CAGS AND ACS EVIDENCE BASED REVIEWS IN SURGERY. 40

Axillary dissection versus no axillary dissection in women with invasive breast cancer and sentinel node metastasis

Steve Latosinsky, MD
Tanya S. Berrang, MD
C. Suzanne Cutter, MD
Ralph George, MD
Ivo Olivotto, MD
Thomas B. Julian, MD
Allen Hayashi, MD
Christopher Baliski, MD
Randall L. Croshaw, MD
Kathleen M. Erb, MD
Jennifer Chen, MD
for the Members of the Evidence Based Reviews in Surgery Group*

*The CAGS/ACS Evidence Based Reviews in Surgery Group comprises N.N. Baxter, K.J. Brasel, C.J. Brown, P. Chaudhury, C.S. Cutter, C.M. Divino, E. Dixon, L. Dubois, G.W.N. Fitzgerald, H.J.A. Henteleff, A.W. Kirkpatrick, S. Latosinsky, A.R. MacLean, T.M. Mastracci, R.S. McLeod, A.M. Morris, L.A. Neumayer, L.K. Temple and M.E. McKenzie.

Correspondence to:
Ms. Marg McKenzie, RN
Administrative Coordinator, EBRS
Mount Sinai Hospital, L3-010
60 Murray St., PO Box 23
Toronto ON M5T 3L9
fax 416 586-5932
mmckenzie@mtsinai.on.ca

DOI: 10.1503/cjs.036011

The term "evidence-based medicine" was first coined by Sackett and colleagues as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients."[1] The key to practising evidence-based medicine is applying the best current knowledge to decisions in individual patients. Medical knowledge is continually and rapidly expanding. For clinicians to practise evidence-based medicine, they must have the skills to read and interpret the medical literature so that they can determine the validity, reliability, credibility and utility of individual articles. These skills are known as critical appraisal skills, and they require some knowledge of biostatistics, clinical epidemiology, decision analysis and economics, and clinical knowledge.

Evidence Based Reviews in Surgery (EBRS) is a program jointly sponsored by the Canadian Association of General Surgeons (CAGS) and the American College of Surgeons (ACS) and is supported by an educational grant from ETHICON and ETHICON ENDO-SURGERY, both units of Johnson & Johnson Medical Products, a division of Johnson & Johnson and ETHICON Inc. and ETHICON ENDO-SURGERY Inc., divisions of Johnson & Johnson Inc. The primary objective of EBRS is to help practising surgeons improve their critical appraisal skills. During the academic year, 8 clinical articles are chosen for review and discussion. They are selected for their clinical relevance to general surgeons and because they cover a spectrum of issues important to surgeons, including causation or risk factors for disease, natural history or prognosis of disease, how to quantify disease, diagnostic tests, early diagnosis and the effectiveness of treatment. A methodological article guides the reader in critical appraisal of the clinical article. Methodological and clinical reviews of the article are performed by experts in the relevant areas and posted on the EBRS website, where they are archived indefinitely. In addition, a listserv allows participants to discuss the monthly article. Surgeons who participate in the monthly packages can obtain Royal College of Physicians and Surgeons of Canada Maintenance of Certification credits and/or continuing medical education credits for the current article only by reading the monthly articles, participating in the listserv discussion, reading the methodological and clinical reviews and completing the monthly online evaluation and multiple choice questions.

We hope readers will find EBRS useful in improving their critical appraisal skills and in keeping abreast of new developments in general surgery. Four reviews are published in condensed versions in the *Canadian Journal of Surgery* and 4 are published in the *Journal of the American College of Surgeons*. For further information about EBRS, please refer to the CAGS or ACS websites. Questions and comments can be directed to the program administrator, Marg McKenzie, at mmckenzie@mtsinai.on.ca.

Reference

1. Evidence-Based Medicine Working Group. Evidence-based medicine. *JAMA* 1992;268:2420-5.

Selected Article

Giuliano AE, Hunt KK, Ballman KV, et al. Axillary dissection vs. no axillary dissection in women with invasive breast cancer and sentinel node metastasis. A randomized controlled trial. *JAMA* 2011;305:569-75.

Abstract

Question: Does a complete axillary lymph node dissection (ALND) affect the overall survival of patients with sentinel lymph node (SLN) metastasis of breast cancer? **Design:** Randomized controlled trial. **Setting:** Multicentre trial that included 115 sites. **Patients:** There were 856 women with clinical T1–T2 invasive breast cancer, with no palpable adenopathy and 1–2 SLNs containing metastases identified histologically. **Intervention:** All patients underwent lumpectomy and tangential whole-breast irradiation. Those with SLN metastases indentified by sentinel lymph node dissection (SLND) were randomly assigned to undergo ALND or no further axillary treatment. Those assigned to ALND underwent dissection of at least 10 nodes. **Main outcome measures:** Overall survival, defined as the time from random assignment until death from any cause. The secondary outcome was disease-free survival. **Results:** Clinical and tumour characteristics were similar among 420 patients assigned to ALND and 436 assigned to SLND alone. The median number of nodes removed was 17 with ALND and 2 with SLND. At a median follow-up of 6.3 years (last follow-up, Mar. 4, 2010), 5-year overall survival was 91.8% (95% confidence interval [CI] 89.1%–94.5%) with ALND and 92.5% (95% CI 90.0%–95.1%) with SLND alone; 5-year disease-free survival was 82.2% (95% CI 78.3%–86.3%) with ALND and 83.9% (95% CI 80.2%–87.9%) with SLND alone. The hazard ratio for treatment-related overall survival was 0.79 (90% CI 0.56–1.11) without adjustment and 0.87 (90% CI 0.62–1.23) after adjusting for age and adjuvant therapy. **Conclusion:** Among patients with limited SLN metastatic breast cancer treated with breast conservation and systemic therapy, the use of SLNB alone compared with ALND did not result in inferior survival.

Commentary

Axillary lymph node dissection (ALND) has been a part of the surgical management of breast cancer since Halsted's popularization of radical mastectomy. At the same time, the morbidity of ALND (lymphedema in particular) has been one of the most feared consequences of breast cancer care.

Surgeons have been acutely aware of the long-term effects of ALND and for a number of years have commented on its declining impact on decision-making regarding systemic treatment for many patients[1] and troubling lack of survival benefit, as suggested in long-term follow-up of trials like the National Surgical Adjuvant Breast and Bowel Project's (NSABP) B-04 protocol (comparison of radical mastectomy with alternative treatments).[2] The adoption of sentinel lymph node dissection (SLND) for patients with clinically negative axillae after studies such as ALMANAC (axillary lymphatic mapping against nodal axillary clearance)[3] and NSABP B-32 (comparison of sentinel node resection to conventional axillary dissection in clinically node-negative breast cancer patients)[4] led to the elimination of completion dissections for patients with negative sentinel node biopsies. The American College of Surgeons Oncology Group (ACOSOG) Z0011[5] study sought to determine whether completion ALND for breast cancer patients with sentinel node metastases was necessary for patients undergoing breast conserving surgery with radiation. The results of this study have attracted much attention and provide hope to many women who wish to avoid the potential morbidity of an ALND.

The Z0011 trial was developed as a noninferiority trial. In superiority trials, the size and precision of the treatment effect between a new and reference treatment are compared. Noninferiority designs, as in Z0011, are less common and seek to determine whether a new treatment is no worse than a reference treatment by not more than a prespecified amount, the noninferiority margin. These trial designs are considered appropriate under the premise that the new treatment has some other advantage, such as greater availability, reduced cost, less invasiveness, greater ease of administration or, as in the study we reviewed, fewer side effects/harms, such that proof of superiority is not necessary.[6] The new treatment would be recommended if it were similar to or better than the existing one or if it were not worse by more than this prestated margin. Once the study is determined to be statistically significant for noninferiority, as in Z0011, the important question to ask is whether the margin of noninferiority was acceptable.

The breast cancer literature gives guidance to what might be a minimally acceptable margin. In a somewhat biased sample, Duric and colleagues[7] found that most women who had completed adjuvant chemotherapy felt that an additional 3% in survival was sufficient to make chemotherapy worthwhile. Reflecting this high acceptance of potentially toxic treatments for small survival gains, the NSABP B-32 sentinel node study was designed such that SLND would be considered to be equivalent to SLND followed by ALND in node-negative patients only if a difference in survival of 2% or less was detected.[4]

The margin of noninferiority in the trial by Giuliano and colleagues is confusing. According to the paper, the margin was based on the SLND group having a 5-year survival of "not less than 75% of that observed in the ALND group." The expected 5-year survival of the ALND group was estimated to be 80%. Thus, based on this statement, an absolute 5-year survival as low as 60% in the SLND group,

a 20% absolute difference from the ALND group would be sufficient to conclude that SLND is not inferior to ALND. This margin of noninferiority would be unacceptable to most clinicians and patients. The paper also states that "Noninferiority of the SLND alone treatment was also considered if the hazard ratio for mortality was less than 1.3 when compared with ALND." A hazard is the rate at which events happen and a hazard ratio is the proportion of the hazard in one group to the hazard in the other group. This would suggest a more reasonable, but still not very conservative, noninferiority margin of 5% if survival in the ALND group was 80%. The final sample size target of 1900 patients would suggest that in fact the investigators were looking for an absolute margin susbtantially smaller than 20% between groups.

Unfortunately, the trial was closed early with less than 50% of the targeted accrual. There were 2 issues that led to the early closure. The first was that accrual was slow, likely related to physicians and patients being biased in favour of standard ALND.[8] The second was that there was a substantially higher than expected survival in the pooled data from the 2 groups. Even with all 1900 targeted patients accrued, it was estimated that it would take 20 years of follow-up to observe the estimated 500 deaths needed to declare noninferiority. Instead the results of the trial were reported with only 92 events.

The final concern related to sample size was that the authors used a 2-sided 90% CI, which corresponds to a 1-sided significance level of $p < 0.05$. This is controversial and not recommended in inferiority trials.[6] The justification for a 1-sided test is that superiority of the new treatment would be a bonus and thus does not have to be shown. The advantage to the researcher is that it requires a smaller sample size to show statistical significance.

In the final analysis, the unadjusted hazard ratio comparing overall survival between the SLND alone group and the ALND group was 0.79 (90% CI 0.56–1.10), which did not cross the specified boundary of 1.3. Even the adjusted hazard ratio did not cross the boundary. It is reassuring that the CI is well away from the noninferiority margin, but these statistics may be difficult for clinicians and patients to interpret. It is reassuring that the survival curves were overlapping and not significantly different on the log rank test. However, failure to show a difference does not mean that a difference does not exist.

Another concern with this study is in data collection. This was one of the first studies conducted by the ACOSOG group, which might account for problems in data collection. Surprisingly, there were many patients entered into the trial who had no positive nodes; 29 (7.0%) patients in the SLND group and 4 (1.2%) patients in ALND group, as would be expected. No explanation is given for the high proportions of patients who failed to meet the inclusion criteria. This is possibly because of the intraoperative random assignment of patients. However, a false-positive result on touch preparation or frozen section should be a rare event.

There were also a large number of women lost to follow-up: 21% in the ALND group and 17% in the SLND group. In comparison, in the NSABP B-32 sentinel node trial, loss to follow-up was less than 1% for a similar study and time frame.[4] Trialists aim to have follow-up on all patients, but when the number lost to follow-up exceeds 10%, the validity of the conclusions may be jeopardized. Follow-up was not the only information that was poorly collected. There are a lot of missing data on demographic and clinical characteristics. Especially concerning is the large number of patients for whom there are no data on the number of positive nodes: 18% in the ALND group.

The radiation techniques in this trial are equally important. The dosing, frequency and field definition guidelines were not described. The protocol stipulated that all women should receive whole-breast opposing tangential-field radiation therapy. Patients undergoing partial breast irradiation were not eligible. No third-field nodal irradiation was allowed for the supraclavicular nodes. There are concerns that unconventional "high tangents" to the axillary area may have been used in patients assigned to the SLND group, which could influence the study findings.

Another interesting issue with noninferiority trials concerns intention-to-treat analyses. There were 43 (5.0%) patients who did not receive their assigned treatment. Of the 420 patients assigned to the ALND group, 32 (7.6%) did not undergo ALND, and of the patients who were assigned to the SLND alone group, 11 (2.5%) had ALND. In superiority trials, intention-to-treat analysis is recommended, as inclusion of these patients tends to decrease differences between groups and is a more conservative analysis. In noninferiority trials, the null and alternative hypotheses are reversed; a type-I error is the erroneous acceptance of an inferior new treatment, whereas a type-II error is the erroneous rejection of a truly noninferior treatment. Thus in noninferiority trials, an intention-to-treat analysis is usually a less conservative approach, as it will often increase the likelihood of falsely concluding noninferiority.[6] As more women in the ALND group did not undergo their assigned treatment, this is likely not the case in this study. There is greater confidence in results when the conclusions from both the intention-to-treat and treatment-received analyses are consistent, as was the case in this study.

It is important to note that the recently presented MA-20,[9] which in contrast to Z0011 showed that comprehensive regional treatment with radiotherapy for patients following segmental mastectomy with 1–3 involved nodes is associated with a small improvement in survival, will likely greatly influence clinical practice and the interpretation and implementation of Z0011. If an ALND is not performed in women who have 1 or more positive sentinel nodes, then radiation oncologists will likely include the

axilla with regional nodal irradiation. Results from the AMAROS (after mapping of the axilla: radiotherapy or surgery) trial,[10] which randomly assigned women with positive sentinel nodes to ALND or axillary radiotherapy, are pending and will also be very relevant.

The authors concluded that "among patients with limited SLN metastatic breast cancer treated with breast conservation and systemic therapy the use of SLND alone compared with ALND did not result in inferior survival." They indicate that ALND as standard practice may no longer be justified in this patient population. The "may" is an important descriptor as, despite enthusiasm for the study results, problems with the methodology, including what noninferiority margin the study can exclude in clinically understandable terms, trial accrual, the inclusion of a number of node-negative patients and loss of follow-up of a large number of patients, lead to difficulties with analysis, interpretation and confidence in the results.

Should the results of Z0011 change practice? Owing to its methodologic limitations, if we had to depend on Z0011 alone the standard of care following a positive sentinel node "may" still be an ALND. However, in light of the new findings of a survival benefit from regional lymph nodal radiotherapy from the MA-20[8] study among patients with positive nodes who meet the criteria of Z0011 (maximum of 2 positive nodes, treated with breast conserving surgery and planning to receive radiotherapy to the breast), the surgeon is advised to seek an early opinion or a joint management discussion with radiation oncology and/or their multidisciplinary breast cancer tumour board to finalize an axillary management strategy. If after breast conserving surgery patients have more extensive axillary disease than would have been eligible for Z0011, ALND should be considered standard of care.

Competing interests: None declared.

References

1. Lin PP, Allison D, Wainstock J, et al. Impact of axillary lymph node dissection on therapy of breast cancer patients. *J Clin Oncol* 1993;11: 1536-44.
2. Fisher B, Jeong JH, Anderson S, et al. 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.
3. Mansel RE, Fallowfield L, Kissin M, et al. Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial. *J Natl Cancer Inst* 2006;98:599-609.
4. Krag DN, Anderson SJ, Julian TB, et al. Sentinel-lymph node resection compared with conventional axillary-lymph node dissection in clinically node-negative patients with breast cancer: overall survival findings from NSABP B-32 randomised phase 3 trial. *Lancet Oncol* 2010;11:927-33.
5. Giuliano AE, Hunt KK, Ballman KV, 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.
6. Piaggio G, Elbourne DR, Altman DG, et al.; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement [published erratum in *JAMA* 2006;296:1842]. *JAMA* 2006;295:1152-60.
7. Duric VM, Stockler MR, Heritier S, et al. Patient's preferences for adjuvant chemotherapy in early breast cancer: What makes AC and CMF worthwhile now? *Ann Oncol* 2005;16:1786-94.
8. Leitch M, McCall L, Beitsch P, et al. Factors influencing accrual to ACOSOG 10011, a randomized phase III trial of axillary dissection vs. observation for sentinel node positive breast cancer. *J Clin Oncol* 2006;24(Suppl):601.
9. Whelan TJ, Olivotto IA, Ackerman I, et al. NCIC-CTG MA20: An intergroup trial of regional nodal irradiation in early breast cancer. *J Clin Oncol* 2011;29(Suppl):abstr LBA1003.
10. Straver ME, Meijnen P, van Tienhoven G, et al. Sentinel node identification rate and nodal involvement in the EORTC 10981-22023 AMAROS trial. *Ann Surg Oncol* 2010;17:1854-61.