=1)

YOUNG P, FINN BC, BRUETMAN JE, et al. El hombre y sus descripciones. Revista médica de Chile. Sociedad Médica de Santiago. 2012; 140(9): 1218-27

## INDICE DE TABLAS y GRÁFICAS

### TABLAS

<b>Tabla 1.</b> Calendario Programa Control de Calidad Diagnóstico de Citología Ginecológica: q-PAP 2022	10
<b>Tabla 2:</b> Valoración estandarizada del índice Kappa	11
<b>Tabla 3.</b> Valoración estandarizada del test de la Mediana Robusta	11
<b>Tabla 4:</b> Valoración estandarizada de Z Score kappa	11
<b>Tabla 5:</b> Clasificación de la potencia de una prueba diagnóstica según los valores de los cocientes de probabilidades	12
<b>Tabla 6:</b> Distribución cuadro de indicadores q-PAP según medida y definición	14
<b>Tabla 7:</b> Distribución casos ejercicio 2022 según historia clínica, identificación caso, ronda y resultado	16
<b>Tabla 8:</b> Tabla simulación calculado grado concordancia, desviación estándar, IC95%, la incertidumbre, el límite de control superior e inferior.	18
<b>Tabla 9:</b> Distribución del valor de referencia según grado de las distintas categorías diagnósticas	19
<b>Tabla 10:</b> Distribución inscripciones según número de instituciones	21
<b>Tabla 11:</b> Distribución concordancia Ronda 1	22
<b>Tabla 12:</b> Distribución concordancia Ronda 2.	20
<b>Tabla 13:</b> Distribución concordancia según valor de referencia del total de casos en ambas rondas.	24
<b>Tabla 14:</b> Distribución de no concordancias según categoría del resultado en la Ronda 1	24
<b>Tabla 15:</b> Distribución de no concordancias según categoría del resultado en la Ronda 2	25
<b>Tabla 16:</b> Valoración estandarizada del índice Kappa	26
<b>Tabla 17:</b> Valoración rango kappa, valor kappa medio y grado de concordancia. Rondas 1 y 2	29
<b>Tabla 18:</b> Distribución mediana robusta por frecuencia y porcentajes ronda 1	30
<b>Tabla 19:</b> Distribución mediana robusta por frecuencia y porcentajes ronda 2	30
<b>Tabla 20:</b> Distribución Z Score kappa por categorías. Rondas 1 y 2	33
<b>Tabla 21:</b> Distribución casos-ronda seleccionados y valor de referencia. Programa q-Pap Molecular. Ejercicio 2022	33
<b>Tabla 22:</b> Distribución inscripciones según número de instituciones	38
<b>Tabla 23:</b> Asignación muestras participantes inscritos q-Pap Molecular	38
<b>Tabla 24:</b> Relación de concordancia de los casos evaluados en la Ronda 1	39
<b>Tabla 25:</b> Relación de concordancia de los casos evaluados en la Ronda 2	39
<b>Tabla 26:</b> Valores Ronda 1 y Ronda 2 Programa q-Pap molecular	39
<b>Tabla 27:</b> Resultados valoración determinación molecular HPV como prueba diagnóstica. Programa q-Pap Molecular Ejercicio 2022	40

<b>Tabla 28:</b> Cuadro de indicadores ejercicio 2022 de q-PAP Morfológico y q-PAP Molecular	46
--	----

## GRÁFICAS

<b>Gráfica 1:</b> Distribución grado concordancia y no concordancia Rondas 1 y 2	23
<b>Gráfica 2:</b> Distribución de no concordancias según categoría de resultado. Ronda 1. Programa Q-Pap Morfológico. Ejercicio 2022	25
<b>Gráfica 3:</b> Distribución de no concordancias según categoría de resultado. Ronda 2 Programa Q-Pap Morfológico. Ejercicio 2022	26
<b>Gráfica 4:</b> Distribución del valor kappa Ronda 1 Programa Q-Pap Morfológico. Ejercicio 2022	27
<b>Gráfica 5:</b> Distribución del valor kappa Ronda 2 Programa Q-Pap Morfológico. Ejercicio 2022	28
<b>Gráfica 6:</b> Distribución grado concordancia inter-observador para la Ronda 1 y 2, según valoración estandarizada del índice Kappa.	29
<b>Gráfica 7:</b> Distribución de los valores de la Mediana robusta Ronda 1 Programa Q-Pap Morfológico. Ejercicio 2022	31
<b>Gráfica 8:</b> Distribución de los valores de la Mediana Robusta Ronda 2 Programa Q-Pap Morfológico. Ejercicio 2022	32
<b>Gráfica 9:</b> Distribución de los valores de la Mediana Robusta total ejercicio 2022 Programa Q-Pap Morfológico.	30
<b>Gráfica 10:</b> Distribución de los valores ZScore atendiendo a resultados de los participantes. Programa Q-Pap Morfológico. Ejercicio 2022	34
<b>Gráfica 11:</b> Dispersión lineal ZScore atendiendo a resultados de concordancia de los participantes. Programa Q-Pap Morfológico. Ejercicio 2022	34

---

## ANEXOS

---

## ***ACTUACIONES A LLEVAR A CABO ANTE LA OBTENCIÓN DE RESULTADOS NO SATISFATORIOS. BENEFICIOS DE LA PARTICIPACIÓN EN UN PROGRAMA DE INTERCOMPARACIÓN<sup>25</sup>.***

Si, tras llevar a cabo la evaluación de acuerdo con el apartado anterior (prestando especial atención a que la evaluación realizada haya sido adecuada al fin pretendido), el laboratorio llega a la conclusión de que el ejercicio de inter comparación es correcto y aun así ha obtenido resultados no satisfactorios, debe realizar una investigación que identifique y explique las causas, para poder así establecer las acciones que eviten la repetición de los problemas encontrados.

La investigación ante un resultado no satisfactorio debería incluir distintas etapas como:

- Comprobar que no se produjeron errores en la expresión de resultados (unidades, adiciones, diluciones...), transcripción de los datos, etc.
- Comprobar que se siguieron las instrucciones del organizador relativas a conservación y manipulación del ítem.
- Verificar que las medidas se realizaron siguiendo el procedimiento que se pretende evaluar.
- Comprobar que el instrumento con el que se realizó la medida funcionó correctamente (Ej.: a partir de una comparación de datos como la sensibilidad con respecto a otros días).
- Comprobar que no se produjeron contaminaciones provenientes del material utilizado, del ambiente, etc.
- Comprobar los resultados de los controles de calidad internos del método de medida (material de referencia, muestra duplicada, controles entre calibraciones...) y su evolución.
- Comprobar resultados de otros ítems similares ensayadas el mismo día que permita descartar, por ejemplo, falsos negativos/positivos, problemas de contaminación, etc.
- Comprobar otros parámetros ensayados por el mismo método, día, operador (Ej.: Si se trata de un ensayo multiparamétrico y únicamente existe un resultado no satisfactorio en uno de ellos, las causas pueden ser diferentes que si todos los resultados son no satisfactorios).
- Verificar si pueden haber afectado al resultado del laboratorio las incidencias o particularidades señaladas por organizador en el informe.
- Comprobar los resultados obtenidos en anteriores ejercicios de inter comparación

Si a pesar de realizar las comprobaciones anteriores el laboratorio, no encuentra el origen del problema, se deberán llevar a cabo estudios más exhaustivos, siempre que sea posible sobre el mismo ítem, que pueden consistir en:

- Si es posible, repetir el ensayo. Si en este segundo ensayo, el resultado es satisfactorio, se podría considerar un error puntual, que el control de calidad interno no hubiera detectado. No obstante, en estos casos, el laboratorio deberá verificar que no se han producido cambios

<sup>25</sup> ENAC. Guía sobre la participación en programas de inter comparaciones. G-ENAC-14. Rev. 1 Septiembre de 2008. Series 1, 2, 3 y 10.

en las condiciones de ensayo (equipos, reparaciones, personal, etc.) que pudieran haber causado el problema. Además, será necesario realizar un seguimiento de resultados posteriores para asegurar que el problema no se vuelva a repetir.

- Realizar modificaciones sobre el método de ensayo o algunas de las etapas claves del mismo (Ej.: digestión, mecanización, tratamiento térmico, extracción, etc).

Una vez considerados todos los aspectos y etapas anteriores, el laboratorio debe establecer acciones correctoras, para solucionar el problema y evitar que se repita en el futuro. Así mismo se deberá valorar la eficacia de dichas acciones correctoras, por ejemplo, mediante la participación en otro ejercicio de inter comparación.

Entre los **beneficios** que los laboratorios pueden obtener de la participación sistemática en ejercicios de inter comparación, pueden destacarse los siguientes:

- Permite confirmar la correcta validación inicial de un método por medio de la comparación de los resultados de exactitud y precisión con los obtenidos en la validación. La participación en inter comparaciones nos permite poder evaluar si la incertidumbre estimada por el laboratorio es realista.
- Así mismo los datos obtenidos de la participación en inter comparaciones pueden ser utilizados para la validación de métodos de medida, siempre y cuando los tipos de ítems y los intervalos de trabajo sean coherentes con el objetivo de la validación.
- Se pueden determinar errores sistemáticos que se producen en el método de medida, por ejemplo, analizando la tendencia de los resultados del rendimiento, se puede identificar la existencia de algún sesgo en el método.
- Permite llevar a cabo mejoras en el método por comparación con otros laboratorios.
- Obtener información de los métodos aplicados por otros participantes, nuevas metodologías, uso de nuevos materiales de referencia, comparación de métodos disponibles en el mercado, etc...
- Permite realizar un seguimiento del comportamiento del método, respecto a la exactitud y precisión del mismo.
- Fomentar la colaboración e intercambio de experiencias entre laboratorios.
- Demostración de competencia técnica frente a terceros

## CRITERIOS PARA LA EVALUACIÓN DE DISCREPANCIAS DIAGNÓSTICAS EN CITOLOGÍA GINECOLÓGICA

### q-PAP Morfológico

Los resultados diagnósticos emitidos por los participantes, en el Programa de Control de Calidad de SEC, se expresan en base a los criterios BETHESDA 2014.

<b>Muestra insatisfactoria</b>
<b>Negativo</b>
<b>ASC-US</b>
<b>ASC-H</b>
<b>L-SIL</b>
<b>H-SIL</b>
<b>Carcinoma escamoso (CAESC)</b>
<b>ACG</b>
<b>Células glandulares atípicas</b>
<b>AIS Adenocarcinoma In Situ</b>
<b>Adenocarcinoma (ADC)</b>

La evaluación de los resultados es realizada por el Programa de Control de Calidad de SEC, en base a criterios de Concordancia/No Concordancia con los resultados de referencia, que son emitidos por expertos acreditados por la norma UNE-EN ISO 15189.

En caso de obtener resultados diagnósticos no concordantes, se establecen categorías en función de su relevancia diagnóstica, teniendo en cuenta la actitud terapéutica posterior sobre las pacientes.

Las categorías se clasifican en:

GRAVE	MAYOR	INTERMEDIA	MENOR
-------	-------	------------	-------

A continuación, se describen las posibles discrepancias y su categoría:

#### NEGATIVO VERSUS INSATISFACTORIAS

Las muestras se catalogan como insatisfactorias cuando las preparaciones con células epiteliales bien preservadas y visualizadas en conjunto cubren menos del 10 % de la superficie de la lámina o bien las células no pueden ser evaluadas por la presencia de abundante material inflamatorio, moco u otro material. Dado que se utiliza citología líquida, puede obtenerse una segunda extensión y valorarla. Si la extensión sigue siendo insuficiente, debe citarse a la paciente y proceder a una nueva toma. La ausencia de células endocervicales o zona de transformación ya no incluye la muestra en categoría insatisfactoria.

Por tanto, la discrepancia, con cualquiera de los diagnósticos patológicos debe ser considerada mayor o grave, dado que se puede perder a la paciente durante al menos 3 años. La discrepancia con un diagnóstico negativo debe ser clasificada como intermedia por el hecho de que no se hayan tenido en cuenta los criterios de Bethesda para evaluar las extensiones insatisfactorias.

### **NEGATIVO VERSUS ASCUS**

El concepto de ASCUS es una atipia en células escamosas de significado incierto. En las recomendaciones actuales, cuando se informa una citología como negativa, la paciente debe ser controlada a los 3 años. Cuando se informa una citología como ASCUS debe practicarse un test de HPV y la paciente debe ser controlada a los 6 meses. Por tanto, una discrepancia diagnóstica de NEGATIVO versus ASCUS, implica que la paciente no será controlada antes de los 3 años. Asociados al diagnóstico de ASCUS, existe entre un 5 y un 10% de casos de lesiones de alto grado (HSIL). En consecuencia, la paciente que padece un ASCUS y no se diagnostica, tiene un riesgo aumentado de desarrollar HSIL-CARCINOMA, que la paciente con un diagnóstico de negatividad. Se clasifica, por tanto, como discrepancia MAYOR.

### **NEGATIVO VERSUS LSIL**

El concepto de LSIL, es una lesión en células escamosas, de bajo grado. En las recomendaciones actuales, la paciente debe ser sometida a control en los meses posteriores. No se indica un test de HPV ya que aproximadamente el 85% de las pacientes muestran positividad para el test, y este test no discrimina dos poblaciones con distinto riesgo. Según la bibliografía, aproximadamente en el 15% de las pacientes con LSIL, existe una lesión histológica de HSIL. En consecuencia, la paciente que padece un LSIL y no se diagnostica citológicamente, tiene un riesgo mayor de desarrollar HSIL-CARCINOMA, que la paciente con un diagnóstico de negatividad. Se clasifica, por tanto, como discrepancia MAYOR.

### **NEGATIVO VERSUS ASCH**

El concepto de ASCH, es una atipia en células escamosas sugestiva, o que no puede excluir, lesión de alto grado. En las recomendaciones actuales se indica colposcopia, mejor con un test de HPV previo ya que en todos los casos de ASCH con test de HPV negativo, la lesión desaparece. En aproximadamente la tercera parte (33%) de los casos de ASCH, se esconde un HSIL. En consecuencia, la paciente que padece un ASCH y no se diagnostica, tiene un riesgo mayor de desarrollar HSIL-CARCINOMA, que la paciente con un diagnóstico de negatividad. Se clasifica, por tanto, como discrepancia GRAVE.

### **NEGATIVO VERSUS HSIL**

El concepto de HSIL, es una lesión escamosa intraepitelial de alto grado, es decir prácticamente una neoplasia intraepitelial. La indicación es colposcopia. No está indicado el test de HPV, ya que prácticamente todas estas pacientes tendrán test de HPV positivo. No todas estas lesiones evolucionan a cáncer invasivo, pero no se puede distinguir entre las que van a desarrollarlo y aquellas que no progresan. En consecuencia, la paciente que padece un HSIL debe diagnosticarse mediante biopsia y tratarse quirúrgicamente. La discrepancia con negatividad se clasifica como GRAVE.

### **NEGATIVO VERSUS CARCINOMA (CARCINOMA ESCAMOSO O ADENOCARCINOMA)**

La imagen citológica de carcinoma, ya sea adenocarcinoma o carcinoma escamoso, es muy típica. La indicación es colposcopia y biopsia. No está indicado el test de HPV porque prácticamente el 100% de los carcinomas escamosos y el 90% de los adenocarcinomas del cérvix (no así los de endometrio) son HPV positivos. La discrepancia es un error GRAVE, dado que el no tratar una lesión neoplásica cuando se diagnostica, implica que esta lesión sigue evolucionando y que en su evolución pone en peligro la vida de la paciente.

**NEGATIVO VERSUS ADENOCARCINOMA IN SITU**

La situación es similar al caso de Negativo versus HSIL. La discrepancia debe clasificarse como GRAVE.

**NEGATIVO VERSUS ACG**

La situación es similar al caso del ASCUS. La discrepancia debe clasificarse como GRAVE.

**ASCUS VERSUS LSIL**

La diferencia es el riesgo de desarrollo de lesión de alto grado, pero la actuación clínica es similar, de manera que la discrepancia debe clasificarse como MENOR.

**ASCUS VERSUS ASCH**

La diferencia es el riesgo de desarrollo de lesión de alto grado y la actuación clínica es distinta, de manera que con el diagnóstico de ASCUS se practica control a los 6 meses y con el diagnóstico de ASCH debe realizarse colposcopia. En cualquier caso, no se pierde el control de la paciente, pero dada la diferencia en la actuación clínica, la discrepancia debe clasificarse como INTERMEDIA.

**ASCUS VERSUS HSIL**

La diferencia es el riesgo de desarrollo de lesión de alto grado y la actuación clínica, de manera que con el diagnóstico de ASCUS se practica control a los 6 meses, y con el diagnóstico de HSIL debe realizarse colposcopia. En cualquier caso, no se pierde el control de la paciente, pero dada la diferencia en la actuación clínica, la discrepancia debe clasificarse como INTERMEDIA.

**ASCUS VERSUS CARCINOMA (CARCINOMA ESCAMOSO O ADENOCARCINOMA IN SITU O INVASOR)**

La diferencia es el riesgo de desarrollo de lesión de alto grado. La actuación clínica es distinta, de manera que con el diagnóstico de ASCUS se practica control a los 6 meses, y con el diagnóstico de carcinoma debe realizarse siempre colposcopia y biopsia. En cualquier caso, no se pierde el control de la paciente, pero dada la diferencia en la actuación clínica, la discrepancia debe clasificarse como INTERMEDIA.

**ASCUS VERSUS ACG**

La situación y la indicación clínica son similar. En consecuencia, la discrepancia debe clasificarse como MENOR.

**LSIL VERSUS ASCH**

Existen diferencias en el riesgo de desarrollo de lesión de alto grado. Además, la indicación clínica es diferente: seguimiento a los 6 meses en LSIL y colposcopia en HSIL. No se pierde a la paciente. La discrepancia debe considerarse como INTERMEDIA.

**LSIL VERSUS HSIL, CARCINOMA (CARCINOMA ESCAMOSO, ADENOCARCINOMA, O ADENOCARCINOMA IN SITU)**

Las diferencias diagnósticas son el diagnóstico de una lesión pre-neoplásica (LSIL), o el diagnóstico de cáncer. La indicación clínica es diferente: seguimiento a los 6 meses en LSIL y colposcopia en el resto. No se pierde a la paciente. La discrepancia debe considerarse como INTERMEDIA.

### **LSIL VERSUS ACG**

Como se ha indicado anteriormente, la situación y la indicación clínica son algo distintas, por el riesgo de desarrollo de lesión de alto grado. La actuación clínica es similar. En consecuencia, la discrepancia debe clasificarse como INTERMEDIA.

### **ASCH VERSUS HSIL**

En un 33% de casos diagnosticados como ASCH existe un HSIL. La actuación clínica es la misma: colposcopia y biopsia. Por tanto, la discrepancia debe ser considerada como MENOR.

### **ASCH VERSUS ACG**

El riesgo de lesión de alto grado es diferente, y la actuación clínica también es diferente, pero no se pierde a la paciente. Por tanto, la discrepancia debe clasificarse como INTERMEDIA.

### **ASCH VERSUS CARCINOMA (CARCINOMA ESCAMOSO, ADENOCARCINOMA O ADENOCARCINOMA IN SITU)**

En un 33% de casos diagnosticados como ASCH existe un HSIL, y las demás lesiones ya lo son. La actuación clínica es la misma: colposcopia y biopsia. Por tanto, la discrepancia debe ser considerada como MENOR.

### **HSIL VERSUS CARCINOMA (CARCINOMA ESCAMOSO, ADENOCARCINOMA O ADENOCARCINOMA IN SITU)**

La actuación clínica es la misma: colposcopia y biopsia. Por tanto, la discrepancia debe ser considerada como MENOR.

### **HSIL VERSUS ACG**

El riesgo de lesión de alto grado es diferente y la actuación clínica también es diferente, pero no se pierde a la paciente. Por tanto, la discrepancia debe clasificarse como INTERMEDIA.

### **CARCINOMA ESCAMOSO VERSUS ACG**

El riesgo de lesión de alto grado es diferente. La actuación clínica también es diferente, pero no se pierde a la paciente. El hecho de volver a citar a la paciente a los 6 meses en caso de AGC, puede retrasar el tratamiento de la lesión neoplásica. Por tanto, la discrepancia debe clasificarse como INTERMEDIA.

### **CARCINOMA VERSUS CARCINOMA**

Las discrepancias entre los diagnósticos de los distintos tipos de carcinoma deben ser clasificadas como discrepancias MENORES, dado que en todos los casos se procederá a colposcopia y el diagnóstico final se basará en el resultado de la biopsia, que, junto con las exploraciones clínicas y radiológicas, determinará el tratamiento en cada caso.