Evidence for sequential effects in reading from a million women's mammograms in the UK

Professor Sian Taylor-Phillips

Professor of Population Health

University of Warwick

David Jenkinson, Matthew Wallis, Aileen Clarke

Contents

NHS
National Institute for Health Research

1. Vigilance hypothesis
2. The CO-OPS trial
3. CO-OPS Results
4. Longer Batch results
5. Summary of findings and implications for future work

1. Vigilance Hypothesis

NHS
National Institute for Health Research

s.taylor-phillips@warwick.ac.uk

2. The CO-OPS Trial

NHS National Institute for Health Research

Changing Case Order to Optimise Patterns of Performance in Screening

- Is there a vigilance decrement?
- Does reverse reading help?

Taylor-Phillips et al. JAMA. 2016;315(18):1956-1965.

NHS National Institute for Health Research

s.taylor-phillips@warwick.ac.uk

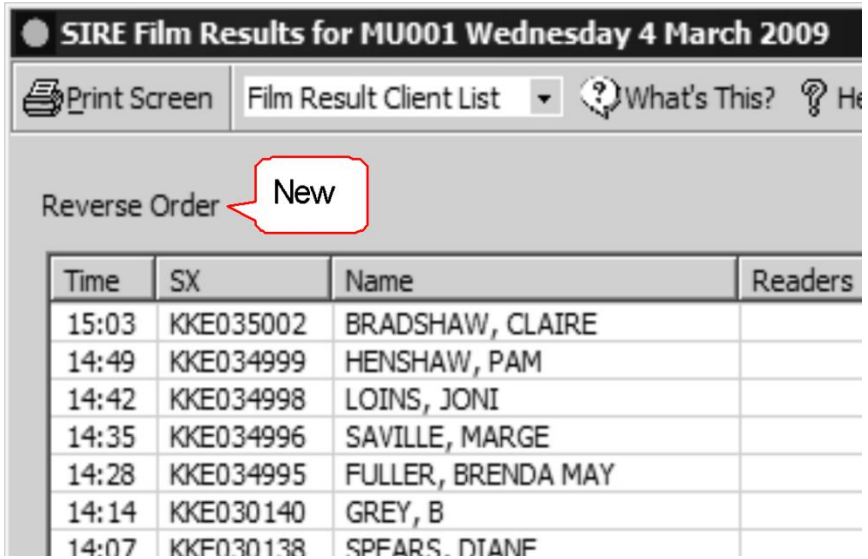
The CO-OPS Trial

NHS
National Institute for Health Research

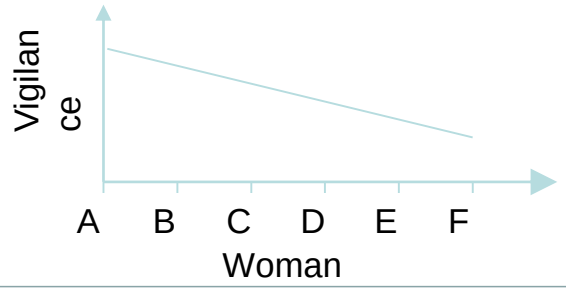
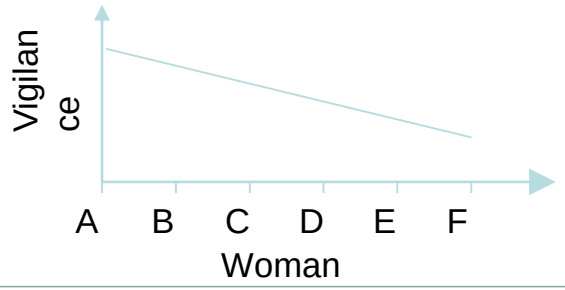
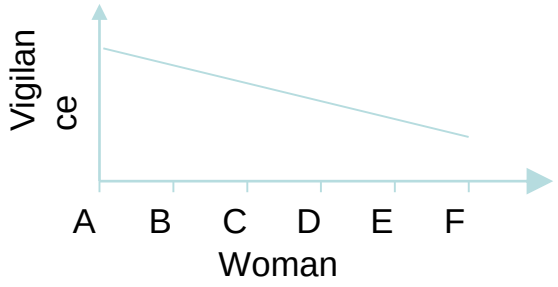
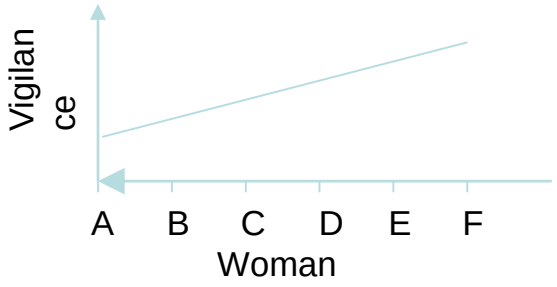
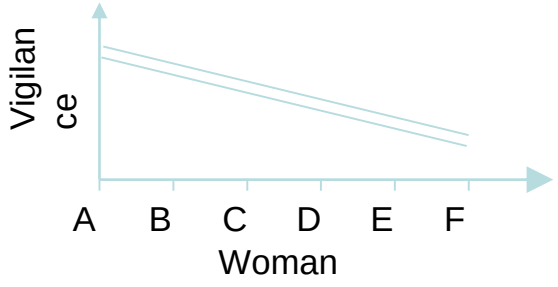
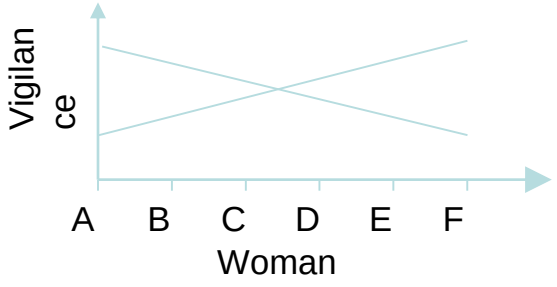
Pragmatic randomised controlled trial of a software intervention to change case order so that any vigilance decrement will occur for the first and second readers when examining different cases

Intervention ↓↑ or ↑↓

Control ↓↓ or ↑↑



s.taylor-phillips@warwick.ac.uk

Control arm		Intervention arm	
Reader 1: Forwards		Reader 1: Forwards	
Reader 2: Forwards		Reader 2: Backwards	
Both readers together		Both readers together	
Number of disagreements = Low Recall Rate = Low Cancer Detection Rate=Low		Number of disagreements = High Recall Rate = High Cancer Detection Rate = High	

The CO-OPS Trial

NHS National Institute for Health Research

Participation:

- 46 English Breast Screening Centres completed the trial
- One year of data collection

s.taylor-phillips@warwick.ac.uk

The CO-OPS Trial

NHS National Institute for Health Research

Primary Analysis

- Multi-level model of the predictors of cancer detection rate

Secondary Analyses

- Multi-level model of the predictors of recall rate and rate of disagreements
- Patterns of performance with time since a break

s.taylor-phillips@warwick.ac.uk

NHS
National Institute for Health Research

Randomized
(n=1,207,633)
in 37,724 batches

Intervention Arm

(n=603,528) in 18,797 batches

Received allocated intervention
(n=524,971, 87%)

Lost to follow-up 0.07%

Analysed
(n=596,642, 98.9%)

Control Arm

(n=604,105) in 18,927 batches

Received allocated intervention
(n=560,760, 93%)

Lost to follow-up 0.06%

Analysed
(n=597,505, 98.9%)

s.taylor-phillips@warwick.ac.uk

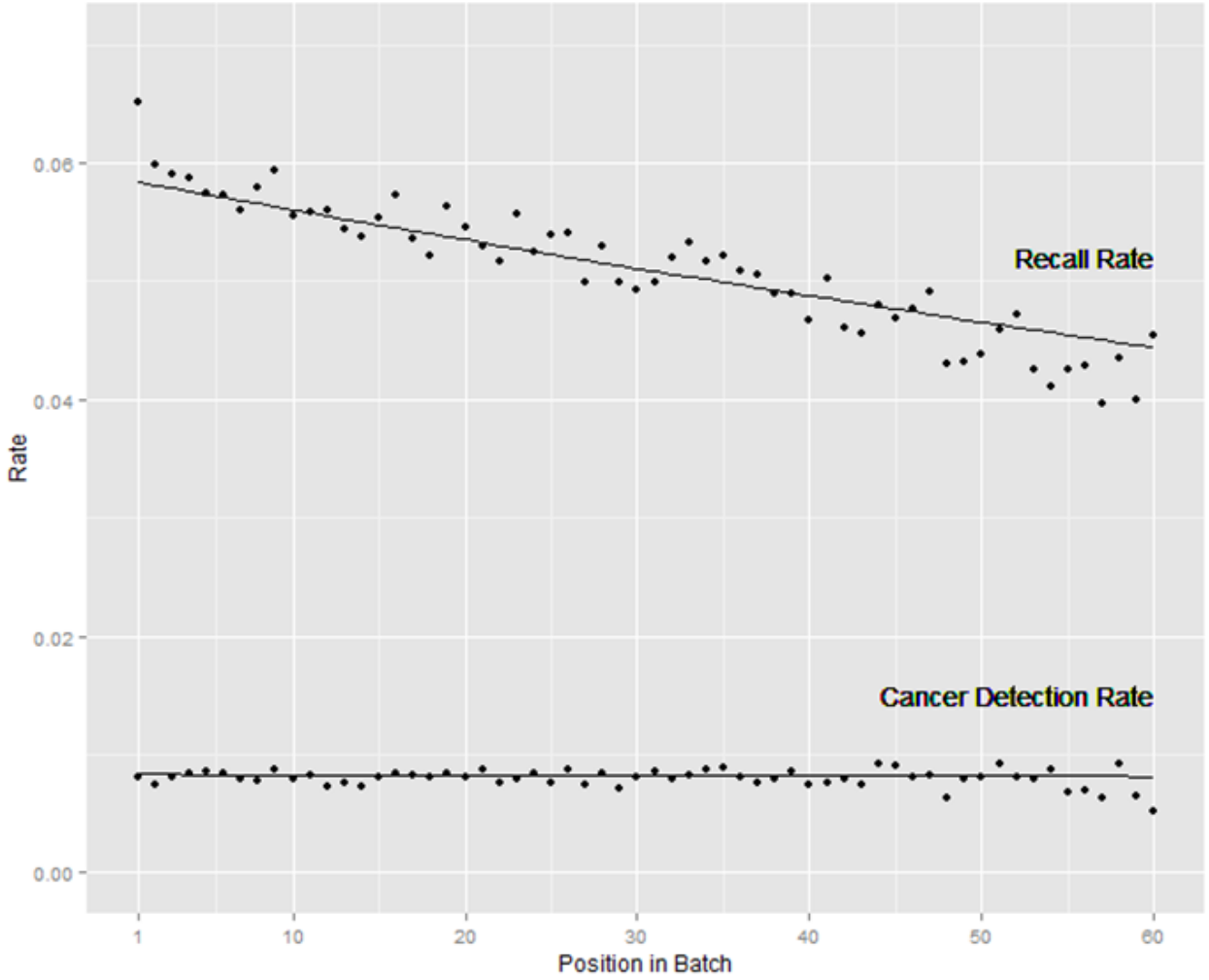
3. CO-OPS trial results

NHS National Institute for Health Research

	Intervention	Control
Cancer Detection Rate (95%CI)	0.89% (0.86% to 0.91%)	0.88% (0.85% to 0.90%)
Recall Rate (95%CI)	4.14% (4.09% to 4.19%)	4.17% (4.12% to 4.22%)
Rate of disagreements (95%CI)	3.43% (3.39% to 3.48%)	3.48% (3.43% to 3.53%)

s.taylor-phillips@warwick.ac.uk

Patterns of recall: Single reader



s.taylor-phillips@warwick.ac.uk

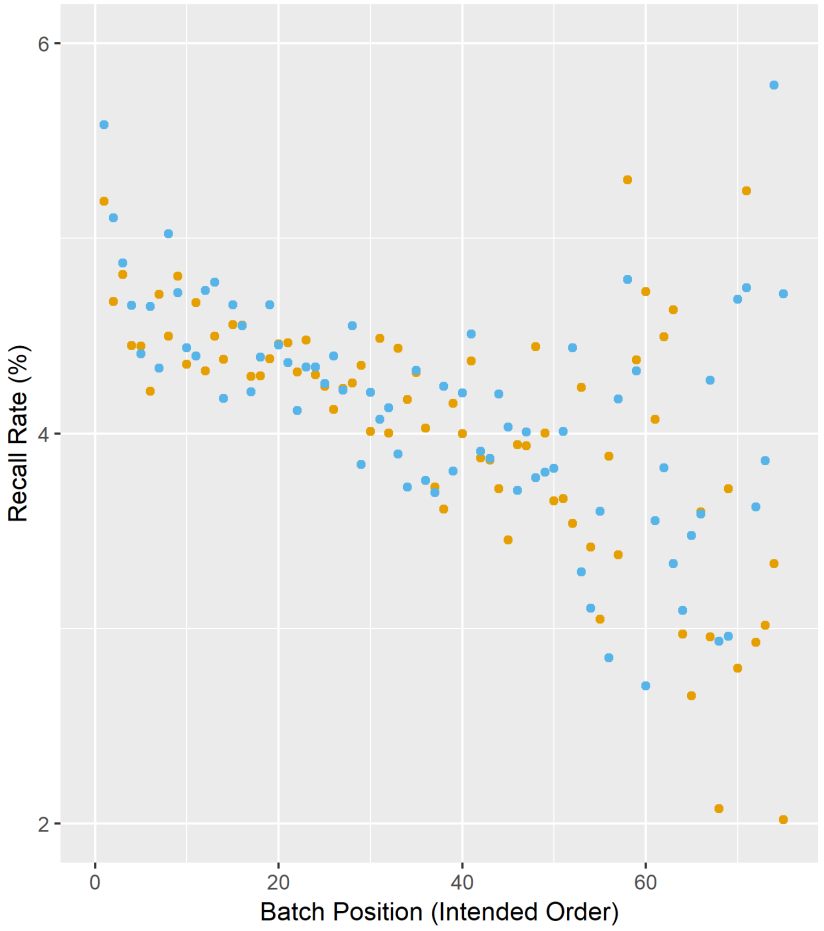
Is this pattern real?

NHS National Institute for Health Research

- Reading order isn't random
- One batch = one day on one machine (30-60)
- Order is alphabetised by GP practice
- Women can rearrange appointment time but rare

s.taylor-phillips@warwick.ac.uk

Patterns of recall: Single reader



Direction Forward Reverse

s.taylor-phillips@warwick.ac.uk

Why doesn't it change overall recall rate?

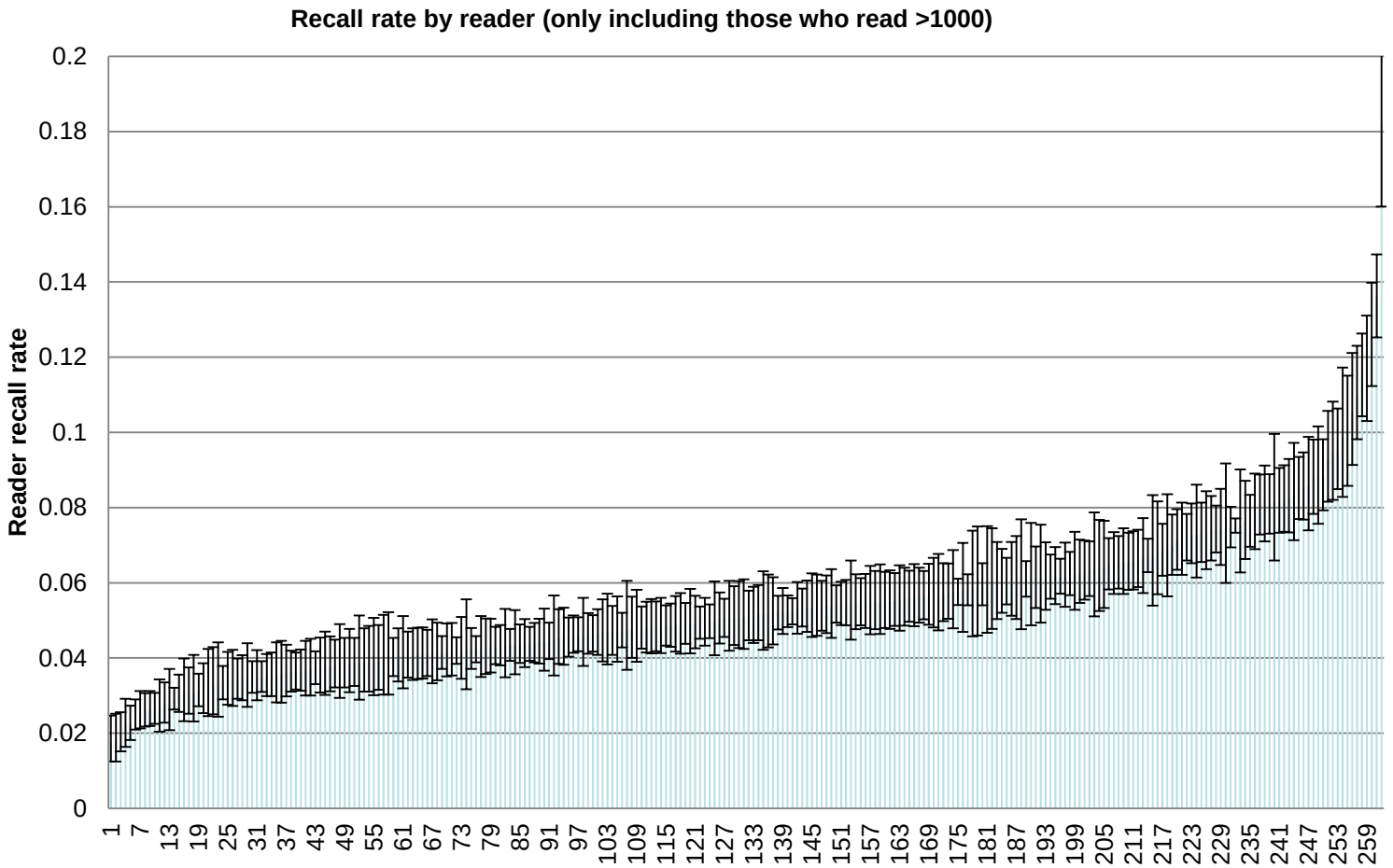
NHS National Institute for Health Research

- Film reader variability?
- Arbitration of discordant cases removing excess recalls at beginning?
- Changes to who is recalled but not number recalled?

s.taylor-phillips@warwick.ac.uk

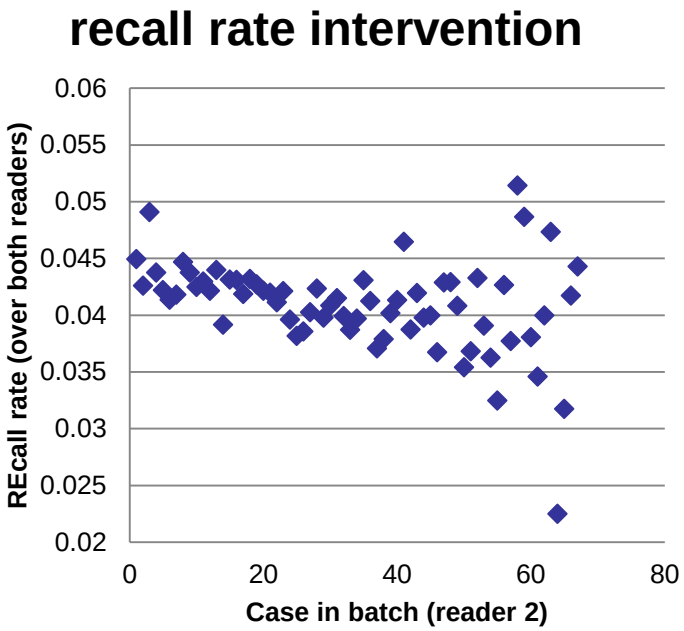
Film Reader Variability

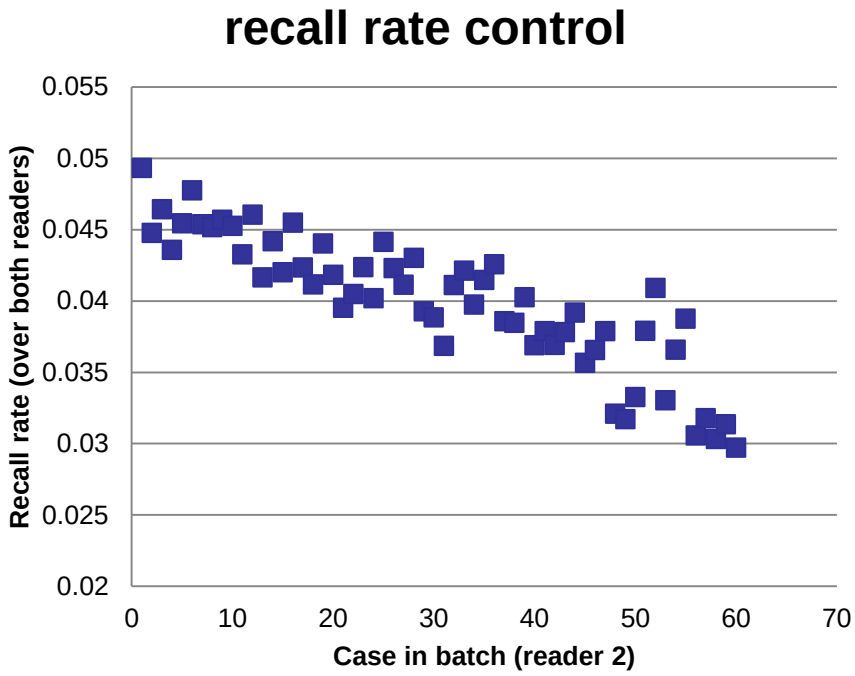
NHS National Institute for Health Research



s.taylor-phillips@warwick.ac.uk

Recall rate combining both readers

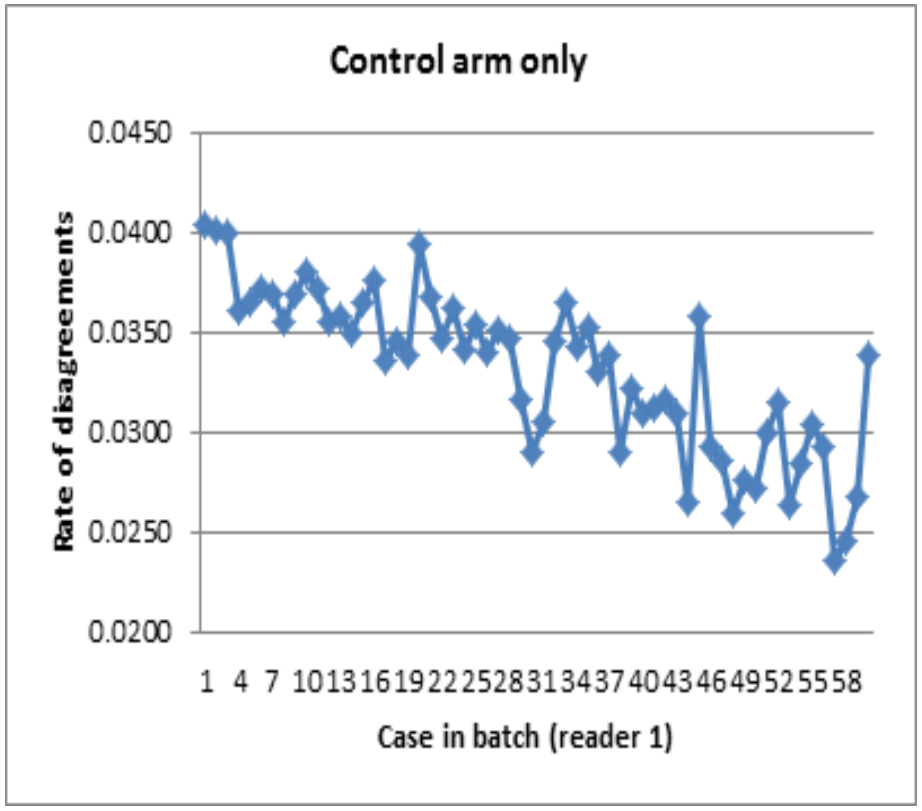




Arbitration of discordant cases removing excess recalls at beginning? No
Changes to who is recalled but not number recalled? Maybe

s.taylor-phillips@warwick.ac.uk

Disagreements



Excess of recall at beginning of the batch, in the intervention arm this is spread over different cases

s.taylor-phillips@warwick.ac.uk

4. Analysis of longer time on task

- What about sessions longer than 60 cases? UK film readers often examine several batches back to back
- Analysed patterns of performance over time for each reader
  - Time stamps for decision on each case
  - Break defined as 10, 20, 60,180, or 480 minutes without a decision
  - Excluded cases moved from intended order
  - Excluded first case after break

s.taylor-phillips@warwick.ac.uk

Follow-on analysis

NHS National Institute for Health Research

- Reader 1 analysed, because reader 2 not blinded
- Multi-level models, levels: woman, reader1, centre
- Adjusted for woman's age and whether she has previously attended screening
- Linear models with knots at 20 and 40 cases
- Distribution: logistic for recall and cancer detection, gamma for time taken

s.taylor-phillips@warwick.ac.uk

Follow-on analysis

NHS National Institute for Health Research

- Outcomes
  - Reader 1 recall yes/no
  - Reader 1 cancer detected yes/no
    - (Reader recalls and cancer detected in follow up biopsy)
  - Reader 1 time taken to examine the case
    - (Time stamp at end of case minus time stamp at end of previous case)
  - Sensitivity and specificity of reader 1
    - (reference standard is cancer detected at screening or symptomatically in 3 years following screening)

s.taylor-phillips@warwick.ac.uk

Summary of findings

NHS National Institute for Health Research

- There appear to be an excess of recalls at the beginning of a reading session
- The reverse reading intervention doesn't affect recall rate overall, but does appear to affect who is recalled.
- Reading speed appears to increase and recall rate appears to decrease with time on task
  - Good fit with evidence that batch reading improves specificity.
- Cautious interpretation of post-hoc analyses of large observational dataset

s.taylor-phillips@warwick.ac.uk

Next Steps in Research

NHS National Institute for Health Research

- What do we want to detect in breast cancer screening?
- Characteristics of cancers/pre-cancers detected.
- Maximise morbidity and mortality benefits, and minimise overdiagnosis.
- What are the long term outcomes after previous changes to breast screening?
  - Two readers
  - Reader test threshold (recall rate)
  - Age of eligibility

s.taylor-phillips@warwick.ac.uk

Further Information

NHS National Institute for Health Research

Email: s.taylor-phillips@warwick.ac.uk

Website:https://warwick.ac.uk/fac/sci/med/research/hscience/pet/screening/breast/

Twitter: @siantphillips2

Warwick Screening

This work was funded by an NIHR Career Development Fellowship for Professor Taylor-Phillips. The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR, or Department of Health.