Breast Screening Unit Size and performance on self-assessment (PERFORMS).

Hazel J. Scott and Alastair G. Gale, AVRC, University of Loughborough, UK

1. Abstract

Introduction

The UK Breast Screening Programme (UKBSP) is comprised of approximately 100 individual Breast Screening Units (BSU's) which vary in size (measured by number of women screened). Previous research using UKBSP real life data (Blanks, Bennett, Wallis and Moss, 2002), attributed differences in performance, related to BSU size, to smaller units' slightly lowered cancer detection rates and PPV scores (comparative to large/medium units). All BSU's on the UKBSP annually take part in the PERFORMS scheme as way of self-assessing their film reading skills. We looked at the performance of all film-readers who had completed the last PERFORMS round (SA08) by BSU size in order to explore any group differences mediated by unit capacity.

Methods

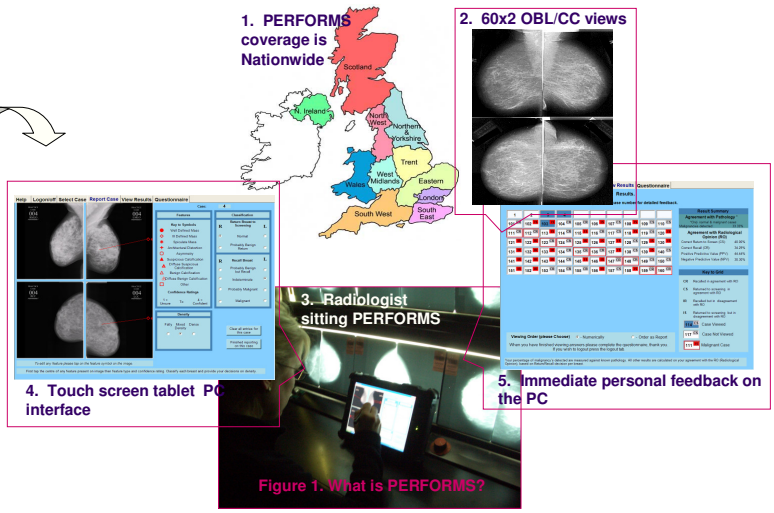
Each BSU's size was approximated by ranking each unit by number of readers who had completed the last PERFORMS round. Subsequently these Units were allocated into three main groups approximating their unit's size: Small =2-6 readers, 33% *(n=174)*, Medium = 7-9 readers, 30% *(n=162)* and Large=10 readers+ 36% *(n=188)*. Several performance measures were compared including Percentage Correct Recall and Malignancies Detected (measures of sensitivity), Percentage Correct Return to Screen (a measure of specificity), NPV and PPV scores and d' values.

Results

Analysis of variance (one-way) did not produce any significant findings (p=n.s) for any of the measures indicating equivocal performance. Descriptive statistics showed smaller units scored circa 1% below medium/large BSU's for Malignancies Detected and Correct Recall only.

Conclusion

Unlike real-life screening, smaller units perform at a similar level to all others on self-assessment.



2. What is PERFORMS?

1. Breast Screening units throughout the UK take part in PERFORMS.
2. Each film reader is sent 60 PERFORMS cases to report (with two views for each case OBLIQUE and CC).
3. Film readers report their decisions on each case using a touch screen PC with miniature images.
4. For each case feature type, classification and density decisions are recorded on the PC.
5. Confidential, immediate individual feedback is given automatically via the PC showing sensitivity/specificity measures.

3. Calculating Unit Size from PERFORMS data.

Over 600 medical specialists in the UK NHS Breast Screening Programme complete PERFORMS on a bi-annual basis, the objective of which is for film readers to report on a series of difficult breast screening mammograms and receive confidential feedback. Once all participants have taken part then anonymous regional and national data are accrued and fed back to each individual. This data is calculated from the National consensus on each case (as well as from pathology). Therefore, all data presented here represents the National Opinion (NO) from the most recently completed PERFORMS set.

Inclusion criteria for this study was the completion of the most recent PERFORMS set, and all those who completed only one half were removed from the analysis . Participants were fractionated into 3 approximately equal groups based on percentage of participants and unit size- see Figure 3.

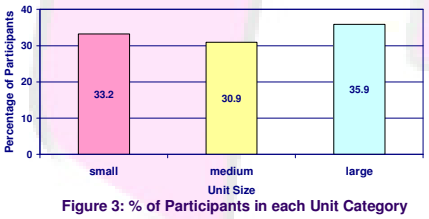


Figure 3: % of Participants in each Unit Category

Figure 3. shows the percentage of PERFORMS participants that formed the three group sizes according to number of film-readers recorded at that unit –as follows:-

- Small = 2-6 readers
- Medium = 7-9 readers
- Large = 10-19 readers

4. Attainment levels on PERFORMS by Unit Size

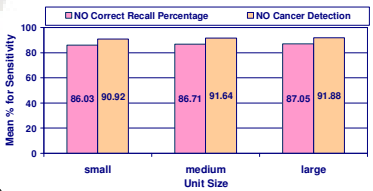


Figure 4. PERFORMS Sensitivity Measures by Unit Size

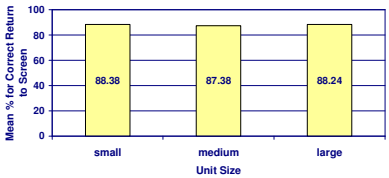


Figure 5. PERFORMS Specificity by Unit Size

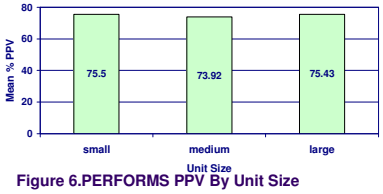


Figure 6.PERFORMS PPV By Unit Size

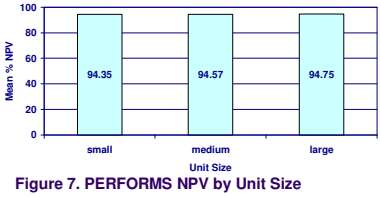


Figure 7. PERFORMS NPV by Unit Size

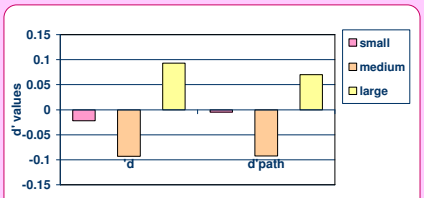


Figure 8. ROC performance measures for d' and d' for pathology by Unit Size

Figures 4-8 show the PERFORMS percentage by unit size for all PERFORMS measures on sensitivity, specificity, PPV, NPV and d'values. One-way ANOVA for each measure revealed non significant differences for unit size categories (p=n.s.). Therefore participants' performance is equivocal regardless of unit size. However a small descriptive trend for sensitivity measures and for ROC performance measures for both d' (on suspicion) and d' for pathology show a very slight advantage for large units who display higher percentages for these measures.

5. Unit reporting 'culture' in real life as recorded on PERFORMS questionnaires by Unit Size.

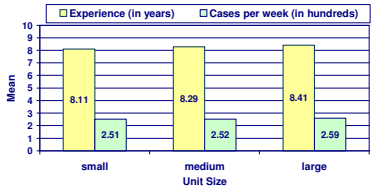


Figure 9. Reported Years of Screening Experience and Cases per Week by Unit Size

Figure 9. shows the reported years of screening experience and reported case volume, a one-way ANOVA showed no significant differences although large units have a slightly higher mean for both measures.

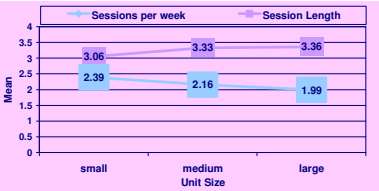


Figure 10. Reported Sessions per Week and Session Length by Unit Size

Key: Session Length 1 unit = 30 minutes (3 units=1hr 30minutes)
Session per Week 1 unit= 1 session per week

Figure 10. shows the reported reading sessions per week and the length of those sessions, a one-way ANOVA showed significant differences for session frequency (p<.01), post-hocs revealed that smaller units read significantly more sessions per week. Session length was approaching significance with large units showing slightly longer reading sessions.

6. Results Summary

- Unit size in this sample had little effect on performance on PERFORMS for any of the performance measures.
- Although when fractionated by unit size performance is equivocal, descriptive statistics show slightly higher mean scores for larger units for sensitivity and d' values.
- Reported real life behaviours show no differences for unit size for years of screening experience or case volume.
- Significant (and approaching) differences for reading sessions show small units reading more often and larger units reading for longer.

7. Discussion

Results seem to indicate that, unlike real life studies, unit size does not affect performance, however unit size in this sample was calculated from the number of readers in each unit and not from real life estimations of unit size i.e. calculated from the number of women actually screened. As number of readers in each unit does not always tally with unit size in terms of number of women screened (approximately 30% of units show different grouping) these results are a preliminary estimate only and should be viewed with caution. Further studies are warranted using real life indicators of unit size. These future studies may show very different results as well as showing a different unit culture for reporting cases.

Acknowledgements: This work is supported by the NHS Breast Screening Programme

avrc applied vision research centre

Loughborough University