

BMJ Open Importance of quality in breast cancer screening practice – a natural experiment in Alberta, Canada

Yan Yuan ,¹ Khanh Vu,¹ Ye Shen,¹ James Dickinson,² Marcy Winget³

To cite: Yuan Y, Vu K, Shen Y, *et al.* Importance of quality in breast cancer screening practice – a natural experiment in Alberta, Canada. *BMJ Open* 2020;**10**:e028766. doi:10.1136/bmjopen-2018-028766

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-028766>).

Received 04 January 2019
Revised 21 October 2019
Accepted 11 December 2019



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹School of Public Health, University of Alberta, Edmonton, Canada

²University of Calgary Cumming School of Medicine, Calgary, Canada

³Stanford University, Stanford, California, USA

Correspondence to

Dr Yan Yuan; yyuan@ualberta.ca

ABSTRACT

Objectives Regular breast cancer screening is a widely used cancer prevention strategy. Important quality indicators of screening include cancer detection rate, false positive rate, benign biopsy rate and post-screen invasive cancer rate. We compared quality indicators of community radiology clinics to those of ‘Screen Test’, which feature centralised batch reading and quality control processes. Both types of providers operated under a single provincial Breast Cancer Screening Programme.

Setting Community radiology clinics are operated by independent fee-for-service radiologists serving large and small communities throughout the Canadian province of Alberta. Launched by the provincial cancer agency, the Screen Test operates two physical clinics serving metropolises and mobile units serving remote regions. Eligible women may self-refer to any provider for screening mammography.

Participants Women aged 50 to 69 years who had at least one screening mammogram between July 2006 and June 2010 in Alberta were included. Women with missing health region information or prior breast cancer diagnosis were excluded.

Results A total of 389 788 screening mammograms were analysed, of which 12.7% were performed by Screen Test. Compared with Screen Test during 2006 to 2008, community radiology clinics had a lower cancer detection rate (3.6 vs 4.6 per 1000 screens, risk ratio (RR): 0.81, 95% CI: 0.67 to 0.98) and a much higher false positive rate (9.4% vs 3.4%, RR: 2.72, 95% CI: 2.55 to 2.90). Most other performance indicators were also better in Screen Test overall and across all health regions. These performance indicators were similar during 2008 to 2010, showing no improvement with time.

Conclusions Screen Test has a quality assurance process in place and performed significantly better. This provides empirical evidence of the effectiveness of a quality assurance process and may explain some of the large differences in breast cancer screening indicators between provinces and countries with formal programmes and those without.

INTRODUCTION

Mammography is a widely used screening test for breast cancer, though estimates of its effectiveness vary widely, from 30% reduction in mortality to possibly none.^{1 2} Recommendations for screening presume that the same

Strengths and limitations of this study

- Our study used population data that includes all women participating in breast cancer screening in one jurisdiction over a 4-year period, which minimises the selection bias.
- The presence of two breast cancer screening models in which women arbitrarily receive screening services in one or the other, provides a natural experiment that allows the assessment of quality between the two service models in Alberta.
- Our study used administrative data which lacks detailed patient information.

benefits and harms occur to women in routine clinical practice as in the randomised clinical trials.^{3–5} Given that the estimated effectiveness is at best moderate, it is critical that quality of mammography is high, otherwise the harms of screening, such as false positives and benign biopsies, may outweigh the benefits from early detection and treatment.^{6–8} Quality of performance is ensured by use of high quality equipment, working under high technical standards, well trained technicians undertaking the mammography and trained, experienced radiologists reading the image with feedback.⁹

In USA, false positive rates in screening programmes tend to be around 9%, while they are below 5% in most European programmes and some Canadian provincial programmes such as Manitoba and Nova Scotia.^{6 10–13} In the province of Alberta, Canada, breast cancer screening is provided via an organised breast cancer screening programme, Screen Test, which uses centralised batch reading and strict quality assurance process, as well as through independently operated community radiology clinics. Operationally, Screen Test is similar to other organised screening programmes in Canada (eg, Manitoba and Nova Scotia) and Europe, while the screening conducted in the community radiology clinics shares

common features of the opportunistic screening typically provided in the private clinics and healthcare organisations in the USA.

In Alberta, mammography was performed on request by many radiology clinics from the mid-1980s, initially without dedicated equipment or specifically trained staff. In 1990, the Alberta Cancer Agency launched Screen Test through two physical clinics, one established in each of Edmonton and Calgary, the two primary cities in the province. In 1991, outreach programmes were launched operating mobile mammography units for rural areas.¹⁴ All mammography images were, and still are, processed and interpreted by salaried sessional radiologists in a central office. Parallel to Screen Test, radiologists working in community radiology clinics and hospitals also perform screening mammograms. Screen-eligible women can self-refer to any service provider. To establish a systematic screening programme which can send screening invitation letters to women who had not been screened, the Alberta Cancer Agency worked on a registry of eligible women who had mammograms. By June 2006, the database of Screen Test and the database maintained by the Alberta Radiology Association to which most community radiologists belong was fully linked as the Alberta Breast Cancer Screening Program (ABCSP).

The purpose of this study was to assess whether there is systematic variation in the quality of breast cancer screening based on standard performance indicators between the two groups of mammography providers and whether the establishment of the ABCSP improved screening quality.

METHODS

Study population

We studied screening results from screen-eligible women from July 2006 to June 2010, at the beginning of the ABCSP, to assess whether differences in quality between the two types of providers existed at the beginning of ABCSP and whether there was improvement in screening mammography quality over time after the initiation of the ABCSP. We split the 4-year period into two 2-year study periods, to assess quality indicators in the two service models over two distinct time periods. Screen-eligible women were defined as those aged 50 to 69 in either study period who had not been diagnosed with breast cancer prior to their screening mammogram.

During the study period, Screen Test had a policy of screening every 2 years, consistent with the Canadian Guidelines at the time,³ while most Alberta radiologists in the community recommended annual screening.

During the study period, about 3.4 to 3.7 million Albertans occupied a land of 662 000 km² in size,¹⁵ about the size of France. Approximately 65% of the total population lived in one of the two urban regions Edmonton and Calgary.¹⁶

Data sources

Every breast cancer screening and diagnostic procedure, corresponding date, patients' birthdate and health zone were obtained from the Screen Test database for procedures conducted by Screen Test. The same data were obtained from the Alberta Health Physician Claims database for procedures performed by radiologists in the community radiology clinics. Separate procedure codes distinguish screening and diagnostic mammograms. These two databases are complementary to each other and their quality is high.¹⁷ They were used to identify all screening mammograms and subsequent diagnostic tests conducted for eligible women. The first screening mammogram for a woman in each study period was defined as the woman's *index* screen, which marks a screening episode.

The Alberta Cancer Registry (ACR) was used to exclude women not eligible for screening due to prior breast cancer diagnosis and to identify breast cancers diagnosed in the screening cohort. ACR is a population-based cancer registry that includes all newly diagnosed cancers in Alberta since 1942. It is regularly awarded the highest degree of certification for data completeness by the North American Association of Comprehensive Cancer Registries.¹⁸

The three data sets were linked at the individual level by the unique provincial healthcare identification number. All data were obtained from the data custodian Alberta Health Services.

Primary outcomes

We used the following standard performance indicators as our primary outcomes to assess screening quality: abnormal recall rates, benign biopsy rates, false positive rates, positive predictive value (PPV), screen-detected cancer rates (invasive and non-invasive) and post-screen invasive cancer rates. These routine performance indicators are used by screening programmes in many countries including Canada.^{5 19-21} In addition, we used overall cancer diagnosed within a specific time frame regardless of mode of detection to approximate the cancer incidence in the underlying population. The derivation of the key quantities using administrative data sources are detailed in the online supplemental materials. Each indicator was stratified by service provider, region of residence and study period. Region of residence was classified into three categories: the two major urban centres Edmonton (Edmonton zone), Calgary (Calgary zone), and the rest of Alberta (comprising the South, Central and North health zones as defined by Alberta Health Services, online supplementary figure 1). Women with missing health zone were excluded (n=267).

Data analysis

Multivariate log-binomial models were used to estimate the rate ratio between providers for abnormal recalls, cancer detection through screening, false positives, benign biopsy and overall cancer. Poisson regression

models were used to estimate the rate ratios for post-screen invasive cancer for specific follow-up periods, accounting for person-year follow-up. PPV was examined using a logistic regression model to estimate the OR between the providers because the log-binomial model did not converge. The 95% confidence intervals (CIs) for these rates and rate ratios were estimated. Stratified analyses by region and study period were performed. All models adjusted for patient age. Statistical analyses were conducted using SAS 9.4 (SAS Institute, Cary, North Carolina).

Patient and public involvement

This project was done without involving screen-eligible women. Screen-eligible women were not invited to comment on the study design and were not consulted to develop relevant outcomes or interpret the results. They were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS

There were approximately 200 000 index screens in each of the two study periods. Over 84% of screens were performed in the community radiology clinics. The total number of screens performed by Screen Test decreased from 2006–2008 to 2008–2010 (online supplementary table 1). Age distributions of screened women were very similar across the two service providers and all regions with a median age of 57 to 58 years and an IQR of 53 to 62 years (online supplementary table 2).

Table 1 shows the rates of abnormal recall, false positives and benign biopsies were higher in community clinics than in Screen Test. The overall abnormal recall rates in the two study periods were 9.7% and 8.5% in community clinics, compared with 3.8% and 3.5% in Screen Test, respectively. The corresponding risk ratio estimates were 2.50 (95% CI: 2.35 to 2.65) for study period A (2006 to 2008) and 2.29 (95% CI: 2.14 to 2.46) for study period B (2008 to 2010). The abnormal recall rates in community clinics ranged widely from 5.9% in ‘other’ areas to 11.2% in Calgary, while Screen Test had fairly consistent abnormal recall rates regardless of residence area (range 3.2% to 4.1%). Similarly, the false positive rates in the two study periods were 9.4% and 8.2% in community clinics and 3.4% and 3.0% in Screen Test, respectively, with corresponding risk ratios (RRs) 2.72 (95% CI: 2.55 to 2.9) and 2.56 (95% CI: 2.38 to 2.76). The benign biopsy rates (per 1000 screens) were higher in community clinics during both study period A (8.5 vs 5.6, RR=1.47, 95% CI: 1.25 to 1.74) and B (6.7 vs 4.8, RR=1.35, 95% CI: 1.11 to 1.65). The community clinics had a lower positive predictive value (PPV) than Screen Test.

Table 2 compares the cancer detection rate (through screening) per 1000 screens. Though community clinics had higher abnormal recall rates than Screen Test, their overall screen-detected cancer rates were lower in both study periods (A: 3.6 vs 4.6, RR=0.81, 95% CI: 0.67 to 0.98,

and B: 3.3 vs 5.0, RR=0.68, 95% CI: 0.55 to 0.83). The performance difference in cancer detection occurred primarily from the detection of invasive cancer. In the two study periods, the community clinics on average detected 2.7 and 2.6 invasive cancers per 1000 screens, while Screen Test on average detected 3.7 and 4.3 invasive cancers per 1000 screens, respectively. The corresponding estimated risk ratios were 0.79 (95% CI: 0.63 to 0.98) and 0.65 (95% CI: 0.52 to 0.81).

The post-screen invasive cancer rates were 6.8 (95% CI: 5.6 to 8.3) and 18.6 (95% CI: 16.0 to 21.7) in community clinics for the 0 to 12 and 12 to 24 months periods, respectively, following a normal screening mammogram in study period A (table 3). The corresponding rates were 3.4 (95% CI: 1.8 to 6.5) and 7.5 (95% CI: 4.7 to 12.1) in Screen Test. When compared, the community clinics had much higher rates for both the first 12 months (RR: 2.36, 95% CI: 1.18 to 4.73) and 12 to 24 months (RR: 2.27, 95% CI: 1.36 to 3.78).

Women screened in the community clinics on average had shorter screening intervals than those screened in Screen Test (online supplementary figure 2). Thus, we carried out a sensitivity analysis and compared service providers by the overall cancer diagnosed within prespecified time frames regardless cancer detection mode (screen or non-screen). Table 4 showed that the overall cancer incidence and the overall invasive cancer incidence were similar between providers in both study periods A and B.

DISCUSSION

Overall the community clinics had lower cancer detection rates, higher abnormal recall rates, higher false positives, lower positive predictive value, higher benign biopsy rate and higher rates of post-screen invasive cancers when compared with Screen Test in both 2-year study periods and across all regions. Since the overall cancer incidence in women attending each type of clinic was similar, these findings are unlikely due to systematic differences in the population. Our results are consistent with a previous study comparing opportunistic and organised screening which showed better sensitivity in the organised screening programme.²²

Screen Test performance measures were consistent across health zones, presumably resulting from its centralised structure. In contrast, community clinics had variation in abnormal recall rates, likely attributable to variation in quality assurance activity across community clinics (online supplementary figure 3). The consistency between Screen Test and the variation in community clinics across the five health zones further support our conclusion that the difference observed in these indicators can be largely attributed to quality assurance differences rather than by differences in the characteristics of women attending the two types of clinics. Measures for community clinics in Edmonton and Calgary improved from the first study period to the second, possibly from



Table 1 Abnormal recall rate, false positive rate, benign biopsy rate (per 1000 screens), positive predictive value (PPV) by service provider group, region and study period. Number in parenthesis is the corresponding 95% CI

Region	Study period	Abnormal recall rate (%)			False positive rate (%)			Benign biopsy rate (per 1000 screens)			Positive predictive value (%)		
		CC	Screen test	Rate ratio*	CC	Screen test	Rate ratio	CC	Screen test	Rate ratio	CC	Screen test	OR*
Edmonton area	A	9.7 (9.5 to 9.9)	4.1 (3.7 to 4.6)	2.35 (2.1 to 2.62)	9.4 (9.2 to 9.6)	3.7 (3.3 to 4.2)	2.51 (2.24 to 2.82)	8.7 (8.0 to 9.4)	6.2 (4.6 to 8.3)	1.38 (1.03 to 1.85)	3.3 (2.8 to 3.7)	9.9 (6.9 to 13.7)	0.30 (0.2 to 0.44)
	B	8.6 (8.4 to 8.8)	3.5 (3.1 to 4.0)	2.29 (2.01 to 2.62)	8.3 (8.1 to 8.5)	3.1 (2.7 to 3.6)	2.50 (2.17 to 2.88)	6.9 (6.3 to 7.5)	4.3 (2.8 to 6.4)	1.62 (1.08 to 2.43)	3.4 (3.0 to 3.9)	11.8 (7.7 to 17.0)	0.30 (0.19 to 0.48)
Calgary area	A	11.2 (10.9 to 11.4)	3.5 (3.1 to 3.9)	3.11 (2.76 to 3.49)	10.8 (10.5 to 11.0)	3.0 (2.6 to 3.4)	3.49 (3.08 to 3.96)	8.9 (8.2 to 9.7)	4.6 (4.2 to 6.3)	1.90 (1.35 to 2.65)	3.5 (3.1 to 4.0)	14.5 (10.6 to 19.3)	0.21 (0.15 to 0.3)
	B	8.6 (8.4 to 8.9)	3.2 (2.7 to 3.7)	2.55 (2.21 to 2.95)	8.3 (8.1 to 8.5)	2.6 (2.2 to 3.1)	2.98 (2.53 to 3.49)	6.6 (6.0 to 7.2)	4.2 (2.7 to 6.4)	1.48 (0.97 to 2.24)	3.8 (3.4 to 4.3)	17.9 (12.5 to 24.5)	0.21 (0.14 to 0.31)
Other areas	A	5.9 (5.7 to 6.2)	3.9 (3.5 to 4.2)	1.53 (1.38 to 1.69)	5.6 (5.3 to 5.9)	3.4 (3.1 to 3.7)	1.62 (1.46 to 1.81)	6.6 (5.7 to 7.7)	5.9 (4.6 to 7.5)	1.09 (0.83 to 1.43)	5.9 (4.7 to 7.2)	11.7 (8.9 to 15.0)	0.46 (0.32 to 0.65)
	B	8.1 (7.8 to 8.4)	3.7 (3.4 to 4.1)	2.12 (1.91 to 2.35)	7.7 (7.4 to 8.0)	3.2 (2.9 to 3.6)	2.35 (2.1 to 2.62)	6.6 (5.8 to 7.5)	5.2 (4.0 to 6.8)	1.16 (0.87 to 1.55)	4.1 (3.4 to 5.0)	13.6 (10.4 to 17.4)	0.28 (0.2 to 0.39)
Total	A	9.7 (9.6 to 9.9)	3.8 (3.6 to 4.1)	2.50 (2.35 to 2.65)	9.4 (9.2 to 9.5)	3.4 (3.2 to 3.6)	2.72 (2.55 to 2.9)	8.5 (8.0 to 8.9)	5.6 (4.8 to 6.6)	1.47 (1.25 to 1.74)	3.7 (3.4 to 4.0)	11.9 (10.0 to 14.0)	0.27 (0.22 to 0.34)
	B	8.5 (8.4 to 8.6)	3.5 (3.3 to 3.8)	2.29 (2.14 to 2.46)	8.2 (8.1 to 8.3)	3.0 (2.8 to 3.3)	2.56 (2.38 to 2.76)	6.7 (6.4 to 7.1)	4.8 (3.9 to 5.8)	1.35 (1.11 to 1.65)	3.7 (3.4 to 4.0)	14.1 (11.7 to 16.8)	0.26 (0.21 to 0.32)

*PPV was examined using a logistic regression model because the log-binomial model did not converge. CC, community clinics.

Table 2 Cancer detection rate via screening (per 1000 screens) by service provider, region and study period. Number in parenthesis is the corresponding 95% CI

Region	Study period	Invasive cancer detection rate				In situ cancer detection rate				Cancer detection rate			
		Community clinics		Screen test		Community clinics		Screen test		Community clinics		Screen test	
		Rate ratio	Screen test	Rate ratio	Screen test	Rate ratio	Screen test	Rate ratio	Screen test	Rate ratio	Screen test	Rate ratio	Screen test
Edmonton area	A	2.3 (2.0 to 2.7)	3.6 (2.4 to 5.2)	0.67 (0.45 to 1.01)	0.8 (0.6 to 1.1)	0.5 (0.1 to 1.3)	1.69 (0.61 to 4.67)	3.2 (2.7 to 3.6)	4.1 (2.8 to 5.8)	0.80 (0.55 to 1.16)			
	B	2.3 (2.0 to 2.7)	3.3 (2.0 to 5.1)	0.75 (0.47 to 1.18)	0.6 (0.4 to 0.8)	0.7 (0.2 to 1.8)	0.88 (0.32 to 2.43)	2.9 (2.6 to 3.3)	4.2 (2.7 to 6.2)	0.77 (0.51 to 1.17)			
Calgary area	A	3.1 (2.7 to 3.5)	3.7 (2.5 to 5.3)	0.89 (0.6 to 1.31)	0.8 (0.6 to 1.1)	1.4 (0.7 to 2.5)	0.61 (0.32 to 1.17)	3.9 (3.5 to 4.4)	5.1 (3.6 to 6.9)	0.81 (0.58 to 1.14)			
	B	2.8 (2.4 to 3.2)	5.2 (3.4 to 7.4)	0.60 (0.4 to 0.88)	0.5 (0.4 to 0.7)	0.6 (0.1 to 1.6)	0.99 (0.3 to 3.2)	3.3 (2.9 to 3.7)	5.7 (3.9 to 8.1)	0.63 (0.44 to 0.92)			
Other areas	A	2.8 (2.2 to 3.6)	3.7 (2.7 to 4.9)	0.79 (0.54 to 1.15)	0.6 (0.4 to 1.0)	0.8 (0.4 to 1.5)	0.73 (0.33 to 1.61)	3.5 (2.8 to 4.3)	4.5 (3.4 to 5.9)	0.79 (0.56 to 1.1)			
	B	2.7 (2.2 to 3.3)	4.3 (3.1 to 5.7)	0.66 (0.46 to 0.94)	0.6 (0.4 to 0.9)	0.8 (0.3 to 1.5)	0.80 (0.35 to 1.81)	3.3 (2.8 to 4.0)	5.1 (3.8 to 6.6)	0.68 (0.49 to 0.94)			
Total	A	2.7 (2.5 to 3.0)	3.7 (3.0 to 4.4)	0.79 (0.63 to 0.98)	0.8 (0.7 to 1.0)	0.9 (0.6 to 1.3)	0.91 (0.59 to 1.39)	3.6 (3.3 to 3.9)	4.6 (3.8 to 5.4)	0.81 (0.67 to 0.98)			
	B	2.6 (2.4 to 2.8)	4.3 (3.5 to 5.3)	0.65 (0.52 to 0.81)	0.6 (0.5 to 0.7)	0.7 (0.4 to 1.1)	0.84 (0.49 to 1.44)	3.3 (2.9 to 3.4)	5 (4.1 to 6.0)	0.68 (0.55 to 0.83)			

**Table 3** Post-screen invasive cancer rate (per 10000 person-years) by follow-up times, service provider and region following a normal screen mammogram in study period A. Number in parenthesis is the corresponding 95% CI

Region	Interval	Provider	N	Rate	Rate ratio	P value
Edmonton area	0–12 months	Community clinics	37	6.7 (4.8 to 9.2)		
		Screen test	3	4.0 (1.3 to 12.4)		
	12–24 months	Community clinics	82	22.7 (18.3 to 28.2)		
		Screen test	5	7.8 (3.3 to 18.8)		
Calgary area	0–12 months	Community clinics	32	5.3 (3.7 to 7.4)		
		Screen test	2	2.6 (0.7 to 10.6)		
	12–24 months	Community clinics	60	16.0 (12.4 to 20.6)		
		Screen test	5	7.9 (3.3 to 19.0)		
Other areas	0–12 months	Community clinics	27	11.2 (7.7 to 16.3)		
		Screen test	4	3.5 (1.3 to 9.3)		
	12–24 months	Community clinics	22	15.2 (10.0 to 23.0)		
		Screen test	7	7.1 (3.4 to 14.8)		
Total	0–12 months	Community clinics	96	6.8 (5.6 to 8.3)	2.36	0.0158
		Screen test	9	3.4 (1.8 to 6.5)	(1.18 to 4.73)	
	12–24 months	Community clinics	164	18.6 (16.0 to 21.7)	2.27	0.0017
		Screen test	17	7.5 (4.7 to 12.1)	(1.36 to 3.78)	

the positive influence of formation of the Alberta Breast Cancer Screening Program, whereas for community clinics in other health zones, both false positive rates and PPV were worse in the second period.

At the time, both community radiology clinics and Screen Test typically sent abnormal screening results to their patients' primary care providers who then coordinated the diagnostic follow-up, and such diagnostic imaging was done through community clinics. One might expect that this practice should produce similar benign biopsy rates between the two types of screening service providers. The rate ratios, however, range from 1.38 to 1.90 comparing community clinics to Screen Test in Calgary and Edmonton in study periods A and B (table 2), showing that women screened in a community clinic had a higher rate of benign biopsy than women screened by Screen Test. This highlights that the higher rate of false positive findings increases the risk of invasive follow-up procedures. Thus, the effectiveness of screening was lower in the community radiology clinics than Screen Test due to both higher false positive rates and higher benign biopsy rates. Additionally, diagnostic tests after a false positive have been shown to reduce the likelihood of future screening participation: benign biopsies have the largest negative impact.²³

The operations and quality assurance practices of Screen Test and community clinics were distinct in several ways during 2006 to 2010. In Screen Test, staff radiologists read batches of 70 to 90 screening mammograms in a 2 to 4 hour screening session. At monthly quality assurance meetings, they reviewed and discussed difficult cases and received personal statistics as a feedback mechanism; these practices are still standard today.

The operation and quality assurance practices of community radiology clinics varied, and still do today. First, many clinics interpret screening mammograms while women are still present, to decide whether further testing is needed. Our previous analysis showed that more than 55% of the follow-up imaging procedures were performed the same day.²³ Although this practice minimises time to diagnosis/resolution, it encourages a diagnostic rather than a screening mindset. Second, screening mammograms are not read separately from diagnostic mammograms thus images from symptomatic and asymptomatic patients are read without clear distinction. Third, it is unknown whether recall statistics are produced to help radiologists self-evaluate against the screening mammography standard.

Another contributing factor to the quality of reading could be reading volume; a high volume has been shown to reduce radiologists' false positive rate.^{24 25} While the average reading volume in Screen Test was about 2300 screening mammograms per year, the minimum requirement for radiologists in the community was 400 mammograms per year during the study period, in accordance with the Canadian Association of Radiologists Mammography Accreditation Program.²⁶

While the Canadian target for PPV is $\geq 5\%$ for initial screen and $\geq 6\%$ for rescans,²¹ PPVs were 3.7% for community clinics in both study periods compared with 11.9% and 14.1% for Screen Test in the respective time periods. The average false positive rate of radiologists in community clinics in Alberta is similar to that of US radiologists in the Breast Cancer Screening Consortium.¹³ False positive rates vary greatly among US radiologists,²⁴ which is likely also true for community radiologists

Table 4 Overall breast cancer incidence per 1000 women in the screening population, by service provider, region, time period and follow-up time since index screen

Region	Study period	Follow-up time	Overall cancer incidence (per 1000 women) (95% CI)			Overall invasive cancer incidence (per 1000 women) (95% CI)		
			Community clinic	Screen test	Rate ratio	Community clinic	Screen test	Rate ratio
Edmonton area	A	1 year	5.1 (4.5 to 5.7)	4.6 (3.2 to 6.4)	1.13 (0.80 to 1.60)	4.1 (3.6 to 4.6)	4.1 (2.8 to 5.8)	1.03 (0.72 to 1.50)
	B	2 year	7.9 (7.2 to 8.7)	7.3 (5.5 to 9.4)	1.12 (0.85 to 1.47)	6.6 (6 to 7.3)	5.9 (4.3 to 7.8)	1.15 (0.85 to 1.56)
Calgary area	A	Up to 31 December 2010*	7.2 (6.6 to 7.8)	6.8 (4.8 to 9.2)	1.16 (0.84 to 1.61)	6.1 (5.5 to 6.7)	5.4 (3.7 to 7.6)	1.24 (0.86 to 1.78)
	B	1 year	4.8 (4.3 to 5.4)	5.5 (4 to 7.4)	0.92 (0.67 to 1.27)	3.9 (3.4 to 4.4)	4.1 (2.8 to 5.7)	1.01 (0.70 to 1.48)
Other areas	A	2 year	7.5 (6.9 to 8.2)	7.9 (6.1 to 10.1)	0.99 (0.76 to 1.29)	6.2 (5.6 to 6.8)	6.1 (4.5 to 8.1)	1.06 (0.79 to 1.43)
	B	Up to 31 December 2010*	6.2 (5.7 to 6.8)	7.4 (5.3 to 10)	0.92 (0.67 to 1.27)	5.3 (4.7 to 5.8)	6.8 (4.8 to 9.4)	0.84 (0.60 to 1.18)
Total	A	1 year	4.6 (3.8 to 5.5)	4.9 (3.8 to 6.4)	0.95 (0.70 to 1.30)	3.8 (3.1 to 4.6)	4.2 (3.1 to 5.5)	0.94 (0.67 to 1.32)
	B	2 year	6.9 (6.0 to 8.0)	6.6 (5.2 to 8.2)	1.08 (0.83 to 1.40)	5.8 (4.9 to 6.8)	5.3 (4.1 to 6.8)	1.12 (0.84 to 1.51)
Total	A	Up to 31 December 2010*	6.3 (5.5 to 7.2)	6.5 (5 to 8.2)	1.01 (0.77 to 1.33)	5.3 (4.5 to 6.1)	5.5 (4.2 to 7.2)	1.00 (0.74 to 1.34)
	B	1 year	4.9 (4.5 to 5.2)	5.0 (4.2 to 5.9)	1.01 (0.84 to 1.21)	4.0 (3.7 to 4.3)	4.1 (3.4 to 4.9)	1.00 (0.82 to 1.22)
Total	A	2 year	7.6 (7.2 to 8.0)	7.2 (6.2 to 8.2)	1.10 (0.94 to 1.27)	6.3 (5.9 to 6.7)	5.7 (4.9 to 6.7)	1.14 (0.98 to 1.35)
	B	Up to 31 December 2010*	6.6 (6.3 to 7.0)	6.8 (5.7 to 8.0)	1.05 (0.94 to 1.18)	5.6 (5.3 to 6)	5.8 (4.8 to 6.9)	1.04 (0.92- to 1.19)

*End of follow-up period; cancer diagnosis data available through 31 December, 2010.



in Alberta. In comparison, the false positive rate of organised screening programmes with strong standard quality assurance programmes, for example, Screen Test and Canadian provinces including Manitoba and Nova Scotia, are comparable to countries in the European Union at less than 5%.⁶

Study limitations and advantages

The main advantages of this study are that the two ways to access breast cancer screening in Alberta provide a natural experiment to assess differences in quality of screening between the two breast cancer screening models. Neither screen-eligible women nor the majority of family physicians in the province are aware that there are two distinct systems offering screening services. Women living in Edmonton and Calgary choose to receive screening at either a community-based clinic or a Screen Test clinic. Their choice is independent of their disease state. The arbitrary selection of screening provider conceptually mimics a single blind randomised experiment.

There are several limitations of the study. We could not distinguish whether a mammogram was an initial or subsequent screen, which influences abnormal recall rate and cancer detection rate. The ages of women are similar among types of service providers, however, which mitigate the potential bias. Women attending Screen Test generally had longer intervals between screens than those attending community radiology clinics. We performed sensitivity analysis for women who had a normal index screen in study period A and were subsequently screened between 18 to 30 months. The same pattern of differences emerged (online supplementary table 3): when compared with Screen Test in the subsequent screening, community clinics had a lower cancer detection rate (2.4 vs 4.0 per 1000 screens) and a higher abnormal recall rate (6.6% vs 2.5%). Additionally, the low cancer detection rate was attributed to the lower invasive cancer rate (online supplementary table 4). There are likely variations between different community clinics and individual radiologists,^{24 25 27} however, we only have aggregated data, so cannot identify individual performance.

CONCLUSION

The majority of quality measures examined were better in the organised screening programme where an effective quality assurance process is in place. This study thus provides empirical evidence that the high false positive rates observed in some provinces of Canada, for example, Quebec and New Brunswick,⁶ and the USA (compared with European Union countries) and significant variation world-wide in effectiveness of screening programmes may be in large part due to lack of rigorous, standardised quality assurance processes (including minimum reading volume requirement) among providers of screening mammography. Second, the differences in quality measures between the community clinics and Screen Test strongly suggest that quality assurance is essential

to reduce harm from false positives and missed cancers. Public health programmes must be carefully implemented to maximise the benefits and minimise potential harms to participants.

Acknowledgements Following individuals assisted in the development of our manuscript: anonymous radiologists for providing insights into our study findings, Jingyu Bu for contributing to the statistical analysis and Zhe Lu for reference management. The authors would also like to acknowledge valuable comments from reviewers which strengthened this manuscript.

Contributors Conception and design: YY, MW and JD. Data acquisition, analysis and interpretation: YS, KV, MW, YY and JD. Manuscript draft and revising: JD, YY, MW, KV and YS. Final approval and accountable for the study: YY, KV, YS, JD and MW.

Funding This research is jointly supported by Canadian Institutes of Health Research (DC0190GP to MW) and the MSI foundation (869 to YY). Dr Yuan's research program is supported by Natural Sciences and Engineering Research Council of Canada (FRN: RGPIN-2019-04862).

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethics approval for the study was obtained from the Health Research Ethics Board of Alberta, Cancer Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Alberta Health Services in the provincial of Alberta Canada, is the data custodian of the administrative data sets analysed in this article. A research protocol and ethics approval are required in order to request the data for research use.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Yan Yuan <http://orcid.org/0000-0002-2073-9924>

REFERENCES

- 1 Tabar L, Gad A, Holmberg L, *et al*. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the breast cancer screening Working group of the Swedish national board of health and welfare. *The Lancet* 1985;325:829–32.
- 2 Miller AB, Wall C, Baines CJ, *et al*. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National breast screening study: randomised screening trial. *BMJ* 2014;348:g366.
- 3 The Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40–74 years. *Can Med Assoc J* 2011;183:1991–2001.
- 4 Ringash J. Preventive health care, 2001 update: screening mammography among women aged 40–49 years at average risk of breast cancer. *CMAJ* 2001;164:469–76.
- 5 US Preventive Services Task Force. Screening for breast cancer: U.S. preventive services Task force recommendation statement. *Ann Intern Med* 2009;151:716–26.
- 6 Canadian Partnership Against Cancer. Organized breast cancer screening programs in Canada: report on program performance in 2007 and 2008, 2013. Available: <http://www.getcheckedmanitoba.ca/files/b-rep-pro-perf-07-08.pdf> [Accessed 10 Sep 2018].
- 7 Törnberg S, Kemetli L, Ascunce N, *et al*. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010;19:87–93.

- 8 Dickinson JA, Pimlott N, Grad R, *et al.* Screening: when things go wrong. *Can Fam Physician* 2018;64:502–8.
- 9 International Agency for Research on Cancer. IARC Handbooks of Cancer Prevention, Volum 15 Breast Cancer Screening [Internet]: International Agency for Research on Cancer, 2016. Available: <http://publications.iarc.fr/Book-And-Report-Series/IARC-Handbooks-Of-Cancer-Prevention/Breast-Cancer-Screening-2016> [Accessed 10 Sep 2018].
- 10 Le MT, Mothersill CE, Seymour CB, *et al.* Is the false-positive rate in mammography in North America too high? *Br J Radiol* 2016;89:20160045.
- 11 Johns LE, Moss SM. False-positive results in the randomized controlled trial of mammographic screening from age 40 ("Age" trial). *Cancer Epidemiol Biomarkers Prev* 2010;19:2758–64.
- 12 Sala M, Domingo L, Macià F, *et al.* Does digital mammography suppose an advance in early diagnosis? Trends in performance indicators 6 years after digitalization. *Eur Radiol* 2015;25:850–9.
- 13 Nelson HD, O'Meara ES, Kerlikowske K, *et al.* Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. *Ann Intern Med* 2016;164:226–35.
- 14 Alberta Health Services. Screen test, 2017. Available: <https://www.albertahealthservices.ca/info/service.aspx?id=1002353> [Accessed 9 Oct 2019].
- 15 Statistics Canada. Table: 17-10-0009-01 Population estimates (quarterly). Available: <https://doi.org/10.25318/1710000901-eng> [Accessed 9 Oct 2019].
- 16 Statistics Canada. Population and dwelling counts, for Canada, provinces and territories, census metropolitan areas and census agglomerations, 2011 and 2006 censuses.
- 17 Yuan Y, Li M, Yang J, *et al.* Using administrative data to estimate time to breast cancer diagnosis and percent of screen-detected breast cancers - a validation study in Alberta, Canada. *Eur J Cancer Care* 2015;24:367–75.
- 18 North American Association of Central Cancer Registries (NAACCR). Certification levels. Available: <https://www.naacrr.org/certification-criteria/> [Accessed 10 Sep 2018].
- 19 Australian Institute of Health and Welfare. *National bowel cancer screening program: monitoring report 2017. cancer series no.104. cat. No. can 103.* Canberra: AIHW, 2017. <https://www.aihw.gov.au/reports/cancer-screening/bowel-cancer-screening-program-monitoring-2017/contents/table-of-contents>
- 20 Chiarelli AM, Halapy E, Nadalin V, *et al.* Performance measures from 10 years of breast screening in the Ontario breast screening program, 1990/91 to 2000. *Eur J Cancer Prev* 2006;15:34–42.
- 21 Canadian Partnership Against Cancer. *The evaluation indicators Working Group report: guidelines for monitoring breast screening program performance.* 3rd edn, 2012. <https://www.cancercare.mb.ca/export/sites/default/screening/.galleries/files/screening-files/b-rep-gdlnmonitor.pdf>
- 22 Bihmann K, Jensen A, Olsen AH, *et al.* Performance of systematic and non-systematic ("opportunistic") screening mammography: a comparative study from Denmark. *J Med Screen* 2008;15:23–6.
- 23 Shen Y, Winget M, Yuan Y. The impact of false positive breast cancer screening mammograms on screening retention: a retrospective population cohort study in Alberta, Canada. *Can J Public Health* 2017;108:e539–45.
- 24 Smith-Bindman R, Chu P, Miglioretti DL, *et al.* Physician predictors of mammographic accuracy. *J Natl Cancer Inst* 2005;97:358–67.
- 25 Théberge I, Chang S-L, Vandal N, *et al.* Radiologist interpretive volume and breast cancer screening accuracy in a Canadian organized screening program. *J Natl Cancer Inst* 2014;106.
- 26 The Canadian Association of Radiologists (CAR). CAR practice guidelines and technical standards for breast imaging and intervention, 2016. Available: <https://car.ca/wp-content/uploads/Breast-Imaging-and-Intervention-2016.pdf> [Accessed 10 Sep 2018].
- 27 Coldman AJ, Phillips N. False-Positive screening mammograms and biopsies among women participating in a Canadian provincial breast screening program. *Can J Public Health* 2012;103:e420–4.