Clinical Radiology 74 (2019) 384–389

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

An analysis of 11.3 million screening tests examining the association between needle biopsy rates and cancer detection rates in the English NHS Breast Cancer Screening Programme

R.G. Blanks [a,*], R. Given-Wilson [b], R. Alison [a], J. Jenkins [c], M.G. Wallis [d]

[a] *Cancer Epidemiology Unit, Nuffield Department of Population Health, Oxford University, Richard Doll Building, Roosevelt Drive, Oxford, OX3 7LF, UK*
[b] *Department of Radiology, St Georges University Hospital Foundation Trust, UK*
[c] *Breast Screening Programme, Public Health England, London, UK*
[d] *Cambridge Breast Unit, NIHR Cambridge Biomedical Research Centre, Cambridge University Hospitals NHS Trust, UK*

ARTICLE INFORMATION

Article history:
Received 7 August 2018
Accepted 23 January 2019

AIM: To examine the association between recall, needle biopsy, and cancer detection rates to inform the setting of target ranges to optimise the benefit to harm ratio of breast screening programmes.

MATERIALS AND METHODS: Annual screening programme information from 2009/10 to 2015/16 for the 80 screening units of the English National Health Service Breast Screening Programme (totalling 11.3 million screening tests) was obtained from annual (KC62) returns. Linear regression models were used to examine the association between needle biopsy rates and recall rates and non-linear regression models to examine the association between cancer detection rates and needle biopsy rates.

RESULTS: The models show and quantify the diminishing returns for prevalent screens with increasing biopsy rates. A biopsy rate increase from 10 to 20 per 1,000 increases the cancer detection rate by 2.13 per 1,000 with four extra biopsies per extra cancer detected. Increasing the biopsy rate from 40 to 50 per 1,000, increases the cancer detection rate by only 0.25 per 1,000, with 40 extra biopsies per extra cancer detected. Although diminishing returns are also seen at incident screens, screening is generally more efficient.

CONCLUSIONS: Increasing needle biopsy rates leads to rapidly diminishing returns in cancer detection and a marked increase in non-malignant/benign needle biopsies. Much of the harms associated with screening in terms of false-positive recall rates and non-cancer biopsies occur at prevalent screens with much lower rates at incident screens. Needle biopsy rate targets should be considered together with recall rate targets to maximise benefit and minimise harm.

* Guarantor and correspondent: R. G. Blanks, Cancer Epidemiology Unit, Nuffield Department of Population Health, Oxford University, Richard Doll Building, Roosevelt Drive, Oxford, OX3 7LF, UK. Tel.: +44 0 1865 289663.
E-mail address: roger.blanks@ndph.ox.ac.uk (R.G. Blanks).

https://doi.org/10.1016/j.crad.2019.01.015
0009-9260/

Introduction

The English National Health Service Breast Screening Programme (NHSBSP) started in 1987 following the Forrest report.[1] The programme currently undertakes two-view mammography on women aged 50–70 years at intervals of 3 years. Following the AgeX trial, invitations now occur from age range 47–73 years as the trial involves randomising women aged 47–49 and 71–73 to receive an invitation to screening.[2]

There has been controversy for many years related to the optimum balance between benefits and harms in screening for breast cancer.[3] In the UK, invasive cancer detection rate targets are informed by the Swedish Two-County randomised controlled trial, which detected nearly all cancers as invasive and are age standardised.[4,5] In Europe, targets are set at three times the underlying incidence rate for first screen and 1.5 for subsequent screens.[6]

Targets for recall rates range from 2% in Holland[7] to a recommended upper threshold of 12% in the United States.[8] Europe and the NHSBSP set separate targets for recall rate. At prevalent (first) screens, <5% in Europe with a minimum standard of <7% (but <7% and <10% respectively in the NHSBSP). At incident (subsequent) screens the European target is <3% with a minimum standard <5% (<5% and <7% in the NHSBSP).[6,9]

Percutaneous needle biopsy has essentially replaced open surgical biopsy for the diagnosis of both palpable and impalpable abnormalities.[10] The NHSBSP has targets to reduce the rate of women referred for benign surgical biopsy, repeat operation rates, and targets to avoid early surveillance/follow-up[9]; however, no target is set for needle (cytology/core) biopsy rates. There has been some upward drift in needle biopsy rates over the years as units try to increase cancer detection and ensure diagnosis of malignancy prior to referral for surgery. In addition, there is marked variation in unit needle biopsy rates by a factor of three times at prevalent screens and a factor of about two times at incident screens.[11]

The aim of the present study was to examine the association between cancer detection and needle biopsy rates and provide information that could potentially enable the setting of evidence-based needle biopsy rate ranges.

Materials and methods

Information on cancer detection, recall, and needle biopsy rates were taken from the national Korner (KC62) returns sent to Public Health England (PHE) annually and published by NHS Digital.[10] Data in this paper are from 80 screening units over the 7 screening years 2009/10 to 2015/16. The KC62 datasets are based on women as the denominator and the needle biopsy rate is interpreted as the number of screened women who have at least one needle biopsy. There is no information on repeat or multiple biopsies in these returns. Similarly the KC62 returns record no information on multiple or bilateral cancers.

The models presented here, mostly look at the association between needle biopsy rates (from referral to cytology/core biopsy) and cancer detection rate where cancer detection rates are based on all cancers. In practice, 97% of all cancers are detected following needle biopsies, the other 3% of cancers are almost all detected via open biopsy. Alternative models can be produced looking at the association between cancers detected only from needle biopsy against needle biopsy rates or the association between any biopsy (including open biopsy) and cancer detection rate. The rationale for the choice of models is discussed later, although it should be noted that in practice the model results are very similar.

Linear models have been used to examine the association between biopsy rate and recall rate and non-linear (two parameter negative exponential) models have been used to examine the association between biopsy rates and cancer detection rates. All models are based on data from each of the 80 screening units and weighted by the number of women screened by each unit. Full details of the rationale behind the use of two parameter negative exponential models are given in Electronic Supplementary Material Appendix A.

The false-positive needle biopsy rate (referred to in this paper for clarity as the non-malignant/benign biopsy rate) is estimated from the non-linear models and defined as the number of biopsies minus the number of cancers (invasive and non-invasive cancers) from the model.

Results

Relationship between biopsy rates and recall rate using linear regression

Prevalent (first) screens (women aged 45–52 years) and incident (subsequent) screens (53–70 years)

The association between needle biopsy rate and recall rate over the range of observed values has been modelled using linear regression weighted by the number of women screened by each unit. The model is biopsy rate (per 1,000 women screened) = 3.44 × recall rate (%) + 6.74, $p<0.001$ with an Adj-R^2 of 50.3% and correlation coefficient of 0.71. Therefore, for every 1% increase in recall rate at prevalent screens the biopsy rate increases by 0.344%. The fitted regression model for incident screens is biopsy rate (per 1,000) = 2.67 × recall rate (%) + 5.41, $p<0.001$ with an Adj-R^2 of 52% and correlation coefficient of 0.73. Therefore, for every 1% increase in recall rate at incident screens the biopsy rate increases by 0.267%. Fig 1 shows scatterplots of the biopsy rate and recall rate for prevalent and incident screens, respectively, with the 95% confidence band. There are no major departures from the line of best fit although a few units have high or low biopsy rates for their recall rate.

The linear regression models are applicable over the range of observed recall rates, but not at lower or higher values. Note that the models predict impossible biopsy rate values of 6.74 per 1,000 and 5.41 per 1,000 for prevalent and incident screens when extrapolated back to 0% recall rates.

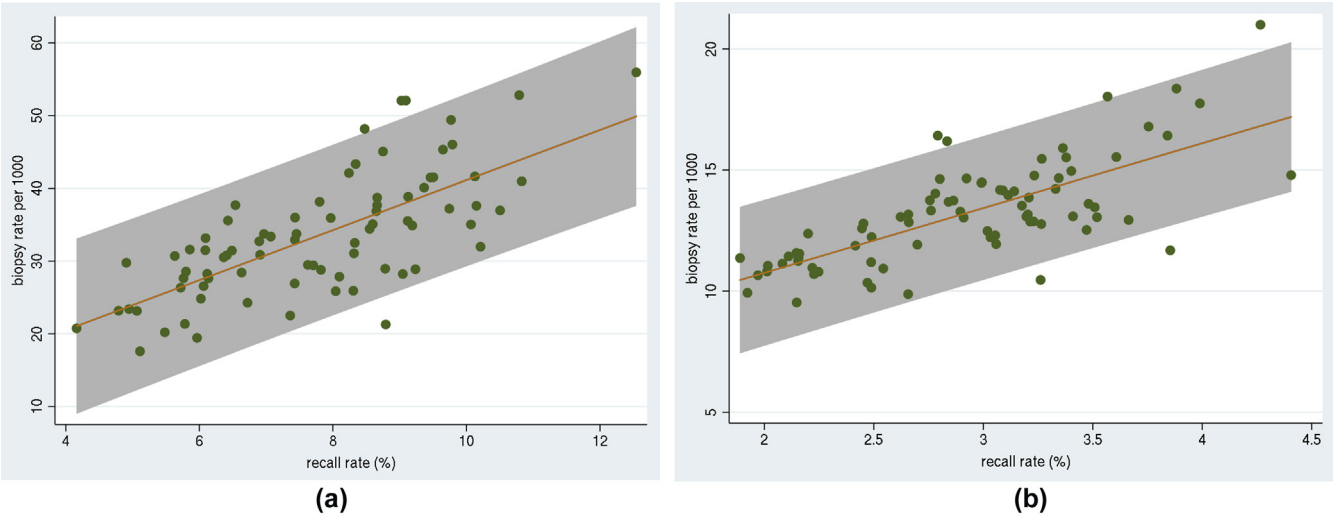

Figure 1 (a) Scatterplot of prevalent (first) screen needle biopsy rate for individual units versus recall rate with 95% confidence band. (b) Scatterplot of incident (subsequent) screen needle biopsy rate for individual units versus recall rate with 95% confidence band.

An alternative non-linear model (used later to examine the association between cancer detection rate and biopsy rates) which goes through the origin (0,0) estimates that biopsy rate (p=1,000) = $0.92(1-0.94^{\text{recall rate (\%)}})$. The model provides meaningful biopsy rates over any recall rate and over the range of observed recall rates the estimated needle biopsy rate using this model is almost the same as for the linear model. This model can be used when we consider the association between cancer detection rates and biopsy rates as non-linear and where the cancer detection rate tends to plateau at increasing biopsy rates.

In summary, the association between biopsy and recall rates over the observed range of data can be modelled with a linear relationship, with increasing numbers of recalls leading to increasing numbers of biopsies. The rate of biopsies and therefore associated assessment workload per woman screened tends to be much greater at prevalent screens.

Invasive cancers detected per women having biopsy at different recall rates

Table 1 and Fig 2 show the number of biopsies per invasive cancer detected at different recall rates. As the recall rate increases there is only a small increase in invasive cancer detection rate, which leads to an increase in the number of women undergoing biopsy per invasive cancer detected, and therefore, a diminishing return with increasing recall rate. As the recall rate is continually reduced, the number of women undergoing biopsy per cancer detected would be expected to decrease to nearly 1.0 as only features with a very high probability of being cancer (e.g., spiculate mass) would be recalled and biopsied. From linear regression models, it can be predicted that at prevalent screens, the number of women undergoing biopsy per invasive cancer detected increases from 4.6 per invasive cancer at a recall rate of 5% up to 8.1 per invasive cancer at 10%. At incident screens, a similar pattern of increasing numbers of biopsies per invasive cancer detected is seen.

Cancer detection rates and biopsy rates and non-linear models

Table 2 shows the results from the non-linear models for the predicted invasive cancer detection rate and the non/micro-invasive cancer detection rate per woman at different levels of biopsy rate for prevalent and incident screens. The non-malignant/benign biopsy rate per 1,000 is simply the rate of women undergoing biopsy minus the cancer detection rate. These models examining the association between cancer detection and biopsy rates are non-linear and if extrapolated go through the origin (0, 0), where no biopsies are associated with no cancers detected. The fitted models (see Electronic Supplementary Material

Table 1
Number of invasive cancers detected per women having biopsy at different recall rates for prevalent (first) and incident (subsequent) screens.

	Recall group	Recall rate (units)	Mean recall rate (%)	Invasive cancers (rate per 1,000)	Women biopsied	No. of biopsies to detect one invasive cancer (95%CI)
Prevalent	Low	<6.5 (24)	5.72	3,550 (5.12)	18,973	5.34 (5.19–5.51)
	Medium	6.5–8.4 (26)	7.47	3,449 (5.13)	21,563	6.25 (6.06–6.45)
	High	8.5+ (30)	9.47	4,866 (5.24)	37,075	7.62 (7.42–7.82)
Incident	Low	<2.7 (29)	2.30	18,899 (5.98)	36,032	1.91 (1.89–1.93)
	Medium	2.7–3.24 (28)	3.01	20,433 (6.26)	45,157	2.21 (2.19–2.23)
	High	3.25+ (23)	3.60	16,492 (6.48)	38,076	2.31 (2.28–2.34)

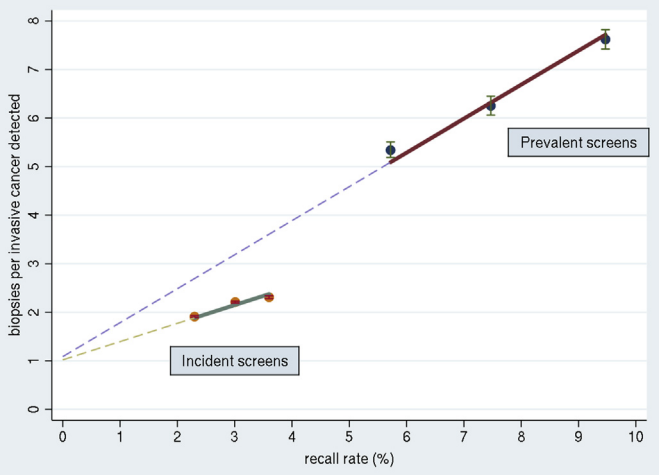

Figure 2 Number of women biopsied per invasive cancer detected by recall rate at prevalent (first) and incident (subsequent) screens.

Appendix B) predict at a prevalent screen biopsy rate of 10 per 1,000 a cancer detection rate of 4.18 per 1,000, and a non-malignant/benign biopsy rate of 5.72 per 1,000 (57% of biopsies are non-malignant/benign). This rises to 86% of biopsies being non-malignant/benign at a biopsy rate of 60 per 1,000. The models also suggest that at a prevalent screen biopsy rate of more than 40 per 1,000 there is little increase in invasive cancer detection rate. The incremental gains at this point become very small. Increasing biopsy rates from 40 to 50 per 1,000 (an increase of 10 per 1,000) only increases the cancer detection rate by 0.25 per 1,000, which is 40 non-malignant/benign biopsies per additional cancer detected. In contrast, at lower biopsy rates, increasing biopsy rates from 10 per 1,000 to 20 per 1,000 detects 2.13 per 1,000 extra cancers, which is just under four biopsies per extra cancer.

Modelled non-malignant/benign biopsy rates for prevalent screens and incident screens are shown in Fig 3a and 3b, respectively. Fig 3a shows how the non-malignant/benign biopsy rate at prevalent screens increases rapidly from about five per 1,000, in contrast to incident screens (Fig 3b) where the non-malignant/benign biopsy rate does not rise above the cancer detection rate until about 20 per 1,000.

Discussion

This study has demonstrated the relationship between cancer detection, recall for assessment, and needle biopsy rates. As previously shown, an increased biopsy rate was associated with a decreased proportion of biopsies demonstrating malignancy.[12] A major strength of the present study is the large and comprehensive dataset from the English national programme covering 11.3 million screening tests over 7 years providing robust information. Although the NHSBSP has targets for recall rates, there has been marked variability in rates between units providing a large observational study on the effects of differing recall rates and biopsy rates.

The balance between benefit and harm is related to concepts of optimality and value in healthcare.[13] As the healthcare intervention (in this case screening recall rate and biopsy rate) is increased, there is an initial rapid increase in benefit (cancer detection). Subsequently, there is a point where near maximal benefit has been reached and as further intervention is made the harms related to that intervention rise with very little increase in benefit. The point of optimality is that where the benefit is near maximal and further increased intervention will mostly increase harm rather than benefit.

Much of the harm associated with screening in the English NHSBSP occurs at prevalent screens. In contrast screening is much more effective at incident screens. This is largely because of higher age associated with increasing sensitivity of mammography and the availability of previous images.[14] Breast screening programmes need to balance the detection of cancers against too many non-malignant/benign biopsies, which cause harm to individual women and societal costs.[15] Fig 3 shows that this is particularly critical for prevalent (first) screens, where the non-malignant/benign biopsy rates rapidly increase beyond a

Table 2
Modelled prevalent (first) and incident (subsequent) screen cancer detection rates by needle biopsy rate using models with results based on all cancers.

Biopsy rate per 1,000 women	Cancer detection rate per 1,000	Invasive cancer detection rate per 1,000	Non/micro-invasive detection rate per 1,000	Non-malignant/benign biopsies per 1,000	Non-malignant/benign biopsies per cancer detected
Modelled prevalent (first) screen cancer detection rates					
10	4.18	3.18	1.00	5.72	1.34
20	6.31	4.56	1.75	13.69	2.17
30	7.30	5.16	2.14	22.70	3.11
40	7.80	5.42	2.38	32.20	4.13
50	8.05	5.53	2.52	41.95	5.21
60	8.18	5.58	2.60	51.82	6.33
Modelled incident (subsequent) screen cancer detection rates					
10	7.27	5.91	1.36	2.73	0.38
15	8.05	6.39	1.66	6.95	0.86
20	8.38	6.54	1.84	11.62	1.39
25	8.53	6.59	1.94	16.47	1.93
30	8.61	6.61	2.00	21.39	2.48

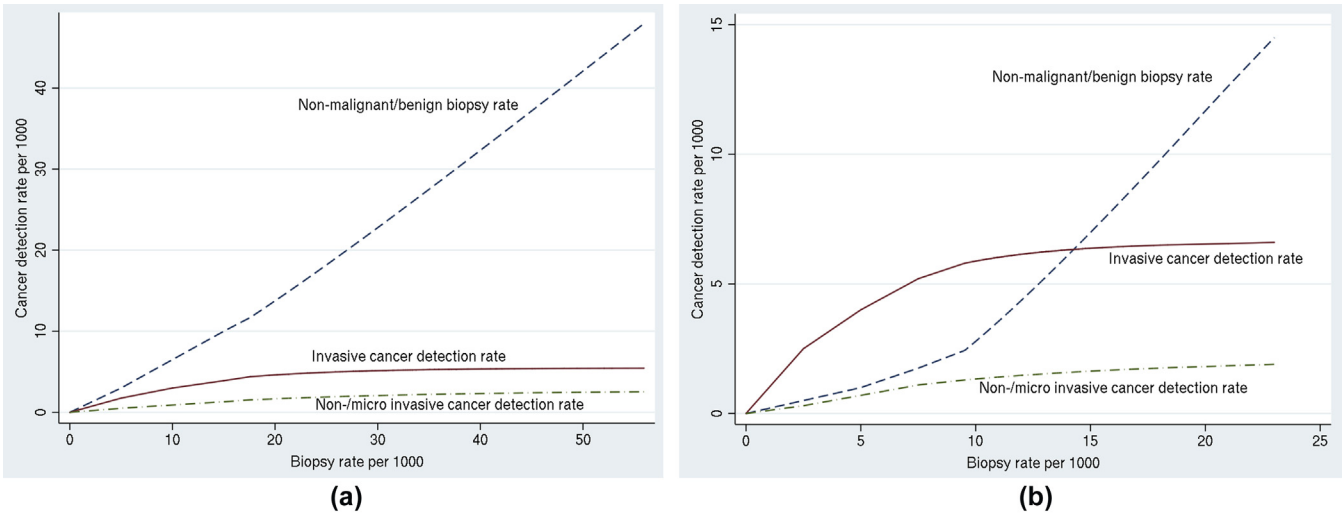

Figure 3 (a) Modelled association between cancer detection and biopsy rates at prevalent screens. (b) Modelled association between cancer detection and biopsy rates at incident screens.

biopsy rate about 20 per 1,000 and almost all biopsies are non-malignant/benign beyond a rate of about 30 per 1,000.

Increased recall and biopsy rates can be driven by a number of factors. These include anxiety about missed cancers and the perception of blame being attributed to radiologists with low cancer detection as well as the legal focus on Duty of Candour.[16] Targets are set for the NHSBSP designed to maximise cancer detection and the rate of non-operative diagnosis as well as minimise the use of follow-up or surgical excision for benign lesions instead of needle biopsy[9] These encourage high needle biopsy rates.

Recalls lead to anxiety, and as shown, the rate of recall is also directly related to the rate of biopsy, an invasive procedure with associated risk of complications. Assessment and biopsy are time and resource consuming. Costs of recall for assessment have been calculated to be seven times those of a screening mammogram.[17] Negative publicity about over-diagnosis in the screening programme is rising and may reduce the acceptability of breast screening to women. Further, there has been a reduction in uptake of screening in all four countries of the UK over the last 10 years.[10]

Although the present dataset is large, there are some limitations due to the nature of the data collected. The number of biopsies per woman or repeat visits for biopsies is not included. This may be significant for women with non-malignant/benign biopsies as repeat biopsy is commonly used to manage women with biopsies that are either inadequate initially or show indeterminate (B3) lesions.[18] The models shown in this paper are of needle biopsy rates versus the cancer detection rate. About 3% of cancers are detected from open biopsy rather than needle biopsy, and these are included in the cancer detection rate even though they are not technically from needle biopsies. The results are similar however the data are analysed. To avoid the complexity of using too many models or more complex models that treat the small number of cancers from open biopsy differently, whilst still including all screen detected cancers, one association was used to develop the methods. This is only justified because the proportion of cancers from open biopsy is very small. A further limitation of the study is the lack of detailed information on size, grade, and nodal status of invasive cancers detected at different biopsy rates.

Any changes to the national programme by setting biopsy rate targets would require careful monitoring to ensure no detrimental effects to the screening programme were occurring. Particular attention needs to be taken to ensure any reduction in biopsy rates does not affect detection rates of small higher-grade invasive cancers. By setting recall rate targets, biopsy rates may be affected and vice-versa. It is therefore probably not advisable to introduce both recall rate and biopsy rates targets or changes in targets at the same time. Any changes to either recall rates or biopsy rates would require close monitoring.

In summary, over the last quarter of a century or more the English NHSBSP has been more focussed on improving sensitivity than specificity. Recall rates higher than targets have tended to be tolerated in preference to low invasive cancer detection rates. This study demonstrates that high recall rates are associated with high biopsy rates and high non-malignant/benign biopsy rates, which contribute to the harms of screening.

Conflicts of interest

The authors declare no conflict of interest

Acknowledgements

R.G.B. and R.A. are funded by Public Health England. The work arose from initial discussions at the Clinical Advisory Group for NHSBSP assessment work (members R.G.B., Claire Borrelli, Sue Cohen, Alison Duncan, R.G.-W., J.J., Olive Kearins, Sarah Pinder, Mark Sibbering, Nisha Sharma, Jim Steel, Anne Turnbull, M.G.W.).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crad.2019.01.015.

References

1. Forrest P. *Breast cancer screening: report to the health ministers of England Wales, Scotland and Northern Ireland*. London: HMSO; 1986.
2. NHS Breast Screening Programme. AgeX trial. Available at: http://www.agex.uk/. Accessed 1 Jan 2018.
3. Marmot MG, Altman DG, Cameron JA, *et al*. The independent UK Review Panel on Breast Cancer screening. The benefits and Harms of breast screening: an independent review. *Br J Cancer* 2013;**108**:2205–40.
4. Tabar L, Fagerberg G, Duffy S, *et al*. Update of the Swedish-Two County programme of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;**30**:187–210.
5. Blanks RG, Day NE, Moss SM. Monitoring the performance of breast screening programmes: use of indirect standardisation in evaluating the invasive cancer detection rate. *J Med Screen* 1996;**3**:79–81.
6. Health & Consumer Protection Directorate. In: N.Perry M, Broeders C, de Wolf, *et al*., editors. *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th edn. Luxembourg: Office for Official Publication of the European communities; 2006.
7. Van Luijt PA, Fracheboud J, Heijnsdijk EAM, *et al*. Nation-wide data on screening performance during the transition to digital mammography: observation in 6 million screens. *Eur J Cancer* 2013;**49**:3517–25.
8. Lehman CD, Arao RF, Sprague BL, *et al*. National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology* 2017;**283**(1):49–58.
9. Consolidated standards for NHS breast screening programme April 2017 public health England Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/589227/Breast_draft_standards_V1.7.pdf. Accessed 20/7/2017.
10. Sanderink WBG. Mann RM advances in breast intervention: where are we now and where should we be? *Clin Radiol* 2018;**73**:724–34 https://doi.org/10.1016/j.crad.2017.10.018.
11. NHS Digital, Breast Screening Programme, England – 2015–16 Available at: https://digital.nhs.uk/catalogue/PUB23376. Accessed 1 Jan 2018.
12. Gur D, Wallace LP, Klym AH, *et al*. Trends in recall, biopsy and positive biopsy rates for screening mammography in an academic practice. *Radiology* 2005 May;**235**(2):396–401.
13. Gray M. Reducing unwarranted variation to maximise the value of healthcare for populations. *BMJ* 2017;**356**:j437.
14. Heidinger O, HeidrichJ, WU Batzler, *et al*. Digital mammography screening in Germany: impact of age and histological subtype on program sensitivity. *Breast* 2015 Jun;**24**(3):191–6.
15. Chubak J, Boudreau D, Fishman P, *et al*. Cost of breast related care in the year following false positive screening mammograms. *Med Care* 2010 Sep;**48**(9):815–20.
16. Given-Wilson R. Commentary on duty of candour and cancer screening. *Br J Radiol* 2017 Jul 14;**91**(1090). 20170457 epub.
17. Rachid R,Madan J,Brennan A, et al. Option appraisal: modelling the effectiveness and cost effectiveness of screening policies for breast cancer in elderly women in England and Wales. Report for NIHR 2010 Available at: https://www.sheffield.ac.uk/polopoly_fs/1.396555!/file/RachidFinal.pdf. Accessed 1 Jan 2018.
18. Public Health England. Clinical guidelines for breast cancer screening assessment, 4th edn. NHSBSP Publication No 49. PHE publications gateway number: 2016426. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/567600/Clinical_guidance_for_breast__cancer_screening__assessment_Nov_2016.pdf. Accessed 1 Jan 2018.