

# Selección de la terapia sistémica adyuvante postoperatoria para cáncer de mama en etapa temprana: Una evaluación crítica de las pruebas de expresión génica disponibles comercialmente

David M. Hyams MD<sup>1</sup> | Eric Schuur PhD<sup>2</sup> | Javier Angel Aristizabal MD<sup>3</sup> | Juan Enrique Bargalló Rocha MD<sup>4</sup> | César Cabello MD<sup>5</sup> | Roberto Elizalde MD<sup>6</sup> | Laura García-Estévez MD<sup>7</sup> | Henry L. Gómez MD<sup>8</sup> | Artur Katz MD<sup>9</sup> | Aníbal Núñez De Pierro MD<sup>10</sup>

<sup>1</sup> Desert Surgical Oncology, Eisenhower Medical Center, Rancho Mirage, CA, EE. UU

<sup>2</sup> VMWA LLC, Palo Alto, CA, EE. UU

<sup>3</sup> Instituto Nacional de Cancerología, Bogotá, Colombia

<sup>4</sup> Instituto Nacional de Cancerología, Ciudad de México, México

<sup>5</sup> Universidade Estadual de Campinas, Campinas, Brasil

<sup>6</sup> Hospital Dr. Ignacio Pirovano, Buenos Aires, Argentina

<sup>7</sup> Centro Integral Oncológico Comprensivo Clara Campal, Madrid, España

<sup>8</sup> Instituto Nacional de Enfermedades Neoplásicas, Lima, Perú

<sup>9</sup> Hospital Sirio-libanés, San Pablo, Brasil

<sup>10</sup> Hospital Dr. Juan A. Fernández, Buenos Aires, Argentina

## Correspondiente

David M. Hyams, MD, Desert Surgical Oncology, Eisenhower Medical Center, Suite K-207, 39000 Bob Hope Drive, Rancho Mirage, CA 92270.

Dirección de correo electrónico: dhyams.md@me.com

La estratificación del riesgo de pacientes con cáncer de mama en etapa temprana puede respaldar la toma de decisiones sobre la quimioterapia adyuvante. Esta revisión analítica detalla el desarrollo y la validación de seis clasificadores multigénicos, cada uno de los cuales afirma brindar información útil de pronóstico y, posiblemente, información predictiva, para pacientes con cáncer de mama en etapa temprana. Se presenta una evaluación meticulosa de la validez analítica, la validez clínica y la utilidad clínica de cada prueba, así como también la calidad de la evidencia que respalda su uso.

## PALABRAS CLAVES

neoplasmas de mama, riesgo genómico, quimioterapia adyuvante, pronóstico, toma de decisiones clínicas, expresión génica, biomarcadores, tumor

## 1 | INTRODUCCIÓN

Durante 20 años, los índices de mortalidad por cáncer de mama han disminuido de manera constante.<sup>1</sup> Aunque la quimioterapia adyuvante (QTx) ha sido un contribuyente eficaz del efecto total, sólo una pequeña proporción de los pacientes se beneficia individualmente del tratamiento que reciben.<sup>2</sup> Las pruebas de expresión génica disponibles comercialmente se promocionan como un abordaje para definir la

biología tumoral intrínseca, lo que permite determinar el pronóstico y los beneficios del tratamiento de manera más precisa que los clasificadores clínico-patológicos tradicionales.

El marco de Haddow y Polamaki para la evaluación de ensayos génicos basado en evidencia, identifica componentes esenciales para toda prueba nueva.<sup>3,4</sup> La *validez analítica* representa la capacidad de un ensayo para medir de manera precisa y reproducible los biomarcadores de interés. La *validez clínica* requiere que la prueba presente

información precisa y confiable sobre los resultados relevantes.<sup>5</sup> La utilidad clínica muestra la habilidad de una prueba para inducir alteraciones favorables en los resultados del paciente.

El Sistema de Calificación de la Utilidad del Marcador Tumoral (TMUGS, por sus siglas en inglés) califica la evidencia según una escala progresiva.<sup>6</sup> La mejor evidencia proviene de los estudios clínicos completamente prospectivos y controlados (Nivel A). Los estudios prospectivos-retrospectivos (Nivel B) conforman un nivel de evidencia inferior (LOE, por sus siglas en inglés) con un diseño de estudio prospectivo sobre tejido archivado de un estudio clínico prospectivo anterior.<sup>7</sup> Dos estudios clínicos de Nivel B concordantes pueden alcanzar evidencia de Nivel I. La evaluación prospectiva de datos del registro de pacientes recopilados anteriormente (nivel C) representa un LOE inferior, pero ofrece la posibilidad de confirmar los hallazgos de estudios provenientes de LOE superior. Ante la falta de terapéuticas definidas, la adquisición de datos o seguimiento, a través de investigaciones post hoc que utilizan muestras por conveniencia, están propensas a sesgos y representan el LOE más bajo (nivel D). Estas solo deberían utilizarse para la confirmación, en términos generales, de los datos de mayor nivel o para formular hipótesis. El TMUGS se resume en la Table 1.

En la actualidad hay seis pruebas de expresión génica disponibles comercialmente a nivel mundial para pacientes con cáncer de mama en etapa temprana (ESBC, por sus siglas en inglés). Esta revisión analítica identifica los datos de validación relevantes para estas pruebas calificadas mediante el LOE de TMUGS.

## 2 | MATERIALES Y MÉTODOS

### 2.1 | Fuentes de artículos

En septiembre de 2016, se realizó una búsqueda sistemática de la literatura científica que identificó los estudios primarios publicados que describían la validación clínica de uno de los seis clasificadores genómicos para cáncer de mama disponibles comercialmente (Table 2). Además, para todo el análisis, las bases de datos consultadas incluyeron Medline y la Biblioteca Cochrane. También se consultaron artículos relevantes publicados anteriormente para garantizar que no se hubiesen pasado por alto informes apropiados.

### 2.2 | Estrategia de búsqueda

Se implementó una estrategia de búsqueda de palabras clave para la búsqueda inicial de bases de datos, que incluyó "BCI", "Breast Cancer Index", "HOXB13/IL17BR", "MGI", "Molecular Grade Index", "IHC4", "Oncotype", "21 gene", "Recurrence Score", "MammaPrint" "70 gene", "Blueprint", "TargetPrint", "PAM 50", "intrinsic subtype", "Endopredict" y "Prosigna".

### 2.3 | Selección de estudios y extracción de datos

Solo se seleccionaron publicaciones en inglés. Los estudios analizaron los datos originales relevantes para la validez clínica o la utilidad clínica

**TABLE 1** Determinación modificada de niveles de evidencia con incorporación de estudios realizados sobre el tipo de biomarcador

Nivel	Categoría	Estudio de tipo de biomarcador realizado	Estudios de validación
I	A	Diseño prospectivo para responder a una pregunta sobre biomarcadores con tratamiento, evaluación y recolección de muestras definidos previamente. La medición y el análisis de biomarcadores son prospectivos con un plan de análisis estadístico contemporáneo (SAP, por sus siglas en inglés).	No se requiere ninguno
I	B	Diseño prospectivo para responder a una pregunta sobre biomarcadores mediante el uso de un estudio prospectivo realizado anteriormente con recolección y archivo de tejido en el estudio original. La medición y el análisis de biomarcadores son prospectivos con un SAP contemporáneo.	Uno o más con resultados sistemáticos
II	B	Diseño prospectivo para responder a una pregunta sobre biomarcadores mediante el uso de un estudio prospectivo realizado anteriormente con recolección y archivo de tejido en el estudio original. La medición y el análisis de biomarcadores son prospectivos con un SAP contemporáneo.	Ninguno o resultados irregulares
II	C	Registro prospectivo o series de casos diseñados para abordar una pregunta sobre biomarcadores con inscripción prospectiva y recolección de tejidos. El tratamiento y el seguimiento suponen el estándar de cuidado. La medición y el análisis de biomarcadores son prospectivos con un SAP contemporáneo.	Dos o más con resultados sistemáticos
III	C	Registro prospectivo o series de casos diseñados para abordar una pregunta sobre biomarcadores con inscripción prospectiva y recolección de tejidos. El tratamiento y el seguimiento suponen el estándar de cuidado. La medición y el análisis de biomarcadores son prospectivos con un SAP contemporáneo.	Ninguno o uno con resultados sistemáticos o resultados irregulares
IV-V	D	Diseño de estudio retrospectivo con muestras por conveniencia de muestras archivadas sin reclutamiento prospectivo de pacientes. No hay tratamiento ni seguimiento definido. La medición de biomarcadores y el análisis de resultados no utilizan un SAP prospectivo contemporáneo.	No aplicable – Nunca satisfactorio

Fuente: Simon RM, et al. *J Natl Cancer Inst.* 2009;101:1446-1452.

**TABLE 2** Resumen de la plataforma de la prueba

Característica	Índice de cáncer de mama	EndoPredict	IHC4	MammaPrint	Prueba de cáncer de mama OncotypeDX	Prosigna
Proveedor	bioTheranostic	Myriad Genetics	Genoptix	Agencia BV	Genomic Health, Inc.	Nanostring, Inc.
Validado en	ESBC con ER+, LN- tratado con TAM	ESBC con ER+ HER2- tratado con TAM de ABCSG-6 (sección de tam únicamente) y ABCSG-8	ESBC con ER+ tratado con TAM o anastrozol	Etapas I, II: ≤ 5.0 cm LN-; ER+ o ER-; HER2: negativo o positivo, <53 años de edad	ESBC con ER+ Ganglio + o -, tratado con TAM	ESBC con ER+, ganglio negativo o positivo en mujeres postmenopáusicas
Genes incluidos	Coefficiente HOX B13/IL17BR, índice de grado molecular (5 genes) (Jenkovitz BCR 2011)	9 relacionados con el riesgo más 3 de normalización; Combinado con indicadores clínicos para el índice de EPClin (Filipits Clin Cancer Res 2011)	4 relacionados con cáncer de mama (Cuzick J Clin Oncol 2011)	70 (van de Vijver N Engl J Med 2002)	5 relacionados con la proliferación, 2 relacionados con la invasión, 4 relacionados con ER, 5 otros, 5 de referencia (Paik N Engl J Med 2004)	50 relacionados con PAM50, 8 de normalización (Nielsen BMC Cancer 2014)
Procesamiento del ensayo	Central por proveedor	Central por proveedor	Laboratorios locales	Central por proveedor	Central por proveedor	Laboratorios locales
Plataforma analítica	RT-PCR	RT-PCR	Inmunohistoquímica	Micromatriz, laboratorio central	RT-PCR	Sistema nCounter de Nanostring Technologies, Inc.
Requisitos de la muestra	Bloque o láminas de FFPE	Bloque o láminas de FFPE	Láminas de FFPE o micromatrices de tejido	Bloque o láminas de FFPE	Bloque, láminas o núcleos de FFPE	Bloques o láminas de FFPE
Lectura del ensayo	Dos categorías de riesgo	Continuo, dos categorías de riesgo	Índice de riesgo continuo	Dos categorías de riesgo	Índice de riesgo continuo, tres categorías de riesgo	Tres categorías de riesgo más índice de riesgo continuo
Pautas de referencia	ASCO (ESBC con ER+ LN-, para pronóstico); St. Gallen (ESBC para pronóstico); St. Gallen (ESBC para ER+ para pronóstico), AGO (ESBC con ER+ LN- para pronóstico),	ASCO (ESBC con ER+ LN-) para pronóstico; St. Gallen (ESBC para pronóstico), ESMO (cáncer de mama primario con ER+ para pronóstico y predicción), AGO (ESBC con ER+ LN- para predicción)	Ninguno	ESMO (cáncer de mama primario para pronóstico y predicción), St. Gallen (ESBC con ER+ para pronóstico), NCCN (ESBC para pronóstico)	ESMO (cáncer de mama primario con ER+ para pronóstico y predicción), St. Gallen (ESBC con ER+ para pronóstico y predicción), NCCN (ESBC con ER+ para pronóstico y predicción), ASCO (ESBC con ER+ LN- para pronóstico y predicción), NICE (ESBC con ER+ LN- para pronóstico y predicción), AGO (ESBC con ER+ LN- para pronóstico)	St. Gallen (ESBC para pronóstico), NCCN (ESBC con ER+ para pronóstico), AGO (ESBC con ER+ LN- para pronóstico)
Disponibilidad comercial	Sí	Sí	Sí	Sí	Sí	Sí

e incluyeron una prueba estadística formal de los criterios de valoración de resultados especificados previamente. Se excluyeron los estudios epidemiológicos, exploratorios y puramente descriptivos. Debido a que esta revisión se centra en la toma de decisiones en el contexto postoperatorio adyuvante inmediato, no hemos incluido la evaluación de las pruebas de expresión génica en el contexto de la terapia neoadyuvante o adyuvante extendida, incluso cuando el LOE era alto (por ej.; Sgroi et al<sup>8</sup>). También se excluyeron publicaciones no relacionadas con el cáncer de mama invasor en mujeres, y estudios sin publicación revisada por colegas. Los informes finales publicados en formato electrónico que cumplieron con los requisitos, también han sido elegibles. Los investigadores analizaron por completo todos los artículos que cumplían con los requisitos para determinar si el estudio satisfacía las condiciones especificadas previamente. El diagrama de flujo PRISMA para esta revisión se presenta en la Figura S1.

## 2.4 | Análisis del estudio

Se identificó un total de 34 artículos para incluir en este análisis (Table S1). Los datos de cada artículo seleccionado se volcaron en un formulario de abstracción estructurado. El marco de Simon y colegas se utilizó para evaluar el LOE de los estudios identificados.<sup>7</sup> A cada elemento de los artículos analizados se le asignó una calificación de acuerdo con el mejor ajuste en relación con los cuatro tipos de estudio del modelo de Simon (Table 1). Si los elementos individuales de un estudio no calificaban en la misma categoría, se utilizó la categoría más baja de un elemento individual dado, como calificación general del estudio.

## 3 | RESULTADOS

Los elementos clave de cada estudio y el LOE de Simon asignado, se resumen en la Table 3.

### 3.1 | Índice de cáncer de mama

El Índice de Cáncer de Mama (BCI, por sus siglas en inglés) es una prueba por reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) con dos componentes: el coeficiente de HOXB13: IL17BR (H/I) y el Índice de Grado Molecular (MGI, por sus siglas en inglés).<sup>9</sup> El coeficiente H/I se concibió al evaluar la expresión génica diferencial en 60 pacientes postmenopáusicas con ESBC y receptor de estrógeno positivo (ER+) tratadas con tamoxifeno (TAM+).<sup>10</sup> El componente de MGI del BCI se desarrolló identificando genes expresados de manera diferenciada entre tumores de grado bajo y alto con enfoque en los que estaban involucrados en el crecimiento invasivo.<sup>9,11</sup> El coeficiente H/I y el MGI han sido objeto de estudios de validación de pronóstico independientes. Debido a que los componentes individuales no se utilizan de manera separada en las pruebas comerciales, sus estudios de validación no se analizan aquí.<sup>9,12,13</sup>

La prueba BCI se enfocó en 314 muestras de pacientes postmenopáusicas con ESBC, TAM+, ER+ y ganglios linfáticos

negativos (LN-), incluidas en el estudio clínico de Estocolmo de tamoxifeno adyuvante (TAM). Se obtuvo como resultado un índice de riesgo continuo de 0 a 10.<sup>14</sup> El estudio clínico de Estocolmo, que se llevó a cabo entre 1976 y 1990, distribuyó aleatoriamente a pacientes postmenopáusicas con LN- a recibir TAM durante 2 a 5 años o a ningún tratamiento.<sup>15</sup> El BCI se validó en 274 pacientes de la rama sin tratamiento con tamoxifeno (TAM-). El BCI resultó significativamente pronóstico para riesgo de recaída a distancia ( $p < 0,001$ ) (DR, por sus siglas en inglés), independiente del tamaño y grado tumoral, la condición del receptor de progesterona (PR, por sus siglas en inglés) y la condición del HER2/neu.

Posteriormente se desarrolló un nuevo algoritmo de BCI a partir de 283 mujeres postmenopáusicas utilizando la rama TAM- del mismo estudio clínico de Estocolmo y, a continuación, se validó en 317 pacientes del estudio clínico con TAM+.<sup>16</sup> El uso de las mismas cohortes de pacientes en estos dos estudios separados con algoritmos diferentes disminuye la independencia de ambos estudios y reduce su valor para la validación clínica.<sup>14,16,17</sup> Con el nuevo algoritmo, se clasificó a las pacientes en grupos de riesgo bajo (64%), intermedio (20%) y alto (16%), con supervivencia sin recidiva a distancia a 5 años (DRFS, por sus siglas en inglés) del 98%, 95,2% y 87,8% respectivamente. Hubo una asociación significativa ( $p = 0,0063$ ) entre el grupo de riesgo y DRFS. La validación externa se realizó en una muestra por conveniencia de varias instituciones de 358 pacientes premenopáusicas (30%) y postmenopáusicas con tumores ER+ LN-.<sup>16</sup> El tratamiento fue heterogéneo en este grupo: el 32% recibió QTx con base ad hoc. En esta muestra, los análisis monofactoriales y de variables múltiples de Cox demostraron que el BCI era el único indicador de pronóstico significativo entre edad, tamaño tumoral, condición del PR o tratamiento con QTx.

Sgroi et al evaluaron 665 muestras de tumores primarios de pacientes postmenopáusicas con ER+ LN- que se incluyeron en el estudio clínico aleatorizado ATAC de TAM vs. anastrozol.<sup>8</sup> El ARN analizado había sido preparado por Genomic Health para otro estudio con el grupo ATAC mediante el uso de su tecnología de dominio privado. En el análisis primario, el uso del modelo cúbico del BCI como variable continua no se asoció con DR (HR 1,39 intercuartil [IC del 95% 0,99-3,70];  $LR-\Delta\chi^2 = 3,70$ ;  $p = 0,054$ ), aunque los grupos de categorías de riesgo previamente especificados sí se asociaron con DR ( $p < 0,0001$ ). En un análisis secundario, se utilizó un modelo lineal modificado del BCI, lo que demostró diferencias significativas en el índice de DR absoluto entre los grupos de riesgo del BCI ( $p < 0,0001$ ). La DR a 10 años para los grupos de riesgo bajo, intermedio y alto fue del 4,8% (IC del 95%, 3,0-7,6), 18,3% (IC del 95%, 12,7-25,8), y 29% (IC del 95%, 21,1-39,1), respectivamente.

El modelo cúbico del BCI se probó nuevamente en 292 muestras de pacientes de ambas ramas del estudio clínico NCIC MA.14 que comparó TAM y octreotida vs. TAM solo.<sup>18</sup> La octreotida demostró ser un agente inactivo en el estudio principal. Todas las pacientes eran postmenopáusicas y el 92% tenía enfermedad ER+. En total, el 51% era LN- y el 35% recibió QTx adyuvante. Los grupos de riesgo del BCI tuvieron una asociación univariada significativa con supervivencia sin recaída (RFS) (Log-Rank estratificado  $p = 0,005$ ). El modelo escalonado

**TABLE 3** Resumen de estudios de validación discutidos con nivel de evidencia asignado por Simon Hayes

Prueba	Enfoque	Estudio	Entorno	Premenopáusica	Postmenopáusica	ER+	ER-	LN+	LN-	Categoría Simon
BluePrint										
Pronóstico										
		Krijgsman Breast Cancer Res Treat 2011	Varios	✓		✓	✓	✓	✓	D
Índice de cáncer de mama										
Pronóstico										
		Goetz Clin Cancer Res 2006	AET		✓	✓		✓	✓	B
		Jerevall Br J Cancer 2011	AET		✓	✓			✓	B
		Sgroi Breast Cancer Res 2016	AET	✓	✓	✓	✓	✓	✓	B
Pronóstico y predictivo										
		Sgroi Lancet Oncol 2013	AET		✓	✓		✓	✓	B
		Zhang Clin Cancer Res 2013	AET		✓	✓		✓	✓	B
EndoPredict										
Pronóstico										
		Filipits Clin Cancer Res 2011	AET		✓	✓		✓	✓	B
		Dubsky Ann Oncol 2013	AET		✓	✓		✓	✓	B
		Martin Breast Cancer Res 2014	AET/ACT	✓	✓	✓		✓	✓	B
		Buus J Natl Cancer Inst 2016	AET	✓	✓	✓		✓	✓	B
IHC4										
Pronóstico										
		Cuzick J Clin Oncol 2011	AET		✓	✓		✓	✓	B
		Park Oncology 2014	AET/ACT	✓	✓	✓		✓	✓	C
		Sgroi Lancet Oncol 2013	AET		✓	✓		✓	✓	B
		Stephen Br J Cancer 2014	AET	✓	✓	✓		✓	✓	B
MammaPrint										
Pronóstico										
		Bueno-de-Mesquita Lancet Oncol 2007	AET/ACT	✓		✓	✓	✓	✓	D
		Buyse J Natl Cancer Inst 2006	Sin Rx	✓		✓	✓		✓	C
		Cardoso N Engl J Med 2016	AET/ACT	✓	✓	✓	✓	✓	✓	A
		Mook Ann Oncol 2010	AET	✓	✓	✓	✓	✓	✓	D
		Mook Breast Cancer Res Treat 2009	AET/ACT		✓	✓		✓		D
		van de Vijver N Engl J Med 2002	AET/ACT	✓		✓	✓	✓	✓	D
		Wittner Clin Cancer Res 2008	AET/ACT	✓	✓	✓	✓	✓	✓	D

(Acontinuación)

TABLE 3 (Continuación)

Prueba	Enfoque	Estudio	Entorno	Premenopáusica	Postmenopáusica	ER+	ER-	LN+	LN-	Categoría Simon
	Predictivo									
		Knauer Breast Cancer Res Treat 2010	ACT	✓	✓	✓	✓	✓	✓	D
Oncotype										
	Pronóstico									
		Dowsett J Clin Oncol 2010	AET		✓	✓		✓	✓	B
		Paik N Engl J Med 2004	AET	✓	✓	✓			✓	B
		Sparano N Engl J Med 2015	AET	✓	✓	✓			✓	A
		Gluz J Clin Oncol 2016	AET	✓	✓	✓		✓	✓	A
	Pronóstico y predictivo									
		Albain Lancet Oncol 2010	AET/ACT		✓	✓		✓		B
	Predictivo									
		Paik J Clin Oncol 2006	AET/ACT	✓	✓	✓			✓	B
PAM50										
	Pronóstico									
		Chia Clin Cancer Res 2012	AET	✓		✓	✓	✓	✓	B
		Dowsett J Clin Oncol 2013	AET		✓	✓		✓	✓	B
		Gnant Ann Oncol 2014	AET		✓	✓		✓	✓	B
		Gnant Ann Oncol 2015	AET		✓	✓		✓	✓	B
	Predictivo									
		Cheang Clin Cancer Res 2012	ACT	✓		✓	✓	✓		B
		Liu Breast Cancer Res Treat 2015	ACT	✓	✓	✓	✓	✓	✓	B

AET, ETx adyuvante; ACT, CTx adyuvante; ER+, receptor de estrógeno positivo; ER-, receptor de estrógeno negativo; LN+, ganglio linfático positivo; LN-, ganglio linfático negativo; LOE, nivel de evidencia (Simon J Natl Cancer Inst 2009).

y estratificado de variables múltiples de Cox sólo demostró un coeficiente de riesgo significativo para T>T2 (HR, 2,22 IC del 95% 1,22-4,07;  $p = 0,01$ ) y para el segmento más alto del continuo del BCI ( $p = 0,004$ ). Los grupos de riesgo bajo (49,7%), intermedio (23,6%) y alto (26,7%) del BCI registraron índices de RFS a 10 años del 87,5%, 83,9% y 74,7%, respectivamente. El BCI fue una herramienta de pronóstico para pacientes con LN- y ganglios linfáticos positivos (LN+).

No hay estudios de BCI con LOE alto que predigan el valor de la ETx o QTx adyuvante postoperatoria inicial en pacientes con ESBC.

### 3.2 | EndoPredict

La prueba Endopredict se desarrolló utilizando datos de expresión génica en micromatrices de 964 tumores de pacientes con ER+, HER2/neu negativo (HER2-), LN- tratadas con TAM+.<sup>19</sup> La expresión de RT-PCR de 8 genes relacionados con el cáncer y 3 genes de estandarización se combinó de manera algorítmica para producir una medición continua del riesgo. Se utilizó un valor de corte para separar los grupos de riesgo alto y bajo, lo que también se combinó con el estado ganglionar y el tamaño tumoral para calcular el índice de EPclin.

La validación clínica con un diseño prospectivo-retrospectivo evaluó 1.702 muestras de pacientes postmenopáusicas con ER+, LN- o LN+ de mujeres con ETx únicamente, en los estudios clínicos ABCSG-6 (378 pacientes) y ABCSG-8 (1.324 pacientes).<sup>19</sup> El estudio ABCSG-8 incluyó pacientes con características de riesgo bajo a moderado, excluyendo tumores de grado 3. El estudio sugirió que los índices de EP y EPclin eran variables pronósticas significativas de DR. Los índices de DR a 10 años para pacientes con EP bajo y EP alto fueron del 8% y 22% en ABCSG-6 ( $p < 0,001$ ) y del 6% y 15% en ABCSG-8 ( $p < 0,001$ ). Más del 51% de la población combinada del estudio tenía EP alto a pesar de las características de riesgo clínico bajo a moderado de las pacientes del ABCSG-8. De todas maneras, los modelos de variables múltiples de Cox demostraron que el índice EP continuo y el estado ganglionar son variables pronósticas independientes para DR.

Cuando se evaluó a las mismas poblaciones de mujeres postmenopáusicas del ABCSG-6 y ABCSG-8 en un estudio de Dubsy, EPclin arrojó una diferencia mayor entre los grupos de riesgo (HR 5,11, IC del 95% 3,48-7,51; Log Rank  $p < 0,001$ ) que las pautas de 2007 de la National Comprehensive Cancer Network (NCCN, por sus siglas en inglés) (HR 2,16, IC del 95% 0,80-5,85;  $p = 0,119$ ), o que las pautas S3 de 2008 de Alemania (HR 2,20, IC del 95% 1,16-4,19;  $p = 0,014$ ), o que las pautas de 2011 de St. Gallen (HR 2,78 IC del 95% 1,50-5,14;  $p < 0,001$ ).<sup>20</sup> EPclin reasignó 58 pacientes (61%) clasificadas como de riesgo alto según las pautas, a riesgo bajo de EPclin; a pesar que que las pautas clínicas de la época. clasificaban al 80% de las pacientes como de riesgo alto. Finalmente, los autores sugirieron que el predominio de pacientes con riesgo verdaderamente bajo en esta cohorte limitaba la potencia estadística del estudio y propusieron que los análisis del resto del subgrupo con riesgo clínico alto se consideraran exploratorios. Los autores sugirieron QTx para el grupo de riesgo alto, pero no se presentó evidencia para respaldar EPclin como variable predictiva de los beneficios de la QTx.

Martin estudió las propiedades pronósticas de EP y EPclin en una cohorte independiente de 555 pacientes con enfermedad ER+ LN+ del estudio prospectivo GEICAM 9906 de 1.246 pacientes tratadas con QTx y ETx.<sup>21</sup> Los grupos de riesgo definidos por EP y EPclin se asociaban significativamente con la supervivencia sin metástasis a distancia (DMFS, por sus siglas en inglés) (Log Rank  $p < 0,0001$ ). Para EP, el 25% (141) de las pacientes evaluables eran de riesgo bajo, con DMFS estimada a 10 años del 93%. Entre el 75% (414) de las pacientes de riesgo alto, la DMFS a 10 años fue del 70%. La diferencia absoluta en términos de riesgo entre grupos de EP de riesgo bajo y riesgo alto fue del 23% (HR, 4,8 (IC del 95% 2,45-9,55),  $p < 0,0001$ ). Con EPclin, solo el 13% era de riesgo bajo con DMFS a 10 años del 100%. Las pacientes de riesgo alto registraron una DMFS a 10 años del 72%.

El estudio GEICAM 9906 incluyó 300 (54%) pacientes premenopáusicas y 255 (46%) pacientes postmenopáusicas con un valor pronóstico evidente en ambos grupos (premenopáusicas, HR, 6,7;  $p < 0,0001$ ; postmenopáusicas, HR, 3,3;  $p = 0,0069$ ). De manera similar, EPclin también sirvió como herramienta de pronóstico en los subconjuntos de pacientes premenopáusicas ( $p = 0,0006$ ) y postmenopáusicas ( $p = 0,0023$ ). Debido a que todas las pacientes del GEICAM 9906 recibieron QTx, EP no puede evaluarse de manera separada como variable predictiva de los beneficios de la QTx. A la fecha, no hay estudios de LOE de EP o EPclin que identifiquen una función para estas pruebas en la predicción de los beneficios de QTx o ETx.

Se comparó la capacidad como herramienta de pronóstico de EP y EPclin con Oncotype DX mediante el uso de 928 muestras de ARN provenientes del estudio ATAC que Genomic Health había preparado anteriormente.<sup>22</sup> En un análisis de variables múltiples, EP, EPclin y la prueba Oncotype DX fueron todas herramientas de pronóstico de la DR a 10 años. EP proporcionó más información de pronóstico que el Índice de Recurrencia (RS) de Oncotype DX en el intervalo de 5 a 10 años. Por otra parte, los ensayos pronosticaron de la misma manera en el intervalo de 0 a 5 años. Los factores clínicos del índice de EPclin arrojaron información de pronóstico que superó a RS y EP. Pero los factores clínicos no fueron generados por la prueba, y estos no se incorporaron de manera similar a los datos del RS, lo que dificultó la comparación directa de EPclin y Oncotype DX.

El componente genómico de estas pruebas clasifica a las pacientes en grupos de riesgo bajo y alto de diferentes tamaños: EP clasificó a más pacientes como de riesgo alto (58,4%) en comparación con la prueba Oncotype DX, que clasificó al 38,3% con un RS "no bajo" ( $RS \geq 18$ ). El coeficiente de riesgo (HR por sus siglas en inglés) para la DRS a 10 años entre las pacientes de riesgo bajo y riesgo no bajo fue de 2,98 (IC del 95% 1,94-4,58,  $p < 0,001$ ) para EP en comparación con HR = 2,73 (IC del 95% 1,91-3,89,  $p < 0,001$ ) para Oncotype DX). Aunque los HR son similares y los intervalos de confianza se superponen, una proporción mayor de pacientes de EP se considerarían como de riesgo alto y quizás podrían recibir indicación de QTx.

Más recientemente, Martin et al, evaluaron a pacientes del mismo estudio GEICAM 9906 y compararon las características como herramienta de pronóstico de EP y EPclin con una versión de investigación no comercial de PAM50.<sup>23</sup> Sin embargo, la comparación

no representa una evaluación del mundo real de la prueba de Prosigna, resultando de poco valor para la toma de decisiones prácticas.

### 3.3 | IHC4

IHC4 es una prueba integrada de inmunohistoquímica (IHC, por sus siglas en inglés) que utiliza cuatro biomarcadores proteínicos de cáncer de mama (ER, PR, HER2 y Ki-67) para evaluar el riesgo en pacientes con ER+ ESBC.<sup>24</sup> El análisis inicial se llevó a cabo en un solo laboratorio con anticuerpos exclusivos para IHC, clasificación cuidadosa del H-score para ER y análisis de imágenes especializado. Los niveles de expresión de estas proteínas se incorporan en el algoritmo de IHC4. Puede agregarse un índice de tratamiento clínico (CTS, por sus siglas en inglés) derivado de las características clínico-patológicas para obtener un índice de riesgo global. La prueba se realizó con tejido fijado en formol e incluido en parafina (FFPE por sus siglas en inglés).

El desarrollo y la validación se basaron en muestras de tejido FFPE del estudio clínico ATAC.<sup>24</sup> Los biomarcadores se evaluaron según la prueba del coeficiente de probabilidad y demostraron aportar información de pronóstico para tiempo a DR (TTDR, por sus siglas en inglés). En la cohorte con ER+, LN- de ATAC, el IHC4 combinado con CTS registró un valor de pronóstico mayor que CTS únicamente. (IHC4 LR- $X^2 = 29,3$ , IC de 95% 27,7 a 30,3).

También se evaluó IHC4 en una muestra por conveniencia de 786 pacientes postmenopáusicas y premenopáusicas de Nottigham con ESBC y ER+ que no recibieron tratamiento o que recibieron tratamiento con ETx únicamente.<sup>24</sup> Debido a la lectura manual y al uso de un anticuerpo diferente, los niveles de Ki-67 fueron casi 2,5 veces más altos que en la cohorte de ATAC, por lo que hubo que reescalarlos. Después del ajuste, el índice de IHC4 modificado fue una herramienta de pronóstico para el resultado oncológico cuando se lo agregó al índice clínico (HR, 3,9;  $p < 0,0001$ ).

Sgroi et al evaluaron IHC4 en la población de ATAC con LN- de Cuzick, la misma población en la que se desarrolló IHC4, para confirmar el valor de la prueba inmunohistoquímica como herramienta de pronóstico.<sup>8,24</sup> En las 915 pacientes, los autores examinaron el índice de probabilidad para IHC4 y BCI; IHC4 resultó ser una herramienta de pronóstico para la recidiva a 5 años en el análisis de variables múltiples. Stephen y colegas también demostraron que IHC4 era una herramienta de pronóstico a 5 años mediante el uso del análisis independiente de Cox en muestras de la serie de BCI de Edimburgo y el estudio clínico TEAM.<sup>25</sup> Sin embargo, otros han demostrado que el nomograma construido a partir de las pautas de St. Gallen y Adjuvant Online (AOL), combinado con IHC4 + CTS, proporciona información de pronóstico más allá de la que proveen IHC4 + CTS solos.<sup>26-28</sup>

IHC4 no se ha validado para predecir los beneficios terapéuticos para los regímenes adyuvantes de ETx o QTx. Además, Cuzick y colegas han reconocido que las variaciones en términos de coherencia y resultado cualitativo entre laboratorios descentralizados han limitado el valor del ensayo.<sup>24</sup> La comercialización de la prueba por parte de Genoptix puede abordar algunas de estas cuestiones. Sin embargo, actualmente no están disponibles los datos de validación para la prueba comercial.

### 3.4 | MammaPrint

MammaPrint es una prueba basada en micromatrices que utiliza los niveles de expresión de 70 genes seleccionados para clasificar tumores como de "mala tipología" o "buena tipología". Los investigadores utilizaron muestras de tumores congelados de 78 mujeres con ESBC de los archivos del Netherlands Cancer Institute (NKI, por sus siglas en holandés) para obtener el ARN usado para desarrollar las tipologías de pronóstico. Las pacientes tenían menos de 55 años e incluían 34 con metástasis a distancia en 5 años. La DR temprana (<5 años) se escogió como discriminador.<sup>29</sup>

La validación clínica se llevó a cabo en 295 pacientes consecutivas del NKI menores de 53 años de edad con cáncer de mama en etapa temprana con ER+ (76,6%) y ER- (23,4%).<sup>30</sup> Había una heterogeneidad considerable en esta población, con 49% con LN+. De este porcentaje, solamente el 44% recibió alguna forma de ETx y/o QTx adyuvante. Casi el 21% (61/295) pertenecía al grupo de pacientes incluidos en el desarrollo original, lo que introdujo un posible sesgo. Globalmente, las pacientes con mala tipología registraron una tasa más alta de metástasis a distancia a los 10 años (HR, 5,1, IC del 95% (2,9-9,0);  $p < 0,001$ ) que las pacientes con buena tipología: DMFS de  $50,6 \pm 4,5\%$  y  $85,2 \pm 4,3\%$ , respectivamente.

Un estudio de validación posterior utilizó 307 muestras de pacientes con ESBC ER+ (71%) y ER- (29%)  $\leq$  de 60 años de edad inscritas en el registro TRANSBIG y que no habían recibido terapia adyuvante.<sup>31</sup> En este estudio, 85 de las 90 pacientes con enfermedad con ER- presentaron una tipología génica de riesgo alto. El HR para la supervivencia global (OS) acorde al grupo de riesgo fue de 2,79 (IC del 95% 1,60-4,87,  $p < 0,001$ ). También fue significativo el TTDR (HR, 2,32, IC del 95% 1,35-4,00,  $p = 0,002$ ) y la supervivencia sin enfermedad (DFS) (HR, 1,50, 1,04-2,16  $p = 0,032$ ). En cada caso, la prueba de 70 genes brindó información de pronóstico más contundente que las evaluaciones clínico-patológicas, incluso AOL.

MammaPrint también se evaluó en sujetos con LN+ de una muestra por conveniencia de varias instituciones, la cual identificó una asociación significativa entre la tipología génica y BCSS (Log Rank  $p < 0,001$ ) y DMFS a 10 años (Log Rank  $p = 0,001$ ).<sup>32</sup> En otra muestra conveniente, Wittner demostró un valor predictivo negativo del 100% en pacientes postmenopáusicas con LN- del Massachusetts General Hospital. Sin embargo, el valor predictivo positivo correspondiente fue solamente del 12%. Además, no hubo una diferencia significativa en términos de TTDR entre los grupos de tipología génica.<sup>33</sup> Mook utilizó una muestra de pacientes postmenopáusicas con LN- del NKI y mostró la existencia de una asociación significativa entre la tipología génica y BCSS (Log Rank  $p = 0,036$ ) pero no con DMFS (Log Rank  $p = 0,07$ ).<sup>34</sup>

MammaPrint es una de las dos únicas pruebas para las que hay evidencia de Nivel I-A publicada que respalda el pronóstico.<sup>35</sup> El estudio clínico MINDACT fue un estudio aleatorizado prospectivo, multicéntrico, que inscribió a 6.693 mujeres con ESBC invasor después de haber explorado 11.288. De los 4.595 fracasos de exploración, se informó que el 26% se debió a fallas técnicas de la prueba. Todas las pacientes inscritas se sometieron a la evaluación de riesgo clínico con una versión modificada de AOL y a todas se les realizó una evaluación

del riesgo genómico con MammaPrint. El riesgo clínico bajo se definió como una DMFS a 10 años superior al 88% para las pacientes con ER+ y del 92% para las pacientes con ER-. La diferencia entre los umbrales de ER+ y ER- fue fijado asumiendo un beneficio absoluto presunto de la ETx adyuvante a 5 años del 4%.

El estudio MINDACT utilizó un diseño complejo con varias ramas, aleatorizaciones y subpoblaciones de estudio. El análisis se hizo más arduo debido a errores en la evaluación del riesgo en 275 pacientes (4% de la población total). La evaluación del riesgo clínico fue defectuosa en 103 casos, principalmente debido a variaciones en la evaluación del sitio local del tamaño y grado tumorales, y del estado de los ganglios. Asimismo, 177 pacientes tuvieron cambios en la evaluación genómica debido a un error en el control de calidad de la prueba. Otra complicación fue que el 15% de las pacientes no cumplió con el tratamiento asignado.

El principal objetivo de valoración de MINDACT fue la supervivencia sin metástasis a distancia a 5 años. Esto se evaluó en una "población de prueba principal" definida previamente conformada por las pacientes con riesgo clínicamente alto y riesgo genómico bajo (cHgL). En la aleatorización hubo 644 pacientes de la población de la prueba principal, equivalente al 9,6% de MINDACT, que tenían cHgL y no recibieron QTx. Se excluyó a 21 pacientes que habían tenido cambios en la evaluación del riesgo y a 85 pacientes que no cumplieron con las recomendaciones de tratamiento. El índice de supervivencia sin metástasis a distancia fue del 94,7% (IC del 95%, 92,5-96,2), lo que superó el criterio de no inferioridad fijado de manera prospectiva del 92% y sugiere un valor como herramienta de pronóstico para la prueba MammaPrint.

Diversos estudios han sugerido que existe una asociación entre la tipología de riesgo de MammaPrint y los beneficios de QTx.<sup>36,37</sup> En un metanálisis retrospectivo realizado por Knauer usando datos agrupados de informes previos sobre terapia adyuvante fuera de estudios clínicos, la DDFS fue del 93% sin QTx y del 99% con QTx en pacientes con tipología buena (HR, 0,26, IC del 95% 0,03-2,02;  $p = 0,20$ ), y del 76% vs. el 88% en pacientes con mala tipología (HR, 0,35, IC del 95% 0,17-0,71;  $p < 0,01$ ).<sup>36</sup> En este estudio con poder estadístico muy bajo, la reducción de HR fue similar para los dos grupos a pesar de las diferencias en significación estadística. Los pequeños tamaños de las muestras y la falta de administración de QTx controlada y aleatorizada, en los estudios que lo integraron, limitan cualquier conclusión. Straver analizó pacientes en estadios II y III con respuesta a la quimioterapia neoadyuvante evaluada como una función de la agrupación de riesgo de MammaPrint. Sin embargo, los estudios neoadyuvantes no pueden extrapolarse directamente para predecir el grado de beneficios a largo plazo dado la ausencia de información sobre resultados tardíos.<sup>37</sup>

Se ha propuesto el estudio observacional prospectivo RASTER, realizado en 16 hospitales comunitarios de los Países Bajos, para respaldar el uso de MammaPrint en la selección de la terapia.<sup>38</sup> Sin embargo, las decisiones sobre el tratamiento las tomaron la paciente y el médico tratante de manera discrecional y se basaron en una combinación de MammaPrint, indicadores de pronóstico, guías de tratamiento y preferencias personales. La heterogeneidad de la terapéutica y la ausencia de MammaPrint como discriminador del tratamiento dificultan la interpretación de los resultados.

MINDACT ha proporcionado los únicos datos de LOE alto disponibles para evaluar la función de MammaPrint en los beneficios del tratamiento. La administración o no de QTx se determinó mediante la aleatorización de 2.745 pacientes en cada una de las ramas con discordancia entre la evaluación del riesgo clínico y genómico a efectos de la intención de tratamiento. El índice de DMFS a 5 años entre las pacientes con cHgL que recibieron QTx fue del 95,9% (IC del 95%, 94,0-97,2) y del 94,4% (IC del 95%, 92,3-95,9) para quienes no recibieron QTx. Se observó una ventaja del 1,5% en la DMFS a 5 años (con intervalos de confianza superpuestos) en quienes sí recibieron QTx. En la población según el protocolo, la reducción del índice de riesgo para QTx en el grupo con cHgL osciló entre el 34 y el 37% para DMFS, DFS y OS. En ese análisis con poder estadístico insuficiente, resulta señalable el impacto negativo, significativo en términos estadísticos, en la DFS por la omisión de QTx (HR 0,64; IC del 95%, 0,43-0,95,  $p = 0,03$ ). Para el grupo discordante con riesgo clínico bajo y riesgo genómico alto (cLgH) no hubo una ventaja clara atribuible a la QTx indicada siguiendo el resultado de la prueba genómica, con una supervivencia a 5 años sin metástasis a distancia del 95% para ambas ramas. En la población según el protocolo, no se registraron ventajas estadísticas en DMFS, DFS u OS por la administración de QTx. Al analizar ambos grupos discordantes, estos datos no pueden excluir una ventaja pequeña para QTx en el grupo con cHgL, y además, no sugieren beneficios para el uso de quimioterapia en el grupo con cLgH. Como resultado, los datos de MINDACT no proporcionan evidencia del valor predictivo de MammaPrint para QTx en pacientes con mal pronóstico.

Dos ensayos complementarios basados en micromatrices se comercializan con MammaPrint. TargetPrint ofrece la evaluación cuantitativa de la expresión de ER, PR y HER2; Blueprint, cuando se utiliza con MammaPrint, clasifica tumores en subconjuntos moleculares.<sup>27,30,39</sup> La concordancia entre Blueprint/MammaPrint y el conjunto génico intrínseco original de Perou et al es del 92%.<sup>39</sup> Sin embargo, el análisis de los datos sugiere que estos enfoques caracterizan poblaciones superpuestas, pero distintas de pacientes luminales.

No hay estudios con objetivos de valoración de resultados a largo plazo para estimar el valor de la combinación de MammaPrint/Blueprint. Whitworth ha sugerido un valor en la reasignación intrínseca de subconjuntos con una mejora resultante en los índices de pCR postneoadyuvancia evaluada en 403 pacientes del registro prospectivo de NBRST.<sup>40</sup> Sin embargo, el estudio no tenía objetivos de valoración de resultados a largo plazo y se basaba solo en la valoración de pCR como subrogante. Las variaciones en el grado de pCR no pueden correlacionarse con las diferencias cuantitativas de resultados en poblaciones variadas, especialmente en cáncer de mama con receptor de estrógeno positivo (HR+) y HER2-.<sup>41</sup>

### 3.5 | Oncotype DX

Oncotype DX es una prueba basada en RT-PCR que mide los niveles de expresión de 21 genes en el ARN de tejido FFPE para evaluar el riesgo de DR en ESBC con ER+. Her2- La prueba se desarrolló a partir de 250

genes de pronóstico para cáncer de mama, en un total de 447 pacientes de tres estudios independientes.<sup>42</sup> Los 21 genes incluyeron 16 genes relacionados con el cáncer y 5 genes de referencia. Un algoritmo ponderado produjo un RS continuo de 0 a 100 que puede usarse para asignar a las pacientes a las categorías de riesgo bajo (0-17), intermedio (18-30) o alto ( $\geq 31$ ). Se han propuesto valores de corte de riesgo alternativos.<sup>43,44</sup>

La validación de Oncotype DX se realizó en un estudio prospectivo-retrospectivo con 668 muestras de tejido archivadas provenientes de la rama TAM+ del estudio clínico NSABP B-14.<sup>42</sup> El índice sin DR a 10 años (DRFR) fue del 93,2% en el grupo de riesgo bajo (51% de las pacientes) y del 69,5% en el grupo de riesgo alto (27% de las pacientes) ( $p < 0,001$ ). El Análisis de Riesgos Proporcionales de Cox con variables múltiples demostró la probabilidad de que el DR estuviera asociado con el RS (HR, 3,21 por 50 unidades, IC del 95% 2,23-4,61;  $p < 0,001$ ), independientemente de la edad o el tamaño tumoral. El RS continuo fue una variable pronóstica de DR y OS ( $p < 0,001$ ).

Un segundo estudio de validación prospectivo-retrospectivo del pronóstico analizó 1.231 muestras de tejido FFPE del estudio clínico ATAC de mujeres postmenopáusicas con tumores ER+, LN- y LN+ que habían recibido 5 años de anastrozol o TAM.<sup>45</sup> Entre las 872 pacientes con LN- se registró una asociación significativa del RS con el riesgo de DR (HR por 50 unidades, 3,92, IC del 95% 2,08-7,39;  $p < 0,001$ ) en un modelo de variables múltiples que se ajustó según los efectos de tamaño tumoral, grado tumoral local, edad y tratamiento. La prueba registró un valor similar al distinguir TTDR para la población con LN+ (HR, 3,47, IC del 95% 1,64-7,38;  $p = 0,002$ ). Cuando el RS era  $< 18$ , los riesgos de DR a 9 años fueron similares entre las pacientes con LN- y 1-3 LN+.

El valor como herramienta de pronóstico de Oncotype DX se analizó además en pacientes postmenopáusicas con LN+ mediante el uso de dos ramas del estudio SWOG 8814.<sup>46</sup> De las 367 pacientes con muestras disponibles, 148 recibieron TAM+ únicamente, mientras que 219 recibieron QTx con ciclofosfamida-doxorrubicina-fluorouracilo (CAF) seguida por TAM+. El RS de Oncotype DX evaluado en el grupo al que se administró TAM+ únicamente fue una herramienta de pronóstico de DFS. El análisis de regresión de Cox mostró que el RS continuo era una variable pronóstica significativa para DFS (HR por 50 unidades, 2,64, IC del 95% 1,33-5,27;  $p = 0,006$ ). El RS continuo también fue una herramienta de pronóstico para OS durante un período de 10 años con HR 4,42 (IC del 95% 1,96-9,97;  $p < 0,001$ ) ajustado por el estado ganglionar.

El nivel más alto de evidencia para pronóstico disponible actualmente proviene del estudio clínico prospectivo del Grupo Cooperativo, TAILORx, que se inició en 2006 y que se diseñó especialmente para evaluar Oncotype DX.<sup>44</sup> La población del estudio de 10.253 mujeres elegibles con ER+, HER2-, LN- estudiadas con Oncotype DX se dividió en tres grupos de riesgo predeterminados con valores de corte de RS en 11 y 25. Se utilizaron valores de corte inferiores a los puntos de corte definidos prospectivamente para los estudios de validación clínica iniciales con el objetivo de evitar la posibilidad de un tratamiento insuficiente de las pacientes, tomando

para ello en cuenta los extremos de los intervalos de confianza del 95% para el riesgo de recidiva a 10 años y no solo las medias. A las pacientes del grupo con RS  $< 11$  se les asignó solo ETx, a las del grupo de nivel intermedio con RS 11-25 se las distribuyó aleatoriamente a QTx seguida por ETx vs. ETx únicamente y a las del grupo con RS  $> 25$  se les asignó QTx seguida por ETx. Los resultados de la aleatorización del grupo de nivel intermedio se esperan para 2017. Ya se han publicado los datos del grupo de pacientes de riesgo bajo con RS  $< 11$  y se ha proporcionado la evidencia clínica de Nivel I-A de TMUGS para fines de pronóstico en esta población.<sup>44</sup> De las 1.626 pacientes con riesgo bajo a quienes se les asignó ETx únicamente, la DFS invasiva a 5 años fue del 93,8% (IC del 95%, 92,4-94,9), la proporción sin DR fue del 99,3% (IC del 95%, 98,7 a 99,6) y la OS fue del 98,0% (IC del 95%, 97,1 a 98,6). Para DFS, los segundos cánceres primarios de mama y las muertes debidas a otras causas superaron el riesgo de recidiva a 5 años. El análisis de variables múltiples que incluyó edad, tamaño tumoral, grado histológico y tipo de cirugía no mostró una asociación significativa con el índice de DFS invasivo o la ausencia de DR. Solamente el grado histológico mostró alguna asociación significativa en términos estadísticos ( $p = 0,02$ ) con la ausencia de recidiva comparando el grado intermedio con el bajo (HR 8,07; IC del 95% 1,06-61,45) y entre el grado alto y grado bajo (HR 4,73; IC del 95% 0,29-76,42).

También están disponibles los datos prospectivos del estudio clínico Plan B del West German Study Group.<sup>47</sup> Este estudio que se inició en 2009 reclutó a 3.198 pacientes. Se diseñó para evaluar dos regímenes alternativos de QTx en pacientes con ESBC, LN+ y LN- de riesgo clínicamente alto. La prueba Oncotype DX se llevó a cabo de manera satisfactoria en 2.577 pacientes, lo que representa el 98% de la población total con HR+ con muestras de tejido tumoral. Después, de una modificación temprana del protocolo, se recomendó y aceptó el tratamiento con ETx únicamente para 348 pacientes con HR+, HER2-, enfermedad pN0-1 y RS  $\leq 11$ . A pesar de que el 31,4% era LN+ y el 20% tenía enfermedad de grado 3, el índice de DFS a 3 años fue del 98,4% (IC del 95% 97,0-99,8). En el grupo de riesgo intermedio (RS12-25), la DFS a 3 años fue del 97,5% (IC del 95%, 95,9-99,0) y en el grupo de riesgo alto (RS  $> 25$ ), la DFS a 3 años fue del 94,9% (IC del 95%, 91,4-98,4). A todas las pacientes de estos dos grupos se les asignó QTx adyuvante. A pesar de que se ha publicado solo el objetivo de valoración de DFS a 3 años, los autores sugirieron que el resultado favorable con ETx únicamente en el grupo de riesgo bajo excluía cualquier beneficio posible de la QTx.

El valor de la prueba Oncotype DX como variable predictiva del beneficio de la quimioterapia sistémica se abordó en dos estudios prospectivos-retrospectivos separados, que arrojaron, en ambos casos, evidencia de validación de Nivel 1-B. Paik analizó muestras de pacientes con ER+ LN- en el estudio clínico NSABP B-20 en el que se había distribuido de manera aleatoria a pacientes con ER+ LN- en TAM vs. TAM más QTx.<sup>48</sup> Las pacientes del grupo de riesgo alto (RS  $\geq 31$ ) presentaron una reducción significativa en la tasa de DR cuando se utilizó QTx (RR, 0,26; IC del 95%, 0,13-0,53). En contraste, no se registraron beneficios significativos con la QTx para los grupos de riesgo bajo (RS  $< 18$ ) (RR, 1,31 IC del 95%, 0,46-3,78) o de riesgo intermedio (RS18-30) (RR, 0,61; IC del 95%, 0,24-1,59). Los modelos

de Cox demostraron una interacción significativa entre el RS continuo y la QTx ( $p = 0,038$ ).

Albain analizó la prueba Oncotype DX en 367 (40%) muestras disponibles del estudio SWOG-8814 que distribuyó aleatoriamente a tratamiento con CAF y TAM+ secuencial vs. TAM+ únicamente.<sup>46,49</sup> La prueba de Log Rank de los beneficios de la QTx estratificada según la cantidad de ganglios positivos reveló, en concordancia con el estudio clínico principal, una tendencia a un beneficio en DFS para toda la población de pacientes analizada (Log Rank estratificado  $p = 0,054$ ). Sin embargo, cuando se analizó mediante la prueba Oncotype DX, la mejora en DFS en las pacientes que recibían CAF, el beneficio se limitó al grupo de riesgo alto con RS>31 (Log Rank  $p = 0,033$ ; HR = 0,59, IC del 95% 0,35-1,01). No se registraron beneficios aparentes en la DFS derivados de la incorporación de QTx en pacientes con RS <18 (Log Rank  $p = 0,97$ ; HR = 1,02, IC del 95% 0,54-1,93) o RS18-30 (Log Rank  $p = 0,48$ ; HR = 0,72, IC del 95% 0,39-1,31). El beneficio en OS se limitó de manera similar al grupo de riesgo alto (Log Rank estratificado  $p = 0,027$ ) y no se registraron mejoras en los grupos de riesgo bajo (Log Rank estratificado  $p = 0,63$ ) o intermedio (Log Rank estratificado  $p = 0,85$ ).

### 3.6 | Prosigna

La prueba de cáncer de mama Prosigna es un ensayo múltiple que mide la expresión de 50 genes de clasificación y 8 genes constitutivos para asignar un índice de "riesgo de recidiva" (ROR, por sus siglas en inglés) a las pacientes con ESBC. Nanostring, Inc. adaptó y comercializó esta prueba, que en un principio se denominó PAM50, para usarlo en su sistema analítico nCounter, que consiste en una reacción de hibridación individual con sondas etiquetadas específicas.<sup>50</sup> Se utilizó la expresión génica de 46 de los genes para comparar una muestra con los subconjuntos biológicos intrínsecos prototípicos descriptos originalmente por Perou.<sup>51</sup> Un algoritmo de dominio privado combina los datos de la expresión génica, un índice de proliferación y el tamaño tumoral para calcular un índice de ROR que corresponde al riesgo de DR a 10 años. Se asigna a las pacientes a grupos de riesgo bajo, intermedio o alto según valores de corte predeterminados. A diferencia de otras pruebas, los valores de corte varían según las características clínico-patológica. Todas las pacientes con LN+>4 se consideran de riesgo alto. Las pacientes con LN+ 1-3 y ROR  $\leq 15$  o las pacientes con LN- y ROR  $\leq 40$  se clasifican como de riesgo bajo. Se ha informado validez clínica satisfactoria cuando la prueba se lleva a cabo con procedimientos estandarizados en laboratorios locales de biología molecular.<sup>52</sup>

Dowsett evaluó Prosigna en 940 muestras de ARN que Genomic Health había extraído del estudio clínico ATAC.<sup>45,53</sup> El estudio prospectivo-retrospectivo demostró que el índice ROR continuo incrementó el valor pronóstico para DR por sobre el provisto por CTS para todas las pacientes de esa población de postmenopáusicas LN- y LN+, ER+, tratada con ETx ( $LR-\Delta X^2 = 33,9$ ,  $p < 0,001$ ). Además, el análisis de Kaplan-Meier mostró que los grupos de riesgo de ROR eran una herramienta de pronóstico para DR a 10 años que variaba según la condición de LN.

Dos informes prospectivos-retrospectivos de Prosigna utilizaron ARN y datos del estudio ABCSG-8 descripto anteriormente.<sup>54,55</sup> Cada estudio analizó la población del ABCSG-8 de pacientes postmenopáusicas con tumores de grado 1-2 solas o con pacientes adicionales del estudio ATAC. En su estudio de validación primario, Gnant y colegas analizaron 1.478 pacientes del ABCSG-8 y desarrollaron una variable predictiva clínica lineal de pronóstico (CLP) basada en grado tumoral, tamaño tumoral y estado de los ganglios.<sup>54</sup> El índice de ROR proporcionó un aumento adicional significativo en la información de pronóstico sobre el CLP (logaritmo-prueba de probabilidad:  $\Delta LR\chi^2 = 53,49$ ;  $p < 0,0001$ ). El análisis de los subgrupos sugirió que ofreció información adicional para todas las pacientes con LN- y LN+, a excepción de aquéllas con enfermedad HER2 positiva. Los grupos de riesgo de ROR especificados previamente también fueron una herramienta de pronóstico para DRFS con una probabilidad a 10 años del 96,7% (IC del 95% 94,6-98,0) para el grupo de riesgo bajo, del 91,3% (IC del 95% 88,1-93,8) para el grupo de riesgo intermedio y del 79,9% (IC del 95% 75,7-83,4) para el grupo de riesgo alto.<sup>54</sup> La clasificación en subtipo luminal A en comparación con luminal B también fue una herramienta de pronóstico para la DMFS a 10 años (HR, 2,85; IC del 95% 2,04-4,00;  $p < 0,0001$ ).

La evaluación del ROR en mujeres postmenopáusicas con LN+ ER+ se llevó a cabo en pacientes pertinentes de ABCSG-8 y ATAC. El análisis reveló un valor de pronóstico para DMFS en el subgrupo global con LN+ (riesgo alto en comparación con bajo: HR, 3,56, IC del 95% 1,62-7,80,  $p = 0,0016$ ) y en el subgrupo con 2-3 LN+ (riesgo alto en comparación con bajo/intermedio: HR, 3,023, IC del 95% 1,462-6,249,  $p = 0,0028$ ) con valores de corte predefinidos.<sup>55</sup> Sin embargo, la cantidad de LN+ en sí misma ejerció influencia en la asignación a subgrupos por encima de la expresión génica real, aun considerando su impacto en el valor de corte del subgrupo del ROR. El análisis de los subtipos intrínsecos reveló una mejora en la DMFS en las pacientes con tumores Luminal A en comparación con Luminal B en los subgrupos N1 y N2-3.

Chia continuó evaluando PAM50 en 398 pacientes postmenopáusicas con HR+ y receptor de hormonas negativo (HR-) en Estadios I-III que recibieron QTx adyuvante como parte del estudio NCIC MA.12.<sup>56</sup> A diferencia de la prueba comercial, se utilizó la tecnología RT-PCR en lugar de la tecnología n-Counter de NanoString. Los individuos con HR+ representaron el 73% de las pacientes mientras que el 94% era de Estadios I o II. Casi el 75% tenía ganglios positivos. En un análisis ajustado de variables múltiples, los subtipos intrínsecos de PAM50 funcionaron como herramienta de pronóstico más allá de las variables clínico-patológicas para DFS ( $p = 0,02$ ) y OS ( $p = 0,02$ ). El ROR continuo, combinado con un índice de proliferación y el tamaño tumoral, se correlacionó con la recaída a 10 años en pacientes con LN- y LN+ mediante el uso del análisis de ROC (índice C de DFS = 57,6%, IC del 95% 52,5-62,1 y para el índice C de OS = 61,1%, IC del 95% 55,8-66,3). Sin embargo, el foco principal del estudio fue la clasificación en los subtipos intrínsecos y no especificó previamente de manera adecuada el modelo de ROR específico que iba a usarse. El estudio sugirió un valor como herramienta de pronóstico para PAM50 en poblaciones de premenopáusicas tratadas con QTx. Sin embargo,

debido a que la QTx era obligatoria en todas las pacientes, su impacto en el resultado no puede apreciarse en su totalidad.

Cheang evaluó las propiedades como herramienta de pronóstico de la versión RT-PCR cuantitativa de PAM50 en 476 pacientes premenopáusicas con LN+ del estudio controlado y aleatorizado NCIC MA.5 que recibieron uno de dos regímenes de QTx.<sup>57</sup> La correlación entre la prueba, expresada como subtipo de ROR, y el resultado clínico mostraron que el clasificador de riesgo de ROR se asociaba con una RFS a 5 años diferenciada y una ventaja de pronóstico en OS (prueba de Log Rank  $p < 0,0001$ ). Sin embargo, debido a que todas las ramas aleatorizadas del estudio original recibieron alguna forma de QTx, no pudo evaluarse el valor predictivo del ROR para seleccionar QTx adyuvante.

Liu y asociados publicaron el único estudio prospectivo-retrospectivo de gran magnitud para evaluar PAM50 en relación con la predicción de los beneficios de la terapia adyuvante con taxano basado en el estudio clínico NCIC CTG MA.21.<sup>58</sup> El estudio original inscribió a 2.104 pacientes con 1.105 disponibles para el análisis de PAM50. Entre 2000 y 2005, se inscribieron mujeres premenopáusicas y postmenopáusicas de hasta 60 años de edad con cáncer de mama con LN+ o LN- de alto riesgo. Se distribuyó de manera aleatoria a las pacientes a recibir doxorubicina, ciclofosfamida y paclitaxel (AC/T); ciclofosfamida, epirubicina y flurouracilo (CEF) en dosis intensa; o epirubicina, ciclofosfamida y paclitaxel (EC/T) en dosis densa o dosis intensa. Ninguna de las pacientes fue asignada de manera aleatoria a ETx sin QTx. CEF y EC/T demostraron ser superiores a AC/T ( $p = 0,01$ ). A pesar de que el ROR continuo se asoció con RFS más bajo ( $p = 0,03$ ), el ROR categórico que comparaba el ROR alto con los grupos de riesgo de ROR bajo e intermedio combinados no fue una herramienta de pronóstico ni predictiva. La falta de rendimiento como herramienta de pronóstico del ROR categórico puede haber estado relacionada con el hecho de que el 78,7% de las pacientes se clasificaba en el grupo de riesgo alto de ROR y solo el 3,4% se identificaba como de riesgo bajo de ROR. Los subtipos intrínsecos basados en PAM50 tampoco proporcionaron valor predictivo para la selección o los beneficios del tratamiento. El subtipo intrínseco y el ROR continuo solo, demostraron asociaciones de pronóstico, pero no predictivas en la evaluación de la RFS a través del análisis de variables múltiples ( $p = 0,03$  y  $p = 0,002$ , respectivamente).

## 4 | DISCUSIÓN

Las pruebas moleculares brindan una oportunidad para evaluar la biología tumoral de manera más precisa que los clasificadores clínico-patológicos tradicionales. Las seis pruebas de expresión génica disponibles comercialmente que se analizaron en esta revisión proporcionan información de pronóstico diseñada para estimar el riesgo de recidiva del tumor y/o mortalidad de un individuo. Algunos fabricantes también sugieren una función predictiva para sus pruebas en la selección de diferentes opciones terapéuticas sistémicas. Sin embargo, cada uno de las pruebas se basa en conjuntos génicos diferentes, se ha desarrollado en una población exclusiva con

heterogeneidad variable y se ha probado en estudios con calidades diversas.

Al evaluar una prueba genómica, es fundamental considerar la validez analítica, la validez clínica y la utilidad clínica. Debe considerarse principalmente la evidencia publicada con revisión de colegas, en lugar de las presentaciones de ponencias o las autorizaciones de comercialización. Los estudios deben ser lógicos y deben poder interpretarse fácilmente con objetivos de valoración primarios y secundarios inequívocos. La evaluación de la predicción de los beneficios de la quimioterapia debe realizarse en base a un estudio clínico aleatorizado que compare pacientes tratadas con quimioterapia con pacientes que no la recibieron. Aunque las diferencias en la pCR neoadyuvante pueden sugerir el valor de la prueba, las diferencias en las tasas de pCR no se correlacionan directamente con el resultado a largo plazo, especialmente cuando se evalúan poblaciones heterogéneas.<sup>59,60</sup> Los datos de los objetivos de valoración clínicos deben captar eventos oncológicos relevantes de corto y largo plazo.

La validación analítica refleja el rendimiento estandarizado y la reproducibilidad de la técnica de una prueba específica. En términos ideales, la metodología de una prueba debe fijarse antes de realizar los estudios de validación clínica definitivos para que la evidencia de validación refleje de manera adecuada el desempeño de la prueba comercial. La disonancia entre una prueba de validación y la prueba disponible en el mercado se ejemplifica con el BCI, que se ha sometido a varias iteraciones. Los primeros dos estudios clave de BCI utilizaron ramas opuestas del mismo estudio de Estocolmo para fines de desarrollo y validación.<sup>14,16</sup> Este tipo de "crosstalk" entre los estudios debilita su independencia y, en consecuencia, su poder de validación. Además, los ensayos desarrollados a partir de cada estrategia fueron, en efecto, pruebas diferentes. El tercer estudio de validación en los pacientes de ATAC utilizó dos enfoques diferentes para realizar la calificación (BCI cúbico y BCI lineal). Con el BCI cúbico (análisis primario) se obtuvo una prueba de pronóstico no significativa y con el BCI lineal (análisis secundario) una herramienta de pronóstico de alto valor. Aunque la evaluación de BCI de ATAC realizada por Biotheranostics utilizó protocolos de las pruebas comerciales, Genomic Health había realizado el aislamiento original del ARN de ATAC con su tecnología de dominio privado. Como resultado de ello, el BCI, así como también las evaluaciones posteriores de Prosigna y EP de cohortes de pacientes de ATAC, incorporaron una técnica que no se reproduce en su prueba comercial final.<sup>8,22,53</sup> En la prueba Prosigna se observan variaciones similares entre el procedimiento del estudio de validación y la prueba comercial. El panel génico final de la prueba Prosigna fue diferente al ensayo PAM50 inicial debido a la omisión y adición de varios paneles génicos. Además, el ensayo se realizó con RT-PCR en algunos estudios de validación y con la tecnología de NanoString en otros.<sup>56,61</sup>

También se ha analizado la validez analítica de IHC4. Debido a la variabilidad en los ensayos de IHC, especialmente Ki-67, realizados localmente con diversos reactivos y valores de corte, se ha desarrollado una versión comercial fija que comercializa Genoptix como Nexcourse Breast. Sin embargo, no hay validaciones publicadas que vinculen los resultados comerciales de Nexcourse Breast de

Genoptix con los datos de validación de IHC4 publicados anteriormente.

Aunque los genes seleccionados y el algoritmo utilizado para la prueba MammaPrint han permanecido intactos durante su desarrollo, se han registrado cambios de procedimiento significativos, incluso un cambio de ARN de tejido preservado/congelado fresco a tejido embebido en parafina fijada en formol y variaciones en cuanto a las placas de micromatrices usadas en los estudios y las de la prueba comercial. Un estudio realizado por Sapino y colegas evaluó el FFPE de MammaPrint en comparación con el tejido fresco de MammaPrint en una serie de 211 muestras de tejidos coincidentes.<sup>62</sup> Las dos técnicas demostraron una concordancia de clasificación categórica de riesgo bajo y riesgo alto del 91,5% (IC del 95% 86,9-94,5); no obstante, las consecuencias clínicas de tener este nivel de discordancia (8.5%) entre plataformas no es claro. Bruner y colegas realizaron un análisis similar al estudiar 552 pares de muestras de MammaPrint FFPE y frescas.<sup>63</sup> Los resultados para los dos tipos de preservación de tejidos presentaron un coeficiente de correlación de Pearson de 0,93 (IC del 95% 0,92-0,94) con la mayoría de las muestras discordantes en ambos estudios dentro de un intervalo del 10% de los puntos de corte. Esta discordancia, que generalmente se refiere a los puntos de corte, cuestiona la estrategia del modelo de informe dicotómico de riesgo alto/riesgo bajo promocionado por esta prueba.

De todas las pruebas disponibles comercialmente, únicamente la prueba Oncotype DX utilizó el procedimiento de la prueba comercial para todo su conjunto de pruebas de validación. La prueba Oncotype DX se desarrolló en tejido FFPE y todos los estudios de validación han utilizado el mismo algoritmo. A pesar de que los parámetros clínicos pueden usarse con la prueba para la toma de decisiones sobre tratamiento clínico, estos no eran obligatorios para la validación de la prueba, y no se incorporan en el cálculo de RS. La validez analítica debe reflejar los controles de calidad relacionados con la patología, la histología y la extracción de ARN y otros procedimientos previos al ensayo y la precisión de la prueba en sí.

EPclin y el índice de ROR de Prosigna le agregan parámetros clínicos al ensayo genómico central para producir una prueba de pronóstico con precisión agregada. Sin embargo, el propósito principal de la prueba genómica consiste en proporcionar información nueva que de otra manera no se encuentra disponible mediante la evaluación de parámetros clínicos comunes. Siempre debe considerarse las características clínicas para la toma final de decisiones en el mundo real. Pero éstas deben agregarse a la prueba genómica, en lugar de ser intrínsecas a la misma. Según lo observado en las comparaciones de EP y EPclin con el resultado de la prueba Oncotype DX, los índices combinados pueden ser engañosos.<sup>22</sup> El análisis de Kaplan-Meier mostró una separación de HR similar entre los grupos de riesgo bajo y no bajo para las pruebas EP y Oncotype DX durante 10 años. Oncotype DX caracterizó un 48% más de pacientes como de riesgo bajo que EP solo, pero solamente un 5% más que EPclin. Sin embargo, el uso unilateral de funciones de pronóstico clínicas reconocidas limita el valor de dichas evaluaciones y brinda poca información sobre la calidad del componente genómico central.

Un componente importante de la validez analítica es la capacidad de la prueba para producir un resultado confiable. En el estudio clínico MINDACT, el 26% de las pacientes que no cumplieron con los requisitos de incorporación o el 10% del total de pacientes examinadas sufrieron una falla analítica de la prueba. Pocos de los ensayos restantes han publicado índices de falla. Sin embargo, se ha informado que el nivel de falla analítica final de la prueba Oncotype DX es inferior al 1% en 103.863 muestras secuenciales.<sup>64</sup> La prueba MammaPrint también ha sufrido una falla de control de calidad (QC) importante que no se detectó durante varios meses. Esto hizo que se informaran resultados erróneos en los laboratorios europeos y estadounidenses. Como consecuencia, hubo una sobreestimación del riesgo y recomendaciones inapropiadas de quimioterapia para cientos de pacientes. Este problema de control de calidad generó que la FDA retirara esas pruebas del laboratorio de California y complicó la realización del ensayo MINDACT en Europa.<sup>35,65</sup> A pesar de que, según lo informado, el error se debió a un cambio de proveedor de un reactivo utilizado para la extracción de ARN, este episodio demuestra lo importante que es realizar verificaciones de control de calidad internas de manera continua en pruebas individuales que pueden tener un impacto dramático en el cuidado de las pacientes.

La validación clínica exige que los estudios clínicos coincidan con la población objetivo de la prueba, y que aborden las preguntas clínicas relevantes. La población del estudio debe estar lo suficientemente bien definida como para permitir la traspolación razonable a los entornos comunes de la enfermedad. Las estadísticas de resultados no deben dar cabida a confusores por la amplia variedad de tratamientos y la heterogeneidad de los grupos. Dado el valor de la ETx en ESBC con HR+, los estudios de validación de pronóstico apropiados deben incluir ETx como base para evaluar las decisiones de quimioterapia.<sup>66,67</sup>

Un ejemplo de confusor por la heterogeneidad de la población es evidente en el estudio de validación original de MammaPrint que incluyó una mezcla de poblaciones diferentes de pacientes, que, además, recibieron diversos tratamientos.<sup>30</sup> De las 295 pacientes del estudio, 144 tenían ganglios positivos y 151 ganglios negativos. ER+ fue una característica en 69 casos, con 226 con ER-. Casi el 37% recibió QTx adyuvante, mientras que el 56% no recibió ningún tratamiento adyuvante sistémico. Los tratamientos de las pacientes no se distribuyeron de manera aleatoria ni se indicaron según un protocolo. En el estudio posterior de Buyse, ninguna de las pacientes recibió tratamiento adyuvante sistémico.<sup>31</sup> En el estudio de validación del BCI con el estudio clínico MA.14, el 35% de las pacientes recibió QTx adyuvante, lo que confundió el valor de la prueba como herramienta de pronóstico. Debido a ello, el impacto de la QTx en el valor de pronóstico de BCI continúa siendo incierto.<sup>18</sup>

El proceso TMUGS para evaluar pruebas genómicas está diseñado para estandarizar la validación mediante la asignación de niveles de evidencia a todos los estudios de validación publicados.<sup>6,7</sup> El mejor respaldo para cualquier prueba de expresión génica proviene de un estudio diseñado y realizado de manera prospectiva. Un solo estudio prospectivo bien diseñado realizado en una población relevante con seguimiento suficiente puede servir como el más alto nivel de evidencia. Sin embargo, el costo de dichos estudios puede ser

prohibitivo y es probable que no siempre sean factibles. El estudio prospectivo-retrospectivo brinda una alternativa para desarrollar un nivel de evidencia alto cuando se confirma con dos o más estudios con resultados concordantes.

A la fecha, solo se han informado tres estudios completamente prospectivos (TAILORx, PlanB y MINDACT). Las publicaciones revisadas por colegas están disponibles para la población de riesgo bajo de TAILORx (RS < 11), el estudio alemán Plan B (RS ≤ 11), y para varios subconjuntos de MINDACT.<sup>35,44,47</sup> Los estudios TAILORx y Plan B se basan en la prueba Oncotype DX, mientras que MINDACT se basa en la prueba MammaPrint de 70 genes. Todos los estudios demuestran resultados excelentes para las pacientes no tratadas con QTx en sus poblaciones de riesgo genómico bajo predeterminado. Los estudios prospectivos-retrospectivos previos han sugerido que el grupo de Oncotype DX de riesgo bajo no obtiene beneficios de la QTx. Sin embargo, MINDACT sugiere que puede extraerse un beneficio pequeño de la QTx (aproximadamente del 1,5% absoluto) en el grupo discordante de riesgo genómico bajo (cHgL) determinado por MammaPrint con una mejora asociada en la DFS. En una editorial que acompaña la publicación de MINDACT, Hudis sugirió que este beneficio pequeño puede ser significativo para algunas pacientes.<sup>68</sup> Más allá de MINDACT, no hay estudios de LOE alto definidos por TMUGS para MammaPrint. A pesar de que sugieren un valor como herramienta de pronóstico, los estudios restantes presentan defectos, como por ejemplo diseño subóptimo y heterogeneidad de la población, así como un sesgo potencial significativo en cuanto al tratamiento y la detección. Las cuatro pruebas genómicas restantes tienen un nivel de evidencia alto de al menos uno o más estudios prospectivos-retrospectivos que sugieren su valor como herramientas de pronóstico. Sin embargo, la validación se ha efectuado en poblaciones variadas de pacientes sin datos de ensayos netamente prospectivos.

En 2016, se publicó un estudio de registro prospectivo de más de 38.000 pacientes con ER+, LN-, HER2- de los Estados Unidos evaluadas con Oncotype DX. Sus resultados coinciden con los de los estudios prospectivos-retrospectivos anteriores.<sup>69</sup> En la base de datos de SEER, 21.023 pacientes con LN- y riesgo bajo (RS < 18) de Oncotype DX, presentaron una mortalidad específica por cáncer de mama a 5 años (BCSM) del 0,4% (IC del 95%, 0,3-0,6) con solo un 7% de las pacientes con constatación de haber recibido QTx.<sup>69</sup> Aún entre las pacientes del SEER con ganglios positivos (micrometástasis o hasta 3 ganglios con metástasis), la BCSM a 5 años fue de solo el 1,0% (IC del 95%, 0,5-2,0) con un 23% de pacientes con constatación de haber recibido QTx. Este estudio basado en el registro SEER, proporciona confirmación del "mundo real" de la utilidad clínica de la prueba Oncotype DX y brinda evidencia de resultados que coinciden notablemente con los de los estudios de validación anteriores.

¿Pero es suficiente la identificación del grupo de riesgo muy bajo mediante una prueba genómica validada de manera adecuada? ¿Debe esperarse beneficio de la QTx en todos los pacientes restantes ubicados por encima del límite de riesgo bajo de la prueba? Los datos prospectivos-retrospectivos utilizados para evaluar la prueba Oncotype DX en el estudio NSABP B-20 mostraron que no hubo beneficios con la QTx en pacientes con resultados de RS bajo o intermedio. Sin

embargo, se registró un beneficio muy importante con QTx en pacientes con resultados de RS de riesgo alto (riesgo relativo 0,26; IC del 95% 0,13-0,53) y una comprobación significativa en términos estadísticos para la interacción entre el tratamiento con quimioterapia y el resultado del RS. Esta relación entre el RS y el beneficio de la QTx también se observó en el estudio prospectivo-retrospectivo SWOG 8814 para LN+ con QTx/ETx en comparación con ETx únicamente.<sup>46</sup> En conjunto, estos estudios sugieren una población de pacientes (RS ≥ 31) para quienes la quimioterapia puede ser extraordinariamente activa y reducir el riesgo de recidiva hasta en un 75%. Como una transición cuántica del beneficio de QTx en RS ≥ 31 es improbable, estos datos sugieren que una parte pequeña de la población con riesgo intermedio puede responder a la QTx, quizá en los niveles de RS más altos de este grupo. Sin embargo, la mayoría de las pacientes del grupo intermedio (RS 18-30) parecen no beneficiarse de la QTx adyuvante.

Ninguna de las pruebas restantes tiene evidencia de TMUGS de nivel alto para predecir de manera convincente los beneficios de QTx en subgrupos de los subconjuntos evaluados. MINDACT brindó la oportunidad de demostrar prospectivamente el valor predictivo de MammaPrint. Sin embargo, un análisis del grupo con cHgL mostró una reducción consistente del riesgo del 35% cuando se utilizó quimioterapia en este grupo. Este beneficio fue significativo en términos estadísticos para DFS, incluso en este subgrupo con poder estadístico insuficiente. Por el contrario, el análisis del subgrupo con cLgH no proporcionó evidencia de beneficios de la quimioterapia, cuando esta indicación se impulsó por el riesgo genómico alto. Por lo tanto, a pesar de que, por el uso de la prueba, el 48% de las pacientes con cHgL pueden omitir la quimioterapia, es probable que al 47% de las pacientes con cLgH se les recomiende recibir quimioterapia innecesaria sin cambios en los resultados. Aunque es probable que un grupo con riesgo muy bajo de cualquier prueba genómica no presente riesgos o beneficios suficientes para justificar la quimioterapia, el valor predictivo sí importa cuando aumenta el riesgo. Oncotype DX brinda el argumento más fundamentado sobre la utilidad clínica para decidir el uso de QTx en pacientes con ESBC y ER+ y HER2-.

Algunos han sugerido que el pronóstico individual y la correlación entre el riesgo de recidiva y la respuesta a la QTx son iguales en todas las pruebas. Sin embargo, las pruebas no clasifican a las pacientes de manera similar y los resultados no pueden extrapolarse de una prueba a otra. Poulet y colegas realizaron una comparación prospectiva entre las versiones comerciales de las pruebas Oncotype DX y MammaPrint en 57 pacientes con ESBC y ER+ LN-.<sup>70</sup> La discordancia fue significativa. Muchas de las pacientes con riesgo alto en MammaPrint se clasificaron como de riesgo bajo en Oncotype DX y registraron una expresión alta de ER. En este caso, debido a que la expresión de ER se correlaciona con el beneficio de ETx, es probable que una paciente de MammaPrint con pronóstico malo evolucione bien con ETx únicamente. También fue evidente la falta de concordancia ( $r = 0,08$ , IC del 95% -0,2 a 0,35; correlación de Spearman) entre las pruebas comerciales Oncotype DX y Prosigna en 52 pacientes estudiadas por Alvarado y colegas. Ellos hallaron una concordancia en la clasificación del riesgo de apenas el 54%.<sup>71</sup>

Más recientemente, el estudio clínico OPTIMA Prelim evaluó 5 pruebas, cuatro de las cuales se discuten en esta revisión (Oncotype

DX, MammaPrint, Prosigna y NexCourse Breast-IHC4). Cada prueba se realizó en el laboratorio comercial de su respectivo fabricante.<sup>72</sup> Es remarkable que se registraron discrepancias significativas entre las pruebas y el nivel general de concordancia fue solamente “moderado” ( $\kappa = 0,40-0,59$ ). Los autores se sorprendieron con el hecho de que varias pruebas, cada una de ellas con su propia validación independiente de pronóstico, pudieran exhibir tal discrepancia. Sugirieron que la concordancia de las pruebas no era mejor que la categorización patológica clásica. Cabe señalar, no obstante, que OPTIMA Prelim no midió los resultados a largo plazo de las pacientes y no pudo proporcionar la validación independiente de cada prueba, ni tampoco identificar a la más precisa de las pruebas dispare. La falta de concordancia entre todas ellas sugiere que una prueba validada no puede reemplazarse de manera confiable con otras que tengan una validación menos sólida.

El desafío en la toma de decisiones médicas consiste en brindar el mejor cuidado para las pacientes actuales y, a su vez, seguir evaluando paradigmas nuevos de diagnóstico y tratamiento para pacientes futuras. La terapia adyuvante permite reducir la posibilidad de la DR y la muerte, pero conlleva su propio riesgo de toxicidad a corto y largo plazo. Las terapias modernas también generan costos significativos a los individuos y los sistemas de cuidado de la salud. Los clasificadores genómicos que proporcionan información de pronóstico precisa y predicción confiable de los beneficios, personalizan la selección del tratamiento. Mejoran así, las probabilidades de obtener resultados favorables en el largo plazo y a su vez preservan los recursos destinados al cuidado de la salud. Sin embargo, la elección debe basarse en el análisis cuidadoso de datos en permanente evolución.

## 5 | CONCLUSIÓN

En la actualidad, existen diversas pruebas de expresión génica que brindan información de pronóstico y predictiva para ESBC. El análisis de los estudios disponibles sugiere que cada una de las seis pruebas tiene algún tipo de evidencia que respalda su uso como herramienta de pronóstico en distintas subpoblaciones. Sin embargo, solo dos ensayos, Oncotype DX y MammaPrint, han publicado evidencia de estudios clínicos prospectivos importantes referida al pronóstico relevante para muchas pacientes con ESBC. Ambas pruebas identifican una subpoblación exclusiva con riesgo bajo a la que la quimioterapia le aportaría beneficios mínimos. Al día de hoy, sin embargo, solo Oncotype DX predice los beneficios de la QTx en una población amplia de pacientes pre y postmenopáusicas con HR+, LN- y LN+ validado a través de estudios clínicos de QTx adyuvante que compararon poblaciones aleatorizadas tratadas con QTx vs. no tratadas con QTx, y con seguimiento a largo plazo. Esta es la combinación de información de pronóstico sólida, y predictiva confiable que le brinda a una prueba la máxima utilidad clínica.

## RECONOCIMIENTOS

Para la realización de este estudio no se recibió ningún apoyo financiero directo.

## DIVULGACIONES DE POSIBLES CONFLICTOS DE INTERESES DE LOS AUTORES

Todos los autores completaron una declaración de divulgación. Los siguientes autores indicaron un interés financiero o de otro tipo que es relevante al tema que se trata en este artículo. Las relaciones marcadas con una “U” son aquellas para las que no se recibió ningún tipo de compensación; las marcadas con una “C” sí lo hicieron: Empleo o cargo de liderazgo: Ninguno Función de Consultor o Asesor: David M. Hyams, Biotheranostics (C), Genomic Health (C), Castlebiosciences (C), Exact Sciences (C); Eric Schuur, Genomic Health (C); Artur Katz, Bristol Myers Squibb (C), Merck Sharp Dome (C); Aníbal Nuñez De Pierro, AstraZeneca (C), Roche (C) Cartera de Acciones: Ninguna Honorario: David M. Hyams, Genomic Health, Castle Biosciences, Exact Sciences; Cesar Cabello, Genomic Health; Aníbal Nuñez De Pierro, Astrazeneca, Novartis, Siemens Fondos de Investigación: Ninguno Testimonio de Peritos: Ninguno Otra Remuneración: Ninguna.

## REFERENCIAS BIBLIOGRÁFICAS

1. International Agency for Research on Cancer WHO. *Breast cancer: Estimated incidence, mortality and prevalence worldwide in 2012*. 2012 [cited 2016 January 9]; Available from: [http://globocan.iarc.fr/Pages/fact\\_sheets\\_cancer.aspx](http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx)
2. Early Breast Cancer Trialists Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687–1717.
3. CDC Center for Surveillance Epidemiology and Laboratory Services (CSELS) Public Health Genomics. *Genomic testing: Acce model process for evaluating genetic tests*. 2010 [cited 2015 October 29]; Available from: <http://www.cdc.gov/genomics/gtesting/ACCE/index.htm>
4. Haddow JE, Palomaki GE. Acce: A model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Burke W, editors. *Human genome epidemiology: A scientific foundation for using genetic information to improve health and prevent disease*. New York: Oxford University Press; 2004:217–233.
5. Hunter DJ, Khoury MJ, Drazen JM. Letting the genome out of the bottle—will we get our wish? *N Engl J Med*. 2008;358:105–107.
6. Hayes DF, Bast RC, Desch CE, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst*. 1996;88: 1456–1466.
7. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009;101:1446–1452.
8. Sgroi DC, Sestak I, Cuzick J, et al. Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the transatac study population. *Lancet Oncol*. 2013;14:1067–1076.
9. Ma XJ, Salunga R, Dahiya S, et al. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. *Clin Cancer Res*. 2008;14:2601–2608.
10. Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell*. 2004;5:607–616.
11. Ma XJ, Salunga R, Tuggle JT, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci USA*. 2003;100: 5974–5979.

12. Goetz MP, Suman VJ, Ingle JN, et al. A two-gene expression ratio of homeobox 13 and interleukin-17b receptor for prediction of recurrence and survival in women receiving adjuvant tamoxifen. *Clin Cancer Res*. 2006;12:2080–2087.
13. Ma X-J, Hilsenbeck SG, Wang W, et al. The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. *J Clin Oncol*. 2006;24:4611–4619.
14. Jerevall PL, Ma XJ, Li H, et al. Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the stockholm trial. *Br J Cancer*. 2011;104:1762–1769.
15. Rutqvist LE, Johansson H, Stockholm Breast Cancer Study G. Long-term follow-up of the randomized stockholm trial on adjuvant tamoxifen among postmenopausal patients with early stage breast cancer. *Acta Oncol*. 2007;46:133–145.
16. Zhang Y, Schnabel CA, Schroeder BE, et al. Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence. *Clin Cancer Res*. 2013;19:4196–4205.
17. Harris LN, Ismaila N, McShane LM, Andre F. Reply to D. Sgroi et al, T. Sanft et al, m.S. Copur et al, and m. Goetz et al. *J Clin Oncol*. 2016;34:3946–3948.
18. Sgroi DC, Chapman JA, Badovinac-Crnjevic T, et al. Assessment of the prognostic and predictive utility of the breast cancer index (BCI): an NCIC CTG MA.14 study. *Breast Cancer Res*. 2016;18:1.
19. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res*. 2011;17:6012–6020.
20. Dubsy P, Filipits M, Jakesz R, et al. Endopredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. *Ann Oncol*. 2013;24:640–647.
21. Martin M, Brase JC, Calvo L, et al. Clinical validation of the endopredict test in node-positive, chemotherapy-treated ER+/HER2- breast cancer patients: results from the geicam 9906 trial. *Breast Cancer Res*. 2014;16:R38.
22. Buus R, Sestak I, Kronenwett R, et al. Comparison of endopredict and eclin with Oncotype DX recurrence score for prediction of risk of distant recurrence after endocrine therapy. *J Natl Cancer Inst*. 2016;108.
23. Martin M, Brase JC, Ruiz A, et al. Prognostic ability of Endopredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. *Breast Cancer Res Treat*. 2016;156:81–89.
24. Cuzick J, Dowsett M, Pineda S, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol*. 2011;29:4273–4278.
25. Stephen J, Murray G, Cameron DA, et al. Time dependence of biomarkers: non-proportional effects of immunohistochemical panels predicting relapse risk in early breast cancer. *Br J Cancer*. 2014;111:2242–2247.
26. Adjuvant Inc. *Adjuvant! Online*. 2013 [cited 2013 February 17]; Adjuvant! Online Home Page]. Available from: <https://www.adjuvantonline.com/>
27. Goldhirsch A, Wood WC, Coates AS, et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the st gallen international expert consensus on the primary therapy of early breast cancer 2011. *Ann Oncol*. 2011;22:1736–1747.
28. Park YH, Im SA, Cho EY, et al. Validation and comparison of CS-IHC4 scores with a nomogram to predict recurrence in hormone receptor-positive breast cancers. *Oncology*. 2014;86:279–288.
29. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415:530–536.
30. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*. 2002;347:1999–2009.
31. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006;98:1183–1192.
32. Mook S, Schmidt MK, Viale G, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. *Breast Cancer Res Treat*. 2009;116:295–302.
33. Wittner BS, Sgroi DC, Ryan PD, et al. Analysis of the Mammprint breast cancer assay in a predominantly postmenopausal cohort. *Clin Cancer Res*. 2008;14:2988–2993.
34. Mook S, Schmidt MK, Weigelt B, et al. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. *Ann Oncol*. 2010;21:717–722.
35. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016;375:717–729.
36. Knauer M, Mook S, Rutgers EJ, et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. *Breast Cancer Res Treat*. 2010;120:655–661.
37. Straver ME, Glas AM, Hannemann J, et al. The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res Treat*. 2010;119:551–558.
38. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (raster). *Lancet Oncol*. 2007;8:1079–1087.
39. Krijgsman O, Roepman P, Zwart W, et al. A diagnostic gene profile for molecular subtyping of breast cancer associated with treatment response. *Breast Cancer Res Treat*. 2012;133:37–47.
40. Whitworth P, Stork-Sloots L, de Snoo FA, et al. Chemosensitivity predicted by blueprint 80-gene functional subtype and mammprint in the prospective neoadjuvant breast registry symphony trial (NBRST). *Ann Surg Oncol*. 2014;21:3261–3267.
41. Earl H, Provenzano E, Abraham J, et al. Neoadjuvant trials in early breast cancer: pathological response at surgery and correlation to longer term outcomes—what does it all mean? *BMC Med*. 2015;13:234.
42. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351:2817–2826.
43. Harbeck N, Sotlar K, Wuerstlein R, Doisneau-Sixou S. Molecular and protein markers for clinical decision making in breast cancer: today and tomorrow. *Cancer Treatment Rev*. 2014;40:434–444.
44. Sparano JA, Gray RJ, Makower DF, et al. Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med*. 2015;373:2005–2014.
45. Dowsett M, Cuzick J, Wale C, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a transatac study. *J Clin Oncol*. 2010;28:1829–1834.
46. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol*. 2010;11:55–65.
47. Gluz O, Nitz UA, Christgen M, et al. West German Study Group Phase III PlanB trial: first prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by

- central and local pathology assessment. *J Clin Oncol*. 2016;34:2341–2349.
48. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol*. 2006;24:p. 3726–p. 3734.
  49. Albain KS, Barlow WE, Ravdin PM, et al. Adjuvant chemotherapy and timing of tamoxifen in postmenopausal patients with endocrine-responsive, node-positive breast cancer: a Phase 3, open-label, randomised controlled trial. *Lancet*. 2009;374:p. 2055–p. 2063.
  50. Geiss GK, Bumgarner RE, Birditt B, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 2008;26:317–325.
  51. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406:747–752.
  52. Nielsen T, Wallden B, Schaper C, et al. Analytical validation of the PAM50-based prognostic breast cancer prognostic gene signature assay and ncounter analysis system using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*. 2014;14:177.
  53. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 risk of recurrence score with Oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol*. 2013;31:2783–2790.
  54. Gnant M, Filipits M, Greil R, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol*. 2014;25:339–345.
  55. Gnant M, Sestak I, Filipits M, et al. Identifying clinically relevant prognostic subgroups of postmenopausal women with node-positive hormone receptor-positive early-stage breast cancer treated with endocrine therapy: a combined analysis of ABCSG-8 and atac using the PAM50 Risk of Recurrence score and intrinsic subtype. *Ann Oncol*. 2015;26:1685–1691.
  56. Chia SK, Bramwell VH, Tu D, et al. A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. *Clin Cancer Res*. 2012;18:4465–4472.
  57. Cheang MC, Voduc KD, Tu D, et al. Responsiveness of intrinsic subtypes to adjuvant anthracycline substitution in the NCIC.CTG MA.5 randomized trial. *Clin Cancer Res*. 2012;18:2402–2412.
  58. Liu S, Chapman JA, Burnell MJ, et al. Prognostic and predictive investigation of pam50 intrinsic subtypes in the NCIC.CTG MA.21 phase iii chemotherapy trial. *Breast Cancer Res Treat*. 2015;149:439–448.
  59. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNOCBC pooled analysis. *Lancet*. 2014;384:164–172.
  60. De Mattos-Arruda L, Shen R, Reis-Filho JS, Cortes J. Translating neoadjuvant therapy into survival benefits: one size does not fit all. *Nat Rev Clin Oncol*. 2016;13:566–579.
  61. Filipits M, Nielsen TO, Rudas M, et al. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clin Cancer Res*. 2014;20:1298–1305.
  62. Sapino A, Roepman P, Linn SC, et al. Mammprint molecular diagnostics on formalin-fixed, paraffin-embedded tissue. *J Mol Diagn*. 2014;16:190–197.
  63. Beumer I, Witteveen A, Delahaye L, et al. Equivalence of mammprint array types in clinical trials and diagnostics. *Breast Cancer Res Treat*. 2016;156:279–287.
  64. Anderson JM, Shak S, Millward C, et al. Molecular characterization of breast cancer core biopsy specimens by gene expression analysis using standardized quantitative RT-PCR. *Cancer Res*. 2009;69:Abstract nr 6021.
  65. United States Food and Drug Administration. *Class 2 device recall Agendia Mammprint*. 2010 [cited 2016 September 27]; Available from: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRes/res.cfm?ID=89513>
  66. Cuzick J. Aromatase inhibitors for breast cancer prevention. *J Clin Oncol*. 2005;23:1636–1643.
  67. Early Breast Cancer Trialists Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet*. 1998;352:930–942.
  68. Hudis CA, Dickler M. Increasing precision in adjuvant therapy for breast cancer. *N Engl J Med*. 2016;375:790–791.
  69. Petkov VI, Miller DP, Howlander N, et al. Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. *NPJ Breast Cancer*. 2016;2:16017.
  70. Poulet B, Jamshidian F, Butler S, et al. Risk classification of early stage breast cancer as assessed by MammPrint and Oncotype DX genomic assays. *Cancer Res*. 2012;72:Abstract nr P6-07-03.
  71. Alvarado MD, Prasad C, Rothney M, et al. A prospective comparison of the 21-gene recurrence score and the PAM50-based Prosigna in estrogen receptor-positive early-stage breast cancer. *Adv Ther*. 2015;32:1237–1247.
  72. Bartlett JMS, Bayani J, Marshall A, et al. Comparing breast cancer multiparameter tests in the OPTIMA Prelim trial: no test is more equal than the others. *J Natl Cancer Inst*. 2016;108:pj:050. <https://doi.org/10.1093/jnci/djw050>

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Hyams DM, Schuur E, Aristizabal JA, et al. Selección de la terapia sistémica adyuvante postoperatoria para cáncer de mama en etapa temprana: Una evaluación crítica de las pruebas de expresión génica disponibles comercialmente. *J Surg Oncol*. 2017;1–17. <https://doi.org/10.1002/jso.24831>

## SINOPSIS DEL CONTENIDO

Las pruebas de expresión génica brindan un abordaje a la biología tumoral intrínseca para poder individualizar mejor la selección de la terapia adyuvante. Se presenta el desarrollo y la evidencia complementaria de las pruebas para cáncer de mama en etapa temprana disponibles comercialmente. El análisis de esta evidencia posteriormente puede usarse para seleccionar pruebas basadas en criterios de validación bien establecidos y niveles de evidencia aceptados.