

Changing screening- What should we be looking for and what not recalling?

Rosalind Given-Wilson

St Georges Hospital

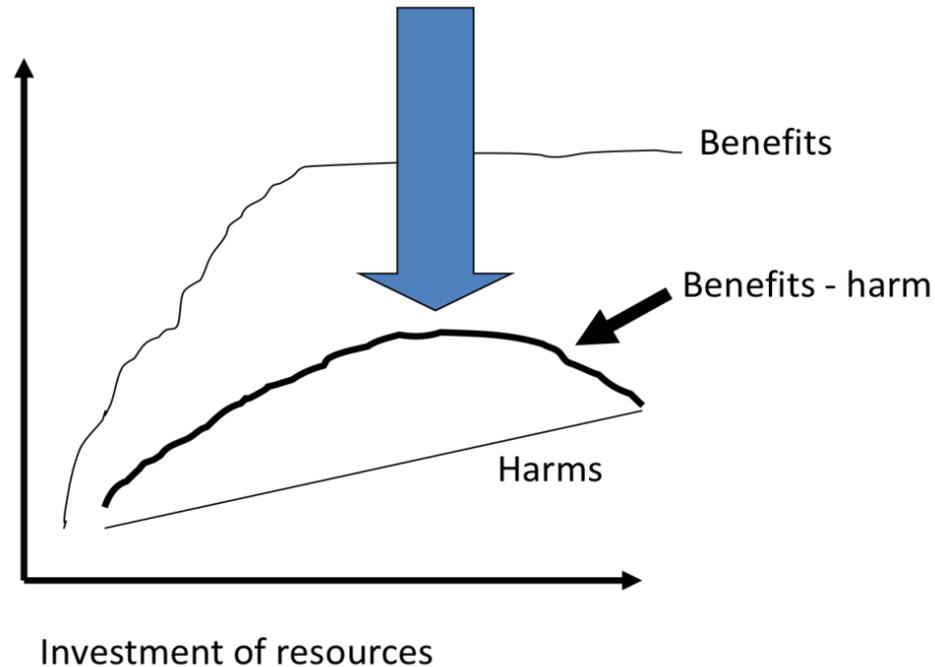
Cambridge International Breast

Cancer Conference 2019

- Harms and benefits of screening
- Why change? Recall
- Why change ? Grade 3 cancers
- How should we change? Recall
- How should we change? Assessment

Sir Muir Gray : Value

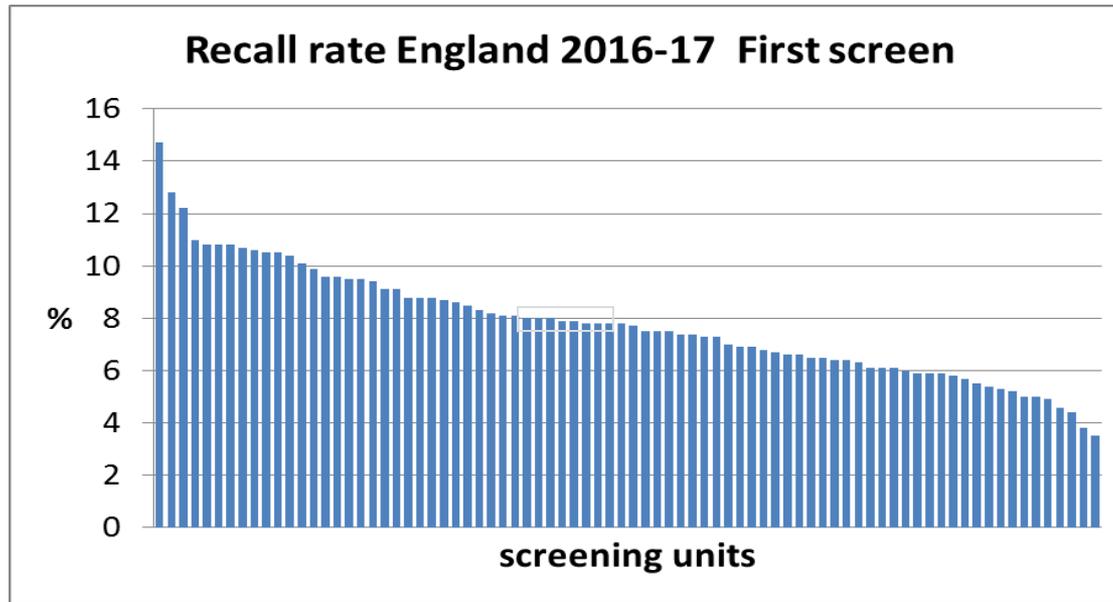
After a certain level of investment the health gain may start to decline:
The point of optimality



- Harms and benefits of screening
- **Why change? Recall**
- Why change ? Grade 3 cancers
- How should we change? Recall
- How should we change? Assessment

Screening challenges : Variation

Recall: Despite targets for upper limits of recall there is marked variation in recall rates within the English National Health Service breast screening programme

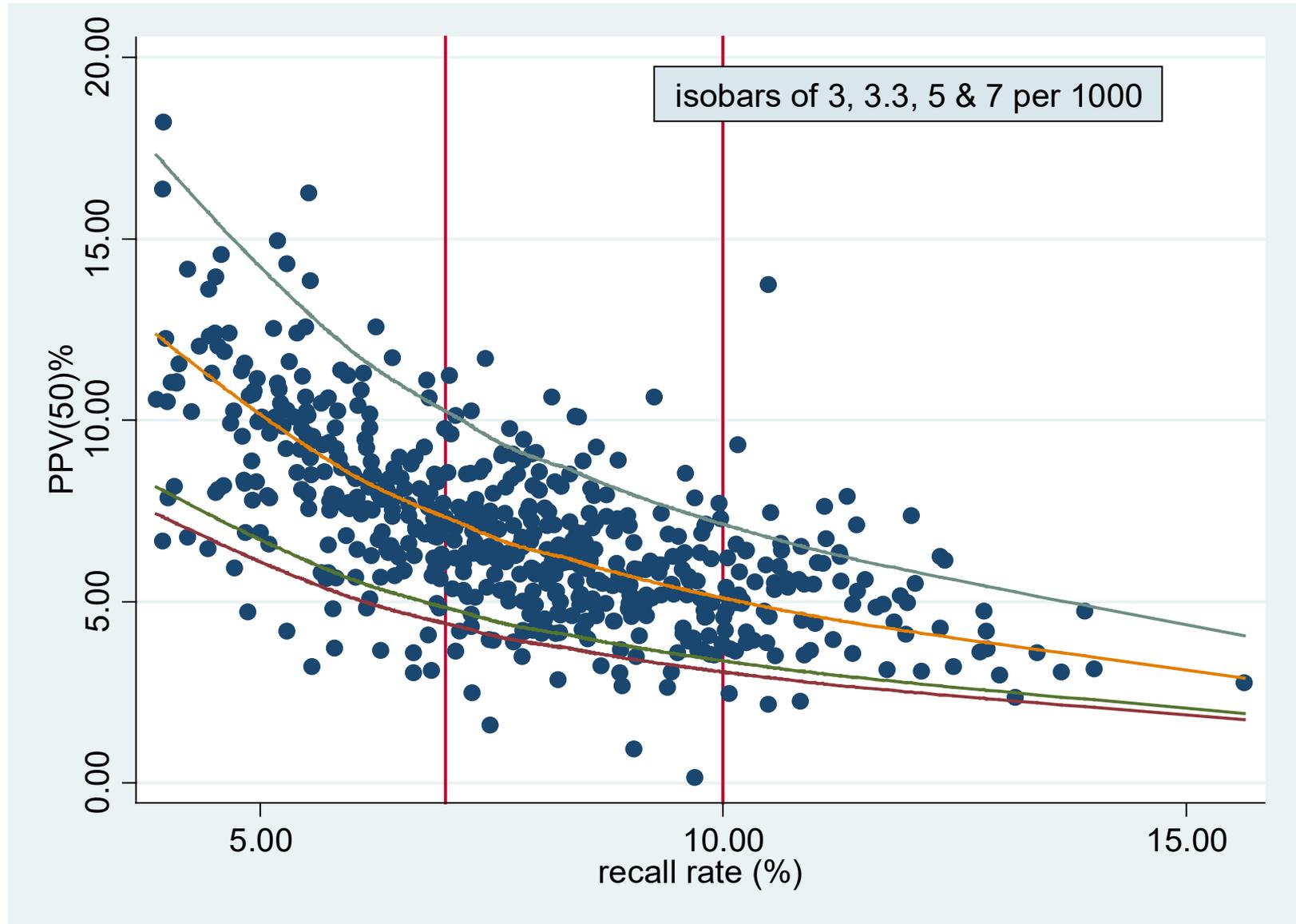


<https://digital.nhs.uk>

Biopsy rates: Marked variation in unit needle biopsy rates by a factor of 3 times at prevalent screens and a factor of about 2 times at incident screens.

Does this variation change cancer detection?

PPV-referral diagram by unit for 2009/10 to 2015/16 by individual years for invasive cancers prevalent screen



Recall rate vs cancer detection: What is the optimum level?

Not too low:

- Low recall is associated low cancer detection (sensitivity)
 - 31 centres 215,665 exams (range 106-47,338)
- *Yankaskas AJR 2001;177:543*
- Low recall is associated with higher interval cancers
 - 5 million screens (UK) 2005 to 2008
- *Burnside ES Radiology 2018; 228:47-54*

Not too High:

- High recall increases harms of anxiety, assessment workload, and benign biopsies, it may contribute to overdiagnosis
- BUT Does high recall bring benefits in detection of important cancers?

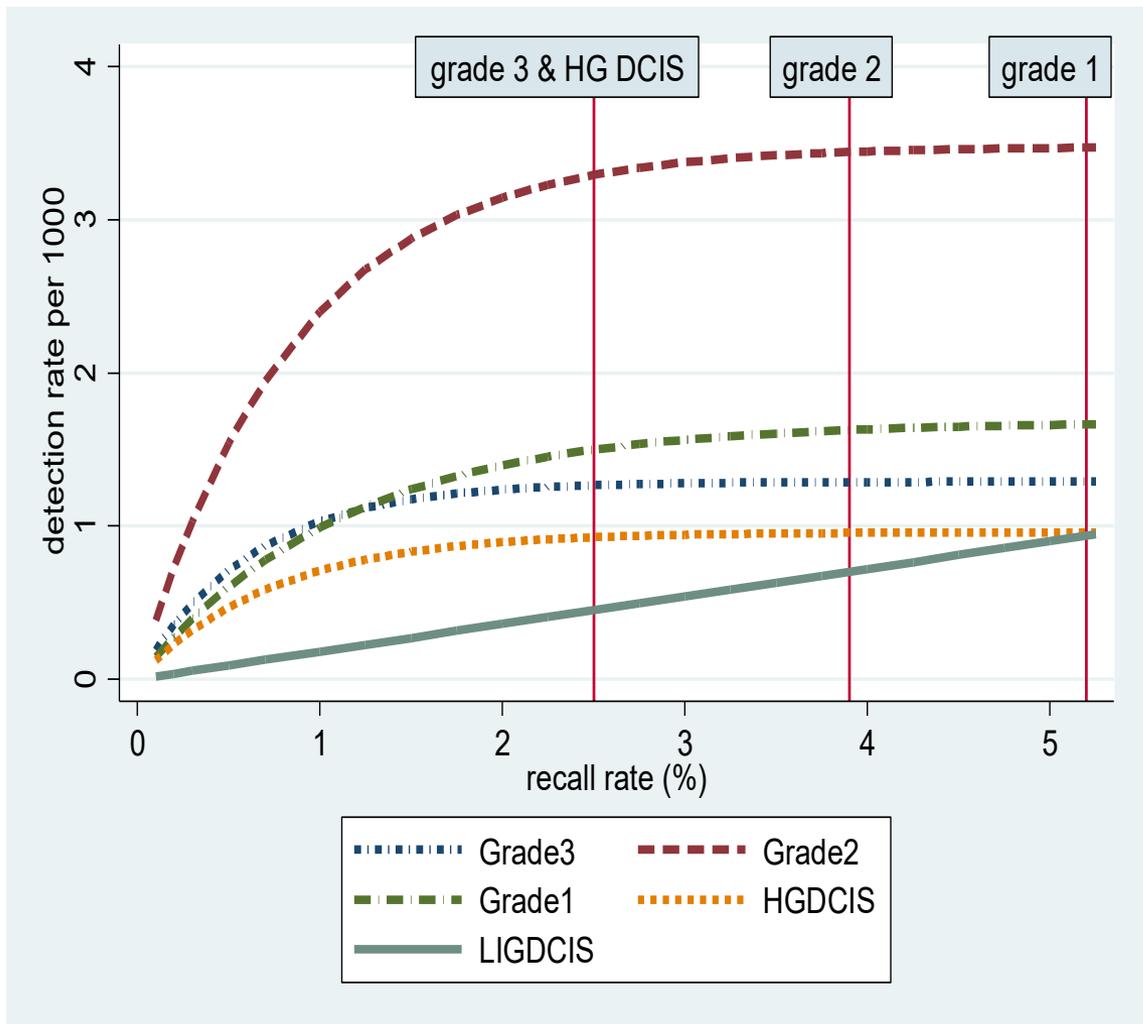
We find the higher grade more important cancers within the lower recall rate ranges, BUT go too low and you miss them

Modelled English data for incident (re) screens by cancer grade indicating P99 recall rate values by grade

P99 Grade 3 & HGDCIS at 2.5%

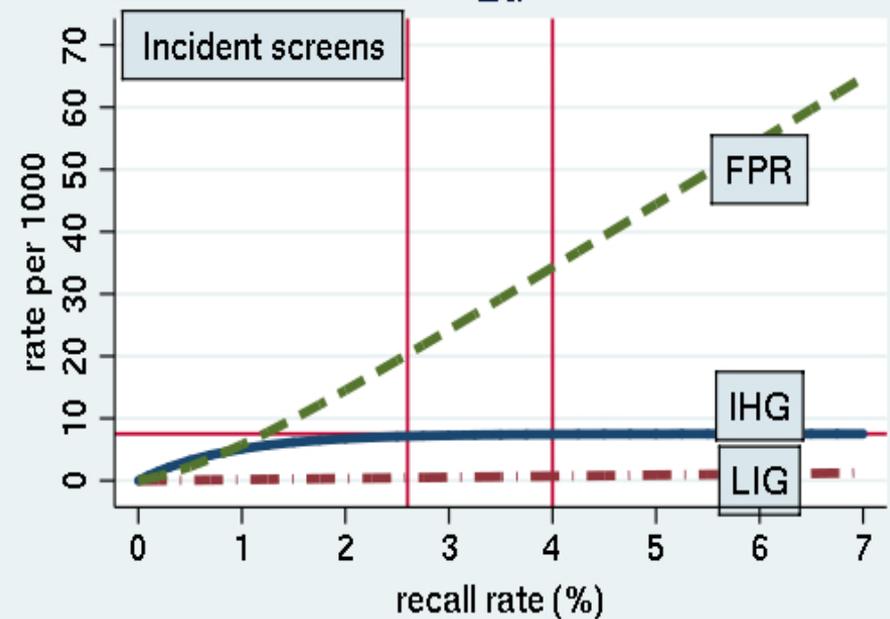
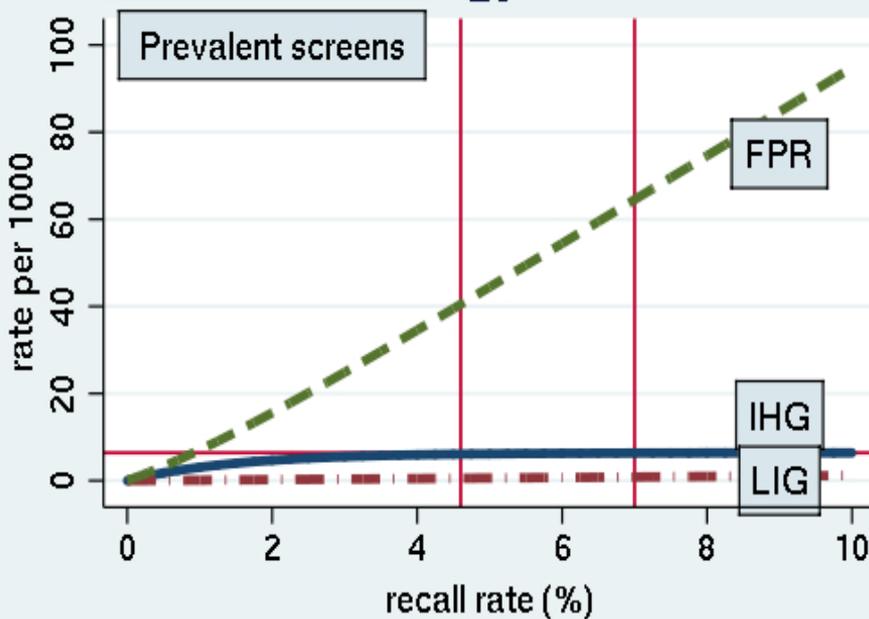
P99 Grade 2 at 3.9%

P99 Grade 1 at 5.2%.



There is a range of recall rates which allow us to predict detection of 95-99% of screen detectable cancers

Modelled false positive recall (FPR), IHG and LIG per 1000 for prevalent and incident screens with P95 and P99 recall rate values



P99 : Rate at which 99% of detectable cancers are found

P95 : Rate at which 95% of detectable cancers are found

Conclusion : Recall rates: Sweet spot

- Our model predicts that there is an optimum range for recall, not too low and not too high that optimises detection of life threatening cancers, whilst minimising harm of false positive recall and benign biopsy and over diagnosis.
- For all cancers apart from L/IG DCIS there is a point above which almost all additional recalls will be false positive
- In the UK recalls should be
 - Prevalent 4.6% to 7%
 - Incident 2.6% to 4%

European Radiology
<https://doi.org/10.1007/s00330-018-5957-2>

BREAST

An analysis of 11.3 million screening tests examining the association between recall and cancer detection rates in the English NHS breast cancer screening programme

R. G. Blanks¹ • R. M. Given-Wilson² • S. L. Cohen³ • J. Patnick¹ • R. J. Alison¹ • M. G. Wallis⁴

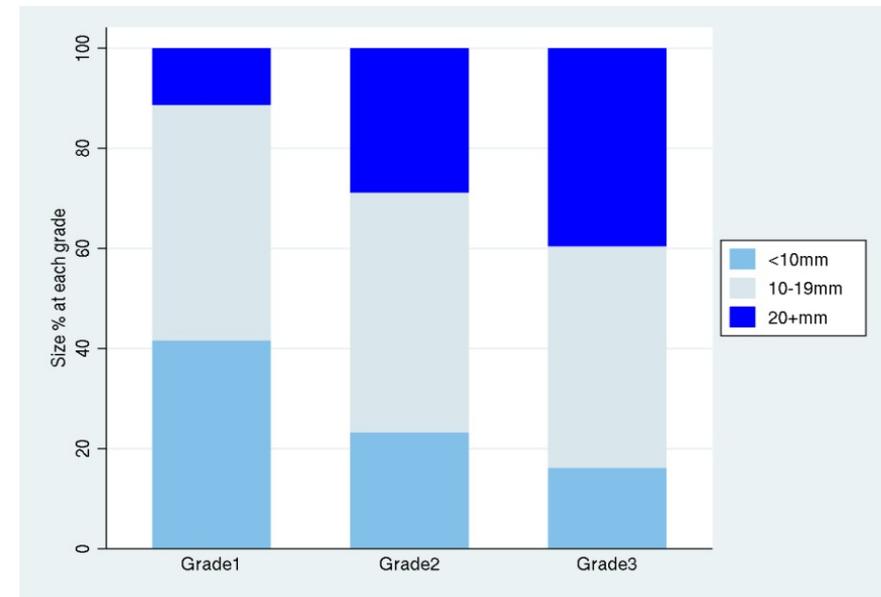
- Harms and benefits of screening
- Why change? Recall
- **Why change ? Grade 3 cancers**
- How should we change? Recall
- How should we change? Assessment

Importance of Grade 3 cancer

- Swedish Two Counties study : In the invited group there was a major reduction in mortality from grade 3 tumours (RR 0.65; 95% CI, 0.53–0.80; P < 0.001),
 - more deaths prevented from grade 3 tumours (n=95) than grade 1 and 2 tumours combined (n=48)
 - *Tabar L. Cancer Epidemiol Biomarkers Prev 2018 27(2) 2010: 154-157*
- Size still matters for Grade 3 cancers
 - *Lannin DR. N Engl J Med 2017; 376(23): 2286-91.*
 - *Saadatmand S. BMJ2015;351:h4901*

How are we doing with finding grade 3 cancer? ...Not very well

- NHSBSP 11.3 million screens over 7 years 2009 to 2016, 65,509 cancers with grade and size recorded
 - Grade 1 25.8%
 - Grade 2 53.8%
 - Grade 3 20.4%
- Size distribution grade 3 larger than lower grades
- Interval cancers 48% grade 3 *



Total screen detected invasive cancers percentage of cancers at various sizes within each grade

* Porter et al AJR 2007: 188: 676-683

- Harms and benefits of screening
- Why change? Recall
- Why change ? Grade 3 cancers
- **How should we change? Recall**
- How should we change? Assessment

What can we do ?

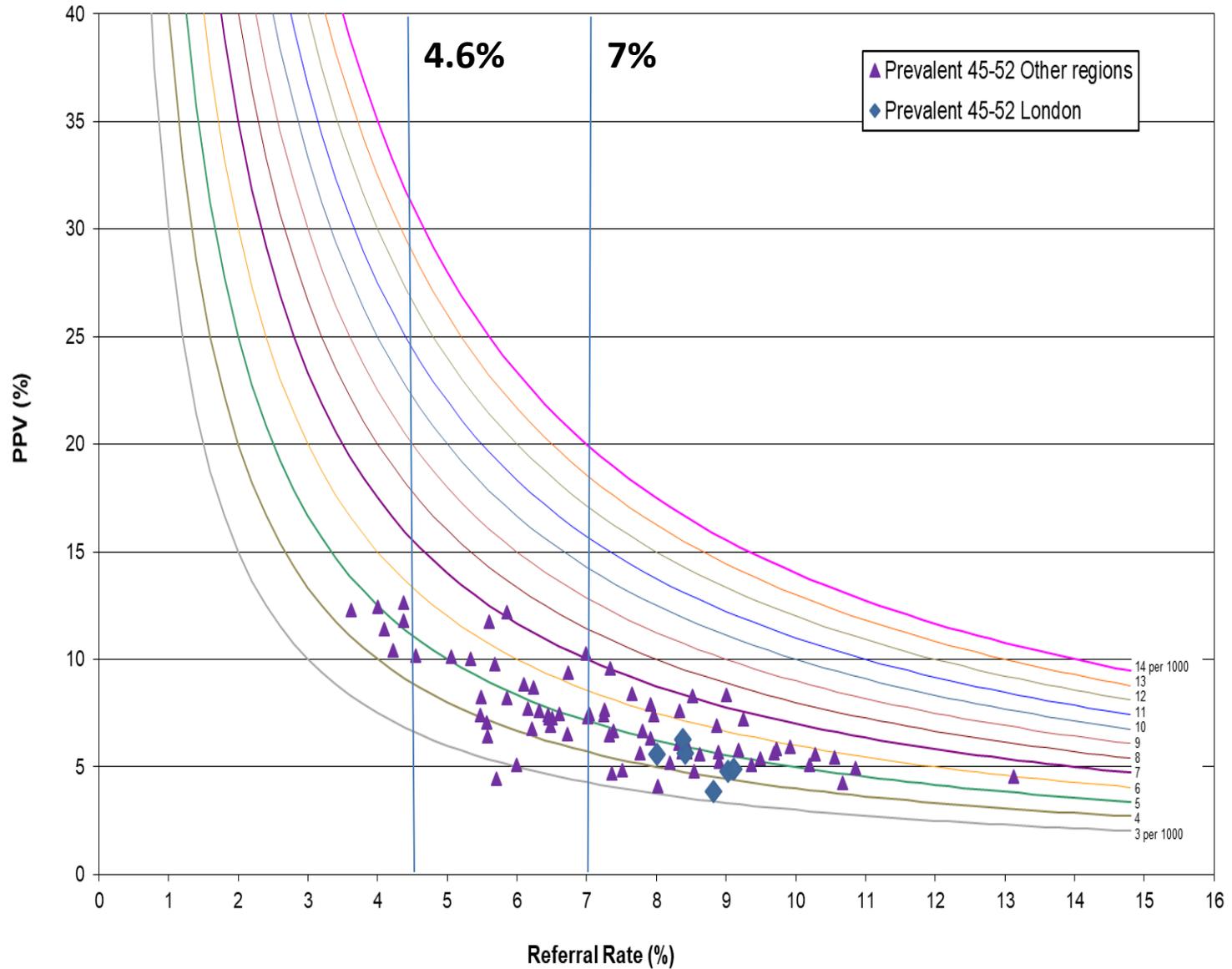
To optimise recall rates so we reduce unnecessary assessments and do not miss grade 3 cancer

- **Know your numbers- agree the target**
- The use of arbitration/ consensus
- Use background knowledge
- Know what grade 3 looks like

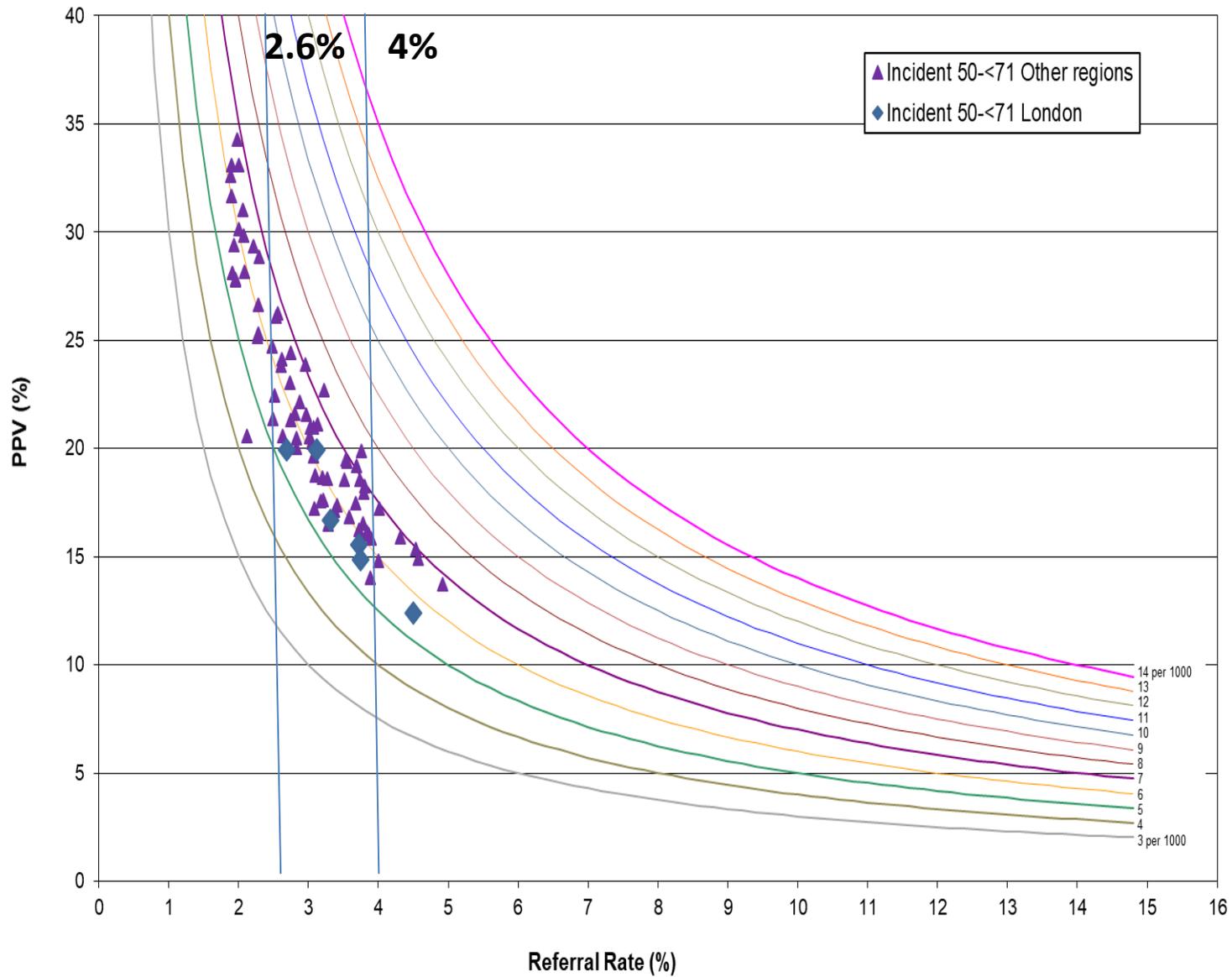
Know your numbers and agree what you want them to be

- The final outcome for the women is determined by the unit rate of recall
- Where do you get the numbers:
 - BSIS Individual reader data
 - SQAS team unit data

London and National PPV vs Referral Rate, 2015-2018, Prevalent



London and National PPV vs Referral Rate, 2015-2018, Incident



What can we do ?

To optimise recall rates so we reduce unnecessary assessments and do not miss grade 3 cancer

- Know your numbers- agree the target
- The use of arbitration/ consensus
- Use background knowledge
- Know what grade 3 looks like

Types of double reading and arbitration

- Double : Unilateral recall (if either reader suggests)
- Double : Arbitration or consensus for discrepancies only
- Double : Arbitration or consensus for all possible recalls

Double reading and arbitration or consensus

Reading type	Recall relative to single read	Sensitivity relative to single read
Double with unilateral recall	Increase recall 7.2-37%	<p style="text-align: center;">In UK Increase 10-14%</p> <p>CO-OPS 8.89%, 95% CI 8.22%-9.55%, p<0.001)</p>
Double with arbitration or consensus of discrepancies	Decrease recall 25-32% arbitration Decrease recall 39-45% consensus	

Pow, J Medical Imaging and Rad Oncology 2016; 60: 306–314
 Taylor-Phillips Radiology. 2018 Jun;287(3):749-757

Arbitration: Individual and team dynamics

'The review also demonstrates the importance of arbitration/consensus in double reading. The introduction of an arbitration step allows the readers to identify cases with minimal signs knowing that they will be reviewed and discussed by colleagues and only a proportion recalled. Eliciting extra assessments for difficult cases in this way allows a more efficient decision threshold to be maintained. A unit staffed by readers with different levels of experience should think carefully about which readers should work together and who should do the arbitration. '

Taylor P, Potts H. 2008 Computer aids and human second reading as interventions in screening mammography: Two systematic reviews to compare effects on cancer detection and recall rate EJC 44:798-807

What about arbitration / consensus for all possible recalls?

- Allows control of recall rates for unit by selected arbitrators
- Widens the pool of cases seen by the arbitrators for education and feedback

What can we do ?

To optimise recall rates so we reduce unnecessary assessments and do not miss grade 3 cancer

- Know your numbers- agree the target
- The use of arbitration/ consensus
- Use background knowledge
- Know what grade 3 looks like

Background knowledge

- Woman's age
- First or subsequent screen (new lesion or unknown, can prior films be obtained, what is prior history eg of biopsy)
- Any risk factors
- Multiple lesions and the rest of both breasts
- Eg new round single soft tissue lesion in a 68 year old carries much higher level of suspicion than a similar lesion in a 48 year old with no prior films

What can we do ?

To optimise recall rates so we reduce unnecessary assessments and do not miss grade 3 cancer

- Know your numbers- agree the target
- The use of arbitration/ consensus
- Use background knowledge
- Know what grade 3 looks like

High grade cancers: What they look like

- Typically masses, round or oval, with circumscribed, microlobulated, indistinct or obscured margins
- Fine linear branching calcification or casting calcification
- Rarely spiculate

Qualitative Radiogenomics: Association between Oncotype DX Test Recurrence Score and BI-RADS Mammographic and Breast MR Imaging Features Woodard, Ray et al Radiology 2018; 286:60–70

Mammographic Tumor Features Can Predict Long-Term Outcomes Reliably in Women with 1–14-mm Invasive Breast Carcinoma Tabar et al CANCER 2004 Volume 101: 8

Mammographic and clinicopathological features of triple-negative breast cancer GAO, ZHANG et al Br J Radiol 2014;87:20130496.

- Harms and benefits of screening
- Why change? Recall
- Why change ? Grade 3 cancers
- How should we change? Recall
- How should we change? Assessment

How should we change? Assessment so we minimize benign biopsy and do not miss cancer

- Use of US
- Use of DBT
- Teamwork

Ultrasound Biopsy : Guidance

- All solitary and/or new masses **recalled for assessment** confirmed as solid on ultrasound and that do not have the typical features of a hamartoma, lipoma, fat necrosis or normal lymph node should undergo needle core biopsy.
- Also biopsy: Atypical cysts, distortion, significant asymmetry not accounted for by normal glandular tissue , soft tissue associated with calcification which is not definitively benign, symptomatic abnormal areas (clinical recall)

NHS Breast Screening Programme

Clinical guidance for breast cancer screening assessment

NHSBSP publication 4 November 2016 PHE

What else may be safely left without biopsy?

- Findings which are NOT the abnormality which was recalled and are incidental, impalpable and **benign** appearing eg
 - dilated ducts without intraductal masses
 - small well circumscribed solid oval masses
 - In the setting of multiple cysts can accept that some will appear non simple
 - Multiple bilateral circumscribed benign appearing masses

Kim et al. Medicine (2016) 95:44

Hooley et al Radiology (2012) 265:1

How should we change? Assessment so we minimize benign biopsy and do not miss cancer

- Use of US
- Use of DBT
- Teamwork

DBT at assessment

- Addition of DBT to standard assessment imaging and omission of biopsy in DBT graded R1 or R2 (RCR Breast Group imaging classification) reduced rate of biopsy at assessment from 69% to 36% of cases with no reduction in cancers detected.
- Still biopsy if US is suspicious (U3-5)

How should we change? Assessment so we minimize benign biopsy and do not miss cancer

- Use of US
- Use of DBT
- Teamwork

Teamwork

- Confidence and consistency is increased when there is access when needed to a second opinion or to the MDT for biopsy (or not to biopsy) decisions

Conclusion

- Optimizing recall rates could reduce high rates of benign assessment and biopsy especially at the first screen and increase lower rates of recall and cancer detection at mainly incident screens
- Optimizing decision making at assessment and using DBT could increase accuracy and reduce benign biopsy
- Grade 3 cancer detection is important