

INSTITUTO NACIONAL DE SALUD PÚBLICA
ESCUELA DE SALUD PÚBLICA DE MÉXICO
MAESTRÍA EN CIENCIAS DE LA SALUD

EVALUACIÓN DE UNA ESTRATEGIA DE
TAMIZAJE PARA CÁNCER
CERVICOUTERINO EN MORELOS

AUTOR: Jorge Eduardo Ortiz Panozo

ÁREA DE CONCENTRACIÓN: Epidemiología

GENERACIÓN: 2007 - 2009

DIRECTOR DE TESIS: Dr. Aurelio Cruz Valdez

EVALUACIÓN DE UNA ESTRATEGIA DE TAMIZAJE PARA CÁNCER CERVICOUTERINO EN MORELOS

AUTOR: Jorge Eduardo Ortiz Panozo

DIRECTOR DE TESIS: Dr. Aurelio Cruz Valdez

JURADO DE TESIS

Presidente: Dr. Bernardo Hernández Prado

Secretario: Dr. Aurelio Cruz Valdez

Primer Sinodal: Dra. Patricia Uribe Zúniga

Segundo Sinodal: Dra. M^a Luisa Gontes Ballesteros

Tercer Sinodal: Dr. Alejandro M. García Carrancá

Última actualización: 24 – Marzo – 2010

INTRODUCCIÓN

A partir del conocimiento de que la infección persistente por el virus del papiloma humano de tipo oncogénico (VPH*) es una causa necesaria en la carcinogénesis cervical¹, se han desarrollado tecnologías para la prevención secundaria del cáncer cervicouterino (CC) basadas en la detección del VPH. La infección por VPH precede por varios años a la aparición de lesiones precursoras (neoplasia intraepitelial cervical de alto grado) de CC¹, por lo que su inclusión en las estrategias de tamizaje se relacionaría con la disminución de la morbilidad y mortalidad asociadas al CC. La detección del VPH ha sido recientemente introducida en las estrategias de tamizaje para el CC², las cuales anteriormente estaban basadas en la citología cervical (Papanicolaou, PAP). En un ensayo aleatorizado realizado en 52 conglomerados de la India, que involucró a más de 130,000 mujeres de 30-59 años, una sola ronda de detección de VPH se asoció con una disminución de la incidencia de CC avanzado así como de la mortalidad por CC³. En un programa de tamizaje, en el que participaron alrededor de 12,500 mujeres de 32-38 años de Suecia, la adición de VPH al tamizaje basado en el PAP redujo la incidencia de lesiones precursoras de CC o CC⁴. Este efecto estuvo relacionado con incremento de la detección de lesiones precursoras de CC, mejora de la sensibilidad de la prueba de tamizaje y disminución en el número de referencias innecesarias⁵.

La muestra para la detección de VPH puede ser tomada por la propia paciente (autotoma), lo cual constituye un beneficio potencial en las tasas de participación y seguimiento en los programas de tamizaje para CC⁶. Algunas de las barreras por las que

* En este trabajo, el acrónimo VPH se refiere exclusivamente a los genotipos oncogénicos (de alto riesgo) del virus del papiloma humano

las mujeres no participan en los programas de tamizaje para CC son: falta de acceso a la atención en salud, incomodidad con el examen pélvico y condiciones religiosas, culturales o personas que impiden que sean examinadas por un proveedor de salud de sexo masculino⁷. Claramente las mujeres que participan poco o nunca en los programas de tamizaje, por las barreras mencionadas, tienen una alternativa en la autotoma. Además, la aceptabilidad de las mujeres hacia la autotoma es más alta que hacia el PAP⁸⁻¹⁰.

En comparación con el PAP, la detección de VPH mediante autotoma tiene mayor sensibilidad (75%-90% vs 55%-75%) y menor especificidad (43%-100% vs 71%-100%) para el tamizaje de lesiones precursoras de CC y CC, aunque menor sensibilidad y similar especificidad que la detección de VPH tomada por el proveedor de salud (toma dirigida)⁸⁻¹⁴. Esto indica que la detección de VPH mediante autotoma podría llevar a evaluaciones y seguimiento innecesarios en una proporción importante de las mujeres, si se usa como único criterio de referencia en un programa de tamizaje de larga escala poblacional⁶. Al respecto, uno de los hallazgos que ha afectado el interés en aplicar la detección de VPH como estrategia primaria de tamizaje ha sido su valor predictivo positivo relativamente bajo en comparación al PAP¹⁵ (7%-18% vs 25-44%). De hecho, se han observado mayores beneficios para el tamizaje del CC en estudios en los que la detección de VPH se usó de manera combinada con el PAP^{5, 16}.

Las condiciones en las que se han realizado las evaluaciones de la detección de VPH como estrategia de tamizaje para el CC han sido diversas¹⁵. Sin embargo, es remarcable que la mayor parte de la información disponible se refiera a estudios controlados realizados en su mayoría en países desarrollados. En efecto, aunque se ha avanzado

mucho en el conocimiento referente a la validez y capacidad predictiva de la detección de VPH para el CC, aún faltan estudios que evalúen la adición de la detección de VPH, como estrategia complementaria al PAP¹⁷, en programas de tamizaje en condiciones no controladas, en especial, en países donde los programas de tamizaje para CC han sido moderadamente eficaces¹⁸. Este es el caso de México que, en cifras absolutas, es el segundo país de Latinoamérica y El Caribe que aporta más casos de CC por año^{19, 20} (evidencia indirecta de las limitaciones del programa de tamizaje en esta región).

La Secretaría de Salud de Morelos, México, incluyó recientemente la detección de VPH mediante autotoma en su programa de tamizaje para CC, como estrategia para incrementar la participación de las mujeres y consecuentemente mejorar la detección, manejo y pronóstico de las lesiones precursoras de CC y CC. El objetivo de este estudio fue evaluar la capacidad diagnóstica y predictiva de las estrategias de tamizaje de CC resultantes de combinar la detección de VPH mediante autotoma con el PAP, en el programa de tamizaje de CC de Morelos.

MATERIAL Y MÉTODOS

Diseño del estudio

Se realizó un estudio de fase 4²¹ en el que se evaluaron las características de la adición de la detección de VPH al programa de tamizaje de CC de Morelos. El programa tuvo tres etapas: 1) Detección de VPH, 2) Seguimiento y referencia y 3) Diagnóstico definitivo [Gráfica 1]. Al final de la tercera etapa, se procedió a la evaluación de la estrategia de

tamizaje. Este estudio fue aprobado por el Comité de Ética del Instituto Nacional de Salud Pública.

Detección de VPH. Durante febrero a diciembre de 2008, se invitó a 123,883 mujeres de 30 o más años, a realizarse la detección de VPH ya sea mediante autotoma o toma dirigida, en los 33 municipios de Morelos. Personal estandarizado dio las instrucciones necesarias para la realización de la autotoma, la cual fue elegida por 92.9% de las mujeres. La recolección, transporte, almacenaje y procesamiento de muestras se efectuó de acuerdo a lo reportado por otros autores¹⁴. La evaluación de la presencia de VPH se realizó mediante captura de híbridos 2 (Digene ®). El punto de corte para clasificar los resultados como positivos fue 1 pg/ml. La prevalencia de VPH fue 11.6%. La confidencialidad de los resultados se mantuvo mediante la asignación de un código de barras a las muestras, en lugar del nombre de la participante. Al momento de tomar la muestra para la detección de VPH se aplicó un cuestionario breve sobre características generales e historia de tamizaje (antecedente de PAP previo y de resultados anormales, percepción sobre el último PAP).

Seguimiento y referencia. Fueron elegibles para el seguimiento 13,557 mujeres (9,603 con resultado positivo y el resto, negativo). Hasta diciembre de 2009 fueron seguidas 6,631 mujeres. La referencia a colposcopia se realizó mediante citatorio entregado durante una visita domiciliaria por personal estandarizado, el cual aplicó un cuestionario breve que solicitaba información de contacto para el seguimiento posterior, características sociodemográficas y antecedentes sobre factores de riesgo para CC. Las mujeres fueron visitadas hasta tres veces antes de declarar el citatorio como negativo. En los casos en los que no se encontró a la mujer en su domicilio y se contaba con número

telefónico, se aplicó el cuestionario breve y se realizó la referencia a colposcopia por vía telefónica. Los entrevistadores estaban cegados al resultado de VPH.

Las colposcopias se realizaron en uno de doce centros disponibles, por ginecólogos adscritos al Colegio de Colposcopistas de Morelos. Al momento de efectuar la colposcopia, se tomó una muestra para PAP con Citobrush® mediante técnica estándar y se recolectó un espécimen para biopsia en los casos en los que se detectó lesión indicativa de CC o CC o hubo otra indicación médica para realizar estudio histopatológico. Los resultados de la colposcopia fueron clasificados como: sin alteraciones, alteraciones inflamatorias inespecíficas, lesión intraepitelial de bajo grado (LSIL), lesión intraepitelial de alto grado (HSIL), lesión sugestiva de invasión o cáncer invasor (CA). Los colposcopistas estaban cegados al resultado de VPH.

Diagnóstico definitivo. Las muestras para PAP se interpretaron de manera independiente de acuerdo al sistema Bethesda como: inadecuado, normal, células escamosas atípicas de significado indeterminado (ASCUS), células glandulares atípicas de significado indeterminado (AGUS), lesiones escamosas intraepiteliales de bajo grado (LSIL), lesiones escamosas intraepiteliales de alto grado (HSIL), adenocarcinoma in situ (AIS), carcinoma escamoso y adenocarcinoma. Las biopsias fueron interpretadas de forma independiente por un patólogo certificado. Los casos confirmados de neoplasia intraepitelial cervical (NIC) 2 y 3 o CC (carcinoma epidermoide o adenocarcinoma) fueron referidos para tratamiento y seguimiento de acuerdo a estándares de la Secretaría de Salud de Morelos. Los resultados del PAP se colapsaron en dos categorías: positivo (ASCUS o HSIL o CC) y negativo (de otro modo). Los resultados inadecuados en el PAP fueron incluidos como resultados positivos en la evaluación de la capacidad diagnóstica y

predictiva de las pruebas de tamizaje⁷ y se realizó un análisis de sensibilidad para comparar el efecto de incluirlo o no el análisis.

Los resultados de la biopsia se colapsaron en dos categorías: positivo (NIC2, NIC3 o CC = NIC23+, categoría que representa la presencia de lesiones precursoras de CC y CC) y negativo (de otro modo). Las biopsias con muestra insuficiente para diagnóstico fueron excluidas de la evaluación de las estrategias de tamizaje.

Población de estudio y diseño muestral

Este estudio incluyó a mujeres de 21 años o más que fueron captadas por el programa de tamizaje de la Secretaría de Salud Morelos, en el cual se aplicó la prueba de detección de VPH en 123,883 mujeres. El programa de tamizaje está orientado a las mujeres de 30 años o más sin antecedente de lesión precursora de CC o CC. Morelos ocupa el quinto lugar entre los estados que tienen mayor mortalidad por CC en México (para el 2005 fue de 12.26, tres puntos más que el promedio nacional)[†]. Morelos está organizado en tres Jurisdicciones Sanitarias, en las cuales hay 208 unidades de salud en total. El laboratorio de referencia para la detección de VPH se encuentra en su capital. Sus características sociodemográficas refieren una heterogeneidad en condiciones de acceso y participación en programas de tamizaje.

Se definió el marco muestral con la información de los cuestionarios cortos recolectados en la etapa de detección de VPH. El tamaño de muestra se calculó con el objetivo de lograr comparaciones confiables entre los valores predictivos de las estrategias de tamizaje²¹, tomando como base los resultados previamente reportados en Morelos^{14, 22}.

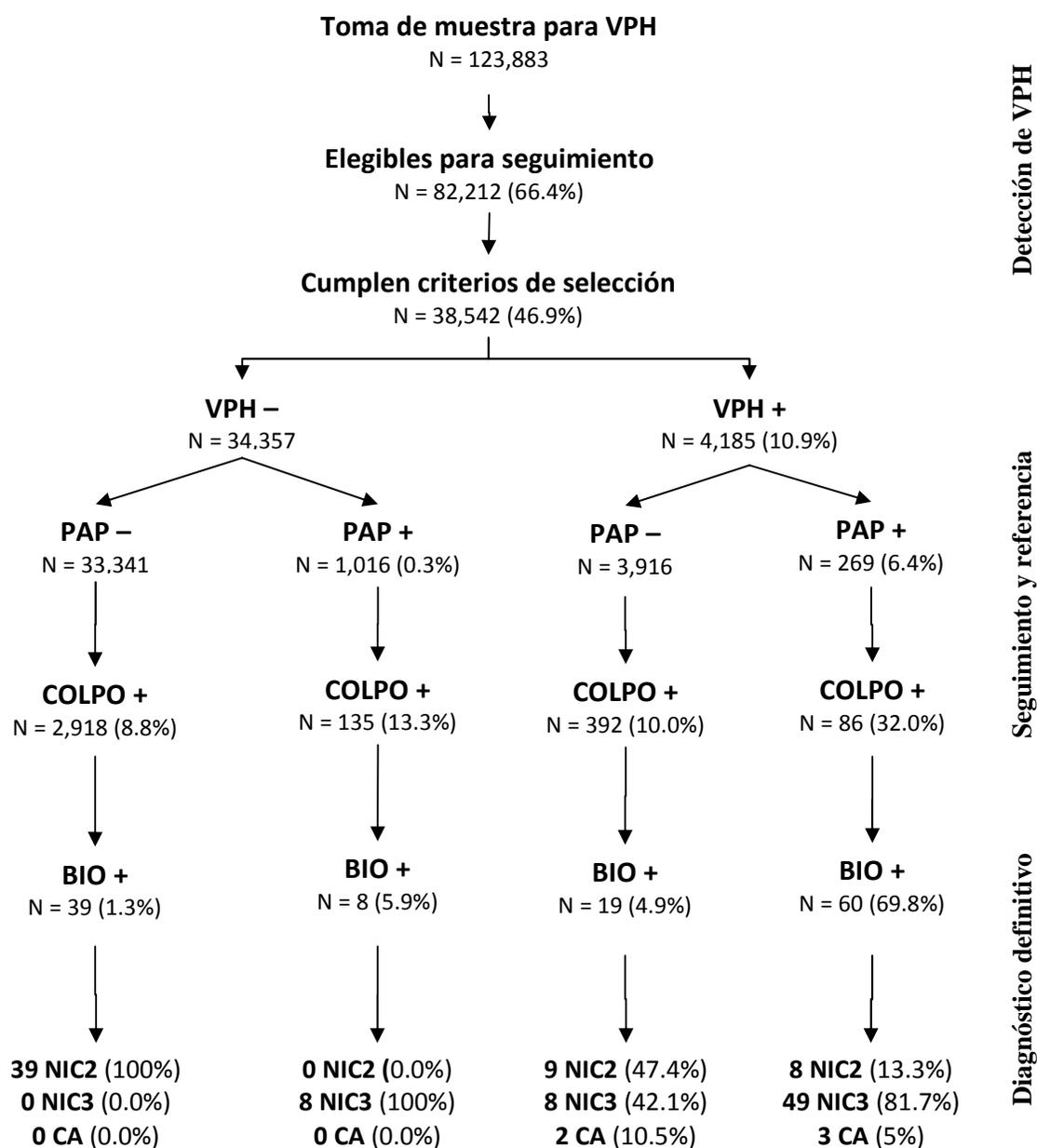
[†] Fuente: Estadísticas Vitales, Registros de Mortalidad, INEGI/SSA

Bajo el supuesto de un error absoluto de 5% en la comparación de valores predictivos positivos y 0.5% en la de valores predictivos negativos, nivel de confianza de 95%, poder de 80%, tasa de respuesta de 60% y un efecto de diseño de 1.5, el tamaño de muestra mínimo fue 1,803 mujeres con VPH positivo y 3,761 con VPH negativo.

Fueron elegibles para la etapa de seguimiento las mujeres que habían sido captadas por unidades de salud con mínimo 60 casos de VPH. La media de casos de VPH en las 68 unidades de salud que cumplieron con este requisito fue 353.5 (Desviación Estándar 196.0, rango: 63 – 662). Las mujeres potencialmente elegibles fueron 82,221 mujeres, de las cuales se seleccionaron todas las mujeres con VPH positivo (N = 9,603) y 3,954 (5%) con resultado negativo, mediante muestreo probabilístico, bietápico. En la primera etapa, se definieron tres estratos con base a las jurisdicciones sanitarias de Morelos, al interior de las cuales se eligieron 68 conglomerados, definidos por las unidades de salud, mediante muestreo proporcional al tamaño. En la segunda etapa, se seleccionaron todas las mujeres con VPH positivo y 3,954 mujeres con VPH negativo por muestreo aleatorio simple en 18 de las 68 unidades de salud seleccionadas.

Hasta diciembre de 2009, se realizó seguimiento a 6,631 de las 13,557 mujeres seleccionadas (con una tasa de respuesta al seguimiento de 73%). Para el presente análisis definieron como elegibles a las mujeres que se realizaron autotoma y no tenían antecedente de histerectomía (n = 5,980), quedando excluidas menos del 10%.

Gráfica 1. Evaluación del programa de tamizaje de cáncer cervicouterino en Morelos



VPH, Virus del papiloma humano de tipo oncogénico; COLPO, Colposcopia; BIO, Biopsia; NIC, Neoplasia Intraepitelial Cervical; CA, Cáncer Invasor; +, positivo; -, negativo.

NOTA: Las etapas de detección de VPH, seguimiento y referencia y diagnóstico definitivo (derecha) corresponden al análisis estadístico realizado, no a la manera en que se instrumentó el programa de tamizaje. En los conteos (N) se han tomado en cuenta los factores de expansión.

Análisis estadístico

El análisis estadístico consistió de: 1) Descripción de las características de las mujeres que eligieron autotoma, 2) Comparación de la sensibilidad, especificidad y valores predictivos de la autotoma y el PAP así como su relación con el grupo de edad y el antecedente de PAP con resultado anormal, 3) Comparación relativa de la autotoma, el PAP y las estrategias combinadas para identificar lesiones precursoras de CC y CC y 4) Evaluación del rendimiento global de la autotoma, el PAP y las pruebas combinadas, mediante la comparación de la probabilidad de detección, la de falsa referencia y la tasa de falsas referencias por cada caso detectado.

Definición de las variables de interés (indicadores). La capacidad diagnóstica fue evaluada mediante la sensibilidad (porcentaje de mujeres con resultado de tamizaje positivo en el grupo de mujeres con biopsia positiva), especificidad (porcentaje de mujeres con resultado de tamizaje negativo en el grupo de mujeres con biopsia negativa) y fracción de falsos positivos (probabilidad complementaria a la especificidad). Adicionalmente se estimaron las probabilidades de detección (número de casos detectados por la estrategia de tamizaje entre el número de pruebas realizadas) y de falsa referencia (número de biopsias con resultado negativo o no concluyente entre el número de pruebas realizadas), así como la tasa de falsas referencias por caso detectado, como medidas de la capacidad discriminatoria de las estrategias de tamizaje. La capacidad predictiva fue evaluada mediante el valor predictivo positivo (VPP, probabilidad de tener una lesión precursora de CC o CC dado que el resultado del tamizaje es positivo) y el valor predictivo negativo (VPN, probabilidad de no tener una lesión precursora de CC o CC dado que el resultado del tamizaje es negativo).

Definición de las estrategias de tamizaje. Tomando en cuenta los resultados binarios de la detección de VPH y el PAP, se definieron cuatro estrategias de tamizaje: 1) PAP solo (proxy de los resultados obtenidos en un programa de tamizaje para CC que no incluya la detección de VPH), 2) Autotoma sola (proxy de los resultados obtenidos al aplicar la detección de VPH como estrategia primaria para el tamizaje del CC), 3) Combinación de VPH y PAP, mediante la regla *ambos positivos*²¹, la cual clasifica como resultado positivo a la presencia de resultados positivos en ambas pruebas (proxy de los resultados obtenidos al realizar PAP sólo a las mujeres con resultado positivo en la detección de VPH por autotoma), y 4) Combinación de VPH y PAP, mediante la regla *ambos negativos*²¹, la cual clasifica como resultado negativo a la presencia de resultados negativos en ambas pruebas (proxy de los resultados obtenidos al efectuar PAP sólo a las mujeres con resultado negativo en la detección de VPH mediante autotoma).

Factores de expansión. La descripción de las características de las mujeres y la evaluación de las estrategias de tamizaje tomaron en cuenta el diseño muestral, aplicando factores de expansión inversamente proporcionales a la probabilidad de selección. Esta probabilidad estuvo determinada por la selección de las unidades de salud y el resultado de la detección de VPH. La muestra incluida en este análisis ($n = 6,580$) fue representativa de 38,542 mujeres participantes del programa de tamizaje de Morelos. Los conteos a lo largo de este reporte se expresan mediante el uso de los factores de expansión.

Estimación y comparación relativa de los indicadores. La estimación de los indicadores se ajustó mediante los factores de expansión (como proxy del recíproco de la probabilidad de llegar al diagnóstico definitivo)²¹. Este procedimiento tuvo el propósito

de ajustar los indicadores por el llamado sesgo de verificación, ocasionado por haber incluido a más mujeres con resultado de VPH positivo que negativo, lo cual podría sobreestimar la sensibilidad y subestimar la especificidad de las estrategias de tamizaje. De esta manera, los indicadores definidos fueron equivalentes a los estimadores de Begg-Greene²¹.

La comparación relativa de la capacidad diagnóstica se efectuó mediante la sensibilidad relativa (rSens) y la fracción relativa de falsos positivos (r(1-Esp)), las cuales fueron definidas, respectivamente, como la sensibilidad y el complemento de la especificidad de cada una de las estrategias de tamizaje entre las del PAP; es decir

$$rSens = \frac{\text{Sensibilidad de otra estrategia de tamizaje}}{\text{Sensibilidad del PAP}} \quad \text{y}$$

$$r(1 - Esp) = \frac{1 - \text{Especificidad de otra estrategia de tamizaje}}{1 - \text{Especificidad del PAP}}. \quad \text{Los valores predictivos}$$

relativos, positivo (rVPP) y negativo (rVPN), se definieron de manera similar. Las comparaciones relativas de la capacidad diagnóstica y predictiva de las pruebas de tamizaje se estimaron mediante modelos de regresión log-lineales²¹, que tomaron en cuenta el agrupamiento (clustering) a nivel de las 68 unidades de salud y estuvieron ajustados por el sesgo de verificación.

Probabilidad de detección y falsa referencia. Estos indicadores fueron empleados para la evaluación global de las pruebas de tamizaje y fueron estimados mediante regresión logística binaria, en la que se compararon cada una de las estrategias de tamizaje con el PAP solo. La tasa de falsa referencia por cada caso detectado para cada una de las estrategias de tamizaje se estimó mediante regresión Poisson. Para el ajuste de estos modelos también se tomaron en cuenta el agrupamiento y el sesgo de verificación.

Covariables. Se examinó el papel de la edad y el antecedente de PAP previo en la capacidad diagnóstica y predictiva de las pruebas de tamizaje solas. La edad fue categorizada en cuatro grupos: Menor de 35 años, 35 a 44 años, 45 a 54 años y mayor de 54 años. Se consideró que el antecedente de PAP previo estaba presente si la mujer reportó que alguna vez le habían realizado una citología cervical. Estos análisis se realizaron mediante la r_{Sens} y $r(1-\text{Esp})$, las cuales fueron estimadas mediante el ajuste de modelos log-lineales²¹, como se explicó anteriormente. No se incluyeron covariables en el ajuste de los modelos de regresión para la comparación de las estrategias de tamizaje debido a que la comparación se realizó entre las mismas mujeres. Los indicadores fueron totalmente comparables en cuanto a características observadas y no observadas debido al diseño pareado de este estudio. El término *pareado* se refiere a que cada una de las mujeres tuvo información sobre todas las estrategias de tamizaje en evaluación²¹.

RESULTADOS

Características de la población de estudio

Un total de 6,631 mujeres fueron seguidas hasta diciembre de 2009 por el programa de tamizaje para CC de Morelos, entre las cuales 5,980 (90.0%) cumplieron con los criterios de inclusión para este análisis. Estas últimas representan a 38,542 mujeres participantes del programa. La mayoría de las mujeres fue residente de las Jurisdicciones Sanitarias I (43.4%) y III (31.5%). El rango de edad estuvo entre 21 y 93 años (media 46.5, IC 95% 45.6 – 47.5). Más de la mitad de las mujeres se encontró entre los 35 y 55 años, un 13.5% fue menor a 35 años (media 31.4, IC 95% 31.3 – 31.4) y 23.5% fue mayor a 55 años

(media 63.6, IC 95% 63.3 – 63.8). Treinta y ocho por ciento de las mujeres refirió que trabajaba fuera de su hogar [Tabla 1].

El antecedente de PAP previo estuvo presente en 92.9% de las mujeres (número medio de PAP previos 5.8, IC 95% 5.5 – 6.1). Entre ellas, el 10.9% reportó haber presentado alguna vez un resultado anormal en la citología. Diecisiete por ciento mencionó haberse sentido avergonzada, incómoda, sin privacidad o maltratada durante su último PAP [Tabla 1].

Tabla 1. Características de la población de estudio (N = 38,542)

VARIABLE	%*	IC 95%
Procedencia		
Jurisdicción 1	43.4	33.8 - 53.4
Jurisdicción 2	25.2	12.9 - 43.4
Jurisdicción 3	31.5	23.4 - 40.8
Edad		
Edad promedio*	46.5	45.6 - 47.5
<35 años	13.5	12.3 - 14.7
35-44 años	36.7	34.5 - 39.0
45-54 años	26.3	25.0 - 27.7
>54 años	23.5	20.4 - 26.9
Trabaja fuera del hogar		
Sí	38.1	36.6 - 39.7
Papanicolaou previo		
Sí	92.9	92.4 - 93.5
Número promedio*	5.8	5.5 - 6.1
Resultado anormal	10.9	9.0 - 13.2
En Papanicolaou previo, se sintió		
Normal	82.7	81.6 - 83.7
Avergonzada	9.4	8.3 - 10.6
Incómoda	6.4	5.7 - 7.3

Faltó privacidad	0.2	0.2 - 0.3
Maltrato	1.3	1.0 - 1.6

*, valores expresados como porcentaje excepto donde se indica
N, conteo ponderado mediante los factores de expansión

Detección de VPH

La prevalencia de VPH en las mujeres que se realizaron autotoma fue 10.9% (IC 95%, 9.0 – 13.1). Esta prevalencia fue más alta en las mujeres menores de 35 años (12.0%, IC 95% 10.3 – 13.6) y en las mayores de 55 años (13.0%, IC 95% 9.6 – 16.5) que en las de 35 a 55 años (9.8%, IC 95% 8.2 – 11.5), $p < .001$. No hubo diferencias en la prevalencia de VPH por Jurisdicción Sanitaria, antecedente de PAP previo ni resultados anormales en alguna citología.

Detección de anormalidades citológicas

La prevalencia de anormalidades citológicas entre las 38,542 mujeres participantes del programa fue 3.1% (IC 95% 2.7 – 3.5). Esta prevalencia varió con la edad, alcanzando 4.3% (IC 95% 3.4 – 5.4) en las menores de 35 años, 2.9% (IC 95% 2.4 – 4.6) en las de 35 – 54 años y 2.5% (IC 95% 1.1 – 2.0) en las mayores de 54 años, $p < .001$. La prevalencia de anormalidades citológicas también varió de acuerdo a la Jurisdicción Sanitaria, siendo 3.1% en la Jurisdicción I, 4.8% en la Jurisdicción II y 2.5% en la Jurisdicción III, $p < .05$. No hubo diferencias entre la prevalencia de anormalidades citológicas respecto a los antecedentes de PAP ni resultado citológico anormal previos

Entre las mujeres que tuvieron alguna anomalía citológica 14.4% (168/1,166) presentó ASCUS; 62.3% (730/1,166) LSIL; 21.4% (249/1,166) HSIL y 1.63% (19/1,166) CA. La edad promedio de las mujeres que recibieron diagnóstico citológico de cáncer fue

60.68 años (rango 42 – 78). Se realizó un análisis de sensibilidad de los indicadores de rendimiento diagnóstico incluyendo o no como positivos 849/38,542 (2.2%) PAP inadecuados⁷. Los resultados de ese análisis [Anexo 1] mostraron que, en general, la dirección y significancia no se modificaban, por lo que la capacidad diagnóstica fue evaluada incluyendo como positivos los PAP inadecuados.

Colposcopia y biopsia

Los resultados de la colposcopia fueron negativos a lesión precursora de cáncer o CC en 86.0% de las mujeres; 7.6% de las mujeres tuvieron un resultado de LSIL, 0.7% HSIL y 0.03 CA. Seis por ciento de las colposcopias fueron inadecuadas para diagnóstico. Se indicó biopsia a 6.8% (86/1,269) de las mujeres con resultados positivos para VPH de alto riesgo y PAP; 10.0% (392/3,916) de las mujeres con VPH de alto riesgo y PAP negativo; 5.9% (8/135) de las mujeres con PAP positivo y VPH de alto riesgo negativo; y 1.3% (39/2,918) de las mujeres con ambos resultados negativos [Gráfica 1].

Entre las mujeres que fueron sometidas a biopsia, 12.8% tuvieron especímenes insuficientes para el diagnóstico, 78.5% tuvo resultado negativo a lesión precursora de cáncer o CC; 5.6% (227/4,047) tuvo NIC1; 1.4% (227/4,047) NIC2; 1.6% (65/4,047) NIC3 y 0.1% (5/4,047) CA. Casi todas las mujeres con diagnóstico de NIC2, NIC3 o CA (79/126) tuvieron VPH positivo. El 69.6% (39/56) de las mujeres con NIC2 y el 12.3% (8/65) de las con NIC3 y tuvieron resultado negativo en la prueba de detección de VPH. La distribución de las lesiones cervicales de acuerdo a las combinaciones de las pruebas de tamizaje se ven en la Tabla 2.

Tabla 2. Distribución de resultados de biopsia de acuerdo a combinaciones de las pruebas de tamizaje

VPH	PAP	Insuficiente	Negativo	NIC1	NIC2	NIC3	CA	Total
-	-	421	2,768	111	39	0	0	3,339
-	+	22	28	20	0	8	0	78
-	?	29	44	35	0	0	0	108
+	-	38	317	56	9	8	2	430
+	+	5	15	4	5	49	1	79
+	?	1	6	1	3	0	2	13
Total		516	3,178	227	56	65	5	4,047

Los conteos están ponderados mediante los factores de expansión

NIC, Neoplasia intraepitelial cervical; CA, cáncer invasor; +, positivo; -,negativo; ?, inadecuado

Capacidad diagnóstica y predictiva de las estrategias de tamizaje

La autotoma tuvo una sensibilidad para detectar NIC23+ de 62.7% (IC 95% 53.1 – 71.4) y una especificidad de 88.3% (IC 95% 85.3 – 90.8). Estos indicadores fueron 54% (IC 95% 46.3 – 61.4) y 95.5% (IC 95% 94.0 – 96.7) para el PAP. El VPP de la autotoma fue 16.5% (IC 95% 14.7 – 18.6) y el del PAP, 30.8% (IC 95% 23.0 – 39.8). Los VPN fueron similares para ambas pruebas, 98.5% (IC 95% 97.6 – 99.0) y 98.3% (IC 95% 97.4 – 98.8), respectivamente [Tabla 3].

No hubo diferencias estadísticamente significativas en la sensibilidad de la prueba de VPH por grupo de edad. En comparación con las mujeres menores de 35 años, la fracción de falsos positivos fue de 0.84 (IC 95% 0.68 – 1.03), $p=0.092$, en las de 35 a 44 años y de 0.73 (IC 95% 0.54 – 0.99), $p<.05$, en las de 45 a 54 años. Al comparar las mujeres con antecedente de PAP anormal con las mujeres sin ese antecedente, la rSens fue 0.08 (IC 95% 0.02 – 0.26), $p <.001$. La sensibilidad del PAP fue superior en las mujeres de 35 a

44 años y en las mayores de 55 años en comparación con las menores de 35 años, rSens 2.76 (IC 95% 1.11 – 6.81), $p < .05$, y 3.29 (IC 95% 1.36 – 7.96), $p < .05$, respectivamente. La $r(1-Esp)$ al comparar las mujeres mayores de 55 años con las menores de 35 fue 0.15 (IC 95% 0.04 - .55), $p < .05$. En las mujeres con antecedente de PAP anormal la sensibilidad del PAP fue menor que en las mujeres sin ese antecedente ($rSens = 0.03$, IC 95% 0.004 – 0.20, $p < .01$). El antecedente de PAP anormal previo no se asoció con cambios en la especificidad del PAP ni de la autotoma.

Tabla 3. Capacidad diagnóstica y predictiva de las pruebas de tamizaje en mujeres con biopsia (N = 3,531)

INDICADOR	PAP	Autotoma
Sensibilidad	54.0 46.3 - 61.4	62.7 53.1 - 71.4
Especificidad	95.5 94.0 - 96.7	88.3 85.3 - 90.8
Valor predictivo positivo	30.8 23.0 - 39.8	16.5 14.7 - 18.6
Valor predictivo negativo	98.3 97.4 - 98.8	98.5 97.6 - 99.0

Los conteos están ponderados mediante los factores de expansión
 Los indicadores se presentan como %, IC 95%. Se consideró biopsia positiva al resultado igual o superior a Neoplasia Intraepitelial Cervical grado II (NIC23+)

En comparación con el PAP solo, la detección de VPH mediante autotoma tuvo menor especificidad y menor VPP ($r(1 - Esp) = 2.61$, $p < .01$, y $rVPP = 0.54$, $p < .05$); la prueba “ambos positivos” tuvo un incremento en la especificidad y en el VPP ($r(1 - Esp) = 0.17$, $p < .001$, $rVPP = 2.27$, $p < .001$). La prueba “ambos negativos” tuvo un incremento en la

sensibilidad, la fracción de falsos positivos y el VPN ($p = 0.055$) acompañado de disminución del VPP ($p < .001$) [Tabla 4].

Tabla 4. Comparación de las estrategias de tamizaje (N = 3,531)

INDICADOR	Autotoma	Ambos positivos	Ambos negativos
	1.16	0.88	1.28**
rSens	0.98 - 1.37	0.77 - 1.01	1.17 - 1.40
	2.61**	0.17***	3.44***
r(1-Esp)	1.72 - 3.96	0.12 - 0.25	2.55 - 4.64
	0.54*	2.27***	0.46***
rVPP	0.41 - 0.70	1.74 - 2.95	0.39 - 0.55
	1.00	0.998	1.004
rVPN	1.00 - 1.00	0.997 - 1.000	1.002 - 1.006

Los conteos están ponderados mediante los factores de expansión
 rSens, Sensibilidad relativa; r(1 – Esp), Fracción relativa de falsos positivos; rVPP, Valor predictivo positivo relativo; rVPN, Valor predictivo negativo relativo
 *, $p < .05$; **, $p < .01$; ***, $p < .001$
 Estrategia de referencia: Papanicolaou solo

Rendimiento global de las estrategias de tamizaje

En la Tabla 5 se muestra la comparación entre el rendimiento global de la autotoma y las pruebas combinadas contra el PAP solo. La tasa de detección de NIC23+ con la prueba combinada “ambos negativos” fue la más alta entre las cuatro estrategias de tamizaje; en comparación con el PAP el OR fue 1.28 (IC 95% 1.20 – 1.36), con 226 casos detectados por cada 100,000 mujeres tamizadas. La probabilidad de detección de NIC23+ mediante autotoma fue la segunda más alta entre las cuatro estrategias de tamizaje; aunque las diferencias no fueron estadísticamente significativas (OR = 1.16, IC 95% 0.89 – 1.50). La

prueba combinada “ambos positivos” tendió a ser inferior que el PAP para detectar NIC23+ (OR = 0.88, IC 95% 0.70 – 1.10).

Tabla 5. Probabilidades estimadas de detección y falsa referencia por 100,000 mujeres tamizadas según estrategia de tamizaje (N=38,542)

Indicador	Probabilidad x	
	100,000	IC 95%
Tasa de detección		
Papanicolaou	176.4	140.2 - 221.9
Autotoma	205.0	158.3 - 265.4
Ambos positivos	155.7	123.7 - 195.9
Ambos negativos **	225.7	175.0 - 291.1
Tasa de falsa referencia		
Papanicolaou	397.8	207.9 - 759.9
Autotoma *	1,036.0	696.5 - 1,538.3
Ambos positivos **	67.5	42.6 - 107.0
Ambos negativos **	1,366.2	948.5 - 1,964.3

Los conteos están ponderados mediante los factores de expansión

*, p <.01 y **, p <.001 para la comparación respecto al Papanicolaou

La tasa de falsa referencia mediante la prueba combinada “ambos positivos” se redujo de manera importante en comparación con el PAP (OR = 0.17, IC 95% 0.12 – 0.24). La autotoma sola y combinada con PAP mediante la regla “ambas negativas” resultaron en incremento de la tasa de falsa referencia (OR = 2.62 y 3.47, respectivamente, p<.001).

El número de colposcopias con biopsia negativa por cada caso de NIC23+ detectado fue el más bajo mediante la prueba combinada “ambos positivos” (0.43, IC 95% 0.38 – 0.48) (Razón de tasas 0.17, IC 95% 0.09 – 0.32, en comparación con el PAP, p <.001). La autotoma (p < .01) y la prueba combinada “ambos negativos” (p < .001) tuvieron las más altas tasas de falsa referencia por un caso detectado, 6.05 (IC 95% 5.05 – 7.04) y 5.05 (IC

95% 4.49 – 5.61), respectivamente (Razones de tasas en comparación con el PAP, 2.60, IC 95% 1.29 – 5.24, y 3.43, IC 95% 2.07 – 5.68, respectivamente).

DISCUSIÓN

En este estudio se encontró que la capacidad diagnóstica y predictiva de la detección de VPH puede modificarse mediante la combinación con el PAP y que la manera en que se las combine es crucial para mejorar la probabilidad de detección de lesiones precursoras de CC y CC. Al comparar las pruebas por separado, la detección de casos mediante la prueba de VPH y PAP fue similar, aunque el VPP y la tasa de falsas referencias por cada caso detectado fueron mejores en el PAP. Al comparar las diversas estrategias, la prueba “ambos positivos” tuvo la mayor disminución en la tasa de falsos negativos así como mejora en la de falsa referencia y ganancia significativa en el VPP, sin pérdida de sensibilidad o VPN.

En general, las características de las mujeres incluidas en el programa piloto de tamizaje de CC de Morelos fueron similares a las de la población objetivo de otros estudios^{9, 14, 23}. Específicamente, la prevalencia de VPH de alto riesgo (aproximadamente 11%) fue similar a la reportada anteriormente en una población derechohabiente de Morelos¹⁴. La mayor parte de las mujeres participantes se encontró entre los 30 y 50 años, lo cual coincidió con la población prioritaria de esta intervención en Morelos y las recomendaciones internacionales para los programas de prevención del CC¹⁷.

De manera interesante, la historia de tamizaje, en cuanto a las percepciones y experiencias al realizarse el PAP se relacionaron con la elección de autotoma. En efecto, aunque la mayoría de las mujeres no reportó molestias durante el PAP previo, se logró realizar seguimiento a una proporción importante de mujeres que tuvieron experiencias

negativas en el PAP previo. Este es un aspecto fundamental para mejorar el acceso a los programas de prevención del CC. En diversos estudios que han evaluado las pruebas de detección de VPH se ha reportado mayor aceptabilidad a la autotoma⁸⁻¹⁰. Esto se debe a que la mayoría de las mujeres se sienten más cómodas si no se les realiza el examen pélvico. En este sentido, la autotoma tiene un potencial único para mejorar la cobertura de tamizaje del CC.

La sensibilidad, especificidad y valores predictivos estimados mostraron concordancia con lo reportado por otros autores^{10, 14, 16, 23-26}. La detección de VPH tendió a mostrar mayor sensibilidad que el PAP para detectar NIC23+, aunque en este estudio no hubo diferencias estadísticamente significativas. A pesar de esto, la tasa de detección de casos por 100,000 mujeres tamizadas al emplear la autotoma fue mayor que la del PAP (205.0 vs 176.4). Sin embargo, la especificidad y el VPP de la autotoma fueron menores a la del PAP.

El intercambio entre ganancia en sensibilidad a expensas de pérdida de la especificidad es un hecho bien documentado en la literatura^{10, 14, 16-18, 23-26}. A pesar de esto, en los programas de tamizaje se prioriza la sensibilidad ante la especificidad; es decir, se enfatiza la detección de mujeres que tengan lesiones precursoras de CC anticipando una tasa importante de falsos positivos. Por ello, se cuenta con fases sucesivas en las estrategias diagnósticas para confirmar lesiones de NIC23+. Entre estas se encuentra el PAP, la colposcopia y la biopsia (como etapa de confirmación final).

En este sentido es fundamental comparar los valores predictivos de la autotoma contra el PAP y otras estrategias combinadas. Al respecto, el VPN fue similar para ambas pruebas, pero se inclinó a favor del PAP en el VPP. Con todo, la probabilidad *a posteriori* de tener

una lesión precursora de CC se incrementó alrededor de 10 veces con el resultado positivo en la autotoma.

Entre las mujeres evaluadas 126/5,980 (2.1%) tuvieron PAP inadecuados. Estas mujeres fueron incluidas en el análisis considerando su resultado como positivo¹⁶ ya que el análisis de sensibilidad mostró que al excluirlas no se modificaba la dirección ni significancia de los indicadores, aunque sí hubo cambio en la magnitud de la $r(1 - \text{Esp})$. En efecto, las fracciones de falsos positivos de la autotoma, la prueba combinada “ambos positivos” y la prueba combinada “ambos negativos” al excluir las mujeres con PAP inadecuado fueron 5.8, 0.28 y 6.50, respectivamente, en comparación con 2.61, 0.17 y 3.44 al considerar que las mujeres con PAP inadecuado tenían resultado positivo [Anexo 1].

El rendimiento global de la autotoma fue superior al del PAP para el resultado positivo pero no para el negativo. Como se ve en la tabla 3, la probabilidad de que una mujer con lesión precursora de CC tenga un resultado positivo fue 11 veces mayor a la de tener un resultado negativo.

En cambio, al evaluar la ganancia diagnóstica al usar PAP como prueba adyuvante de la detección de VPH se observó que el rendimiento global de la estrategia de tamizaje disminuyó para el resultado positivo pero mejoró para el resultado negativo. Esto está dentro de lo esperable ya que la estrategia de tamizaje empleada en el programa piloto de Morelos fue secuencial, por lo que se espera un incremento de la especificidad con disminución concomitante de la sensibilidad. De esta manera, la ganancia diagnóstica para el resultado negativo fue superior para la autotoma que para el PAP.

Lo último sugiere que hay un mecanismo de balance en la estrategia de diagnóstico ya que con las pruebas solas (sin PAP), la tasa de falsos positivos fue mayor para la autotoma. En otras palabras, parece ser que con el uso de PAP como prueba adyuvante de la autotoma, el exceso de falsos positivos podría equilibrarse a través de la mayor detección de verdaderos negativos en la segunda etapa del diagnóstico.

La información obtenida en este estudio es relevante para la implementación de programas de tamizaje para CC a gran escala, ya que permitirá tomar decisiones sobre qué pruebas usar, o en qué combinaciones, tomando en cuenta qué es lo que se requiere priorizar: la detección de verdaderos positivos en la primera etapa y segunda etapa (lo cual sería más eficiente con una prueba combinada) o la detección de verdaderos positivos en la primera, y la identificación verdaderos negativos en la segunda (en cuyo caso sería mejor usar emplear la autotoma).

El efecto de la autotoma sobre la detección oportuna de lesiones precursoras de CC fue equivalente al del PAP. Este es un hallazgo que no ha sido documentado previamente en una población grande y no controlada, como la que fue motivo de este estudio. Esto representa un potencial enorme para la mejora de los programas de tamizaje ya que indicaría que se puede llegar a tamizar a más mujeres, evitando la incomodidad del examen pélvico y sorteando las dificultades de acceso persistentes en los entornos de gran marginación de diversos países, pero conservando un nivel confiable en la capacidad diagnóstica.

En suma, los resultados de este estudio aportan evidencia de que la autotoma es una alternativa eficaz y válida para el tamizaje del CC y que la realización de PAP sólo a las personas con VPH positivo es la estrategia de tamizaje más eficiente. No obstante, estos

hallazgos deben ser evaluados durante un mayor tiempo de seguimiento; en efecto, es importante conocer cuál es el valor predictivo de esta autotoma en el largo plazo ya que en la historia natural del CC intervienen diversos factores que no fueron considerados en este estudio. Al respecto, parecen ser importantes la edad y la historia de tamizaje. Esto contribuirá a la toma de decisiones acerca de la implementación de programas de tamizaje basados en la detección de VPH como prueba inicial y el uso del PAP como prueba adyuvante. Adicionalmente, es necesario conocer la capacidad de los servicios de salud para instaurar un programa de tamizaje basado en la detección de VPH, ya que la implementación eficiente de esta estrategia depende de las características de los servicios y del personal involucrado así como de los recursos disponibles.

Anexo 1. Análisis de sensibilidad respecto al rendimiento diagnóstico del Papanicolaou y comparaciones relativas con otras estrategias de tamizaje

INDICADOR	PAP inadecuado	
	Excluido N = 3,440	Incluido como positivo N = 3,531
PAP		
Sensibilidad	52.1	54.0
Especificidad	98.0	95.5
Valor predictivo positivo	48.5	30.8
Valor predictivo negativo	98.3	98.3
rSens		
Autotoma	1.20	1.16
Ambos positivos	0.84	0.88
Ambos negativos	1.33	1.28
r(1-Esp)		
Autotoma	5.80	2.61
Ambos positivos	0.28	0.17
Ambos negativos	6.50	3.44
rNPV		
Autotoma	1.002	1.002
Ambos positivos	0.997	0.998
Ambos negativos	1.005	1.004
rPPV		
Autotoma	0.34	0.54
Ambos positivos	1.53	2.27
Ambos negativos	0.34	0.46

Los conteos están ponderados mediante los factores de expansión
rSens, Sensibilidad relativa; r(1 – Esp), Fracción relativa de falsos positivos; rVPP, Valor predictivo positivo relativo; rVPN, Valor predictivo negativo relativo

REFERENCIAS

1. Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol.* Apr 2002;55(4):244-265.
2. Wright TC, Jr., Schiffman M. Adding a test for human papillomavirus DNA to cervical-cancer screening. *N Engl J Med.* Feb 6 2003;348(6):489-490.
3. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med.* Apr 2 2009;360(14):1385-1394.
4. Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med.* Oct 18 2007;357(16):1589-1597.
5. Naucler P, Ryd W, Tornberg S, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst.* Jan 21 2009;101(2):88-99.
6. Stewart DE, Gagliardi A, Johnston M, et al. Self-collected samples for testing of oncogenic human papillomavirus: a systematic review. *J Obstet Gynaecol Can.* Oct 2007;29(10):817-828.
7. Bingham A, Bishop A, Coffey P, et al. Factors affecting utilization of cervical cancer prevention services in low-resource settings. *Salud Publica Mex.* 2003;45 Suppl 3:S408-416.
8. Dannecker C, Siebert U, Thaler CJ, Kiermeir D, Hepp H, Hillemanns P. Primary cervical cancer screening by self-sampling of human papillomavirus DNA in internal medicine outpatient clinics. *Ann Oncol.* Jun 2004;15(6):863-869.
9. Sellors JW, Lorincz AT, Mahony JB, et al. Comparison of self-collected vaginal, vulvar and urine samples with physician-collected cervical samples for human papillomavirus testing to detect high-grade squamous intraepithelial lesions. *CMAJ.* Sep 5 2000;163(5):513-518.
10. Wright TC, Jr., Denny L, Kuhn L, Pollack A, Lorincz A. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. *JAMA.* Jan 5 2000;283(1):81-86.
11. Garcia F, Barker B, Santos C, et al. Cross-sectional study of patient- and physician-collected cervical cytology and human papillomavirus. *Obstet Gynecol.* Aug 2003;102(2):266-272.

12. Lorenzato FR, Singer A, Ho L, et al. Human papillomavirus detection for cervical cancer prevention with polymerase chain reaction in self-collected samples. *Am J Obstet Gynecol.* May 2002;186(5):962-968.
13. Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ, et al. Primary screening for high risk HPV by home obtained cervicovaginal lavage is an alternative screening tool for unscreened women. *J Clin Pathol.* Jun 2002;55(6):435-439.
14. Salmeron J, Lazcano-Ponce E, Lorincz A, et al. Comparison of HPV-based assays with Papanicolaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes Control.* Aug 2003;14(6):505-512.
15. Castle PE. Invited commentary: is monitoring of human papillomavirus infection for viral persistence ready for use in cervical cancer screening? *Am J Epidemiol.* Jul 15 2008;168(2):138-144; discussion 145-138.
16. Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. *JAMA.* Oct 9 2002;288(14):1749-1757.
17. Meijer CJ, Berkhof J, Castle PE, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer.* Feb 1 2009;124(3):516-520.
18. World Health Organization. *Control integral del cáncer cervicouterino : guía de prácticas esenciales.* Ginebra: Organización Mundial de la Salud; 2007.
19. Lewis MJ, Pan American Health Organization. *Situational analysis of cervical cancer in Latin America and the Caribbean.* Washington, D.C.: Pan American Health Organization; 2004.
20. WHO Consultation on Cervical Cancer Screening, World Health Organization. Dept. of Reproductive Health and Research., WHO Programme on Cancer Control. *Cervical cancer screening in developing countries : report of a WHO consultation.* Geneva: World Health Organization; 2002.
21. Pepe MS. *The statistical evaluation of medical tests for classification and prediction.* Oxford ; New York: Oxford University Press; 2003.
22. Flores Y, Bishai D, Lazcano E, et al. Improving cervical cancer screening in Mexico: results from the Morelos HPV Study. *Salud Publica Mex.* 2003;45 Suppl 3:S388-398.
23. Schiffman M, Herrero R, Hildesheim A, et al. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. *JAMA.* Jan 5 2000;283(1):87-93.

24. Bergeron C, Jeannel D, Poveda J, Cassonnet P, Orth G. Human papillomavirus testing in women with mild cytologic atypia. *Obstet Gynecol.* Jun 2000;95(6 Pt 1):821-827.
25. Lee KJ, Lee JK, Saw HS. Can human papillomavirus DNA testing substitute for cytology in the detection of high-grade cervical lesions? *Arch Pathol Lab Med.* Mar 2004;128(3):298-302.
26. Ronco G, Segnan N, Giorgi-Rossi P, et al. Human papillomavirus testing and liquid-based cytology: results at recruitment from the new technologies for cervical cancer randomized controlled trial. *J Natl Cancer Inst.* Jun 7 2006;98(11):765-774.