Variation in performance in low-volume mammography screening programmes: Experience from Switzerland

J.-L. Bulliard [a,*], C. Ducros [b], E. Dayer [c], B. Arzel [d], F. Levi [a,e]

[a] Unité d'épidémiologie du cancer, Institut de médecine sociale et préventive, Centre Hospitalier Universitaire Vaudois et Université de Lausanne, Lausanne, Switzerland
[b] Fondation pour le dépistage du cancer du sein, Lausanne, Switzerland
[c] Centre valaisan de dépistage du cancer du sein, Sion, Switzerland
[d] Fondation genevoise pour le dépistage du cancer du sein, Geneva, Switzerland
[e] Registres Vaudois et Neuchâtelois des Tumeurs, Centre Hospitalier Universitaire Vaudois et Université de Lausanne, Lausanne, Switzerland

ABSTRACT

Background: Reading volume and mammography screening performance appear positively correlated. Quality and effectiveness were compared across low-volume screening programmes targeting relatively small populations and operating under the same decentralised healthcare system. Except for accreditation of 2nd readers (restrictive vs non-restrictive strategy), these organised programmes had similar screening regimen/procedures and duration, which maximises comparability. Variation in performance and its determinants were explored in order to improve mammography practice and optimise screening performance. *Methods*: Circa 200,000 screens performed between 1999 and 2006 (4 rounds) in 3 longest standing Swiss cantonal programmes (of Vaud, Geneva and Valais) were assessed. Indicators of quality and effectiveness were assessed according to European standards. Interval cancers were identified through linkage with cancer registries records. *Results*: Swiss programmes met most European standards of performance with a substantial, favourable cancer stage shift. Up to a two-fold variation occurred for several performance indicators. In subsequent rounds, compared with programmes (Vaud and Geneva) that applied a restrictive selection strategy for 2nd readers, proportions of *in situ* lesions and of small cancers (≤1 cm) were one third lower and halved, respectively, and the proportion of advanced lesions (stage II+) nearly 50% higher in the programme without a restrictive selection strategy. Discrepancy in second-year proportional incidence of interval cancers appears to be multicausal. *Conclusion*: Differences in performance could partly be explained by a selective strategy for second readers and a prior experience in service screening, but not by the levels of opportunistic screening and programme attendance. This study provides clues for enhancing mammography screening performance in low-volume programmes.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Most western countries have established national or regional breast cancer screening programmes since randomised controlled trials demonstrated that mammography screening was effective in reducing breast cancer mortality in females aged 50 and over [1,2]. In Switzerland, where healthcare delivery is organised at a regional (cantonal) level, the first screening programmes were implemented in 1999, following a successful pilot project [3,4]. The development of organised programmes concomitantly to pre-existing opportunistic screening was facilitated by the decision of the Federal Office of Social Insurance to reimburse from 1999 biennial mammography screening for women aged 50 and above through the Swiss compulsory health insurance coverage system, when performed within a quality-assured programme.

Wide variation in performance indicators has been documented across service screening programmes in Europe and similar variation is expected in the impact of screening on breast cancer mortality [2]. This discrepancy has broadly been ascribed to differences in screening organisation and procedures, and to individual programme size and volume of activity [5,6]. Most comparisons of short-term performance indicators have additionally been subjected to potential age confounding since screening performance is also sensitive to programme duration and the age groups screened [1,7].

The region-oriented healthcare system in Switzerland enables to compare performance indicators across programmes initiated at

* Corresponding author at: Unité d'épidémiologie du cancer, Institut de médecine sociale et préventive, Centre Hospitalier Universitaire Vaudois et Université de Lausanne, Route de la Corniche 2, Bâtiment Biopôle 1, 1066 Epalinges, Switzerland. Fax: +41 21 314 7373.

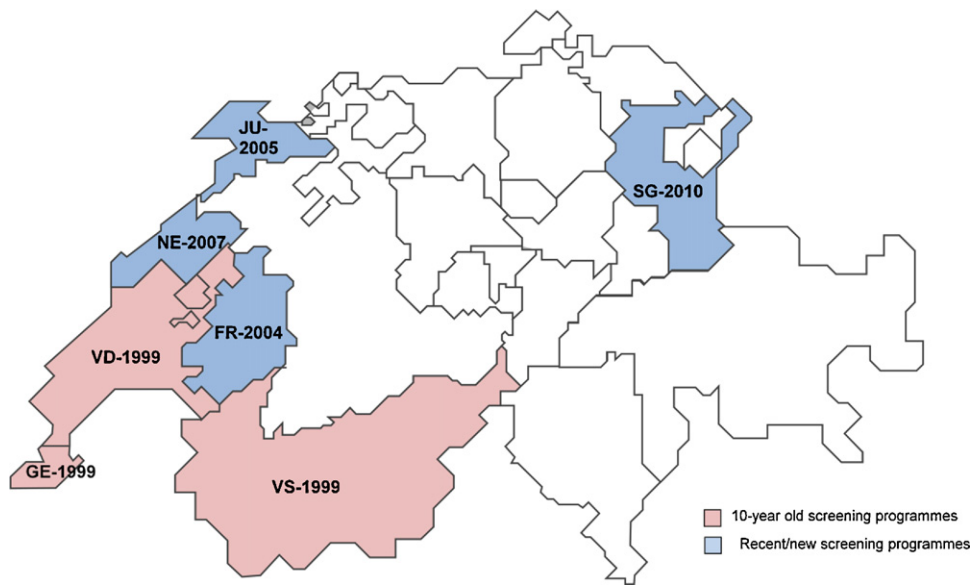E-mail address: Jean-Luc.Bulliard@chuv.ch (J.-L. Bulliard).

Fig. 1. Organised breast cancer screening programmes in Switzerland, 2010.

the same time, with similar screening organisation and procedures, but with substantial differences in multiple reading strategy, population size, historical screening habits and prevalence of opportunistic screening [3,4]. The aim of this study was (1) to examine the variation in quality and effectiveness of mammography screening in low-volume Swiss regional programmes after 8 years of operation (4 screening rounds), and (2) to explore determinants of this variation in order to improve mammography practice and optimise screening performance.

2. Materials and methods

2.1. Organised mammography screening in Switzerland: current state

Five mammography screening programmes are operating in Switzerland, covering six (all French-speaking) cantons and about 25% of the Swiss female resident population aged 50–69 (Fig. 1). A sixth programme will start in 2010 in German-speaking Switzerland. All programmes share the same operating software and quality control procedures, and have joint educational material and similar screening procedures that include biennial double reading with arbitration and two-view mammography. The invitation and reading processes in Swiss programmes have been described elsewhere [3].

This study focussed on the three longest standing programmes (Vaud (VD), Geneva (GE) and Valais (VS)) which all were launched in 1999 (Fig. 1). Programmes in the cantons of Fribourg (FR-2004) and Jura/Neuchâtel (JU-2005/NE-2007) were too recent to have reached their full potential. The three programmes considered took place in cantons with different population characteristics with respect to breast cancer screening, but with similar age structure of their 50–69 female populations.

2.2. The Vaud screening programme

The pioneer activities in breast cancer screening in the Vaud canton contributed to the development of service screening programmes in Switzerland [3]. This largest Swiss programme – it encompassed over 40% of all programme-based mammograms – elicited stringent criteria for second reader radiologists in order to comply with the 5000 minimum reading requirement recommended in the European guidelines (Table 1) [8]. Several accredited radiologists had prior experience with mammography reading in an

organised setting from the pilot trial [3]. Exposure to organised screening from the pilot programme, but not accessibility, influenced participation in this canton [9]. In this region with a dense offer of screening facilities, overall attendance was about 50% but varied locally: the denser the population, the lower was participation [9].

2.3. The Geneva screening programme

The Geneva programme applied a similar second reader strategy to that in Vaud. Difficulties in collaborating with private radiologists temporarily hindered the development of a programme (participation rate: about 25%) set in a population with one of the highest incidence of breast cancer worldwide [10]. Opportunistic screening prevails in this highly urbanised canton and has increased since the inception of the programme. Conflicting evidence has been reported about attraction by the programme of women from predominantly lower socioeconomic classes [11,12].

2.4. The Valais screening programme

In Valais, the second reading approach favoured volunteering, self-motivation and self-responsibility of radiologists. This led to

Table 1
Screening activity in the 3 main Swiss regional programmes, 1999–2006.

Parameter	GE	VS	VD
Target population invited, 50–69 years old women[a]	72,655	44,237	100,853
Mean annual number of mammography	5069	8242	13,598
Number of accredited radiology centres	11	11	29
Mean annual number of accredited radiologists	27	22	34
Mean annual number of 2nd readers	6	11	4
Annual volume (all readers)			
Mean	188	746	364
Median	194	646	852
Annual volume (2nd readers only)			
Mean	813	1086	3390
Median	758	939	3334

GE: Geneva; VS: Valais; VD: Vaud.
[a] Includes ineligible women.

Table 2
Indicators of screening performance (quality, timeliness and effectiveness) in Swiss regional programmes, 1999–2006.

Performance indicator	GE $n = 40,553$	VS $n = 52,889$	VD $n = 104,166$	European norms
Referral rate (%)				
Initial screen	8.0	7.1	6.4	<7
Subsequent screen	4.7	3.3	3.9	<5
Benign to malignant biopsy ratio				
Initial screen	n.a.	0.55	0.47	≤0.5
Subsequent screen	n.a.	0.24	0.28	≤0.25
Timeliness (%)				
Re-invitation within 24 months	65.0	83.2	91.8	>95
Re-invitation within 30 months	92.0	94.6	98.0	>98
Detection rate (/1000)				
Initial screen	6.5	6.2	7.6	(3× I.R.)
Subsequent screen	6.2	4.8	5.9	(1.5× I.R.)
In situ cancers (%)				
Initial screen	18.9	12.2	18.2	10–20
Subsequent screen	18.9	12.5	17.1	10–20
Invasive cancers ≤ 10 mm (%)				
Initial screen	27.5	25.2	39.6	≥25
Subsequent screen	40.2	21.2	44.7	≥25
Stage II+ (%)				
Initial screen	31.5	35.4	29.7	<30
Subsequent screen	29.6	39.3	28.9	≤25
Node-negative cancers (%)				
Initial screen	75.4	74.3	72.3	≥70
Subsequent screen	74.7	75.9	73.8	≥75
Interval cancers (initial screen, proportional incidence)	$n = 56$	$n = 59$	$n = 77$	
0–11 months after screening	15.0	27.3	24.0	≤30
12–23 months after screening	67.0	53.2	41.9	≤50

GE: Geneva; VS: Valais; VD: Vaud; n.a.: not available; I.R.: incidence rate.

more second readers with lesser annual throughput. Participation has been the highest among Swiss service screening programmes, almost reaching the European recommendation (66% vs 70%) [13]. Participation however differed within this alpine canton, being higher in the French than the German-speaking community [13].

2.5. Opportunistic screening

Self-reported biennial mammography (screening and diagnostic) coverage was in 2002 about 90% in the 3 cantons, above the Swiss average for 50–69 year olds females (71.5%) [12]. The wide difference in participation across programmes (GE: 26%, VD: 49%, VS: 66%) [14] reflected different prevalences of opportunistic screening [4].

2.6. Analyses

Analyses included 197,608 screens performed during the first 8 years (1999–2006) of operation among 50–69-year-old women residing in these 3 cantons. Performance and prognostic indicators as defined in the European Guidelines were calculated for prevalent (initial) and incident (subsequent) screening rounds for each programme [8]. Interval cancers were identified by linkage between screening and cancer registries records, performed separately for the Vaud, Geneva and Valais Cancer Registries. To estimate the proportional incidence of interval cancers (the ratio of observed to expected incidence in the absence of screening), age-specific breast cancer incidence rates for females aged 50–69 in the 5-year period preceding the screening programmes were used.

3. Results

The volume of activity for each regional programme is described in Table 1. The average throughput ranged between circa 5000 and 13,600 yearly readings between 1999 and 2006. Both the overall and individual volume increased over time as participation raised (data not shown). The number of accredited second readers was inversely related to the population size of the programme. The combined effect of different participation rates and strategies for second readers led to a four-fold variation across programmes in the annual volume interpreted by second readers.

Performance indicators for the 3 programmes, ranked by increasing size, are presented by type of screening round in Table 2 and the relationship between some key indicators is illustrated in Fig. 2. The proportion of subsequent screens ranged from 44% in GE to 57% in VD, and was 50% in VS (data not shown). In terms of quality, the biopsy yield and the referral rate met European standards in subsequent screens and, for one programmme only, in initial screens. Overall, screening performance improved between prevalent and incident rounds (higher PPV, lower false-positive and referral rates, Fig. 2). This improvement was accompanied by less heterogeneity in performance across programmes in subsequent screens (Fig. 2).

The fraction of women reinvited within 2 years was too low for the 3 programmes (65–92% vs 95% recommended, Table 2), although the reinvitation rate at 30-month substantially improved and reached the European norm for one programme. The median time between two screening invitations was 719 days (VD), 739 days (VS) and 742 days (GE), respectively, and these delays decreased over screening rounds (data not shown).

Most indicators of screening effectiveness fulfilled European requirements for all programmes (Table 2). The proportion of advanced cancers was consistently high, irrespective of screening round. Substantial regional variations were observed for (1) *in situ* cancers whose rates were 50% lower in VS than elsewhere, (2) small cancers (≤1 cm) in subsequent screens whose proportion was in VS half that in VD and GE, and (3) interval cancers for which
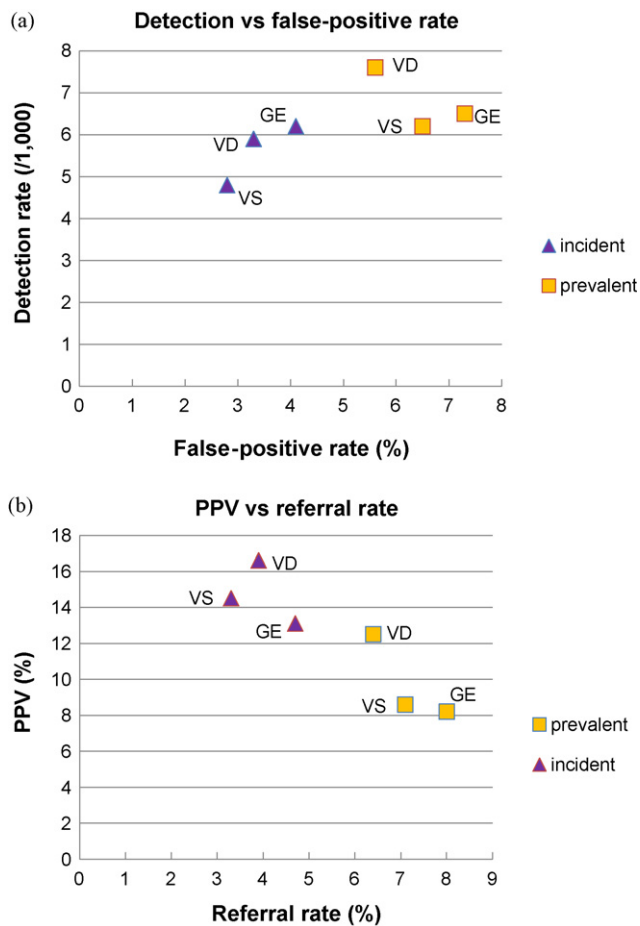
(a)

Detection vs false-positive rate

(b)

PPV vs referral rate

Fig. 2. Performance of mammography screening programmes in Switzerland, (a) detection rate vs false-positive rate, (b) positive predictive value vs referral rate, by type of screening round.

the second-year proportional incidence was 50% greater in GE (67%) than in VD (42%).

4. Discussion

Swiss mammography screening programmes met most international standards of quality and effectiveness [8]. Performance was in line with other European service screening programmes [2,15]. However, up to a two-fold variation was observed in interim indicators of performance across these programmes despite the same screening regimen (double reading with arbitration, two-view biennial mammography, same length of operation) and common physical/technical quality control of radiological equipment.

The goal of any cancer screening initiative is to obtain a high cancer detection rate while minimising interval cancer and avoiding unnecessary diagnostic evaluation following false-positive results, which are costly and associated with psychological morbidity [1]. Consequently, a programme with a low referral rate and high interval cancer rate (high specificity/low sensitivity), like in The Netherlands [16], or a high referral rate and low interval cancer rate may reflect two screening policies or strategies with similar overall performance.

Screening policy and criteria are uniform in Swiss programmes [4] and indicators of programmes' sensitivity and specificity were weakly related (Fig. 2). For instance, one programme (VD) had *in the initial round* the lowest referral and false-positive rates, the lowest 2-year interval cancer rate (1.9/1000 vs 2.3/1000 (VS) and 2.5/1000 (GE)) but the highest detection rate. Regional variations

in cancer detection appear unlikely to be due to differential background risk since pre-screening levels of incidence differ by less than 5% in these regions (ranging from 290 to 303 per 100,000 for women aged 50–69) and self-selection of participants seems to occur on common factors across Swiss programmes (Swiss nationality, married and younger women) [17,18].

Although random fluctuations cannot be discarded, our results are compatible with a genuine difference in performance across Swiss programmes, at least in initial round. Higher performance was observed in the region (VD) that had both a prior pilot trial [3] and the highest annual volume per second reader. Our results show that expertise has successfully been transferred from an experimental setting to a service programme and confirm that quality improved in service programmes with duration of activity [7].

Performance indicators were closer for the two programmes (VD and GE) that elicited a restrictive second reading strategy. In Switzerland, accredited radiologists interpret screening mammograms both within and outside the setting of a service programme. Although the total annual throughput of radiologists was unknown, those reading less mammograms within a programme were likely to interpret more mammograms outside a programme as overall screening coverage and number of radiologists per capita were roughly comparable in the three regions [12]. These results confirm that performance *per se* is independent from the level of attendance [19]. In presence of substantial, concomitant opportunistic screening, the relationship between volume and accuracy remains difficult to measure in screening programmes.

Comparisons are generally more reliable for subsequent rounds since differences in women's prior screening practices (i.e. opportunistic screening) are largely attenuated after an initial round. Performance appears to vary mostly in subsequent rounds between the regional programme (VS) that favoured self-motivation and self-responsibility of radiologists to perform second readings and those programmes (VD and GE) that elicited a selective strategy for accreditating 2nd readers. The VS programme showed the highest participation but the lowest detection rate, with both fewer small cancers (<1 cm or *in situ*) and more advanced lesions being detected for rather similar proportional incidence of interval cancers as GE and VD. A restrictive accreditation of second readers led to a median throughput four times higher for second than first readers. Despite this measure, no Swiss programme reached the reading threshold of 5000 cases per year for second readers. Such a threshold has rarely been met in low-volume screening programmes run in decentralised healthcare systems [20]. To our knowledge, there is no specific recommendation aimed at optimising accuracy of mammography interpretation for low-volume screening programmes. Strong restriction on second readers' accreditation and/or a centralisation of breast assessment in liberal healthcare system can improve performance in low-volume programmes [21].

Even though indicators of early effectiveness (% of cancers *in situ*, ≤1 cm, N0) met the criteria of the European community, the proportion of screen-detected cancers of stage II+ was consistently higher than the reference value [8]. Similar proportions of advanced cancers have been observed in two high-quality Scandinavian service programmes, and a revision of the European criterion for the stage distribution has been proposed [22,23]. A trend analysis of advanced breast cancer rates by detection mode in these Swiss populations is warranted prior to any speculative interpretation.

The relative incidence of first-year but not second-year interval cancer met the levels set by the European guidelines [8]. Similar findings were reported in a pooled analysis of interval cancers in six European countries, as well as in Norway [15,24]. Apart from random variability due to few numbers in our series, an update of the current guidelines (based on the Swedish Two-County trial) has been suggested since they do not consider changes in

background incidence unrelated to the screening programme [22]. A high completeness of cancer registration, an accurate linkage procedure to identify interval cancers and a substantial prevalence of concomitant opportunistic screening have been evidenced to contribute to a greater frequency of interval cancer, regardless of screening quality [25]. All these factors apply to Switzerland.

A temporal bias has been advanced to explain the sharp increase between first and second-year proportional incidence [26]. The date of incidence mostly recorded by cancer registries is the date of the first positive histology report or hospital admission, and not that of the onset of symptoms. This shift in time of diagnosis led some first-year interval cancers to appear as second-year cancers in non-research setting.

To our knowledge, evaluation of timeliness in screening programmes has not been documented in the scientific literature. No Swiss programme reinvited 95% of screen-eligible women within 24 months, as recommended in the European guidelines [8]. Invitation letters do not specify a time or place in Switzerland and subsequent invitations are automatically sent 22-month after the prior screening test. Waiting time for a screening appointment in a radiology centre is a major contributor to this delay [9]. Delays do not affect screening performance *per se* but, if too lengthy, may incite women to be screened outside the programme and decrease attendance.

In conclusion, variation of performance across Swiss service programmes was related to difference in multiple reading strategies and a prior experience in service screening, but could not be explained by levels of opportunistic screening and programme attendance. The lesser second readers, the higher the screening performance. Performance improved over time in all programmes but variability in process indicators subsisted between programmes with and without a selective 2nd readers strategy. Centralisation of second readings to fewer radiologists should further improve mammography practice and screening performance in low-volume programmes.

Conflict of interest

None.

Acknowledgements

References

[1] International Agency for Research on Cancer. In: Vainio H, Bianchini F, eds. Breast cancer screening. IARC handbooks of cancer prevention, vol. 7. Lyon: IARC Press, 2002. 229 pp.

[2] Lynge E, Olsen AH, Fracheboud J, Patnick J. Reporting of performance indicators of mammography screening in Europe. Eur J Cancer Prev 2003;12:213–22.

[3] Bulliard J-L, De Landtsheer J-P, Levi F. Results from the Swiss mammography screening pilot programme. Eur J Cancer 2003;38:1760–8.

[4] Bulliard J-L, Ducros C, Jemelin C, Arzel B, Fioretta G, Levi F. Effectiveness of organised versus opportunistic mammography screening. Ann Oncol 2009; 20:1199–202.

[5] Blanks RG, Bennett RL, Wallis MG, Moss SM. Does individual programme size affect screening performance? Results from the United Kingdom NHS breast screening programme. J Med Screen 2002;9:11–4.

[6] Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, et al. International comparison of performance measures for screening mammography: can it be done? J Med Screen 2004;11:187–93.

[7] Giordano L, Giorgi D, Piccini P, Ventura L, Stefanini V, Senore C, et al. Time trends of process and impact indicators in Italian breast screening programmes: 1998–2007. Epidemiol Prev 2009;33:29–39.

[8] Europe Against Cancer. In: Perry N, et al., eds. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed., Luxembourg: Office for Official Publications of the European Communities, 2006. 416 pp.

[9] Bulliard J-L, Levi F. Evaluation épidémiologique du programme vaudois de dépistage du cancer du sein, 1999–2006. In: Raisons de santé. Lausanne: IUMSP; 2007. 56 pp.

[10] Curado M, Edwards B, Shin H, Storm H, Ferlay J, Heanue M, et al., editors. Cancer incidence in five continents. IARC Sci. Publ. No. 160 ed., vol. IX. Lyon: International Agency for Research on Cancer; 2007.

[11] Chamot E, Charvet AI, Perneger TV. Who gets screened, and where: a comparison of organised and opportunistic mammography screening in Geneva, Switzerland. Eur J Cancer 2007;43:576–84.

[12] Schmidlin K. Wie häufig und wie unterschiedlich werden Krebsvorsorgeuntersuchungen von der Schweizerischen Bevölkerung verwendet?—Eine Auswertung der Schweizerischen Gesundheitsbefragung 2002. Bern: Social and Preventive Medicine Institute; 2006. 57 pp.

[13] Bulliard J-L, Levi F. Evaluation épidémiologique du programme valaisan de dépistage du cancer du sein, 1999–2006. In: Raisons de santé. Lausanne: IUMSP; 2006. 38 pp.

[14] Swiss Cancer League and Oncosuisse. Request "reimbursement of screening mammography as a service of basic health insurance" to the Federal Office of Health. Bern: Swiss Cancer League and Oncosuisse; 2007. 127 pp.

[15] Törnberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B, et al. A pooled analysis of interval cancer rates in six European countries. Eur J Cancer Prev 2010;19:87–93.

[16] Otten JD, Karssemeijer N, Hendriks JH, Groenewoud JH, Fracheboud J, Verbeek AL, et al. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. J Natl Cancer Inst 2005;97:748–54.

[17] Bouchardy C, Fioretta G, Neyroud-Caspar I. Evaluation épidémiologique du programme genevois de dépistage du cancer du sein, 1999–2006. Geneva: Geneva Cancer Registry; 2009. 31 pp.

[18] Bulliard J-L, De Landtsheer J-P, Levi F. Profile of women not attending in the Swiss Mammography Screening Pilot Programme. Breast 2004;13:284–9.

[19] Broeders MJ, Scharpantgen A, Ascunce N, Gairard B, Olsen AH, Mantellini P, et al. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. Eur J Cancer Prev 2005;14:107–16.

[20] Autier P, Ouakrim DA. Determinants of the number of mammography units in 31 countries with significant mammography screening. Br J Cancer 2008;99:1185–90.

[21] Autier P, Shannoun F, Scharpantgen A, Lux C, Back C, Severi G, et al. A breast cancer screening programme operating in a liberal health care system: the Luxembourg mammography programme, 1992–1997. Int J Cancer 2002;97:828–32.

[22] Sarkeala T, Anttila A, Saarenmaa I, Hakama M. Validity of process indicators of screening for breast cancer to predict mortality reduction. J Med Screen 2005;12:33–7.

[23] Sarkeala T, Anttila A, Forsman H, Luostarinen T, Saarenmaa I, Hakama M. Process indicators from ten centres in the Finnish breast cancer screening programme from 1991 to 2000. Eur J Cancer 2004;40:2116–25.

[24] Hofvind S, Geller B, Vacek PM, Thoresen S, Skaane P. Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. Eur J Epidemiol 2007;22:447–55.

[25] Bulliard JL, Sasieni P, Klabunde C, De Landtsheer JP, Yankaskas BC, Fracheboud J. Methodological issues in international comparison of interval breast cancers. Int J Cancer 2006;119:1158–63.

[26] Bucchi L, Ravaioli A, Foca F, Colamartini A, Falcini F, Naldoni C. Incidence of interval breast cancers after 650,000 negative mammographies in 13 Italian health districts. J Med Screen 2008;15:30–5.