PERGAMON

European Journal of Cancer 39 (2003) 1746–1754

European Journal of Cancer

www.ejconline.com

Quantifying the potential problem of overdiagnosis of ductal carcinoma *in situ* in breast cancer screening

M.-F. Yen[a,b], L. Tabár[c], B. Vitak[d], R.A. Smith[e], H.-H. Chen[f], S.W. Duffy[a,*]

[a]*Cancer Research UK Department of Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK*
[b]*Department of Statistical Science, University College London, UK*
[c]*Department of Mammography, Falun Central Hospital, S-79182 Falun, Sweden*
[d]*Department of Radiology, Mammography Department, Medical Radiology, University Hospital, S-58185 Linköping, Sweden*
[e]*Department of Cancer Control, American Cancer Society, Atlanta, GA, USA*
[f]*Institute of Preventive Medicine, National Taiwan University, Taipei, Taiwan*

Received 7 October 2002; received in revised form 9 January 2003; accepted 4 February 2003

Abstract

The relevance of detection of ductal carcinoma *in situ* (DCIS) in a breast cancer screening programme, and the extent of overdiagnosis of non-progressive lesions, remains controversial. It was the purpose of this paper to estimate the incidence of non-progressive, 'overdiagnosed' DCIS. We defined non-progressive DCIS ($DCIS_0$) as DCIS which could not have progressed to invasive disease if left untreated. Progressive DCIS ($DCIS_1$) was defined as DCIS which has the propensity to progress to invasive disease. We fitted a Markov process model of the incidence of progressive and non-progressive DCIS, the transition of the former to preclinical invasive disease and the subsequent progression to clinical symptomatic cancer. We used data from the Swedish Two-County Trial and from service screening programmes in the UK, Netherlands, Australia and the USA to estimate the incidence of progressive and non-progressive DCIS, and the detection rates of each at the first and subsequent screening. Average incidence of non-progressive DCIS was 1.11 per 100 000 per year. Average incidence of progressive DCIS was 2.1 per 1000 per year. At prevalence screen, 37% of DCIS cases were estimated to be non-progressive. A woman attending prevalence screen has a 19 times greater chance of having a progressive DCIS or an invasive tumour diagnosed than of having a non-progressive DCIS diagnosed. At incidence screen, only 4% of DCIS cases were estimated to be non-progressive. A woman attending an incidence screen has a 166 times higher probability of having a progressive DCIS or invasive lesion diagnosed than of having a non-progressive DCIS diagnosed. There is an element of overdiagnosis of DCIS in breast cancer screening, but the phenomenon is small in both relative and absolute terms.

Keywords: Ductal carcinoma *in situ*; Breast screening; Overdiagnosis

1. Introduction

Since the inception of mammographic screening programmes for breast cancer, concerns have been expressed about the possibility of overdiagnosis of breast cancer, in particular of ductal carcinoma *in situ* (DCIS) [1,2]. High percentages of DCIS among screen-detected tumours have been reported as a basis for this concern, as has an increased incidence of DCIS since the advent of screening [2,3]. Various interpretations of these results have appeared in the literature ranging from the opinion that this mainly represents overdiagnosis and is likely to cause more harm than benefit [4], to the position that detection of DCIS is the ideal target of early detection and that a high rate of DCIS represents a large number of invasive cancers avoided [5]. As a consequence, percentages of DCIS among screen-detected cases are often quoted in comparisons of screening programmes as an indicator of benefit or harm, depending on opinion.

It is not clear what the prevalence of DCIS should be, or to what extent variability in the percentages of DCIS

* Corresponding author. Tel.: +44-20-7882-6202; fax: +44-20-7882-6096.
E-mail address: stephen.duffy@cancer.org.uk (S.W. Duffy).

0959-8049/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0959-8049(03)00260-0

between screening programmes is due to variation in the absolute detection rate of DCIS per person screened or to variation in the absolute detection rate of invasive carcinoma, which also affects the percentage. For example, a prevalence screen which results in 2 DCIS cases per thousand and 4 invasive carcinoma cases per thousand has 33% DCIS, whereas a prevalence screen resulting in 2 cases of DCIS per thousand and 7 invasive cases per thousand will have 22% DCIS. The former is not diagnosing any more DCIS than the latter; it is diagnosing fewer invasive cancers.

In our companion paper [6], we found that the detection of DCIS in the Swedish Two-County Trial accounted for 5–12% of the deaths prevented, whereas shifting from invasive stage II or worse to invasive stage I accounted for more than 60% of the deaths prevented. In addition, we noted that despite variation between programmes in the percentages of cases of DCIS, with some programmes reporting particularly high percentages, absolute detection rates were modest, around 1 per 1000 at prevalence screen and around 0.5 per thousand at incidence screens. While these observations suggest that the percentage of cases diagnosed with DCIS should be interpreted with caution, the question remains, however, as to what *rates* of detection of DCIS should ideally be observed?

In principle, one would expect that overdiagnosis could be assessed from the randomised trials of breast cancer screening. There are, however, serious problems of interpretation. First, the trials were powered for differences in breast cancer mortality rather than differences in incidence of DCIS. Secondly, for any excess of DCIS observed in the study arm of a given trial, it is not clear how much is due to stage-shifting from early detection and how much to overdiagnosis, except where results are presented in sufficient detail to address this issue, as in our companion paper [6]. Otherwise, overdiagnosis would have to be estimated on the basis of total cancers (invasive plus *in situ*). Thirdly, the issue of overdiagnosis is complicated by lead time, in that there is an observed excess in a screened population which is due to anticipated rather than overdiagnosed tumours. This is further complicated by design issues such as whether the trial included a screen of the control group at closure.

It is generally agreed that a proportion of DCIS would not have progressed to invasive carcinoma of breast in the absence of screening. This proportion, however, is unknown. It is known that certain clinicopathological features of DCIS are established as predictors of recurrence, either as DCIS or as invasive disease, after treatment [7]. Clearly, cases which recur after treatment are all progressive cases and cannot be regarded as overdiagnosed in any sense. However, of those that do not recur after treatment, a proportion would have done so, and a proportion would not. This proportion is not identifiable, since one cannot tell for any excised DCIS lesion what would have transpired if it had not been excised. Therefore, the only way to estimate the proportion of progressive DCIS cases is by statistical modelling based on the numbers of DCIS and invasive cases detected at screening and on the numbers of breast cancers arising clinically between screenings.

In this paper, we review the rates of DCIS from the Swedish Two-County Trial and from service screening programmes in order to:

1. derive tentative estimates of DCIS detection rates that should be typically observed;
2. describe the typical range of absolute detection rates of DCIS in mammographic screening programmes; and
3. estimate the proportion of DCIS detected at prevalence and incidence screens that truly represents ''overdiagnosis'', i.e. which would not have progressed to invasive disease if left untreated.

2. Patients and methods

2.1. Data

For purposes of estimation, we required data including absolute numbers screened and absolute numbers of invasive and DCIS cases detected at both the prevalence and incidence screens. Data on interval cancers was desirable, but not essential. We also required that the data pertained to mass population screening. Programmes providing such data included our own study, the Swedish Two-County Trial [6], and population service screening programmes from the UK, [8,9], the Netherlands [10], south Australia [11] and New York [12]. These programmes all reported incidence and prevalence screening data in sufficient detail for analysis. The Swedish Two-County Trial also provided interval cancer data, and we were able to impute the number of interval cancers in the Netherlands programme from rates of 0.96 per 1000 reported in the English abstract of a later paper in Dutch [13].

The Swedish Two-County trial was a randomised controlled trial with 77 080 women aged 40–74 years randomly assigned to screening invitation and resulting in 1426 breast cancers during the trial, and 55 985 women assigned to no invitation and resulting in 1042 breast cancers. In the present analysis, only the data on DCIS and invasive breast carcinomas detected at prevalence and the first subsequent incidence screen, interval cancers detected between prevalence and the first subsequent incidence screen and carcinoma-free cases for women aged 40–69 years were used. The corresponding

figures from the service screening programmes in the UK, Netherlands, South Australia, and New York were extracted from published papers (Table 1).

2.2. Statistical methods

We assume that of the DCIS detected at screening, there are two types: $DCIS_0$, or non-progressive DCIS, which has no propensity to progress to invasive disease during the lifetime of the host; and $DCIS_1$, or progressive DCIS, which has the propensity to progress to invasive breast cancer. For treated DCIS which does progress despite treatment, we know that such a case must be of type $DCIS_1$. For a case which does not progress after treatment, we do not know its type, since we cannot tell what would have occurred in the absence of treatment. Thus, we cannot allocate each individual case to one of the two types. We can, however, estimate rates and therefore proportions of the two types from data on detection rates of DCIS and invasive breast cancer at prevalence and incidence screens, and from interval cancer rates (almost universally invasive).

We consider the six-state continuous-time Markov model depicted in Fig. 1. In this model, progressive and non-progressive DCIS are taken into account simultaneously. Note that state (5), no tumour apparent after non-progressive DCIS regression is not treated as a return to normal, but as a separate absorbing state. This is done partly to assist in estimation, but also to reflect the fact that the natural history of non-progressive DCIS is not yet established. For example, those in

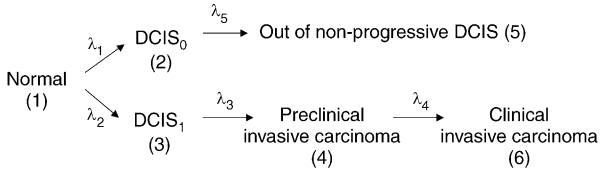

Fig. 1. Natural history of breast cancer incorporating progressive and non-progressive DCIS.

Table 1
Numbers of DCIS, invasive carcinoma and subjects carcinoma-free at prevalence and the first incidence screen in Swedish Two-County, UK, Netherlands, South Australia and New York programmes

Programme	Detection mode	Prevalence screen	First incidence screen
Swedish Two-County	DCIS	8	7
40–49 years	Invasive carcinoma	31	39
	Carcinoma-free	18 456	16 396
	Interval cancer[a]	20	
Swedish Two-County	DCIS	15	12
50–59 years	Invasive carcinoma	87	49
	Carcinoma-free	21 457	18 731
	Interval cancer*	25	
Swedish Two-County	DCIS	17	8
60–69 years	Invasive carcinoma	167	81
	Carcinoma-free	20 395	16 372
	Interval cancer[a]	38	
UK	DCIS	2767	173
	Invasive carcinoma	12 323	925
	Carcinoma-free	2 520 526	227 503
Netherlands	DCIS	908	383
	Invasive carcinoma	5548	2223
	Carcinoma-free	1 077 844	791 790
	Interval cancer[a]	760	
South Australia	DCIS	94	12
	Invasive carcinoma	439	61
	Carcinoma-free	75 573	21 433
New York	DCIS	42	17
	Invasive carcinoma	230	67
	Carcinoma-free	52 378	45 839

DCIS, ductal carcinoma *in situ*.

[a] Numbers of interval cancers in the Swedish Two-County trial and the Netherlands programme are adjusted by the first incidence screen coverage to represent the subgroup attending the first subsequent screen.

whom a non-progressive DCIS has regressed might be at increased risk of a new primary tumour thereafter. In this model, DCIS, whether progressive or not, denotes ductal carcinoma *in situ* which has given rise to calcifications and is therefore screen-detectable. It does not include mammographically-undetectable DCIS.

The transition rates of the six-state model above can be expressed as an intensity matrix:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} -(\lambda_1+\lambda_2) & \lambda_1 & \lambda_2 & 0 & 0 & 0 \\ 0 & -\lambda_5 & 0 & 0 & \lambda_5 & 0 \\ 0 & 0 & -\lambda_3 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_4 & 0 & \lambda_4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (1)$$

λ_1 and λ_2 represent the DCIS incidence rates for non-progressive and progressive DCIS, respectively. λ_3 and λ_4 are the transition rates from progressive DCIS to invasive preclinical phase, and from preclinical invasive to clinical disease, respectively. λ_5 is the annual transition rate from non-progressive DCIS to no apparent tumour.

Given the transition intensity matrix in Eq. (1), transition probabilities of progressing from state i to state j in a time interval, x, can be expressed as $P_{ij}(x)$. The derivation of transition probabilities is based on the solution of the Kolmogorov equations [14,15]. For example, the probability of having preclinical invasive cancer at a prevalence screen at a given age is

$$\begin{aligned} P_{14}(\text{Age}) &= \int_0^{\text{age}} \lambda_2 e^{-\lambda_2 x} e^{-\lambda_1 x} \int_0^{\text{age}-x} \lambda_3 e^{-\lambda_3 y} e^{-\lambda_4(\text{age}-x-y)} \mathrm{d}y \mathrm{d}x \\ &= \frac{\lambda_2\lambda_3}{\lambda_4-\lambda_3} \\ &\times \left\{ \frac{e^{-(\lambda_1+\lambda_2)\text{age}} - e^{-\lambda_3 \text{age}}}{\lambda_3-\lambda_1-\lambda_2} - \frac{e^{-(\lambda_1+\lambda_2)\text{age}} - e^{-\lambda_4 \text{age}}}{\lambda_4-\lambda_1-\lambda_2} \right\} \end{aligned} \quad (2)$$

The probabilities used for maximum likelihood estimation in this six-state model are shown in Table 2. Note that in this more formal representation, it is necessary to take account of the fact that those with a history of clinical breast cancer prior to the start of screening are excluded. This is the reason for the denominators in the first screen probabilities in Table 2. The probabilities are conditional on being either normal, with DCIS, or with preclinical invasive disease at first screen.

Suppose N_1 women attend the prevalence screen, and n_{11} and n_{12} are detected as DCIS and invasive carcinoma, respectively. At the first subsequent screen, then the numbers of attending women, of screen-detected DCIS and invasive carcinoma cases are N_2, n_{21}, and n_{22}. The number of interval cancers between prevalence screen and the first subsequent screen is n_{ic}. Based on the probability formulae in Table 2, the total likelihood, including data on interval cancers is:

$$\begin{aligned} L = &\left(\frac{P_{12}(\text{Age}) + P_{13}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})} \right)^{n_{11}} \\ &\times \left(\frac{P_{14}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})} \right)^{n_{12}} \\ &\times \left(\frac{P_{11}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})} \right)^{N_1 - n_{11} - n_{12}} \\ &\times (P_{12}(x) + P_{13}(x))^{n_{21}} \times (P_{14}(x))^{n_{22}} \times (P_{11}(x))^{N_2 - n_{21} - n_{22}} \\ &\times (P_{16}(x))^{n_{ic}} \end{aligned} \quad (3)$$

In most of the service screening programmes, there were no data on interval cancers available. In this case, we used a complement probability including total likelihood from women not attending the first subsequent screen (including women not yet invited thereto) and all interval cancers appearing between the prevalence screen and the first subsequent screen. Coverage is $N_2/(N_1-n_{11}-n_{12})$, the proportion of those attending the

Table 2
Transition probabilities by detection mode at prevalence and the first subsequent screen in the six-state Markov model

Detection mode	Transition probability for models with interval cancer	Transition probability for models without interval cancer
Prevalence screen		
DCIS	$\frac{P_{12}(\text{Age}) + P_{13}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}$	$\frac{P_{12}(\text{Age}) + P_{13}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}$
Invasive carcinoma	$\frac{P_{14}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\textit{Age})}$	$\frac{P_{14}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}$
Carcinoma-free	$\frac{P_{11}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}$	$\frac{P_{11}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}$
Subsequent screen		
DCIS	$P_{12}(x) + P_{13}(x)$	$[P_{12}(x) + P_{13}(x)] \times \text{Coverage}$
Invasive carcinoma	$P_{14}(x)$	$P_{14}(x) \times \text{Coverage}$
Carcinoma-free	$P_{11}(x)$	$P_{11}(x) \times \text{Coverage}$
Interval Cancer	$P_{15}(x)$	–
Non-attender	–	$[P_{11}(x) + P_{12}(x) + P_{13}(x) + P_{14}(x)] \times (1 - \text{Coverage}) + [P_{15}(x) + P_{16}(x)]$

first screen who also attend the second. Then, the total likelihood for screening programmes without interval cancers is:

$$L = \left(\frac{P_{12}(\text{Age}) + P_{13}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}\right)^{n_{11}} \times \left(\frac{P_{14}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}\right)^{n_{12}} \times \left(\frac{P_{11}(\text{Age})}{P_{11}(\text{Age}) + P_{12}(\text{Age}) + P_{13}(\text{Age}) + P_{14}(\text{Age})}\right)^{N_1 - n_{11} - n_{12}} \times [(P_{12}(x) + P_{13}(x)) \times \text{Coverage}]^{n_{21}} \times [P_{14}(x) \times \text{Coverage}]^{n_{22}} \times [P_{11}(x) \times \text{Coverage}]^{N_2 - n_{21} - n_{22}} \times \{[P_{11}(x) + P_{12}(x) + P_{13}(x) + P_{14}(x)] \times (1 - \text{Coverage}) + [P_{15}(x) + P_{16}(x)]\}^{n_{11} - N_2}$$

The maximum likelihood estimates were obtained by using Newton–Raphson optimisation. SAS/IML 6.12 software was used. 95% confidence intervals (CIs) were calculated using variance estimated from the inverse Hessian matrix.

Note that the coverage at second screen is not to be confused with compliance. At the time of the reports in several of the service screening programmes, the first screen was completed, but the second is still underway. Thus, those who had not attended the second screen were a mixture of those choosing not to attend and those who had not yet been invited to a second screen.

3. Results

Table 3 shows results from estimations based on the six-state Markov model. Note that the exponential distribution of time to transition implies that the inverse of the estimated transition rate from a state is the estimated mean time spent in the state (mean sojourn time). Thus, for example, from the Swedish Two County trial experience, for women aged 50–59 years, the estimated average sojourn time of invasive preclinical cancer from the time of becoming invasive to clinical symptomatic disease is $1/0.40 = 2.50$ years, and the mean sojourn time of non-progressive DCIS before regression is estimated as $1/0.0617 = 16.21$ years.

The annual incidence rates for non-progressive DCIS ranged from 7.22 per million per year in the Swedish Two-County trial in women aged 40–49 to 72.7 per 100 000 per year in the UK programme. Most rates, however, were of the order of 1–3 per 100 000. The annual incidence rates for progressive DCIS ranged from 1 to 2.7 per thousand. The shortest estimated average time of progression of DCIS to invasive disease was 2 months in the Swedish Two-County trial for women aged 60–69 years, and the longest was 5.22 months in the New York programme. The estimated average sojourn time of non-progressive DCIS ranged from 6 years in the UK programme to 37 years in the Swedish Two-County trial for women aged 60–69 years. Weighted averages were used to create pooled estimates of the annual incidence rates for non-progressive DCIS (1.11 per 100 000) and progressive DCIS (2.1 per 1000), the sojourn time of non-progressive DCIS (30 years) and progressive DCIS (3 months), and the sojourn time of preclinical invasive carcinoma (2.5 years). Variance estimates on parameters pertaining to non-progressive DCIS could not be obtained, partly due to the very low incidence, and therefore to the lack of data on which to base the estimation. The profile likelihood was therefore used to calculate 95% CI on parameters pertaining to non-progressive DCIS. This strategy also failed for the rate of transition from non-progressive DCIS to no apparent tumour in the Swedish Two-County Trial for women aged 40–49 years and aged 50–59 years because of an almost flat likelihood in this dimension.

Table 4 shows the probability of non-progressive DCIS, progressive DCIS, preclinical invasive carcinoma and the proportion of non-progressive DCIS at prevalence and at the first subsequent screen derived from the estimates in the six-state Markov model. The larger the proportion of non-progressive DCIS, the more serious the problem of overdiagnosis. Overall, approximately 20–50% of DCIS cases (average 37%) at the prevalence screen were estimated as non-progressive. At first incidence screen, the corresponding estimate of the proportion of DCIS that is non-progressive ranged from 3 to 7%, with the exception of an estimate of 21% in the UK programme, with an overall average of 4%.

Goodness-of-fit tests for the Markov models in the Swedish Two-County trial, for women aged 50–59 years and 60–69 years at randomisation, show very good model fitting. The Pearson chi-squared statistics (2 degrees of freedom) were 0.33 ($P = 0.8$) and 0.17 ($P = 0.9$) for models with women 50–59 years and with women 60–69 years, respectively. However, the model seems to fit poorly in the Swedish Two-County trial, for women 40–49 years with a Pearson Chi-square 8.12 ($P = 0.02$), overestimating the numbers of invasive cancers at the prevalence screen and underestimating the corresponding numbers at the incidence screen.

For service screening programmes, the chi-squared figures were 12.01 ($P = 0.003$), 19.20 ($P < 0.001$), 1.24 ($P = 0.5$) and 1.51 ($P = 0.5$) for the UK, Netherlands, Australia, and New York programmes, respectively. The models for the UK and Netherlands programmes showed a statistically significant lack of fit, again overestimating the numbers of invasive cancers at the prevalence screen and underestimating the corresponding

Table 3
Estimated results of the six-state Markov model

Programme	Parameter	Estimate	95% CI
Swedish Two-County 40–49 years	Normal→$DCIS_0$	7.22×10^{-6}	0–3.89×10^{-5}
	Normal→$DCIS_1$	0.0017	0.0013–0.0021
	$DCIS_1$→Inv	4.93	1.24–8.61
	Inv→Clinical	0.80	0.57—1.04
	$DCIS_0$→Out of DCIS	0.0857	N/E[a]
Swedish Two-County 50–59 years	Normal→$DCIS_0$	9.86×10^{-6}	0–3.54−10^{-5}
	Normal→$DCIS_1$	0.0016	0.0013–0.0019
	$DCIS_1$→Inv	2.99	1.22–4.76
	Inv→Clinical	0.40	0.30–0.49
	$DCIS_0$→Out of DCIS	0.0617	N/E[a]
Swedish Two-County 60–69 years	Normal→$DCIS_0$	1.18×10^{-5}	1.5×10^{-6}–2.60×10^{-5}
	Normal→$DCIS_1$	0.0027	0.0023–0.0032
	$DCIS_1$→Inv	6.13	1.92–10.33
	Inv→Clinical	0.33	0.27–0.40
	$DCIS_0$→Out of DCIS	0.0273	0–0.2080
UK	Normal→$DCIS_0$	7.27×10^{-5}	6.57×10^{-5}–7.93×10^{-5}
	Normal→$DCIS_1$	0.0026	0.0023–0.0028
	$DCIS_1$→Inv	3.87	3.20–4.53
	Inv→Clinical	0.52	0.46–0.58
	$DCIS_0$→Out of DCIS	0.1693	0.1546–0.1859
Netherlands	Normal→$DCIS_0$	1.07×10^{-5}	1.06×10^{-5}–1.09×10^{-5}
	Normal→$DCIS_1$	0.0021	0.00199–0.00212
	$DCIS_1$→Inv	4.19	3.83–4.54
	Inv→Clinical	0.39	0.38–0.41
	$DCIS_0$→Out of DCIS	0.0253	0.0235–0.0254
South Australia	Normal→$DCIS_0$	2.89×10^{-5}	1.68×10^{-5}–4.27×10^{-5}
	Normal→$DCIS_1$	0.0021	0.0015–0.0029
	$DCIS_1$→Inv	3.12	2.08–6.22
	Inv→Clinical	0.37	0.25–0.52
	$DCIS_0$→Out of DCIS	0.0511	0.0304–0.0892
New York	Normal→$DCIS_0$	1.41&times10^{-5}	5.1×10^{-6}–2.54×10^{-5}
	Normal→$DCIS_1$	0.0010	0.0008–0.0013
	$DCIS_1$→Inv	2.30	1.44–3.17
	Inv→Clinical	0.24	0.17–0.31
	$DCIS_0$→Out of DCIS	0.0388	0.0119–0.1092
Overall[b]	Normal→$DCIS_0$[c]	1.11×10^{-5}	1.098×10^{-5}–1.129×10^{-5}
	Normal→$DCIS_1$[c]	0.0021	0.00202–0.00213
	$DCIS_1$→Inv[c]	4.0030	3.7242–4.3025
	Inv→Clinical[c]	0.3990	0.3857–0.4127
	$DCIS_0$→Out of DCIS[c]	0.0332	0.0321–0.0344

95% CI, 95% Confidence Interval; N/E, not evaluable.
[a] Likelihood almost flat and a 95% CI so wide as to be meaningless.
[b] Weighted average from Swedish Two-County, women aged 50–69, and screening programs of UK, Netherlands, South Australia and New York.
[c] Weighted by the precision of parameters.

numbers at the incidence screen. The model also slightly overestimated the numbers of interval cancers in the Netherlands programme.

4. Discussion

Before considering the implications of our results, we should consider their reliability. There are two major assumptions in the model: first, the Markov process assumption which implies an exponential distribution of time to any transition; second, the only heterogeneity of average rates of transition modelled is the distinction between non-progressive and progressive DCIS. The exponential distribution has been shown to be a good model for breast cancer progression in the past [14,16], and for three out of the seven programmes and age groups here it gave a good fit. It is evident, however, that a relaxation of the assumption of a common rate of transition within the progressive DCIS cases is

desirable. Propensity to progress and rapidity of progression are likely to be continuous phenomena, dependent on various host, tumour and treatment factors. We plan to develop such a model for use in the context of DCIS treatment studies in the future. In the meantime, because there has been much recent publicity around the issue of overdiagnosis of DCIS, we believe our estimates of such overdiagnosis are useful, provided the above limitations are borne in mind.

For the most part, the assumed Markov process model fits reasonably, but not perfectly. The estimated rates of progressive and non-progressive DCIS, and of progression to invasive disease, reflect the raw data from these programmes—thus, the programmes with high detection rates of DCIS tend to have high estimated incidence in Table 3. The Swedish 60–69 year age group, which has the highest detection rate of invasive tumours at prevalence, also has the highest estimated rate of progression from DCIS to invasive disease.

One can never know with certainty what would have happened to a particular case of DCIS if left untreated, and there is a substantial range of uncertainty around the estimated rates of progressive and non-progressive DCIS. However, the estimates are derived from empirical observations on the detection of DCIS and invasive breast cancer at screening, and from interval cancer incidence where available. They do not rely on *post hoc* extrapolation of outcome in treated DCIS cases for assumptions about the natural history if treatment had not taken place.

On a technical point, the complementary probability in the likelihood function is based on the assumption that the prevalence and incidence screens apply to the same populations. This is approximately true for all studies except New York [12], for which the two screens are from separate screening programmes. The likelihood approximation still seems to work for the New York data, in that estimates compatible with the other programmes are obtained, but it should be remembered that it is an approximation nevertheless.

The rates of progression in Table 3 can be readily transformed to average times spent in the relevant states simply by inversion. Taking the combined estimates at the bottom of the table, the average time a case of screen-detectable progressive DCIS remains in that state before progression to invasive cancer is $1/4.0030 = 0.25$ years, or three months. The time is short since it refers to the window of opportunity for early detection, not from DCIS inception to invasion, but from the appearance of the associated calcifications that render it screen-detectable to the time of invasion. The estimated duration of the preclinical phase once the disease has become invasive is $1/0.3990 = 2.5$ years, or 30 months. This gives a total combined preclinical phase of 33 months, which is similar to estimates of around 3 years found elsewhere [14]. What is evident from these data is that the window of opportunity in the sojourn time to detect DCIS that will become invasive breast cancer is very short, and therefore detection of progressive DCIS at rates similar to invasive cancer detection rates is an unrealistic goal in a screening programme.

Note that the harvest of progressive DCIS is estimated to be very similar at the prevalence and incidence screens (Table 4). This is consistent with the similar detection rates of DCIS and DCIS:invasive ratios observed in the screening programmes (Table 1), and results from the rapid transition to invasive cancer once the DCIS becomes calcified and therefore screen-detectable. If there is only around three months to the transition to invasive disease, the pool of preclinical tumours at an incidence screen two or three years after the last screen will be much the same as at a prevalence screen.

Detailed interpretation of the results in Tables 3 and 4 gives rise to some interesting implications. Table 4 indicates that on average 37% of DCIS cases diagnosed at a prevalence screen are non-progressive. The corresponding proportion at an incidence screen is 4%. This is compatible with previous findings that overdiagnosis

Table 4
Estimated rate per 100 000 of $DCIS_0$, $DCIS_1$ and preclinical invasive carcinoma, with the proportion of $DCIS_0$ at prevalence and the first subsequent screen, under the six-state Markov model

Programme	Prevalence screen				First subsequent screen			
	$DCIS_0$	$DCIS_1$	Preclinical INV	$DCIS_0$/DCIS	$DCIS_0$	$DCIS_1$	Preclinical INV	$DCIS_0$/DCIS
Swedish Two-County								
40-49 years	8	35	216	19%	1	35	164	3%
50–59 years	16	54	411	23%	2	54	251	4%
60–69 years	38	44	821	46%	3	44	473	6%
UK	43	66	490	39%	17	65	369	21%
Netherlands	34	49	520	41%	2	49	259	4%
South Australia	56	67	578	46%	5	67	263	7%
New York	33	45	439	42%	3	45	135	6%
Overall	30	52	520	37%	2	52	279	4%

INV, invasive.

and length bias are largely a phenomena of the prevalence screen [14]. The results suggest that some overtreatment is inevitable, but it does not necessarily follow that 37% of all treatment of DCIS cases diagnosed at the prevalence screen confers no benefit. In the first place, there is uncertainty about the estimate. In the second, it is not clear that those with non-progressive DCIS have the same risk status as women free of breast carcinoma. Clinical experience suggests that women with a prior diagnosis of breast cancer are at greater risk of new primary breast cancers. If this applies to women with non-progressive DCIS, treatment of these may in some cases still forestall future new primaries.

Accepting that there is some diagnosis of non-progressive DCIS and therefore some overtreatment, we can obtain an estimate of the relative burden of this in comparison with the benefit of early treatment of progressive lesions. On the basis of the estimates in Table 4, a woman attending for screening for the first time has a 1 in 3300 chance (30 per 100 000) of being diagnosed with a non-progressive DCIS. This is a 19 times smaller chance than the 1 in 175 probability of being diagnosed with a progressive *in situ* or invasive lesion (520 + 52 per 100 000). At an incidence screen, the chance of having a non-progressive DCIS lesion diagnosed is 1 in 50 000, whereas the chance of having a progressive lesion, whether invasive or *in situ*, diagnosed is 166 times higher at 1 in 302.

Neither our estimates, nor the raw detection rates of DCIS, are consistent with the findings of autopsy studies. The latter have typically observed high prevalences of DCIS at autopsy, even in women who were not diagnosed with breast cancer during their lifetime. Welch and Black [17] found a prevalence of 9% and Nielsen and colleagues [18] a prevalence of 14%. The screening programmes considered here typically observed around 0.1% prevalence at first screen and considerably lower at subsequent screens. It therefore seems that the dead population is a very different in this respect from the living, and that mammographically-detectable DCIS *in vivo* is either a different clinical entity or a small and special subgroup of occult DCIS detectable at autopsy. It is unclear whether the DCIS described in the autopsy studies actually would have been detectable by conventional screening techniques.

We have specifically avoided using the clinical and pathological aspects of treated DCIS in the above exercise. As stated, this was to avoid the necessity of extrapolation of results in treated DCIS to the estimation of what would happen if the disease were left untreated. It is, however, useful to consider our results in the light of observed behaviour after treatment. After treatment with local excision alone, it is estimated that 18% of DCIS cases recur [19]. Lower rates are observed after mastectomy or after local excision with irradiation [20]. Probability of recurrence is also affected by size and grade of lesion, presence of necrosis and resection margin width in the case of local excision [7]. Opinions vary as to the likely course of disease if left untreated. McCready [20] suggests that progression of untreated DCIS to invasive disease would occur in 25–35% of subjects, whereas Frykberg and Bland suggest that such progression would happen in the majority of cases [21]. Our results are more consistent with the latter. Given the expected detection rates of progressive and non-progressive DCIS in Table 4, one would expect 50–80% of DCIS lesions to progress if left untreated. This is also consistent with the results of Evans and colleagues [22], who found that 61% of screen-detected DCIS cases were of high grade, which is a strong risk factor for progression or recurrence in treated DCIS [7]. In addition, Evans and colleagues also point out that the evidence for a minority of lesions being progressive derives mainly from retrospective studies of DCIS cases originally misdiagnosed as benign [23], which are not representative of screen-detected DCIS [22].

Our results, indicating a small element of overdiagnosis, are largely consistent with the evidence from the randomised trials of screening. The only trials reporting a substantial excess of *in situ* or invasive tumours in the arm invited to screening were the Edinburgh trial [24], the Malmö trial [25] and the Canadian National Breast Screening Study-1 [26]. In the case of the Edinburgh and Malmö trials, the excesses may be due to lead time rather than overdiagnosis. In the case of the Canadian trial, at least a large proportion of the excess is in invasive, node-positive tumours [27,28], so this is also unlikely to be a product of overdiagnosis.

If the results above are accepted, the implication is that the majority of screen-detected DCIS cases would progress if not treated. This does not mean that early detection should be particularly focused on the diagnosis of DCIS. Our companion paper showed that the majority of deaths prevented in screening are due to detection of invasive disease at an early stage [6]. It also noted that the calcifications which are the usual mammographic sign of DCIS are often easier to observe than the more subtle asymmetric densities which frequently accompany a small invasive tumour. Thus, the emphasis in screening should be on achieving the quality required to detect small invasive lesions. A programme which has this quality will have no difficulty in detecting DCIS.

The likelihood that there are a number of non-progressive screen-detected DCIS cases indicates that it is important to target therapy in a way that reflects the risk the lesion poses to the patient. Of course, if a majority of lesions progress, resection is indicated in all cases, but it may be reasonable to reserve mastectomy and use of adjuvant therapies for high-risk lesions. A consolidation of current knowledge to form the basis of practice and further therapeutic research are both indicated.

The implications of our results are therefore:

1. there is an element of overdiagnosis and overtreatment of DCIS in mammographic screening programmes;
2. this element is modest compared with the likely benefit of early diagnosis and treatment of progressive lesions; and
3. the increasing diagnosis of DCIS poses a challenge to therapy as much as to early detection.

Acknowledgements

The American Cancer Society provided financial support for this study through a gift from the Longaberger® Company.

References

1. Ernster VL, Barclay J. Increases in ductal carcinoma in situ (DCIS) of the breast in relation to mammography: a dilemma. *Monogr Natl Cancer Inst* 1997, **22**, 151–156.
2. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996, **275**, 913–918.
3. Barchielli A, Paci E, Giorgi D. Recent trends of in situ carcinoma of the breast and mammographic screening in the Florence area, Italy. *Cancer Causes Control* 1999, **10**, 313–317.
4. Mittra I, Baum M, Thornton H, Houghton J. Is clinical breast examination an acceptable alternative to mammographic screening? *Br Med J* 2000, **321**, 1071–1073.
5. Cady B. How to prevent invasive breast cancer: detect and excise duct carcinoma in situ. *J Surg Oncol* 1998, **69**, 60–62.
6. Duffy SW, Tabar L, Vitak B, *et al*. The relative contributions of screen-detected in situ and invasive carcinomas in reducing mortality from the disease. *Eur J Cancer* (in press).
7. Silverstein MJ, Lagios MD, Craig PH, *et al*. A prognostic index for ductal carcinoma in situ of the breast. *Cancer* 1996, **77**, 2267–2274.
8. Chamberlain J, Moss SM, Kirkpatrick AE, Michell M, Johns L. National Health Service Breast Screening programme: results for 1991–92. *Br Med J* 1993, **307**, 353–356.
9. Moss SM, Michel M, Patnick J, Johns L, Blanks R, Chamberlain J. Results from the NHS breast screening programme 1990–93. *J Med Screening* 1995, **2**, 186–190.
10. Fracheboud J, de Koning HJ, Beemsterboer PMM, *et al*. Nationwide breast cancer screening in the Netherlands: results of initial and subsequent screening 1990–1995. *Int J Cancer* 1998, **75**, 694–698.
11. Robinson JI, Crane CEB, King JM, Scarce DI, Hoffman CEJ. The South Australian Breast X-ray Service: results from a statewide mammographic screening programme. *Br J Cancer* 1996, **73**, 837–842.
12. Dershaw DD, Loring L, True SJ, Momrow DC, Liberman L, Morris EA. A comparison of screening mammography results from programs for women of different socioeconomic status. *Cancer* 1998, **82**, 1692–1697.
13. Nationwide breast cancer screening fully accomplished; results from the implementation phase 1990–1997. National Evaluation Team for Breast Cancer Screening. *Ned Tijdschr Geneeskd* 2000, **144**, 1124–1129.
14. Chen HH, Duffy SW, Tabar L, Day NE. Markov chain models for progression of breast cancer part I: tumour attributes and the preclinical screen-detectable phase. *J Epidemiol Biostat* 1997, **2**, 9–23.
15. Cox DR, Miller HD. *The Theory of Stochastic Processes*. London, Methuen, 1965.
16. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 1984, **40**, 1–14.
17. Welch HG, Black WC. Using autopsy series to estimate the disease 'reservoir' for ductal carcinoma in situ of the breast: how much more breast cancer can we find? *Ann Intern Med* 1997, **127**, 1023–1028.
18. Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA. Breast cancer and atypia among young and middle-aged women: a study of 100 medicolegal biopsies. *Br J Cancer* 1987, **56**, 814–819.
19. Vezeridis MP, Bland KI. Management of ductal carcinoma in situ. *Surg Oncol* 1994, **3**, 309–325.
20. McCready DR. Ductal carcinoma in situ. In Souhami RL, Tannock I, Hohenberger P, Horiot J-C, eds. *Oxford Textbook of Oncology*. Oxford, Oxford University Press, 2002, 1717–1724.
21. Frykberg ER, Bland KL. Overview of the biology and management of ductal carcinoma in situ of the breast. *Cancer* 1994, **74**, S350–S561.
22. Evans AJ, Pinder SE, Ellis IO, Wilson AR. Screen-detected ductal carcinoma in situ (DCIS): overdiagnosis or an obligate precursor of invasive disease? *J Med Screen* 2001, **8**, 149–151.
23. Page DL, Dupont WD, Rogers LW, *et al*. Continued local recurrence of carcinoma 15–25 years after a diagnosis of low grade ductal carcinoma in situ treated only by biopsy. *Cancer* 1995, **76**, 1197–1200.
24. Alexander FE, Anderson TJ, Brown HK, *et al*. The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. *Br J Cancer* 1994, **70**, 542–548.
25. Andersson I, Aspegren K, Janzon L, *et al*. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *Br Med J* 1988, **297**, 943–948.
26. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. *Ann Intern Med* 2002, **137**, 305–312.
27. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study-1: breast cancer detection and death rates among women aged 40–49 years. *Can Med Assoc J* 1992, **147**, 1459–1476.
28. Duffy SW. Commentary on the Canadian National Breast Screening Study-1. *Evidence Based Obstet Gynecol* (in press).