Title: Case mix does not fully explain variation in rates of non-surgical treatment of older women with operable breast cancer.

Running Head: Case mix doesn't explain variation in elderly breast cancer treatment

Authors:
Jenna Morgan
Paul Richards
Sue Ward
Matthew Francis
Gill Lawrence
Karen Collins
Malcolm Reed
Lynda Wyld

Author affiliations:
1. Academic Unit of Surgical Oncology, University of Sheffield Medical School, Beech Hill Road, Sheffield, UK.
2. Department of Health Economics and Decision Science, School for Health and Related Research, University of Sheffield, Sheffield, UK.
3. Public Health England, Knowledge and Information Team (West Midlands), 5 St Philip’s Place, Birmingham, UK.
4. Centre for Health and Social Care Research, Sheffield Hallam University, Collegiate Crescent, Sheffield, UK.

Corresponding author: Miss Jenna Morgan, Academic Department of Surgical Oncology, University of Sheffield Medical School, Beech Hill Road, Sheffield, S10 2JF, UK. Email: j.morgan@sheffield.ac.uk
Tel: +44 (0)114 271 3611. Fax: +44 (0)114 271 3314.

Funding:
This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-1209-10071). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Article type: Original article.

Previous communication to a society or meeting: None.

Key words: Breast cancer, elderly, primary endocrine therapy, surgery, treatment variation
Abstract:

Introduction: Non-surgical management of older women with ER positive, operable breast cancer is common in the UK with up to 40% of over 70s receiving primary endocrine therapy. Whilst this may be appropriate for frailer patients, for some it may result in treatment failure, contributing to the poor outcomes seen in this age group. Wide variation in the rates of non-operative management of breast cancer in older women exists across the UK. Case mix may explain some of this variation in practice.

Methods: Data from two UK regional cancer registries were analysed to identify whether variation in treatment observed between 2002 and 2010 at hospital and clinician level persisted following adjustment for case mix. Expected case-mix adjusted surgery rates were derived by logistic regression using the variables age, proxy Charlson Co-morbidity Score, deprivation quintile, method of cancer detection, tumour size, stage, grade and nodal status.

Results: Data on 17154 women over 70 with ER+ operable breast cancer were analysed. There was considerable variation in rates of surgery at both hospital and clinician level. Despite adjusting for case mix, this variation persisted at hospital level, although not at clinician level.

Conclusion: This study demonstrates variation in selection criteria for older women for operative treatment for early breast cancer, meaning that some older women may be under or over treated and may partly explain the inferior disease outcomes associated with this age group. It emphasises the urgent need for evidence based guidelines for treatment selection criteria in older women with breast cancer.

(249/250 word)
Introduction.

One third of all breast cancers occur in women over 70 years in the UK. With increasing age, levels of co-morbidity and frailty increase, resulting in deaths from other causes exceeding breast cancer mortality in older women with breast cancer. Additionally, tolerance of some breast cancer therapies also decreases. Consequently, older women with operable breast cancer may be offered alternative treatment schedules when compared to younger women. One such treatment strategy is primary endocrine therapy (PET), where for women with potentially operable, oestrogen receptor positive (ER+) cancers surgery may be omitted in favour of endocrine therapy alone. Primary endocrine therapy gained popularity in the 1980s for the management of older women after Tamoxifen was shown to be effective in this setting and a succession of randomised controlled trials comparing its efficacy against surgery followed. A subsequent Cochrane review comparing PET with surgery in the over 70s demonstrated superior rates of local control with surgery but no difference in survival rates. However the studies included in the review were flawed by modern standards because tumour ER status was not always tested and the age range of included younger, healthy women. Recent studies have advocated the use of PET only in the very old or frail and the most recent NICE guidelines issued in 2009 state that PET should only be offered to patients if “significant comorbidity precludes surgery”, and that age should not be a factor in itself.

In the UK there is considerable variation in the non-operative management of women over 70, with regional rates varying between 12 and 40%. However these studies did not adjust for case mix which may account for some of this variation. Similarly, variation in socio-economic status may also impact on levels of co-morbidity, education, screening uptake and stage at presentation which are all factors that may contribute to the treatment decision. Therefore it is important to correct for differences between populations by adjusting for patient and tumour characteristics to understand whether these explain variations in treatment.

Several studies have used registry data to identify factors affecting the receipt of surgery in older breast cancer patients, but none have examined the variation in treatment assignment according to individual hospital and clinician level. The present study aimed to analyse UK practice in older women with operable, ER-positive breast cancer to establish whether the variation observed at hospital and clinician level persists following adjustment for the patient and tumour characteristics of the cases managed.
Methods.

Records on new invasive breast cancers diagnosed in women aged 70 years and over between the years of 2002 and 2010 were acquired for two UK cancer registration regions (West Midlands, Northern and Yorkshire). Data on patient and tumour characteristics and deprivation were included. Deprivation was recorded as quintiles of the income domain of the English Indices of Multiple Deprivation 2010\(^{(16)}\), derived from the patient’s postcode. Data were also obtained from a linked, matched Hospital Episode Statistics (HES) dataset. Hospital Episode Statistics are a record-based system that collects data on all admissions, outpatient appointments and A&E attendances at NHS hospitals in England, including diagnostic codes\(^{(17)}\). A proxy Charlson Comorbidity Index \(^{(18)}\) score (excluding cancer) was calculated for each patient using the diagnostic codes recorded for any in-patient or day case hospital admission in the 18 months before diagnosis of their breast cancer. The cancer component of the Charlson Comorbidity Index was derived from the registry data, in a method consistent with other similar registry data analyses \(^{(14, 19)}\). The Charlson Index includes the following co-morbidities: myocardial infarct, congestive cardiac insufficiency, peripheral vascular disease, dementia, cerebrovascular disease, chronic pulmonary disease, diabetes, liver disease, kidney disease, and other cancers. Higher scores indicate higher levels of comorbidity.

Analyses were restricted to patients with operable, oestrogen receptor positive (ER+) disease at diagnosis. Patients with oestrogen receptor negative (ER-) disease, metastatic disease at diagnosis or pre-invasive disease (ductal carcinoma in situ or pure Paget’s disease of the nipple) were excluded. Patients who died within 91 days of diagnosis were also excluded from the analysis as they were likely to have had advanced disease or other terminal illness which would have influenced treatment decision making. Oestrogen receptor status was only recorded for 43.5% \((n=10\,429)\) of the population, due in part to this information not being routinely collected in the Northern & Yorkshire registry until 2009. However the completeness of data regarding receipt of hormone therapy was more comprehensive and reliably documented for these patients \(^{(13)}\). As such, it was assumed that patients with unknown ER status who received hormone therapy were ER+ (as hormone therapy should only be used in these patients). Patients with unknown ER status who did not receive hormone therapy were assumed to be ER- and were excluded.

Primary treatment was dichotomised as surgery or no surgery according to whether or not the patient had an episode of breast surgery recorded within 6 months of diagnosis, this cut-off was chosen in order to attempt to differentiate between those patients having neo-adjuvant endocrine therapy and those having surgery following failed PET. The proportion of patients undergoing surgery was calculated for each clinician and hospital. Only hospitals and clinicians that treated 10 or
more patients over the period studied were included in the analysis, (excluding 2.9% of hospitals and 3.1% of clinicians) leaving 68 hospitals and 167 clinicians in the final analysis. The final number of patients included in the hospital analysis were 16 654, with 16 606 included in the clinician analysis.

Multivariable logistic regression was used to estimate the probability of a woman undergoing surgical treatment based on patient level factors, including age, proxy Charlson co-morbidity score, level of socioeconomic deprivation, tumour detection method, size, grade, TNM stage and nodal status. Missing data on disease characteristics and co-morbidity was handled using the method of multiple imputation by chained equations (MICE)\(^{(20)}\) to produce 25 imputed data sets and combining the results\(^{(21)}\). Covariates with over 50% missing data, such as HER2 status, were not included in the regression models.

Expected rates of surgical treatment were calculated for each clinician and hospital by summing the individual patient probabilities estimated from the logistic regression model. Risk adjusted rates of surgery were produced by dividing the observed rate by the expected rate for each clinician and hospital and multiplying this by the national rate\(^{(22)}\).

Both unadjusted and adjusted rates of surgery at clinician and hospital level were displayed graphically as funnel plots to allow examination of the variability at each level and identification of outlying practice. Funnel plots contain two limits; under the hypothesis that treatment choice is randomly determined and independent of clinician or hospital, 95% of units would lie within the inner limits (2 standard deviations from the mean) and 99% within the outer limits (3 standard deviations from the mean).

Logistic regressions were performed in IBM SPSS Statistics version 21 and multiple imputations were performed using the open source statistical programming language R (version 3.0.1), with the remaining data handling and analysis performed in Microsoft Excel for Windows 7.
Results.

Cancer registration records were obtained for 23 960 patients over the age of 70 years diagnosed with invasive breast cancer between the years 2002 and 2010. After applying the exclusion criteria (as described above) 17 129 records remained for analysis (see Figure 1). On the basis of the assumptions made to define ER status, it was estimated that 77% of women with non-metastatic disease had ER+ tumours. This is lower than observed in previous cohort studies; for example Diab et al\textsuperscript{(1)} reported 90% of women over 75 diagnosed with breast cancer in the US had ER+ disease.

The median age of the included population was 79 years (70-103 years). Of the 17 129, 9 955 were treated with surgery, giving an overall rate of 58.1%. Once again this is in keeping with other published data from the UK \textsuperscript{(13)}. Patient and disease characteristics are shown in Table 1. The proportion of older women being treated with surgery varied with patient and disease characteristics, with a woman being more likely to undergo surgery if she was younger, living in a less deprived area, having fewer or no co-morbidities, presenting through screening and having a smaller, node negative, Stage I or grade III cancer.

The unadjusted rates of surgery varied substantially between hospitals (Figure 2(a)) and clinicians (Figure 3(a)), with 25 of 68 (36.8%) hospitals and 36 of 167 (21.6%) clinicians falling outside of the outer 99% limits, and 39 of 68 hospitals (57.4%) and 73 of 167 (43.7%) clinicians falling outside of the inner 95% limits on the funnel plots, meaning that they statistically differ from the expected norms.

Taking account of patient level characteristics and adjusting for case mix did not significantly reduce the variation in surgery rates between hospitals, with 15 of 68 (22.1%) still falling outside of the outer 99% limits and 30 of 68 (44.1%) falling outside of the inner 95% limits on the funnel plot (Figure 2(b)).

However, at clinician-level, adjusting for case mix did appear to reduce the variation in surgery rates, with 7 of 167 (4%) falling outside of the outer 99% limits and 17 of 167 (10.2%) falling outside the inner 95% limits on the funnel plot (Figure 3(b)).
Discussion.

Between 2002 and 2010, 17,129 women were treated for assumed operable, ER+ breast cancer in the West Midlands and Northern & Yorkshire cancer registration regions. This represents a quarter of all breast cancer cases in the UK and the populations covered by these registries are demographically representative of the UK as a whole, making it reasonable to extrapolate these findings to the UK population generally. Of these, 9,955 were treated surgically, with the remaining 7,174 (41.8%) having non-surgical management – this figure is in keeping with the rate of PET for these two registration regions in previous studies(13).

The analysis demonstrates that increasing age at diagnosis is associated with a reduced likelihood of receiving surgical treatment which is consistent with other similar studies (5, 15, 23-26). Deprivation level was also associated with treatment type, with the most deprived group being less likely to undergo surgical management, a finding also described by Lavelle and colleagues in their prospective cohort of 800 women (26). This may be due to the fact that affluence is associated with lower levels of comorbidity and smoking, and greater longevity and education (27), thereby promoting better health and discussion of treatment options. Higher levels of comorbidity were also associated with non-surgical treatment, which is also consistent with other published studies, where co-morbidity is stated as a major reason for choosing PET over surgery (28-30). Tumour factors were also associated with treatment type, with larger, node positive tumours being less likely to be treated surgically which may represent patients and clinicians trying to avoid more major surgery, such as mastectomy and axillary node clearance. These results corroborate and update those found by Lavelle and colleagues in their study of 23,038 women aged 65 years and over between 1997 and 2005 (14).

There was considerable variation in the rates of surgical treatment across the 68 hospitals and this variation persisted, despite case-mix adjustment, with 44.1% of units remaining outside the 95% limits in funnel plot analysis. Sixteen hospitals had significantly higher and 14 hospitals had significantly lower rates of surgery than could be explained by the case mix information available. There was also substantial variation in rates of surgical treatment between 167 clinicians, although this variability lessened with case-mix adjustment, with only 10.2% falling outside the 95% limits on funnel plot analysis. However, this still showed that 12 clinicians had significantly higher and 4 had significantly lower rates of surgery than could be explained by case mix alone. It should be noted that there were much smaller numbers available for analysis at clinician level and so these results are less reliable than the hospital level data. It is also possible that the persistence of variability in treatment at hospital level but not at clinician level is a result of a "cluster effect" – in that clinicians working within the same hospital are likely to have trained locally, will work together within a multi-
disciplinary team and may subscribe to a local protocol, thereby having similar practices, resulting in magnified effect at hospital level when the data from individuals is combined.

This persistence of variation in the treatment of older women with operable, ER+ breast cancer at hospital level is due to factors not included in the case-mix adjustment. One possible cause is clinician preference for either treatment. Current guidelines on the use of PET in the older breast cancer population state it should only be used in patients with a short life expectancy (less than 2-3 years), when significant comorbidities preclude surgery, or in patients who refuse surgery\textsuperscript{11, 31}. It is left to the treating clinicians’ judgement as to which patients should be offered PET as an alternative treatment option to surgery. Patient preference or refusal of surgery is also often stated as a possible reason for variation in treatment, which may reflect clinician preference and how the treatment options are presented, as was proposed by Hamaker et al \textsuperscript{32}. Qualitative research in this older group of patients has suggested that they are more passive decision-makers, relying on the advice of healthcare professionals \textsuperscript{33, 34}. Lavelle and colleagues found, in their cohort of 800 women over the age of 70, that lower rates of surgery among elderly patients are unlikely to be due to patient choice \textsuperscript{26}.

Cancer registration data allows analysis of large cohorts of women treated in everyday, normal clinical practice. The routine nature of data collection through hospital coding teams makes this type of observational data less prone to selection bias. However, this method is hampered by missing data and potential coding inaccuracies which is a limitation of this study. A strength of this analysis is the use of multiple imputation which is less prone to bias than other commonly used methods to account for missing data, such as complete case analysis or inclusion of missing as a category in factor variables \textsuperscript{35}. However, whilst exploratory analysis of the imputed data suggested that the values were plausible, it is not possible to verify the extent to which the distribution of the imputed data accurately represents that of the missing values. By using 25 imputations, uncertainty around the missing data is incorporated into the probabilities used to adjust for case mix which mitigates against any small biases due to problems with the imputation model. Despite this model containing many clinically-relevant variables, not all covariates could be included due to missing data, e.g. HER2 and progesterone receptor (PR) status. Additionally, assumptions had to be made regarding the ER status of the patients, with the resulting proportion of ER+ patients in the population being considerably smaller that reported in other studies \textsuperscript{1} indicating that some eligible patients may have been missed from the analysis. However, we do not anticipate that this would affect the main findings of this analysis.
Another limitation of this analysis is the proxy Charlson score using HES data. Data are only available from HES if a patient who has had a hospital in-patient or day case admission in the year preceding their cancer diagnosis and relies heavily on coding of the relevant co-morbidities, and the accuracy of coding within HES. This method may under-score patients who have chronic co-morbidities which are well-controlled and managed in the community or outpatient setting, such as diabetes or dementia, as these alone are unlikely to precipitate a hospital admission.

The case-mix adjustment may also have been inadequate, due to lack of data on important covariates, such as frailty, which are not captured by cancer registration data. Detailed data on every aspect of a patient’s care that could influence treatment choice cannot be collected in this setting, so factors such as frailty, patient choice, family input, social circumstances and clinician preference have not been included but may all play a part when deciding on a treatment modality in the elderly population. It is therefore possible that some other variables are confounding the results presented in this analysis.

Many factors influence treatment choice, as discussed above and examining how these vary in relation to treatment may provide evidence to help explain the variability in treatment of older patients across the UK. Whilst this study has identified outlying practice, it is not clear why they are out-with normal practice, nor whether this outlying practice is unreasonable. Outlying status could be explained by data quality or confounders as previously discussed. However this variation should not be ignored, but further research to determine why they vary significantly should be undertaken.

This study demonstrates that whilst the majority of UK hospitals and clinicians have similar decision making practices, there are some units where practice varies substantially from this norm and is not compensated for by case mix adjustment. Such significant variation in practice is important, particularly in view of the literature on this topic suggesting that patients who are treated with PET may have inferior rates of local control if selected poorly such that they were fit enough for surgery from the outset. It also suggests that in some units, very frail and unfit women who would be adequately managed non-surgically, may be having surgery unnecessarily with its attendant morbidity. This highlights the urgent need for evidence based guidelines for decision making in this age group.

Acknowledgments:

The authors would like to acknowledge Catherine Lagord from Public Health England, Knowledge and Information Team (West Midlands) and Sarah Lawton from Public Health England, Knowledge and Information Team (Northern & Yorkshire) for assisting with providing the data for this analysis.
References: