Radiological audit of interval breast cancers: Estimation of tumour growth rates


Leeds Teaching Hospital NHS Trust, Beckett Street, Leeds, LS9 7TF, UK
Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Mile End Road, London, E1 4NS, UK
Cambridge Breast Unit, And NIHR Cambridge Biomedical Research Centre, Cambridge University Hospitals NHS Trust, 277 Hills Road, Cambridge, CB2 0QQ, UK
University Hospitals of Derby and Burton NHS Foundation Trust, Uttoxeter Road, Derby, DE22 3NE, UK
Oxford Breast Imaging Centre, Old Road, Oxford, OX3 7LE, UK
King’s College Hospital NHS Foundation Trust, Denmark Hill, London, SE5 9RS, UK
University of Oxford, Wellington Square, Oxford, OX1 2JD, UK

Original article

Radiological audit of interval breast cancers: Estimation of tumour growth rates

Abstract

Introduction: This multicentre, retrospective study aimed to establish correlation between estimated tumour volume doubling times (TVDT) from a series of interval breast cancers with their clinicopathological features. The potential impact of delayed diagnosis on prognosis was also explored.

Materials and methods: Interval cancers, where screening mammograms demonstrated changes that were retrospectively classified as either uncertain or suspicious, were reviewed from five screening units within the UK NHS Breast Screening Programme (NHSBSP). Data collected included the time interval between screening mammogram and cancer diagnosis, the size of the initial mammographic abnormality and of the subsequent cancer, demographics, mammographic density and tumour biology. We estimated volume doubling times and the estimated change in size and node status, which would have followed if these cancers had been detected at the previous screen.

Results: 306 interval cancers meeting the inclusion criteria were identified. Average time from screening to diagnosis was 644 days (SD 276 days). 19% were diagnosed in the first twelve months, 42% in the subsequent twelve months and 39% thereafter. Overall average estimated TVDT was 167 days (95% CI 151–186). Significant differences were noted with age (p = 0.01), grade (p < 0.001) and ER status (p < 0.001) with women under 60, grade 3 cancers and ER negative cancers having shorter TVDTs. HER2 positive tumours had shorter doubling times than HER2 negative, but this difference was not statistically significant. It was estimated that diagnosing these cancers at the previous screen would have increased ten-year survival from 82% to 86%.

Conclusion: High grade, ER negativity and younger age were associated with shorter durations of TVDT. The role of HER2 status on interval cancer growth rate requires further assessment. It is likely that the delayed diagnosis of interval cancers confers a 4% reduction in ten-year survival.

© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. **Introduction**

An interval cancer is defined as having occurred after negative mammography, where the screening mammogram did not result in further investigation and cancer was diagnosed before the next scheduled screen [1]. In England there are 18,000 cancers diagnosed on screening mammography annually with a further 6000 presenting with interval cancers [1]. Audit of interval cancers is routine in the NHSBSP. Previous screening mammograms are examined and classified as 1 = normal, 2 = uncertain or 3 = suspicious.

1. Normal. Newly detectable cancers that have developed since the last screening appointment (common).
2. Uncertain. Cancers that were visible at the last screen but not recalled for further tests because the signs of cancer were very subtle and thought to be normal (less common).
3. Suspicious. Cancers that were visible at the last screen but not recalled for further tests because the signs were missed (rare).

In most cases no evidence of the diagnosed cancer on retrospective review of screening images is seen, but in 20% of cases, the cancer can be identified, representing the ‘uncertain’ and ‘suspicious’ categories of mammogram review, and it is this group that we studied.

Interval cancers that have developed since the preceding mamagram are likely to have been faster growing with an anticipated negative impact on prognosis, although this remains unclear due to the challenge in retrospectively distinguishing between cancers that were mammographically occult and those that were not present at the time of screening. Of note, interval cancers contain a higher proportion of grade 3 tumours than screen detected cancers [2,3].

Tabar et al. analysed data from the Swedish Two-County trial [4] and found that the reduction in breast cancer mortality associated with screening was particularly strong in grade 3 tumours. It is therefore worth considering the grade of interval cancers, which potentially constitute screening failures. Patients diagnosed with interval cancers have a poorer prognosis than those with screen-detected cancers, partly due to detection at a later stage, and possibly partly due to a length bias phenomenon whereby screening will tend to detect a higher proportion of slow growing tumours [5–7]. The observation of less favourable grade and biological characteristics is consistent with some length bias [7,8].

It would be helpful to understand the implications of potentially missed cancers in terms of what is ‘lost’ in clinicopathological terms as a result of these tumours remaining undetected at screening and progressing on to symptomatic disease. This entails estimation of their growth rate in terms of tumour volume doubling time (TVDT) and from this, imputation of the likely pathology at the time of the previous screen.

In other programmes, TVDT has been previously reported with an average of 191 days (95% CI 158–230) and no significant difference in growth rate observed by age between 50 and 74 years [9]. A recent Canadian study found that triple negative cancers (TNBC) and ER-negative/HER2-positive tumours are diagnosed more frequently as interval cancers than as screen-detected. The authors hypothesised that the increased proportion of HER2 positive and TNBCs seen as interval cancers in their study was due to the more rapid growth of these tumours and not to a failure to identify these tumours mammographically [10].

Nakashima and colleagues identified significant differences in short-term pre-operative TVDTs according to histological grade, molecular subtype and Ki67. Shorter TVDTs were noted in high grade and triple negative cancers and where Ki67 indices were high, than in ER positive and/or HER2 positive tumours with lower grades or Ki67 indices [11]. The association of higher grade with reduced doubling time has also been reported in another study [12].

Radiological audit of interval cancers provides an opportunity to derive approximate estimates of growth rates, by comparison of radiological sizes at diagnosis and at the original screen, for those cancers where there is a potential lesion in the area of interest on the previous screening mammogram. Here we report the estimated growth rates from audit of interval cancers from five units in the NHS Breast Screening Programme.

This paper does not attempt to evaluate screening per se. Rather, the aims of this work are threefold:

1. For interval cancers which were at least potentially detectable at the previous screen, to estimate the clinical implications of earlier detection, had they been diagnosed at screening.
2. To estimate the growth of tumours in terms of volume doubling time by comparing radiological sizes at diagnosis and on the previous screening mammogram.
3. Identification of associations of clinical and pathological features with volume doubling time.

2. **Materials and methods**

Data were collected retrospectively from five NHS breast-screening units across the UK, Cambridge, Leeds, St George’s Hospital London, Derby and King’s College Hospital London. All interval cancers known to the units with a previous screening mammogram classified as suspicious and uncertain, diagnosed between 2008 and 2016 in women aged 47–90 at diagnosis were included. Our motivation for focussing on these tumours for which there was a possible sign on the previous screening mammogram was first that these are interval cancers that could potentially be avoided by screen detection and that they afford two serial measures of the size of the tumour, facilitating estimation of growth rates. We included only invasive tumours, as it was likely that there would be substantial heterogeneity between growth rates of in situ and invasive disease. A pragmatic approach was taken to include as many cases as met the inclusion criteria but the data do not represent a complete consecutive series between these dates. Interval cancers are identified locally by the team and by the Breast Screening Quality Assurance team and are reviewed as they are identified and therefore are not consecutive cases. The interval cancer review involves the film readers reviewing the screening mammogram and giving it a score of 1 (normal), 2 (uncertain) or 3 (suspicious) and then the screening mammogram and diagnostic mammogram are reviewed together and the scores collated. Only those scored as 2 or 3 were included in the study. We collected information on diagnosis, demographics and details of the previous screening mammogram, all of which had been double-reported. Radiologic features included categorisation of the lesion (mass, asymmetry, calcification or architectural distortion) and diameter on the previous screening and diagnostic mammograms. Where possible, ultrasound size at the time of diagnosis and BIRADS mammographic density were also noted. Pathology data, including tumour size, node status, grade, ER status and HER2 status, were extracted from local pathology reports.

We fitted a linear regression of pathological size (invasive) on radiological size at diagnosis. Using this, we imputed the pathological size at the time of the original screen based on radiological measurement. We also fitted a multinomial regression of node status (negative, 1–3 nodes positive with macro or micro-metastases, 4+ nodes positive) on invasive pathological size at diagnosis. This gave us estimates of proportions in node status categories for given tumour size. Using this and the imputed
pathological size at original screen, we then imputed the likely distribution in the study population of node status at the time of the screen (although not individual node status at screen). Assuming grade did not change[13], we were then able to estimate the likely average Nottingham Prognostic Index (NPI) at the time of the screen [14].

Tumour volume was assumed to be proportional to the cube of the radiological diameter. If for a given case we denote the diameter at original screen as $d_1$, diameter at diagnosis as $d_2$, time between screen and diagnosis as $x$, and doubling time as $T$, we have:

$$kd_1^3 = 2kd_2^3$$

The constant of proportionality $k$ cancels out. Taking logarithms, we have:

$$\frac{1}{T} = \frac{3(\ln(d_2) - \ln(d_1))}{x \ln(2)}$$

We therefore estimated the inverse of the TVDT as above for each case and carried out statistical analysis on the inverse rather than the TVDT, as the inverse was more regularly distributed. We estimated means and 95% confidence intervals on these for all tumours and for subgroups by age, histological type, grade, oestrogen receptor (ER) status and HER2 status. Thereafter, we transformed these to TVDT in days. Significance of associations with the inverse of TVDT were assessed using linear regression. No adjustments were made for screening location.

3. Results

311 patients with invasive interval cancers met the eligibility criteria. In 248 (80%) cases, ultrasound size at the time of diagnosis was included and for 278 (89%) cases, at least partial pathology data were available.

Five cases were excluded on the following grounds:

- Not an interval cancer but asymptomatic screen detected on additional surveillance due to family history.
- Mass at original screen, asymmetry at interval diagnosis, no pathology information.
- No invasive component size, 60 mm in situ component.
- No pathology data, no ultrasound size.
- No pathology data, no ultrasound size.

This left a total of 306 for final analysis.

Three subjects had radiological size missing at the time of interval cancer diagnosis, and for these we used the size on ultrasound. For two subjects with radiological size at interval cancer diagnosis smaller than at the original screen, we replaced the radiological size with the ultrasound size, and for two we replaced the radiological size with that expected based on the regression relationship between radiological size and pathological size (see below).

The mammographic features of the 306 interval cancers were: 232 (76%) mass with or without calcifications, 40 (13%) asymmetry with or without calcifications, 21 (7%) stromal deformity with or without calcifications, 11 (4%) calcifications alone, and 2 (<1%) other. Table 1 shows basic patient and tumour characteristics. Of the 306 subjects retained for analysis, 58 (19%) had interval cancers after a (first) prevalent screen and 248 (81%) after an incident re-screen. The average age at diagnosis was 62 years (SD 7 years), and ages ranged from 47 to 90 (the latter being a self-referral), with 264 (86%) cases aged 50–70. The average time from screen to diagnosis was 644 days (SD 276 days). There were 60 cases (19%) diagnosed in the first year after screening, 128 (42%) in the second, 109 (36%) in the third, and 9 (3%) more than three years after screening.

Mean radiological tumour diameter at diagnosis was 26.6 (SD 15) mm with imputed size at screening 12.8 (SD 8) mm. Pathological tumour diameter at diagnosis was 27.7 (SD 19.9) mm and imputed at screening 17.3 (SD 6.1) mm. Size, node status, grade and NPI at diagnosis and imputed size, node status, grade and NPI at the time of the original screen are shown in Table 2. We estimated that had these cancers been diagnosed at the screen, they would have been on average 10 mm smaller in histological size, and 32% would have been node positive, rather than the 42% at diagnosis. The corresponding estimated NPI was 3.92 compared to 4.29 at diagnosis, corresponding to a ten-year survival of 86% if diagnosed at screening, compared to 82% from the observed NPI at diagnosis [14].

The overall estimated TVDT was 167 days (95% CI 151–186) (Table 3). The corresponding median doubling time was 196 days. There was no significant difference between average doubling times of cancers classified as uncertain (164 days, 95% CI 147–186) and those classified as suspicious (179, 95% CI 138–250). There was a significant increase in doubling time with age ($p < 0.01$). In women aged less than 60 years the average TVDT was 145 days (95% CI 126–170) and in women aged 60 or over, the average TVDT was 185 (95% CI 163–213). There were significant differences in TVDTs by grade ($p < 0.001$) and ER status ($p < 0.001$), grade 3 and ER negative cancers having shorter TVDTs. This is reflected in larger radiological sizes at interval cancer diagnosis for grade 2 and 3 cancers (respectively mean 28 mm, (SD 16 mm) and mean 28 mm, (SD 14 mm) compared to grade 1 tumours (mean 21 mm, SD 11 mm). ER negative cancers were also radiologically larger (mean 35 mm, SD 22 mm) compared to ER positive (mean 26 mm, SD 14 mm). The average estimated pathological size at original mammogram was 17 mm for all three grade categories. This was due to the difference in times between original screen and interval cancer diagnosis, with average times of 729 (SD 207), 663 (SD 267) and 589 (SD 280) days for grade 1, 2 and 3 cancers respectively. Fig. 1 shows the projected growth over time for grade 1, 2 and 3 tumours based on the TVDT estimates.

There was no significant difference in doubling time by tumour type ($p = 0.6$), with similar doubling times observed for ductal and lobular carcinomas. However 114 (37%) of the 306 tumours had type unrecorded, and for these the estimated doubling time was longer than for those with known type. The 37 HER2 positive cancers had shorter TVDTs than HER2 negative cancers but this difference was not statistically significant ($p = 0.4$). Table 3 shows the doubling times by pathological/biological subgroups.

BIRADS breast density information was available for 219 women. TVDT was not significantly related to BIRADS breast
density where this was known ($p = 0.2$). However, those in the lowest density category had rather shorter doubling times than the other three categories, and those with unknown density had substantially longer doubling times than those with density recorded.

### 4. Discussion

This study used data on patients with who developed interval cancers, for which some sign was perceptible on the previous screening mammogram to allow an estimation of tumour growth.
rates, the likely difference in size and node status and the consequent survival benefit if the cancers had been detected at the original screen. We found a modest but potentially worthwhile survival benefit of detecting these cancers at screening.

Overall, we estimated an average TVDT of 167 days, just under six months. This is similar to that estimated by others [9,15,16]. We found a clear and significant gradient of increasing TVDT (i.e., slower growth) with increasing age. Otten et al. (2018) found a non-significant increase in doubling times with age [9]. Our results are consistent with findings in the past that tumour progression is faster at younger ages [17].

Of note, we found a significant monotonic reduction in doubling time (i.e., faster growth) with worsening grade, which has not invariably been observed in the past [18]. Grade, as originally described by Elston and Ellis [19] is the summation of three components; the frequency of cell mitosis (mitotic score), the tubule formation score, and the nuclear pleomorphism score. Thomas went on to suggest that Grade 2 might be a mixture of higher and lower grade tumours distinguished by high and low mitotic rate/scores [20]. Whatever the underlying biology, our results show a very clear relationship of grade with growth.

Our finding of faster growth in ER negative tumours has been observed by others [11,12]. Interestingly, we found no significant association of doubling time with HER2 status. This may be due to low statistical power due to relatively small numbers of HER2 positive cases.

In our data, there was no significant association of mammographic density with doubling times, although the doubling time was short in those with very non-dense breasts. This is consistent with findings of faster growth with higher body mass index [21] and poorer prognosis in the patients with very fatty breast tissue [22]. However, it is not clear what are the causal effects in the complex interplay between body mass index, mammographic density, screening sensitivity and tumour progression. Indeed, Chiu et al. (2010) found a faster progression in patients with dense breasts [23].

The results in Table 2 suggest that a moderate benefit in terms of earlier detection with a consequent reduction in breast cancer deaths could be achieved if these potentially screen detectable cancers were detected at screening instead of subsequently arising symptomatically. The question remains as to how practice might change to achieve this. A starting point will be a careful analysis of the features suggestive of malignancy on the previous mammogram, and consideration of possible amendments to guidelines. Other possibilities include additional imaging for younger screenees or those with denser breasts. Reducing the screening interval is likely to result in earlier diagnosis for those currently presenting with interval cancers with uncertain or suspicious findings on retrospective review, and for a proportion of those without prior mammographic abnormality. An advantage of examining TVDT in this cohort of tri-annually screened women is the inclusion of the slower growing tumours, providing a more complete picture upon which to base decisions practice change.

The findings of this study are subject to some limitations in that we were unable to present a full cohort from each unit due to difficulty in accessing complete data. The cases are exclusively interval cancers with prior mammograms retrospectively classified as suspicious or uncertain. However, this is the population with a clear potential for earlier diagnosis at screening. The estimation also required assumptions, specifically that the grade did not change with time, and that volume was proportional to the cube of the maximum diameter on pathology. However, any estimation of tumour growth from serial imaging requires these or similar assumptions [9,16,24]. Another limitation is our assumption that grade does not change with time. It is at least possible that some of the grade 3 tumours at diagnosis were grade 2 or grade 1 at the time of the previous screen. There is some evidence that breast tumours can progress with respect to grade, although there is not universal agreement about this [5,13]. There is evidence that if such progression does occur, it does so in a relatively small minority of cases, so the assumption may be a reasonable approximation [25].

In spite of these limitations, this study demonstrates the association of TVDT with age, tumour grade and oestrogen receptor status: The higher the grade, the shorter the doubling time, similarly, significantly increased growth rates were seen in ER negative cancers, in keeping with some but not all findings published by others [11,12,18,21,24]. In terms of association of growth rates with age and with histological grade, our results support clear trends of faster growth at younger ages and in grade 3 cancers. It is likely that technological innovations such as different or supplemental imaging and/or adoption of machine intelligence technology will reduce interval cancers. It is not known at the moment which particular types of interval cancers will be affected, however.

5. Conclusion

High grade, ER negativity and younger age are associated with shorter durations of tumour doubling time, i.e., accelerated growth. There was no correlation between breast density and growth rate.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Formatting of funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors have no conflicts of interest to declare.
Acknowledgements

We would like to thank the administrative teams in all breast screening units for helping collate the data.

List of abbreviations

NHSBSP National Health Service [of the United Kingdom] Breast Screening Programme
SD Standard deviation
CI Confidence interval
ER Oestrogen receptor
TVDT Tumour volume doubling time
TNBC Triple negative breast cancer (oestrogen, progestogen and HER2 negative)
BIRADS Breast Imaging Reporting and Data System
NPI Nottingham Prognostic Index

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.breast.2020.03.006.

References