Lack of uniform diagnostic criteria for inflammatory breast cancer limits interpretation of treatment outcomes: a systematic review

Kim T, Lau J, Erban J

CRD summary
The authors concluded that, although pathologic complete response rates appear to improve with increasing intensity of neoadjuvant treatment in patients with inflammatory breast cancer, these findings are not based on high-quality data and are only tentative. The review had a number of methodological flaws but the authors' conclusion, that further research is required, is appropriate.

Authors' objectives
To describe patient characteristics, diagnosis definitions, treatment regimens and intensity of neoadjuvant chemotherapy in relation to clinical response and survival rates, and to identify potential explanations for any differences, in cohort studies of people with inflammatory breast cancer (IBC).

Searching
MEDLINE and PREMEDLINE were searched from inception to November 2005; the search terms were listed. The Cochrane Library, Best Evidence, DARE and Dissertation Abstracts were also searched but dates were not given. The reference lists of identified articles were screened. A further literature search was carried out to identify relevant studies found in abstracts and letters to journal editors. Only English language or translated articles were eligible for inclusion.

Study selection
Study designs of evaluations included in the review
The authors stated that there were no published randomised controlled trials or systematic reviews of IBC-only patients, therefore only cohort studies were included. Case-control studies and review articles were excluded.

Specific interventions included in the review
Studies where neoadjuvant systemic chemotherapy was used for some or all of the patients, and where established treatment protocols with chemotherapy and doses were clearly described, were eligible for inclusion. The included studies used differing treatment regimens (incorporating surgery, radiotherapy, both, or neither) and different chemotherapy drugs.

Participants included in the review
Studies that included only people with IBC were eligible for inclusion. Studies that included any patients who did not have IBC were excluded, as were studies of people with additional diagnoses or patients with stage IV disease at presentation and studies including people with IBC who had been previously treated. The included participants were aged from 24 to 82 years.

Outcomes assessed in the review
Pathologic and clinical response rates, 3- and 5-year disease-free survival (DFS) and overall survival (OS) status, and treatment-associated mortality and morbidity were assessed in the review.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
The authors did not state that they assessed validity.
Data extraction
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

The outcome rates were calculated using the total number of evaluable patients or specimens in each study. Three- and 5-year DFS and OS were reportedly based on the entire treatment regimen. The pathologic complete response rates (pCRs) were determined from studies based on reports of mastectomy or biopsy specimens and axillary node assessment after neoadjuvant treatment. Overall response rates were defined as the sum of complete clinical response rates plus partial responses, determined before local surgery. Partial response was defined as an objective measured response of more than 49% but less than 100%, as reported in each study.

Methods of synthesis
How were the studies combined?
The studies were combined for each of four neoadjuvant induction regimens: no anthracycline induction, low-dose anthracycline-based induction, moderate-dose anthracycline induction, or high-dose chemotherapy (HDC). If HDC occurred after breast surgery, studies were grouped by pre-operative anthracycline dose for pCR outcome and by HDC for survival outcomes. Response rates in individual studies were expressed as proportions and logit transformed to maintain normal distribution of the data during pooling. The logits were then combined using a random-effects model.

How were differences between studies investigated?
The authors stated that the random-effects model incorporated within- and between-study heterogeneity. The sources of clinical heterogeneity (treatment regimens and diagnosis) were discussed in the text.

Results of the review
Twenty-seven cohort studies (n=1,232) were included in the review.

The authors noted variability between the studies in terms of reporting of signs and symptoms, definition of IBC, treatment regimens, and method of determining and reporting pCR. If an assessment of statistical heterogeneity was carried out as part of the meta-analysis, its results were not reported.

No anthracycline induction: the pCR was 4% (95% confidence interval, CI: 1, 18; 2 studies, n=42); 3- and 5-year DFS were 38% (95% CI: 15, 69; 3 studies, n=57) and 44% (95% CI: 29, 60; 2 studies, n=39), respectively; 3- and 5-year OS were 47% (95% CI: 9, 88; 2 studies, n=43) and 60% (95% CI: 39, 79; 1 study, n=25), respectively.

Low-dose anthracycline-based induction: the pCR was 11% (95% CI: 7, 17; 8 studies, n=317); 3- and 5-year DFS were 45% (95% CI: 40, 50; 14 studies, n=698) and 37% (95% CI: 31, 43; 11 studies, n=622), respectively; 3- and 5-year OS were 61% (95% CI: 52, 68; 12 studies, n=662) and 47% (95% CI: 38, 56; 11 studies, n=630), respectively.

Moderate-dose anthracycline-based induction: the pCR was 14% (95% CI: 8, 22; 3 studies, n=97); 3- and 5-year DFS were 45% (95% CI: 34, 56; 2 studies, n=76) and 38% (95% CI: 30, 50; 2 studies, n=76), respectively; 3- and 5-year OS were 69% (95% CI: 59, 78; 3 studies, n=91) and 52% (95% CI: 36, 68; 3 studies, n=91), respectively.

HDC before surgery: the pCR was 32% (95% CI: 24, 41; 3 studies, n=123); 3- and 5-year DFS were 58% (95% CI: 39, 75; 3 studies, n=133) and 38% (95% CI: 18, 62; 1 study, n=21), respectively; 3- and 5-year OS were 71% (95% CI: 62, 78; 3 studies, n=133) and 62% (95% CI: 38, 82; 1 study, n=21), respectively.

HDC after surgery: 3- and 5-year DFS were 62% (95% CI: 53, 70; 4 studies, n=116) and 56% (95% CI: 42, 69; 3 studies, n=86), respectively; 3- and 5-year OS were 72% (95% CI: 59, 82; 4 studies, n=116) and 53% (95% CI: 31, 73; 3 studies, n=86), respectively.

HDC, combined: 3- and 5-year DFS were 59% (95% CI: 50, 68; 7 studies, n=249) and 52% (95% CI: 39, 65; 4 studies, n=107), respectively; 3- and 5-year OS were 70% (95% CI: 64, 76; 7 studies, n=249) and 55% (95% CI: 38, 71; 4 studies, n=107), respectively.
Authors' conclusions
There was notable variation in the diagnostic criteria and reporting of IBC, with great potential heterogeneity in participants. The pCRs seemed to be related to the intensity of neoadjuvant treatment, but this finding was not based on randomised data and should be used only as a hypothesis for future trials.

CRD commentary
The review's stated objective made the assumption that no randomised or controlled trials exist of patients with IBC, so only cohort studies could be included. It was not stated whether any controlled trials were found in the literature search, although we assume that this was the case. The inclusion criteria for the participants were clear and the outcomes to be extracted were well-defined. The literature search seemed reasonable in that more than two large electronic databases were searched in addition to other sources. However, only English language publications were eligible for inclusion in the review, which might have led to some studies being missed.

No details were reported of how the review was carried out (e.g. how many reviewers performed the study selection and data extraction processes). If only one reviewer was involved this may have made it more likely that subjective decisions would be made at these stages, possibly introducing bias. No validity assessment was carried out, which is a major flaw as cohort studies can vary considerably in their susceptibility to bias and, without a validity assessment, the reader cannot judge how much credence to give to the findings of the included studies. The review authors acknowledged that none of the included studies had a comparison group so, even though pooling of the studies was carried out, the combined results for each of the neoadjuvant regimens cannot be compared with one another as they have come from different studies. This note of caution is also sounded by the review authors. A random-effects model was used to pool the studies, to take account of acknowledged clinical heterogeneity between the studies, but statistical heterogeneity does not seem to have been assessed.

The authors’ conclusions were appropriately cautious given the evidence presented.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors stated that future studies should try to investigate and report the following variables: duration of symptoms before diagnosis; degree of breast involvement at diagnosis; complete pathologic analysis of breast and nodal tissue to generate an overall pCR after systemic therapy; complete reporting of oestrogen receptor, PgR and HER2/neu status; method of and criteria for IBC diagnosis.

Bibliographic details

PubMedID
17239263

DOI
10.3816/CBC.2006.n.055

Indexing Status
Subject indexing assigned by NLM

MeSH
Anthracyclines /administration & dosage /therapeutic use; Antineoplastic Agents /administration & dosage /therapeutic use; Breast Neoplasms /diagnosis /mortality /therapy; Female; Humans; Neoadjuvant Therapy; Treatment Outcome

AccessionNumber
12007000701

Date bibliographic record published
31/01/2008

Date abstract record published
31/01/2008

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.