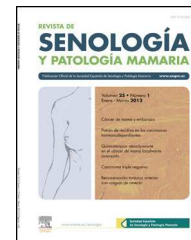




Revista de Senología y Patología Mamaria

www.elsevier.es/senologia



ARTÍCULO DOCENTE

Lectura crítica de la literatura biomédica (II): cómo valorar la calidad de la evidencia y establecer recomendaciones



Jaime Latour-Pérez^{a,*} y Álvaro Rodríguez-Lescure^b

^a Servicio de Medicina Intensiva, Hospital General Universitario de Elche, Elche, España

^b Servicio de Oncología Médica, Hospital General Universitario de Elche, Elche, España

Recibido el 29 de julio de 2014; aceptado el 3 de octubre de 2014

Disponible en Internet el 18 de marzo de 2015

PALABRAS CLAVE

Lectura crítica;
Ensayos
aleatorizados;
Calidad de la
evidencia;
Guías de práctica
clínica;
Fuerza de la
recomendación

KEYWORDS

Critical appraisal;
Randomized clinical
trials;
Quality of the
evidence;
Clinical practice
guidelines;
Strenght of
recommendation

Resumen El ejercicio de una práctica clínica basada en la evidencia exige unas competencias básicas para valorar e interpretar la literatura biomédica. En este artículo se revisan los puntos críticos para determinar la calidad de la evidencia global y formular recomendaciones.

© 2014 SESPM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Critical appraisal of the biomedical literature (II): How to assess the quality of evidence and the strength of recommendations

Abstract Clinicians interested in evidence-based clinical practice need some basic competencies in critical appraisal of the biomedical literature. This article reviews key concepts for the assessment of the quality of the evidence and the strength of recommendations.

© 2014 SESPM. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

En la primera parte de esta serie¹ se presentó el escenario de una mujer de 53 años con adenocarcinoma ductal infiltrante de mama subsidiario de cirugía conservadora, con

* Autor para correspondencia.

Correo electrónico: jlatour@coma.es (J. Latour-Pérez).

2/2 biopsias de ganglio centinela positivas, en la que se planteaba la opción de vaciamiento ganglionar frente a cirugía axilar conservadora. Una vez leído detenidamente el estudio, se concluía que el ensayo ACOSOG-Z0011 presenta debilidades metodológicas importantes (en especial las pérdidas en el seguimiento y el desequilibrio en las variables pronósticas basales) que le restan credibilidad.

Por otro lado, somos conscientes de que la calidad de la evidencia no se puede valorar con un único estudio, y hemos decidido examinar el cuerpo de evidencia disponible. En este punto, seguiremos los pasos de la aproximación *Grading of Recommendations Assessment, Development and Evaluation* (GRADE) para establecer el grado de evidencia y la fuerza de la recomendación sobre la seguridad de evitar la disección ganglionar en nuestra enferma.

La aproximación *Grading of Recommendations Assessment, Development and Evaluation*

GRADE es un sistema formal de jerarquización del grado de evidencia y de la fuerza de la recomendación, que ha sido asumido por un gran número de organizaciones².

El sistema GRADE define la calidad de la evidencia en el contexto de una revisión sistemática como la confianza que tenemos en que la estimación del efecto es correcta. GRADE clasifica la calidad de la evidencia en 4 grados: alta, moderada, baja o muy baja. En gran medida el grado de evidencia está condicionado por el tipo de diseño: así, los estudios aleatorizados parten de un nivel de evidencia alto, sin embargo la evidencia procedente de ensayos clínicos puede degradarse en uno o 2 niveles (a evidencia moderada o baja) en función de ciertas características que comprometen la credibilidad de los resultados. De igual forma, los estudios observacionales parten de nivel de evidencia bajo, pero este nivel puede aumentar (generalmente en un grado) o reducirse si concurren otras características.

Respecto al grado de recomendación a favor o en contra de una intervención, GRADE lo clasifica en 2 niveles (fuerte o débil), según el grado de confianza en que los efectos beneficiosos superan los efectos adversos, lo cual depende de la relación (*trade-off*) entre los efectos beneficiosos y los efectos adversos de la intervención, del nivel de evidencia, de las preferencias del paciente y, en ocasiones, de los costes.

En las páginas que siguen intentaremos seguir los pasos de la metodología GRADE para intentar resolver el escenario que nos planteábamos al principio del artículo.

Formulación de la pregunta y búsqueda de la evidencia

El primer paso es formular correctamente la pregunta, según el formato PICO, especificando el tipo de Paciente, la Intervención, la alternativa de Comparación y desenlace clínico (*Outcome* en inglés). En el escenario que nos ocupa estos 4 elementos podrían definirse de la siguiente manera:

- Paciente: mujer con cáncer de mama sin indicaciones de mastectomía, en estadio T1-T2, sin adenopatías palpables, con metástasis en la biopsia del ganglio centinela.

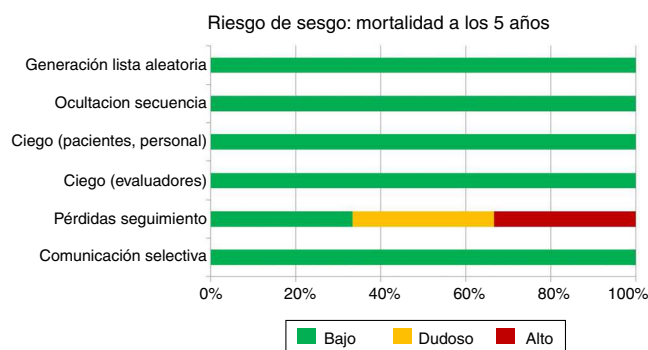


Figura 1 Valoración del riesgo de sesgo global.

- Intervención: biopsia de ganglio centinela sin disección de adenopatías axilares.
- Comparación: vaciamiento axilar.
- Desenlace clínico a evaluar: supervivencia global, supervivencia libre de enfermedad, recidivas locorreccionales, linfedema, etc.

Estos desenlaces se jerarquizan en una escala de 9 grados, agrupados como desenlaces críticos para la decisión (niveles 7-9), desenlaces importantes pero no críticos para la decisión (niveles 4-6) y desenlaces poco importantes (niveles 1-3). Por razones de claridad expositiva, en nuestro escenario consideramos solo 3 desenlaces, que ordenamos de la siguiente forma: supervivencia global (9), recidivas locorreccionales (7) y morbilidad quirúrgica (linfedema) (7).

Las preguntas que nos formulamos (una para cada desenlace) son preguntas de efectividad, por lo que inicialmente buscaremos para contestarlas ensayos aleatorizados o revisiones sistemáticas de ensayos aleatorizados. Afortunadamente encontramos un metaanálisis³ con 3 ensayos aleatorizados: el estudio ACOSOG-Z0011^{4,5}, el IBCSG 23-01⁶ y el EORTC 10981-22023⁷⁻⁹.

El siguiente paso es valorar la calidad de la evidencia, lo cual se establece a partir de 5 parámetros¹: el riesgo de sesgo global², inconsistencia³, imprecisión⁴, evidencia indirecta⁵ y sesgo de publicación.

Riesgo de sesgo del conjunto de los estudios

El riesgo de sesgo de los 3 ensayos encontrados se resume mediante el instrumento de la colaboración Cochrane (fig. 1, tabla 1). Este instrumento nos muestra que, con la excepción del estudio ACOSOG-Z0011, la evidencia respecto a la mortalidad a los 5 años procede de estudios con bajo riesgo de sesgo. Es decir, el riesgo de sesgo de los estudios no constituiría por sí mismo una razón que nos haga dudar de la estimación del efecto.

Inconsistencia

La obtención de resultados parecidos en diferentes estudios aumenta la confianza en la estimación del efecto. La figura 2 muestra el diagrama de Forest correspondiente al riesgo relativo de mortalidad a los 5 años. La I^2 —un indicador de la heterogeneidad relativa— es del 0%, lo cual indica que las diferencias observadas en la mortalidad se explican

Tabla 1 Riesgo de sesgo

	ACOSOG-Z0011	IBCSG 23-01	EORTC 10981-22023
Generación de la lista aleatoria	Bajo riesgo: Un artículo previo del grupo ⁶ afirma: «a computer would randomly assign her to either group 1 (SLND_ALND; n_445) or group 2 (SLND alone; n_446) in a 1:1 ratio»	Bajo riesgo: «Randomisation was done with permuted blocks generated by a congruence algorithm»	Bajo riesgo: No se especifica
Ocultación secuencia de aleatorización	Bajo riesgo: Asignación central ⁶	Bajo riesgo: «Participating centre staff accessed the central randomisation system»	Bajo riesgo: «Randomization was accomplished centrally by the EORTC headquarters»
Enmascaramiento (participantes y personal)	Bajo riesgo: No enmascaramiento de pacientes/personal, pero es dudoso que esto influya en los resultados de mortalidad	Bajo riesgo: No enmascaramiento de pacientes/personal, pero es dudoso que esto influya en los resultados de mortalidad	Bajo riesgo: No enmascaramiento de pacientes/personal, pero es dudoso que esto influya en los resultados de mortalidad
Enmascaramiento (evaluadores)	Bajo riesgo para mortalidad, dudoso para edema linfático o movilidad	Bajo riesgo para mortalidad, dudoso para edema linfático o movilidad	Bajo riesgo para mortalidad, dudoso para edema linfático o movilidad
Manejo pérdidas	Alto riesgo: Número de pérdidas alto, capaz de condicionar sesgo de selección	Bajo riesgo (pocas pérdidas)	Riesgo dudoso (datos no disponibles)
Notificación selectiva	Bajo riesgo: El artículo incluye todos los desenlaces esperables	Bajo riesgo: Registrado en clinicaltrials.gov (NCT00072293)	Bajo riesgo: Registrado en clinicaltrials.gov (NCT00014612)
Otros	Grupos desbalanceados	No analiza por separado las metástasis > 2 mm	Datos no disponibles

Pregunta: ¿Se debe realizar disección ganglionar axilar en el cáncer de mama (no avanzado)?
Bibliografía: Rao et al³.

por error de muestreo. Por lo tanto, no tenemos argumentos de peso para degradar la calidad del estudio por inconsistencia. No obstante, si calculamos —mediante el software adecuado— el intervalo de confianza al 95% del I^2 comprobamos que este se extiende entre el 0 y el 67%; por tanto, tampoco se puede afirmar en rigor que los estudios sean consistentes.

La consistencia de los estudios respecto a la recidiva ganglionar es similar a la encontrada para la supervivencia. En cambio, respecto al desenlace del linfedema, aunque los 3 estudios muestran resultados claramente favorables al tratamiento conservador, la I^2 es del 70% ($p=0,04$), lo que indica que un porcentaje importante de la dispersión de efectos observados se debe a verdadera heterogeneidad, quizá por las diferentes definiciones operativas utilizadas por los diferentes estudios.

Imprecisión

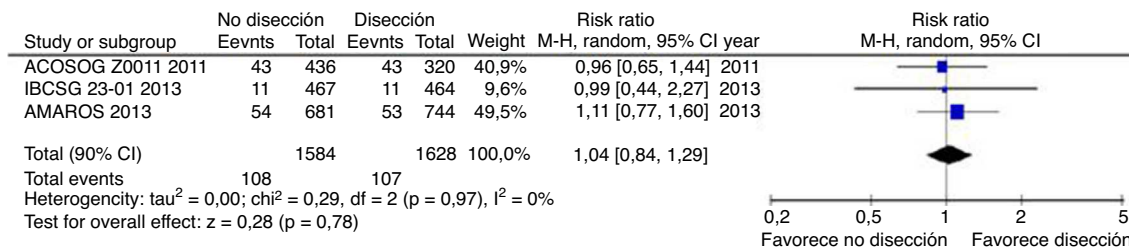
Aquí el punto crítico es valorar si el intervalo de confianza es suficientemente estrecho². En el ámbito del escenario, en el que el interés final es establecer una recomendación sobre si

se debe realizar disección ganglionar axilar, esto se traduce en comprobar si el límite superior del intervalo de confianza al 90% del riesgo relativo de mortalidad supera el umbral de decisión fijado en el estudio de Giuliano (un aumento relativo del riesgo del 30%). Como se muestra en las figuras 2 y 3, el límite superior del IC 90% es de 1,29, prácticamente en el límite de la no-inferioridad. Y si se excluye el estudio de Giuliano et al. (de mayor riesgo de sesgo) el límite superior del IC 90% se eleva hasta 1,45. Este intervalo de confianza cuestiona nuestra convicción de que el tratamiento conservador es no-inferior al vaciamiento ganglionar, por lo que los autores decidimos reducir el grado de evidencia en un escalón (desde evidencia alta a moderada).

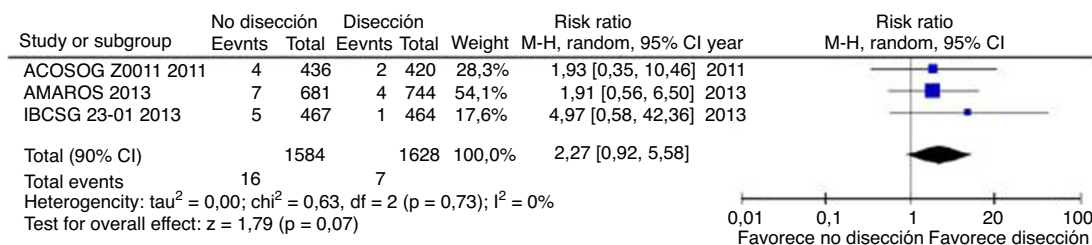
Respecto a la recidiva axilar, en cambio, el límite superior del IC 95% es un aumento absoluto del riesgo de solo un 1%. Si consideramos que este nivel es aceptable para la mayoría de las pacientes (dado el mayor riesgo de linfedema asociado al vaciamiento ganglionar), entonces no habría razón para degradar el nivel de evidencia para este desenlace por imprecisión¹⁰.

Y respecto al riesgo de edema linfático, todos ellos presentan un aumento importante del riesgo en el grupo tratado con vaciamiento axilar. Por lo tanto, se debería degradar la

A Mortalidad a los 5 años



B Recidiva axilar a los 5 años



C Edema linfático

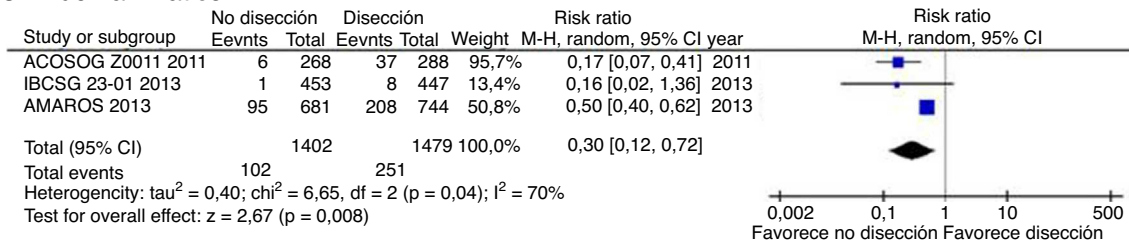


Figura 2 Diagramas de Forest.

evidencia respecto a este *outcome* por inconsistencia, pero no por imprecisión.

Volveremos sobre este tema al hablar de las recomendaciones.

Evidencia indirecta (indirectness)

Cuando la población en estudio, la intervención o los desenlaces evaluados son distintos de los de la población de interés, la confianza en la estimación del efecto disminuye¹¹. Sin embargo, hay que tener en cuenta que, habitualmente, las condiciones en que se realizan los ensayos difieren en importantes aspectos de las condiciones locales del paciente individual. Por tanto, la aplicación de la evidencia científica al paciente individual exige, además de competencias de lectura crítica, un adecuado conocimiento de las circunstancias locales y buen juicio clínico.

Por ejemplo, algunos autores¹² han criticado que la población finalmente reclutada en el estudio de Giuliano et al.⁴ es una población de bajo riesgo, no representativa del espectro de pacientes con criterios de inclusión. Asimismo, el estudio de Galimberti⁶ incluye una muy baja proporción de pacientes con macrometástasis. Y el estudio AMAROS^{7,8} evalúa 2 estrategias (radioterapia versus vaciamiento axilar) que no coinciden exactamente con las del escenario (en

el que las 2 alternativas incluyen la radioterapia). Por lo tanto, se podría considerar degradar un escalón el grado de evidencia debido a datos indirectos. En nuestro caso, esta posibilidad se desestimó al considerarse que existe un importante cuerpo de evidencia indirecta que sugiere que la presencia de metástasis en la biopsia del ganglio centinela no es más que un factor pronóstico¹³, y que el tratamiento coadyuvante actual garantiza una baja tasa de recaída axilar

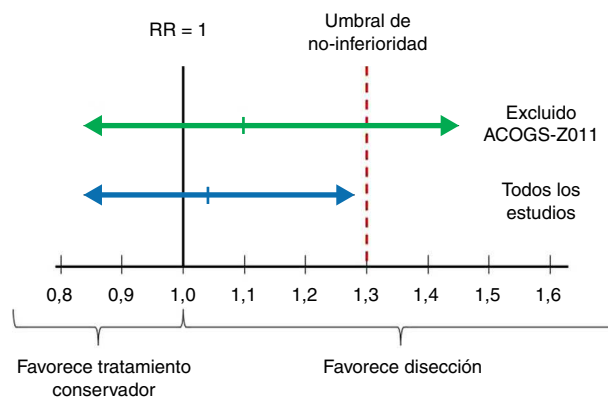


Figura 3 Valoración de la imprecisión.

Tabla 2 Perfil de evidencia para los principales desenlaces clínicos

N.º de estudios	Valoración de la calidad						Efecto relativo (IC 95%)	Calidad
	Diseño	Riesgo de sesgo	Inconsistencia	Datos indirectos	Imprecisión	Otras		
Supervivencia (mediana de seguimiento 5-6,3 años)								
	Ensayos aleatorizados	No serio riesgo de sesgo ^a	No seria inconsistencia	No ^b	Seria ^c	Ninguna	RR 1,09 (0,78 a 1,53)	Moderada
Recidiva regional								
	Ensayos aleatorizados	No serio riesgo de sesgo ^a	No seria inconsistencia	No ^b	No seria imprecisión	Ninguna	RR 2,27 (0,92 a 5,58) ^d	Alta
Edema linfático								
	Ensayos aleatorizados	No serio riesgo de sesgo ^a	Seria ^e	No ^b	No seria imprecisión	Ninguna	RR 0,30 (0,12 a 0,72)	Moderada

^a Estudio ACOSOG-Z0011 sensible a las pérdidas en el seguimiento. Grupos desbalanceados a favor del grupo tratado con biopsia sola.
^b Bajo número de metástasis (Galimberti). Perfil pronóstico favorable (Giuliano).
^c Tras excluir el estudio con mayor riesgo de sesgo, LS IC 90% del riesgo relativo de mortalidad = 1,45.
^d Diferencia absoluta de riesgos: 0,01 (IC 95%: 0-0,01).
^e I² = 70% (p = 0,04).

en pacientes con afectación ganglionar tratados de forma conservadora¹⁴.

Sesgo de publicación

Es bien conocido que los estudios con resultados positivos tienen más probabilidad de ser publicados que los estudios con resultados menos favorables. Ello es especialmente frecuente en estudios pequeños, lo que se puede detectar mediante diversas herramientas gráficas y estadísticas. Como consecuencia se genera un sesgo de selección que favorece la sobreestimación del efecto. En el caso de metaanálisis de Rao et al.³ se hizo una búsqueda exhaustiva, y dado el escaso número de estudios las herramientas de diagnóstico no son de utilidad, por tanto se decidió no degradar la calidad de evidencia por este concepto.

Perfil de evidencia y calidad global de la evidencia (tabla 2)

De forma general GRADE propone como grado de evidencia global el asignado al desenlace crítico de peor calidad (en nuestro caso, la mortalidad a los 5 años). Para la pregunta que nos formulábamos en el escenario, la calidad de la evidencia sería por tanto moderada. Ello significa que tenemos moderada confianza en que el tratamiento conservador no es inferior a la disección ganglionar en pacientes con cáncer de mama no avanzado con metástasis axilares con intención de tumorectomía y radioterapia..

Grado de recomendación

El siguiente paso consiste en determinar el grado de confianza en que los efectos beneficiosos del tratamiento conservador (la reducción de la morbilidad quirúrgica) superan a los efectos adversos (aumento de las recidivas axilares, posible reducción de la supervivencia) comparación con el vaciamiento ganglionar. Ello depende de la calidad de la evidencia de los efectos importantes, la valoración de los riesgos con respecto a los beneficios, las preferencias de los pacientes y, en ocasiones, los costes¹⁵.

Cuando el grado de evidencia no es alto, las preferencias del paciente desempeñan un papel importante. En efecto, Galper et al.¹⁶ entrevistaron cuidadosamente a un grupo de mujeres con cáncer de mama encontrando una gran variabilidad en las preferencias. Mientras que el grupo de mujeres con cáncer invasivo exigía una mejoría mínima de la supervivencia del 3% para someterse a una disección ganglionar axilar, en pacientes con carcinoma ductal *in situ* este porcentaje subía al 10%. Por lo tanto, la recomendación adecuada para una enferma con un umbral del 5% sería la de hacer tratamiento conservador, mientras que para una enferma con mayor aversión al riesgo de muerte o menor rechazo del edema linfático, la recomendación adecuada sería la de hacer vaciamiento ganglionar.

Cuando se puede asumir que la práctica totalidad de los pacientes bien informados elegirían una de las opciones, la recomendación sería fuerte. Si, como ocurre en nuestro caso, una parte sustancial de pacientes discrepa respecto a la decisión a elegir, entonces la recomendación sería débil. Ello obligaría al clínico a dedicar un mayor tiempo a la decisión compartida, para asegurarse que la recomendación final es coherente con los valores y preferencias de la enferma.

Con matices, esta recomendación es consistente con la de las guías más recientes. Así, las guías SIGN de 2013¹⁷ afirman que «(las) pacientes tratadas con cirugía conservadora de mama y radioterapia para T1 o T2, ganglios axilares clínicamente negativos y con uno o 2 ganglios positivos en la biopsia del ganglio centinela se pueden considerar para no tratamiento axilar ulterior». Y de acuerdo con las guías ASCO de 2014¹⁷ «(los) clínicos no deberían recomendar la disección ganglionar axilar en mujeres con cáncer de mama en estadio precoz y solo una o 2 metástasis del ganglio centinela programadas para cirugía conservadora de la mama con radioterapia mamaria fraccionada convencional».

Conclusiones

Aunque con limitaciones, el cuerpo de evidencia apoya la hipótesis del estudio ACOSOG Z0011, y sugiere que el beneficio de evitar la iatrogenia (linfedema) justifica la omisión del vaciamiento axilar en pacientes seleccionados con cáncer de mama no avanzado y metástasis en la biopsia del ganglio centinela.

La aproximación GRADE puede facilitar la valoración de la calidad del cuerpo de la evidencia científica, y establecer recomendaciones de acuerdo con dicha evidencia y las circunstancias locales.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que en este artículo no aparecen datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

- Latour-Perez J, Rodríguez-Lescure A. Lectura crítica de la literatura biomédica (i): cómo valorar el riesgo de sesgo de un estudio. *Rev Senol Patol Mamar*. 2015;28:34–8.
- Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence-impresión. *J Clin Epidemiol*. 2011;64:1283–93.
- Rao R, Euhus D, Mayo HG, Balch C. Axillary node interventions in breast cancer: A systematic review. *JAMA*. 2013;310:1385–94.
- Giuliano AE, Hunt KK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz PW, et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: A randomized clinical trial. *JAMA*. 2011;305:569–75.
- Lucci A, McCall LM, Beitsch PD, Whitworth PW, Reintgen DS, Blumencranz PW, et al. Surgical complications associated with sentinel lymph node dissection (SLND) plus axillary lymph node dissection compared with SLND alone in the American College of Surgeons Oncology Group Trial Z0011. *J Clin Oncol*. 2007;25:3657–63.
- Galimberti V, Cole BF, Zurrada S, Viale G, Luini A, Veronesi P, et al. Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (IBCSG 23-01): A phase 3 randomised controlled trial. *Lancet Oncol*. 2013;14:297–305.
- Rutgers EJ, Donker M, Straver ME, Meijnen P, Van De Velde CJH, Mansel RE, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer patients. En: *Final analysis of the EORTC AMAROS trial (10981/22023)*. 2013, editors.
- Bean J. EORTC study shows radiotherapy and surgery provide excellent regional control for breast cancer patients. 2013 [consultado 26 Feb 2014]. Disponible en: <http://www.eortc.org/news/eortc-study-shows-radiotherapy-and-surgery-provide-excellent-regional-control-breast-cancer-pat>
- Goodman A. Axillary radiotherapy: New standard of care in node-positive breast cancer? *The ASCO Post*. 2013;4(10).
- Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7 rating the quality of evidence-inconsistency. *J Clin Epidemiol*. 2011;64:1294–302.
- Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8 Rating the quality of evidence-indirectness. *J Clin Epidemiol*. 2011;64:1303–10.
- Gatzemeier W, Mann GB. Which sentinel lymph-node (SLN) positive breast cancer patient needs an axillary lymph-node dissection (ALND)-ACOSOG Z0011 results and beyond. *Breast*. 2013;22:211–6.
- Fisher B, Jeong JH, Anderson S, Bryant J, Fisher ER, Wolmark N. Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. *N Engl J Med*. 2002;347:567–75.
- Giuliano AE, McCall L, Beitsch P, Whitworth PW, Blumencranz P, Leitch AM, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. *Ann Surg*. 2010;252:426–32.
- Galper SR, Lee SJ, Tao ML, Troyan S, Kaelin CM, Harris JR, et al. Patient preferences for axillary dissection in the management of early-stage breast cancer. *J Natl Cancer Inst*. 2000;92:1681–7.
- Scottish Intercollegiate Guidelines Network (SIGN). Treatment of primary breast cancer. Edinburgh: SIGN; 2013. (SIGN publication no. 134). [September 2013]. Disponible en: <http://www.sign.ac.uk>
- Lyman GH, Temin S, Edge SB, Newman LA, Turner RR, Weaver DL, et al. Sentinel lymph node biopsy for patients with early-stage breast cancer: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol*. 2014;32:1365–83.