

# Results of a breast cancer surgery trial compared with observational data from routine practice.

**Marubini, E, Mariana L, Salvadori B, Veronesi U, Saccozzi R, Merson M, Zucali R, Rilke F.**

*The Lancet 1996, 347:1000-3.*

## **AIM:**

The aim of this trial was not clearly stated in the abstract but was defined at the end of the introduction as "to compare the characteristics and the course of disease in out-trial and MTI patients, so as to establish how well research findings translated into clinical practice."

In other words to determine whether the practice of quadrantectomy, axillary dissection and radiotherapy for early stage breast cancer was as effective when applied to the general clinic population as a special trial population. This paper is a follow-on from the MII randomised control trial, commenced by the same institute in 1973. Endpoints were not clearly stated.

This is an important question to answer both from an ethical and a logical standpoint as trial patients are often not a representative sample of the general population.

## **TRIAL SUMMARY:**

The study is a single institution study from the National Tumour Institute (NTI) of Milan. 1760 T1 and T2 N0M0 patients were studied of whom 352 were part of the MII study. The study is an audit of these patients who were treated with quadrantectomy, axillary dissection and radiotherapy (QUART) from 1980-1984. The MII population is used as a control group. Demographic, prognostic and clinical information has been collected prospectively and followup and events added as they occurred.

## **TRIAL DESIGN:**

### **Patient factors:**

The referral base of the patients is not defined. We cannot be sure that this population was not selected. Indications for treatment and exclusion and inclusion criteria are stated.

In the MII trial only tumours up to 2cm were included. From the methods section tumours greater than 2cm in the outtrial population were to be excluded.

Only those with unilateral breast cancer were included. It is practical that only patients with unilateral breast cancer be included. It would be difficult to know whether a second breast cancer was a metastasis or a second primary if the histology was similar.

Patients with other cancers were excluded. Some cancers such as superficial skin cancers, or in those with highly curable tumours treated years previously may be unlikely to have any effect on mortality and morbidity. It could be argued that providing histological proof of the tumour of recurrence was available then these patients could be included. In practice this is often not attainable or impractical to pursue. However, these patients (including those patients with genetically linked second cancers) are part of the out-trial population and as such it would thwarting the aim of the study "to determine how well research findings translate into clinical practice" by excluding these patients. The converse could be argued that how the biology of a second cancer may interact with another is not known, and to include these patients is to potentially confound the results. If the treatment of the prior tumour compromised the treatment of the breast cancer then it would be wise to exclude them from the study. On balance it would have been possible to include in the analysis those patients with previous malignancy in whom histological proof of the recurrence was sought if it occurred. I suspect these numbers would have been small and would have made little difference to the results.

Patients greater than 70 years of age were excluded. There is no justification for this. When testing the applicability of a treatment to the general population extreme age and intercurrent morbidity are common factors. It is not stated whether these patients were not treated with this regimen at all or whether their data was not included in the study. If they were given the treatment, then they should be included in the study. Statistical methods for survival and events should be censored and make no difference to the

survival data.

Exclusions of non-infiltrating disease and Paget's disease is reasonable.

Patient age distribution was similar in the two groups.

**Disease factors:**

Apart from selection for staging there is no stratification for disease factors. When the two groups were compared, the outtrial patients had larger tumours, more multifocal tumours, higher numbers of positive nodes and therefore higher numbers of patients receiving chemotherapy. Despite these differences, there is no statistical measure of the differences between the populations. Hormone receptor status was not reported and this is an omission as this is a prognostic indicator. The authors state that this information was collected and entered into the database.

**Treatment factors:**

For the MII patients chemotherapy was not given until 1976 to women with histologically positive nodes and so although a lower proportion of the MII patients had positive nodes, a lower proportion of those with positive nodes received chemotherapy (60% v 84%). Other adjuvant treatments were entered into the database, but not documented in the study. These should have been mentioned.

Surgery of the primary is well defined. The surgery to the axilla does not state the level of the axillary dissection, nor the usual or median yield of nodes. Radiotherapy is inadequately described. The field arrangement is not given, the prescription point, the beam energy, the use of wedges are all not stipulated. The boost dose is given only as 10 Gy, and the energy as orthovoltage. In addition, the field borders or tissues to be included in the radiation portals are not included. The simulator technique is not described. The doses to tolerance tissues - in this case lung is not given. Radiotherapy was started within one month of surgery, but the authors do not state in how many patients this was actually achieved, and the reasons for not achieving this.

Chemotherapy is described only as CMF (cyclophosphamide, methotrexate and 5-fluorouracil). The doses are not given, nor the dosing interval, and the mode of administration. The guidelines for treatment delays and treatment exclusions are also not given. There is no mention of monitoring, nor of assessment of treatment toxicities - an important point to measure when extrapolating the results of a highly selected group to the general population. There is no mention of compliance and treatment protocol violations.

**Followup:**

The patients were followed up quarterly for 5 years and then yearly. Chest x-rays were done every 6 months. This is a reasonable time frame for followup. The authors do not state by whom the followup was done and what parameters measured. The clinical and radiographic evidence required to document recurrence is not defined.

**Data Collection:**

The information gathered was stored on a database. Some of the parameters eg adjuvant treatments, receptor status are not mentioned in the results, and would be prudent to do so. The authors state that at 31/12/94 5% of patients were lost to followup and that "for all of them vital status was ascertained by a search for death certificates". This statement is ambiguous in that it could be taken to mean that all patients lost to followup were confirmed to be deceased, or that all patients lost to followup were searched in registers of deaths, and those patients not confirmed to be deceased were assumed to be alive. The second is not a valid way to ascertain survival. The authors should have made further efforts to contact patients not confirmed to be deceased. Information of death certificates is notoriously unreliable and a cause of death noted as breast cancer cannot reliably be taken to mean that the patient died of breast cancer. Further efforts to clarify these issues should have been sought. Only 5% of patients lost to followup is a reasonable number however.

**Statistical Design:**

The authors should calculate the power of their study. Cumulative incidence of death was calculated using the Kaplan-Meier method, adapted by Marubini and Valsecchi. Endpoints measured were IBTR (intra-breast tumour recurrence), distant metastasis, contralateral breast cancer, and death.

Nodal recurrence is either omitted, or included in either local or distant recurrence. The authors should

preferably state regional recurrence rates in the nodes, or indicate in which category they have been included. The date of operation was used as a starting date, which is reasonable.

#### **TRIAL CONDUCT:**

Despite their intentions only to include the tumours up to 2cm in the outtrial population, 10% of the patients had tumours greater than this size and were included in the analysis. A further 25 (7%) of the MII group and 77 (5%) of the outtrial population had no size of tumour documented.

There is no mention on the quality control measures in the data collection or data entry. There is no quality control of the extent of surgery or of the adequacy of the radiotherapy. There is no data presented on the mean doses received of radiotherapy, or of the average field sizes. There is no data presented of withdrawals from treatment, or of patients who received a deviation from the treatment protocols. There is no independent assessment of the histological diagnosis. Ascertainment of the patients lost to followup or deaths were inadequately checked as mentioned previously.

#### **RESULTS:**

It is not stated how many patients were excluded from the analysis. The authors only state that 5% were lost to followup, but their vital status was determined.

The authors present results to show that the rate of QUART compared to surgical interventions for breast cancer in the region increased with time. This is meaningless. "Breast cancer interventions" are not defined; they could include diagnostic biopsies, second operations for inadequately excised primaries, definitive surgery (as a 1 or 2 step procedure) and reconstruction. The change in the operation rates could merely reflect a change in the stage at presentation of the breast cancers, resulting in increased numbers of patients now eligible for lumpectomy who would not formerly be so. It would be better for the authors to present the definitive surgical rates for T1 and T2 patients only, or if this data were not available, to present the data on stage distribution in the population.

The authors state one of their aims is to compare the characteristics of outtrial patients and the MII trial patients. This is presented first in the methods section and reiterated in the results section. The table is quite clear and would suffice. The table should be in the results section. The authors have grouped tumour size into 5 categories. Four of these categories are divisions of T1 tumours and only one category (>2cm) is given for T2 tumours. Since subanalysis of the different sizes of tumours within stage is not a stated aim of this study this seems superfluous. The authors could merely state that there was a higher proportion of Stage 2 tumours in the outtrial population. The authors do not give any statistical analysis of the difference between the two groups, an omission as they later go on to draw conclusions based on the difference between them.

The authors give crude overall death rates for the populations, and crude relapse rates. Although this is not clear, this probably means unadjusted for prognostic variables. Mortality and incidence curves are presented, with 5 and 10 year survivals with their standard errors in Table 3. The authors comment that outtrial women had worse survival, IBTR, and distant metastasis than MII patients, which would appear so from the graphs, however no statistical support is given. They do present the adjusted analysis for covariate effect and this confirms an effect only in the IBTR results. The results are given as single hazard ratios (which is appropriate as the proportional hazards assumption has been confirmed) and 95% confidence intervals with a p value calculated from the likelihood ratio test which is appropriate.

The results do not address the first aim of the study i.e. do the characteristics of the disease in the two groups differ? The data is inadequately presented without statistical support so that the reader cannot make any conclusions. Because of this it is difficult to interpret the results of the second issue regarding outcome in the two groups.

#### **STATISTICAL ANALYSIS:**

The authors use the proportional hazards (Cox) regression to adjust for baseline differences in the two populations (1). This method can be used when known prognostic factors can be allowed for but will not adjust for the influence of unmeasured confounders. Only randomisation will adjust for these.

The authors state that the proportional hazards model was verified by log minus log survival plots prior to using the Cox regression models. This is appropriate and the method used to verify this is robust. The other assumption to be verified before using the Cox regression analysis is the multiplicative assumption. By the statement that "Possible interactions between the cohort and the prognostic factors were also investigated, but no evidence was obtained in this direction." one may infer that this was also tested and found to be true. The Cox regression takes prognostic factors into account, and the authors have stated that they used factors "such as" age, tumour size, number of nodes involved and histology, all of which are reasonable. It is not stated whether these are the only prognostic factors taken into account. Other prognostic factors which could have been considered include receptor status, lymphatic and vascular invasion, and degree of differentiation (2). If other prognostic factors had been considered and found not to add to the discrimination of the model, then this should have been stated.

Use of the likelihood ratio test is an appropriate measure of how good the regression equation fits the data (3).

#### **CONCLUSIONS:**

The authors conclude that the outtrial patients did worse in terms of IBTR than the trial population. They were similar in outcome for the other measures of disease status. The authors concluded that this is due to the differences between the MII and the outtrial population.

While I agree with the results, the inference is not correct. Firstly, the authors do not present an analysis that the outtrial population were significantly different from the trial population. In the second instance, the occurrence of the two results does not imply causation. The authors have neglected to consider other well known significant prognostic factors. The authors also state that the surgery performed in the outtrial population may have changed over the time period. A subanalysis on the results for surgeons involved in the trial and those not involved in the MII trial showed no significant difference in the incidence of local failure. The authors also point out that in the time period following the trial the extent of surgery has decreased from a full quadrantectomy to a segmentectomy. While the authors state that they are unable to verify this because of incomplete information regarding the adequacy of excision, it may have also been informative to present the information as survival rates for two populations: QUART v segmentectomy.

#### **PRESENTATION:**

The paper is clear and well written. Illustrations are appropriate, although the crude survival curves are superfluous, and it would have been better to present the survival curves adjusted for prognostic factors.

#### **ETHICS:**

This is an ethical study and one of interest to the oncology community. The treatment arms are ethical and the authors approach this topic from an unbiased viewpoint. This was the standard treatment in the community at the time, and the means of study (retrospective) and the lack of consent (as this was not experimental) is appropriate.

#### **RECOMMENDATIONS:**

I would recommend that this paper be published. I would suggest the following alterations:

1. Patient referral base should be defined.
2. The authors should define the treatment received by those patients who did not take part in the trial and who were over 70 years of age.
3. The authors should define how nodal recurrence is handled in the data.
4. Statistical test for difference between the characteristics of the two treatment populations should be given.
5. More information should be presented for QUART as a percentage of the treatments for breast cancer in this population. Alternatively, the stage distribution of the population should be presented.
6. Better description of the chemotherapy needs to be given.
7. Description of the radiotherapy needs to include the following: fields arrangement, prescription and prescription points, beam energy, wedges, boost prescription point, tissues compensated for, simulator technique, tolerance tissue doses or volumes.
8. Actual doses achieved needs to be included in the results.
9. For all treatment modalities, variations in the treatment given and violations of the treatment protocol need to be given.
10. Followup needs to be defined in more detail to include: definitions of recurrence and the diagnostic criteria required to be certain of recurrence.
11. The main prognostic factors not included in the regression analysis need to be mentioned, and why they were not included.
12. The survival curves adjusted for prognostic features should be presented.

#### **SUMMARY:**

This is a retrospective analysis of the application of a trial protocol to a patient population in the community to assess the response as measured by

survival, local recurrence, distant recurrence and second primary breast cancer. This study found that there was an increase in the rate of local recurrence in the breast in the community population. The reasons for this are discussed.

#### **REFERENCES:**

1. Katz M.H and Hauck W.W, Proportional Hazards (Cox) Regression. J. Gen. Int. Med. 1993, 8:702-711
2. De Vita V.T, Hellman S, Rosenberg S.A, Cancer: Principles and Practise of Oncology, Lippincott-Raven Publishers 1997 5th Edn.
3. Myers R.H, Classical and Modern Regression with Applications, 1990 2nd Edn: Duxbury Press.