

Non-consensual Double Reading in the Singapore Breast Screening Project: Benefits and Limitations

B H Kwek,*^{MBBS, FRCR}, T N Lau,**^{MBBS, FRCR, MRCP}, F C Ng,***^{MBBS, FRCR}, F Gao,****^{MSc}

Abstract

Introduction: The aims of this study were to assess the efficacy of non-consensual double reading (against single reading) in the Singapore Breast Screening Project (SBSP) and to compare the benefits (increase in cancer detection) and limitations (increase in recall and biopsy) with published data. **Materials and Methods:** The data from the SBSP was retrospectively analysed and the recommendations of the first and second readers were evaluated separately with regards to rate of recall, biopsy and cancer detection. The mean second screener contribution (MSSC) was also calculated for double reading. **Results:** In the SBSP, double reading detected 7 additional cancers (5.2% of cancers detected) compared with single reading and the MSCC was 5.5%. Double reading also resulted in 632 additional recalls, with a decrease in the positive predictive value (PPV) of cancer in the recalled women from 8.2% for single reading to 6.1% for double reading. An additional 30 biopsies were performed with double reading which represented a small decrease in PPV (41% compared to 42.9% for single reading). **Conclusion:** In the SBSP, non-consensual double reading (compared to single reading) resulted in a modest increase in cancer detection (MSSC, 5.5%) with a modest decrease in PPV of recall (2.1%) and biopsy (1.9%). These findings were consistent with that of published data and non-consensual double reading is thus recommended.

Ann Acad Med Singapore 2003; 32:438-41

Key words: Breast cancer, Cancer detection, Mammography

Introduction

Double reading of mammogram is not the standard practice in the United Kingdom National Health Service Breast Screening Programme (NHSBSP). Rather, single reading of the mammogram remains the basic policy, as was recommended in the Forrest report.¹ In the Singapore Breast Screening Project (SBSP), 28,231 women underwent 2-view mammography with non-consensual double reading. There were 135 cancers detected in 132 women with a recall rate of 8% and a biopsy rate of 1%.²

The aims of this study were to assess the efficacy of non-consensual double reading (against single reading) in the SBSP and to compare the benefits (increase in cancer detection) and limitations (increase in recall and biopsy) with published data from other similar breast screening programmes,³⁻⁶ which also employed non-consensual double reading. These programmes were conducted in

Italy, Sweden and the United Kingdom and involved a large number of women.

Materials and Methods

The data from the SBSP² was retrospectively analysed in this observational study. In the SBSP, 28,231 women received 2-view screening mammography over a period of 2½ years. The mammograms were read independently by 2 of 6 radiologists (all of whom had prior experience with mammography) and they were strictly blinded to the findings of the other radiologists. Non-consensual double reading was performed and all the radiologists were assigned as first or second reader according to the work schedule without any special consideration. A recall decision was made whenever recommended by either radiologist, that is, the worst of the 2 outcomes was acted upon. During the recall, a surgeon examined the women and the mammograms

* Registrar

** Consultant

*** Senior Consultant

Department of Diagnostic Radiology

Singapore General Hospital

**** Biostatistician

Department of Clinical Trials & Epidemiology

National Cancer Centre

Address for Reprints: Dr Kwek Boon Han, Department of Diagnostic Radiology, Singapore General Hospital, Outram Road, Singapore 169608.

Email: boonhankwek@yahoo.com

were reviewed. Cone compression or magnification mammography and ultrasound examinations were performed as required. Biopsy of the breast lesion was offered to the women whenever there is any feature of malignancy (suspicious microcalcifications, mass and all stellate lesions).

For the purpose of this study, the recommendations of the first and second readers were analysed separately and the rate of recall, biopsy and cancer detection determined for each. To quantify the positive effect of double reading, the mean second screener contribution (MSSC) was obtained.^{4,7} The MSSC is defined as the proportional increase in detected cancers because of the use of 2 screeners instead of 1.⁴ This can also be understood as the average of cancers detected by only 1 of the radiologists divided by the average number of cancers found by each radiologist.⁷

The simplified calculation is as follows:

$$\text{MSSC} = (b + c) / (2a + b + c),$$

where *a* are the number of cancers detected by both readers,

b are the number of cancers detected only by the first screener (missed by the second reader), and

c are the number of cancers detected only by the second screener (missed by the first reader).

The MSSC is an efficient estimate of the contribution of the second reader only when the readers are blinded and the worst case recall is instituted, as was the case in the SBSP.⁷

Results

Adopting non-consensual double reading, 28,231 women in the SBSP were screened with 135 breast cancers detected in 132 women, that is, a mean of 4.8 cancers per 1000 women. There were 2185 (7.7%) recalls and 422 (1.5%) biopsies.²

The first reader detected 130 cancers, but missed 5 cancers. The second reader detected 126 cancers, but missed 9 cancers (Table I). If mammograms were read by only 1 reader, there would have been an average of 128 [(130 + 126)/2] cancers detected and an average of 7 [(9 + 5)/2] missing cancers. Thus, double reading resulted in the detection of 7 additional cancers (5.2% of cancers detected) compared with single reading. The MSSC was calculated as follows:

$$(5 + 9) / [(2 \times 121) + (5 + 9)] = 5.5\%$$

Double reading resulted in 2185 recalls compared to an average of 1553 recalls [(1669 + 1437)/2] for single reading (Table II). The positive predictive value (PPV) of cancer detected amongst the recalled women was 6.2% (135/2185) for double reading and 8.2% (128/1553) for single reading.

With double reading, an average of 30 [329 – (300 + 297)/2] more biopsies were needed to detect 7 additional cancers, which represented a small decrease in PPV from 41% compared to 42.9% [1/2 (43.3% + 42.4%)] for single reading (Table II).

Discussion

Initial examination of the SBSP data showed that double reading resulted in a 5.2% increase in cancer detection compared to single reading. This meagre increase in cancer detection paled in comparison when compared to the data of Anderson et al⁴ and Ciatto et al,⁵ who reported between 10% and 15% increased cancer detection with double reading (Table III).

TABLE I: NUMBER OF CANCERS DETECTED AND MISSED BY THE FIRST AND SECOND READERS

	Detected by first reader	Missed by first reader	Total
Detected by second reader	121	5	126
Missed by second reader	9	0	9
Total	130	5	135

TABLE II: NUMBER OF RECALLS AND BIOPSIES FOR SINGLE AND DOUBLE READINGS

	Double reading	First reader only	Second reader only
Number of recalls	2185	1669	1437
Number of biopsies	329	300	297
Malignant biopsy results	135	130	126
Benign biopsy results	194	170	171
PPV (%)	41	43.3	42.4

PPV: positive predictive value

TABLE III: BENEFITS OF DOUBLE READING IN THE SBSP COMPARED WITH REPORTED DATA

	Thurfjell et al ³	Anderson et al ⁴	Ciatto et al ⁵	SBSP
Country	Sweden	UK	Italy	Singapore
No. of women screened	11,343	31,146	18,817	28,231
Age group screened	41-75 y	>50 y	50-70 y	50-64 y
Increased cancer detection by second reader (%)	15	10	15	5.2
MSSC (%)	15	5.8	4.6	5.5

MSSC: mean second screener contribution which is the average of cancer detected by one of the radiologists divided by the average number of cancers found by each radiologists

SBSP: Singapore Breast Screening Project; UK: United Kingdom

TABLE IV: PPV FOR CANCER FOLLOWING RECALLS AND BIOPSIES FOR SINGLE AND DOUBLE READINGS

Study	Recall		Biopsy	
	Double reader (%)	Single reader (%)	Double reader (%)	Single reader (%)
SBSP	6.1	8.2	41.0	42.9
Anderson et al ⁴	47.4	51.9	71.5	73.1
Thurfjell et al ³	Not reported	Not reported	58	59
Ciatto et al ⁵	15.5	17.0	Not reported	Not reported

PPV: positive predictive value; SBSP: Singapore Breast Screening Project

However, when MSSC is considered,⁷ the SBSP figure of 5.5% is comparable to that of Anderson et al (5.8%) and Ciatto et al (4.6%). Hence, the benefit of non-consensual double reading in the SBSP was within the expectations of published data. The modest gain of about 5% in these 3 studies was, however, much lower than the MSCC of 15% achieved by Thurfjell et al (Table III). In the study by Thurfjell et al,³ only 62% of the mammograms underwent double reading and this may introduce a possible bias.

In the SBSP, 7 more cancers were detected with an additional 632 recalls and 30 more biopsies. Non-consensual double reading, would, in all certainty, detect additional cancers, but it would also incur additional recalls and biopsies when compared with single reading. The 29% increase in recall rate with the second reading in the SBSP is consistent with the 31% increase reported by Anderson et al.⁴

The absolute numbers would be of limited use in themselves and do not allow for comparison between studies. On the other hand, the PPV of cancer following recall and biopsy could theoretically increase, decrease or remain constant following the introduction of double reading. This would allow comparison of the “efficiency of cancer work-up” in single and double readings.

In the SBSP, the PPV of cancer detection following recall and biopsy were small for single and double readings; 8.2% and 6.1% for recall and 42.9% and 41.0% for biopsy, respectively. Compared to other studies, the SBSP’s statistics were modestly lower than that of Thurfjell et al and Ciatto et al (Table IV). The impressive numbers reported by Anderson et al (PPV, 51.9%/47.4% for recall and 73.1%/71.5% for biopsy of single and double readings respectively)⁴ did not seem to be the norm. The authors noted that double reading was often not independent, that is, the second reader had access to the opinion of the first reader.

The other observation was that the PPV of single reading was always higher than that of non-consensual double reading (Table IV) for both recalls and biopsies. This may be inherent in non-consensual double reading, but there was no apparent explanation for this.

Additional costs in employing the second radiologist and increase in cost per cancer diagnosed (with decreased PPV of recall and biopsies) were to be expected as limitations of double reading. In addition, well women who were recalled (for false positive findings) would also be subjected to psychological stress.⁸

As long as resources are available, non-consensual double reading is recommended as it has shown an increase in cancer detection in the SBSP. In the SBSP, the majority of cancers detected by only one reader were advanced in stage (9 out of the 14 or 64%) with metastasis to the axillary lymph nodes. This figure is contrary to the findings of Thurfjell et al³ (85% of cancers were stage 0 or 1) and Anderson et al⁴ (29% of cancers were invasive), where the majority were early stage cancers. The benefit of detecting cancers that would otherwise be missed with single reading is not known. Evaluating breast cancer mortality reduction would require a long period of follow-up.

Double reading may take other forms as by consensus (recall decision to be made by two readers) or by arbitration (recall decision to be made by a third party who is usually more experienced). They were reported to detect more cancers (MSCC of 8.8% for Anttinen et al⁹ and 7.1% for Warren and Duffy¹⁰) and to decrease the number of recall and biopsy. Anttinen et al⁹ reported a 45% decrease in recall with double reading. Compared to non-consensual double reading, consensual double reading and arbitration are complex and difficult to organise with the additional requirement for a third radiologist and the associated costs. They are, however, invaluable as an educational tool as the experienced radiologist may interact with and provide feedback to less experienced staff.

While double reading by consensus and arbitration will certainly detect more cancer compared to non-consensual double reading, further study is needed to determine the cost effectiveness of these strategies and whether breast cancer mortality will be significantly affected. Besides double reading, there have been recent interest in adding ultrasound to breast cancer screening within a selected population of women with dense breasts.¹¹

Conclusion

In the SBSP, non-consensual double reading (compared to single reading) has shown a modest increase in cancer detection (MSSC, 5.5%) with a modest decrease in the PPV for recall (2.1%) and biopsy (1.9%). These findings are consistent with published data. Thus, non-consensual double reading is recommended in screening mammography so long as resources are available. Other forms of double reading, namely, consensus and arbitration, warrant further investigation.

Acknowledgements

We thank the Singapore Breast Screening Project for access to their data.

REFERENCES

1. Breast Cancer Screening. Report to the Health Ministers of England, Wales, Scotland and Northern Ireland by a Working Group chaired by Sir Patrick Forrest. London: HMSO, 1987.
2. Ng E H, Ng F C, Tan P H, Low S C, Chiang G, Tan K P, et al. Results of intermediate measures from a population-based, randomized trial of mammographic screening prevalence and detection of breast carcinoma among Asian women: the Singapore Breast Screening Project. *Cancer* 1998; 82:1521-8.
3. Thurfjell E L, Lernevall K A, Taube A A. Benefit of independent double reading in a population-based mammography screening program. *Radiology* 1994; 191:241-4.
4. Anderson E D, Muir B B, Walsh J S, Kirkpatrick A E. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994; 49:248-51.
5. Ciatto S, Del Turco M R, Morrone D, Catarzi S, Ambrogetti D, Cariddi A, et al. Independent double reading of screening mammograms. *J Med Screen* 1995; 2:99-101.
6. Denton E R, Field S. Just how valuable is double reporting in screening mammography? *Clin Radiol* 1997; 52:466-8.
7. Williams L J, Hartwood M, Prescott R J. Methodological issues in mammography double reading studies. *J Med Screen* 1998; 5:202-6.
8. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. *Br J Cancer* 1989; 60:781-4.
9. Anttinen I, Pamilo M, Soiva M, Roiha M. Double reading of mammography screening films—one radiologist or two? *Clin Radiol* 1993; 48:414-21.
10. Warren R M, Duffy S W. Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *Br J Radiol* 1995; 68:958-62.
11. Zonderland H M. The role of ultrasound in the diagnosis of breast cancer. *Semin Ultrasound CT MR* 2000; 21:317-24.