Problems, solutions, and perspectives in the evaluation of interval cancers in Italian mammography screening programmes: a position paper from the Italian group for mammography screening (GISMa)

Problemi, soluzioni e prospettive nella valutazione dei cancri d'intervallo nei programmi italiani di screening mammografico: un position paper del Gruppo italiano screening mammografico (GISMa)

Lauro Bucchi,[1] Alfonso Frigerio,[2] Manuel Zorzi,[3] Chiara Fedato,[4] Giovanni Angiolucci,[5] Daniela Bernardi,[6] Cinzia Campari,[7] Emanuele Crocetti,[8] Stefano Ferretti,[9] Daniela Giorgi,[10] Francesca Marchisio,[11] Doralba Morrone,[12] Carlo Naldoni,[13] Marco Petrella,[14] Antonio Ponti,[15] Alessandra Ravaioli,[1] Gianni Saguatti,[16] Dolores Santini,[11] Priscilla Sassoli de Bianchi,[13] Monica Serafini,[17] Viviana Vergini,[15] Livia Giordano[15]

This article is dedicated to the memory of Stefano Ciatto

[1]Registro tumori della Romagna, Istituto scientifico romagnolo per lo studio e la cura dei tumori (IRST) IRCCS, Meldola, Forlì

[2]Centro di riferimento regionale per lo screening mammografico, Torino

[3]Registro tumori del Veneto, Regione Veneto, Padova

[4]Coordinamento regionale screening, Regione Veneto, Venezia

[5]Senologia diagnostica, Azienda unità sanitaria locale, Arezzo

[6]Unità operativa di senologia clinica e screening mammografico, Azienda provinciale servizi sanitari, Trento

[7]Centro screening, Azienda unità sanitaria locale, Reggio Emilia

[8]Unità di epidemiologia clinica e descrittiva, Istituto per lo studio e la prevenzione oncologica, Istituto tumori toscano, Firenze

[9]Registro tumori di Ferrara, Ferrara

[10]Unità di epidemiologia, Azienda sanitaria locale, Istituto tumori toscano, Lucca

[11]Centro screening mammografico, Azienda unità sanitaria locale, Modena

[12]Unità operativa di senologia, Istituto per lo studio e la prevenzione oncologica, Istituto tumori toscano, Firenze

[13]Assessorato alle politiche per la salute, Regione Emilia-Romagna, Bologna

[14]Unità operativa di epidemiologia, Azienda unità sanitaria locale 2 dell'Umbria, Perugia

[15]Unità di epidemiologia dei tumori, Centro di riferimento per l'epidemiologia e la prevenzione oncologica in Piemonte, AOU Città della salute e della scienza, Torino

[16]Unità operativa di senologia, Azienda unità sanitaria locale, Bologna

[17]Centro di prevenzione oncologica, Azienda unità sanitaria locale, Ravenna

Corresponding author: Lauro Bucchi, lauro.bucchi@irst.emr.it

Abstract

In this position paper, a self-convened team of experts from the Italian Group for Mammography Screening (Gruppo italiano screening mammografico, GISMa) pointed out the problems that increasingly hamper the feasibility and validity of the estimate of the proportional incidence of interval breast cancer (IBC) in Italy, suggested potential solutions and an agenda for research, and proposed that the question of the sensitivity of mammography be viewed in a larger perspective, with a greater attention to radiological review activities and breast radiology quality assurance programmes.

The main problems are as follows: the coverage of cancer registration is incomplete; the robustness of using the pre-screening incidence rates as underlying rates decreases with time since the start of screening; the intermediate mammograms performed for early detection purposes may cause an overrepresentation of IBCs; the classification of many borderline screening histories is prone to subjectivity; and, finally, the composition of cohorts of women with negative screening results is uncertain, because several mammography reports are neither clearly negative nor clearly positive, and because of the limitations and instability of the electronic mammography records.

Several possibilities can be considered to cope with these issues: standard methods for using the hospital discharge records in the identification of IBCs should be established; for the calculation of regional estimates of the underlying incidence, a suitable mathematical model should be identified; the definition of IBC according to the 2008 GISMa guidelines needs to be updated, especially with respect to in situ cancers and to invasive cancers with borderline screening histories; a closer adherence to standard screening protocols, with a simplified patient management, would make it easier to objectively identify IBCs; alternative methods for estimating the sensitivity of mammography should be taken into consideration; and, finally, analysis could be restricted to the absolute incidence rate of IBC, which would make comparison of the risk between neighbouring populations possible.

Epidemiologists must extend their attention to the prevention of the risk of IBC and the implementation of breast radiology quality assurance practices. Epidemiologists and radiologists can share common objectives: it is necessary to promote the idea that the availability of a registry-based series of IBCs is not a prerequisite for their radiological review; radiological review of breast cancers greater than 20mm in size detected at second and subsequent screens, that are potential substitutes for IBCs, needs radiological and epidemiological validation studies; the advent of digital mammography brings about the possibility to create libraries of mammograms accessible online, which enables the conduct of large studies of the diagnostic variability of radiologists; and, finally, epidemiologists and radiologists have the responsibility to monitor the effects that a loss of cumulative professional experience in screening centres, due to the imminent retirement of a substantial proportion of healthcare workforce, could cause on their performance.

Riassunto

In questo position paper, un team spontaneo di esperti associati al Gruppo italiano screening mammografico: (1) puntualizza i limiti metodologici e i fattori distorsivi che compromettono la valutazione dell'incidenza dei cancri d'intervallo nei programmi di screening in Italia, (2) suggerisce le possibili soluzioni e un'agenda per la ricerca, e (3) propone che il problema dei cancri d'intervallo sia inserito in una prospettiva più ampia, con una maggiore attenzione per le attività di revisione radiologica e per i programmi di quality assurance in radiologia senologica.
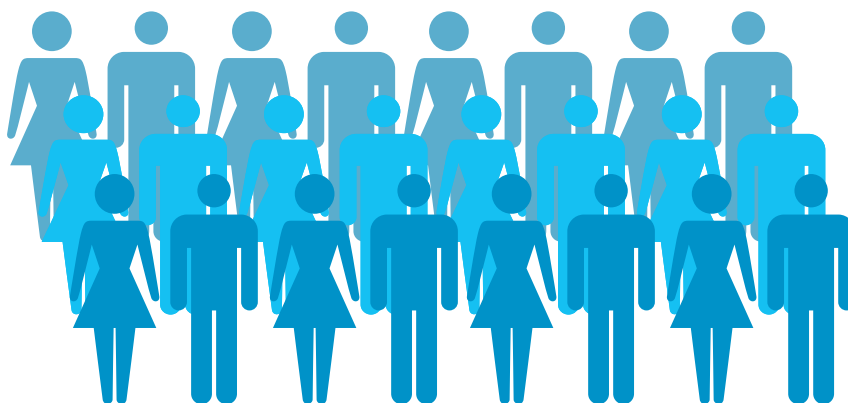
INTRODUCTION

The sensitivity of mammography is a major factor for the effectiveness of a breast screening programme. The reference method to evaluate the sensitivity of mammography is based on the estimate of the proportional incidence of interval breast cancer (IBC).

IBCs are cancers diagnosed after a negative mammography result and before next invitation to screening, or within two years if the woman has reached the age for screening cessation. The proportional incidence of IBC is the incidence observed during the screening interval as compared to the incidence that would be expected in the absence of screening, or underlying incidence. This proportion gives an approximation of the rate of mammography failures in abolishing the incidence of breast cancer during the screening interval. In other words, the proportional incidence of IBC is equal to 1 – the sensitivity of mammography.

According to the 2008 guidelines from the Italian group for mammography screening (Gruppo italiano screening mammografico, GISMa),[1] the scientific society that gathers all professionals involved in any aspect of mammography screening in the country, the performance indicators of every screening programme must include the absolute and proportional rates of IBC, as well as the rate of IBCs interpreted to be visible on retrospective radiological review. In the epidemiological guidelines chapter of the *European guidelines for quality assurance in breast cancer screening and diagnosis*,[2] the estimate of the proportional incidence of IBC is among the impact indicators, although it is stated that it suffers from «several limitations».

This position paper originated from an initiative of members of GISMa's Coordinating Committee, who drafted a working document and asked for amendments and proposals from epidemiologists and radiologists members of the society. The paper aims at:

■ pointing out the problems that increasingly hamper the feasibility and validity of the estimate of the proportional incidence of IBC in Italy;

■ suggesting potential solutions and an agenda for research;

■ proposing that the question of IBC be viewed in a larger perspective, with a greater attention to radiological review activities and breast radiology quality assurance programmes.

The authors of this paper will submit a set of essential proposals to the incoming Coordinating Committee of the GISMa.

PROBLEMS

The problems that affect the estimate of the proportional incidence of IBC can be summarized as follows.

■ With respect to the identification of IBCs, the main limitations are the incomplete coverage of cancer registration and the delay – of a few years – by which the annual case series are completed. The only available alternative is to create efficient special breast cancer registries, whether based on standard methods of cancer registration or hospital discharge records. This can also be done by the screening centres themselves. GISMa guidelines accepted the use of hospital discharge records, although they stated that developing standard methods was an urgent need.[1] To this end, they proposed the formation of a workgroup.
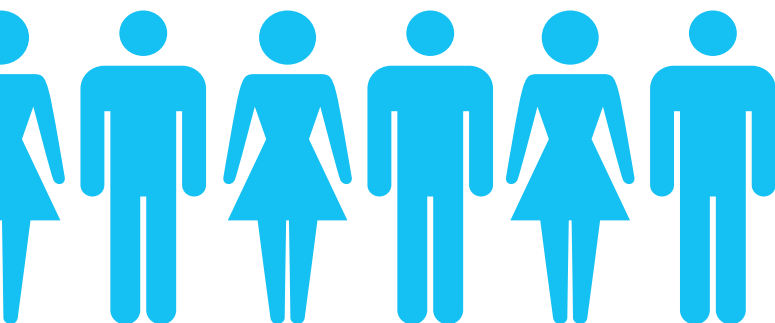
■ The robustness of using the pre-screening incidence rates as underlying rates decreases with time since the start of the screening programme. It is unsafe both to assume that those rates, if not modified by screening, would have been stable over time, and to linearly extrapolate them to the present time. This limitation is mentioned in the epidemiological guidelines chapter of the European guidelines.[2] The 2008 GISMa guidelines suggested the calculation and use of regional incidence estimates.[1] These too were defined as an urgent need.

■ Intermediate mammograms performed at clinical radiology facilities for early detection purposes may cause an overrepresentation of IBCs. They lead to the detection of asymptomatic cancers that cancer registries, if lacking information on their actual clinical status, inevitably classify as IBCs. The same may happen following intermediate mammograms actively offered within the screening programmes (early rescreen), if they are recorded as diagnostic examinations rather than true screening examinations. It is an epidemiological paradox that the practice of performing intermediate mammograms, while increasing the sensitivity of mammography for early breast cancer, causes apparently the opposite effect.

■ GISMa guidelines took into consideration the question of whether the definition of interval cancer may include the cancers diagnosed during the third interval year or later, or after a negative or an inconclusive assessment, or after a woman's refusal to undergo assessment, or after discontinuation of participation in the programme, or after a previous diagnosis of breast cancer.[1] The definition of IBC was expanded to include some of these screening histories, but their interpretation in a real-world screening setting remains prone to subjectivity.

■ Another source of variability is the eligibility of in situ breast cancers, which is interconnected with the problem of their registration. GISMa guidelines suggested excluding in situ breast cancers from the estimate of the proportional incidence of IBC, given that they are incompletely registered and given their benign and generally non-progressive behaviour. Nevertheless, the guidelines recommended that interval in situ breast cancers known to the screening centres be subject to radiological review.[1]

■ Along with the diffusion of mammography screening into widespread use, the procedure has become increasingly het-

erogeneous and complex. This change is connected to the emerging idea of an individually tailored screening.[3] One of the most notable consequences of this is that the classic dichotomous classification of mammography results has been abandoned in certain screening programmes and in certain circumstances. More and more often there are borderline mammography reports that are neither clearly negative nor clearly positive. In the estimate of the proportional incidence of IBC, this introduces a degree of uncertainty both in the composition of cohorts of women with negative screening results and in the detection mode of incident breast cancers.

■ The composition of cohorts of women with negative screening results is also uncertain because of the limitations and instability of the electronic mammography records. Screening centres are equipped with a variety of computer systems and softwares. Many of these are designed solely for the delivery of the service, not for the evaluation of results.
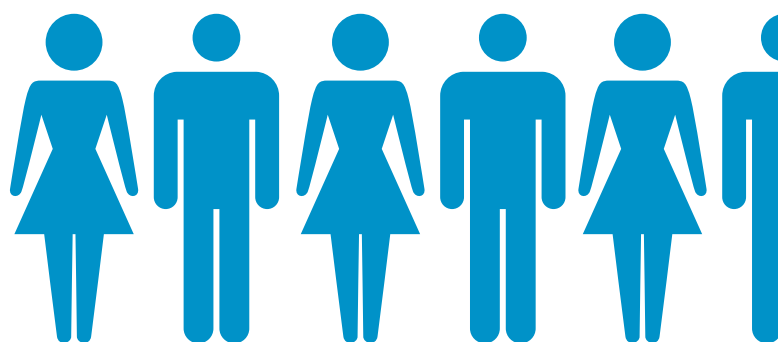
SOLUTIONS

Several possibilities can be considered to resolve these issues, at least to a certain extent.

■ The GISMa guideline recommending that a workgroup be appointed to establish standard methods for using the hospital discharge records in the identification of IBCs[1] should be implemented. The workgroup can be comprised of those epidemiologists who are currently using the hospital discharge records as a basis for registration.

■ As far as the underlying incidence rates are concerned, the GISMa guideline recommending the calculation and use of regional estimates[1] remains valid. It can be suggested to GISMa's Coordinating committee to formally ask the National centre for screening monitoring (Osservatorio nazionale screening, ONS) to examine the mathematical models that are being used to estimate breast cancer incidence, and to select the most suited one.

■ Certain issues of the 2008 GISMa guidelines[1] need to be updated. In particular, it would be advisable to re-examine the eligibility of in situ cancers and of invasive cancers diagnosed during the third interval year or later, or after a negative or an inconclusive assessment, or after a woman's refusal to undergo assessment, or after discontinuation of participation in the programme, or after a previous diagnosis of breast cancer. Moreover, the chapter on the definition of IBC should include a definition of what a negative mammography result is, taking the problem of borderline screening histories into consideration. Epidemiologists with previous experience in the classification of IBC detection modes should compare their methods with each other and with the radiologists' point of view.

■ Theoretically, an option to objectively classify IBC detection mode is to draw the attention of screening units to the opportunity of adhering more closely to standard screening protocols. A simplified patient management would make it easier to identify IBCs and – no less important – limit the diffusion of unplanned forms of individually tailored screening. This could be coupled with an effort to standardize the nomenclature used in mammography reports as well as their format, at least on a regional scale.

■ An innovative approach to the evaluation of the sensitivity of mammography, which is commonly referred to as the unbiased set method,[4] is not to use estimates of the underlying incidence nor pre-screening incidence rates. The method requires the availability of a general or a special cancer registry and of information on the detection mode of registered breast cancers. However, it uses only screen-detected cancers (except those detected in the prevalence screen) and IBCs. The method was explicitly proposed for screening programmes of long duration, which is the case for most programmes in Italy. It could be suggested to GISMa as well as ONS to consider adopting the unbiased set method as a reference method.

■ A minimalist approach to the evaluation of the incidence of IBC, which has already been advised by European guidelines,[2] would be to restrict analysis to the absolute incidence rate. On the one hand, this would mean neglecting the estimate of mammography sensitivity. On the other hand, however, it would allow the risk of IBC to be compared between neighbouring populations (for example, those living in different health care districts of an administrative region) who can be assumed to have the same underlying breast cancer incidence. This would also provide radiologists with a practical self-evaluation tool.

■ Until workable and effective solutions are found, the limitations in estimating the proportional incidence of IBC need to be well understood across the health system. The present paper aims at preventing the use of currently available estimates for legal and administrative purposes.

■ The same caution should be used in public communication concerning the harms of mammography screening, which is recommended by European guidelines.[5] In the presentation of screening programmes (public advertising campaigns and invitation letters), information on false-negative mammography results is insufficient. However, the information material should simply state that false-negative results are possible, and should describe the radiology facility characteristics that may influence the accuracy of diagnosis (for example, the range of annual screening mammogram reading volume of local radiologists). Numerical estimates of the sensitivity of mammography, which are poorly reliable and difficult to communicate, must be avoided.

■ Lastly, we suggest a change in the scientific paradigm that has
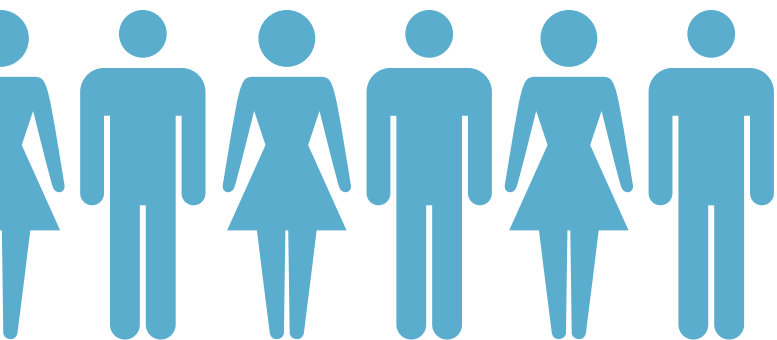
so far underlain IBC evaluation. GISMa guidelines recommend not only to estimate the proportional incidence of IBC, but also to retrospectively review the mammograms.[1] More attention and resources should be devoted to the reviewing process. The value of radiological review, both for quality assurance and continuing education purposes, is repeatedly emphasized by European guidelines.[2,6]

■ Approaching the problem of IBCs from the perspective of breast radiology quality assurance would give practical implementation to a 2008 document from the Ministry of Health (Direzione generale della prevenzione sanitaria del Ministero della salute) in which it was stated that the registration of IBCs should be accompanied by actions aimed at increasing the levels of quality of the screening process.[7] The document suggested that the review process be done in a semi-informed manner, which has a greater educational impact in that it focuses on criteria for women's recall and not on medico-legal evaluations. For medico-legal purposes, the reference method is a blinded review of mammograms performed by a group of non-informed expert consultants from a national certified register.

PERSPECTIVES

Epidemiologists must extend their attention to the prevention of the risk of IBC, and consider that the implementation of breast radiology quality assurance practices can be a common point of interest with radiologists. The proportional incidence of IBC, which is generally calculated at the screening programme level and not at the single radiologist level, does not provide clues to improve the sensitivity of mammography, because it has no specific feedback on the diagnostic performance. Conversely, radiological review of IBCs has a direct educational impact.[8] The estimate of the proportional incidence of IBC is a good example of how the descriptive epidemiology of cancer provides valuable information about the size of problems, but often without the capacity to make a real contribution to cope with them.[9] Unfortunately, in the long run, a descriptive work that fails to promote appropriate actions loses its rationale.

Following this line of reasoning, it must be noted that the annual GISMa surveys of results of mammography screening in Italy have shown for years a situation where the recall at second and subsequent screens is above the acceptable standard of 5% for one-third of local programmes, and where the prac-

tice of performing intermediate mammograms is widespread.[10] This would require regular training programmes that are currently insufficient.

The following are some suggestions on how epidemiologists and radiologists can interact positively and fruitfully.

■ The absence of radiological review activities in those screening centres that are served by a general or a special cancer registry is an unacceptable situation, in addition to being an original type of underuse of cancer registration.[9] Where this occurs, epidemiologists and radiologists should work together to find a solution.

■ At the same time, both epidemiologists and radiologists should promote the idea that the availability of a complete, registry-based series of IBCs is a prerequisite only for estimating their proportional incidence, not for their radiological review. Besides, this should be done as soon as an IBC is detected or becomes known to the screening centre.

■ GISMa guidelines[1] and a study from the screening unit of Trento[11] have supported the radiological review of breast cancers greater than 20 mm in size detected at second and subsequent screens. GISMA's Workgroup on diagnosis (Gruppo di lavoro area diagnosi) has proposed, in particular, that these cancers be used as substitutes for IBCs in radiological review activities at those screening centres where reviewing IBCs is problematic.[12] The radiological review of screen-detected breast cancers greater than 20 mm in size is potentially feasible on a national scale and would make it possible to set up true national standards for all screening programmes in the country. This approach, however, requires radiological and epidemiological validation studies.

■ Screening units that have already estimated the proportional incidence of IBC should be encouraged to determine whether, in their data, there is a relationship between the sensitivity of mammography and the prevalence of breast cancers greater than 20 mm in size detected at second and subsequent screens.[13]
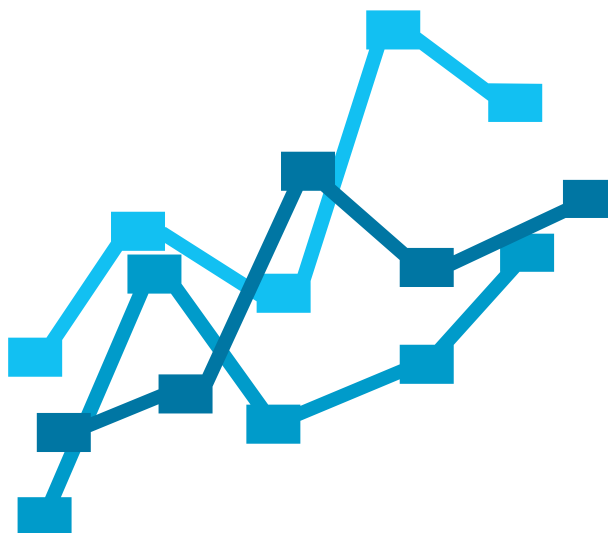
■ The advent of digital mammography has brought about the possibility to create, in conjunction with central radiological review activities, libraries of mammograms accessible online. Although there remains the problem of obtaining the informed consent of patients, radiologist access to reviewed mammograms would represent an important opportunity for research and training. Online libraries could be completed with images representing a larger spectrum of mammographic abnormalities. Epidemiologists could contribute to these developments by designing studies of the radiologist variability in interpretation of mammography findings.

■ Between mid-2013 and early 2014, GISMa's Coordinating committee carried out a national questionnaire survey of radiologist's experience-related characteristics (for example, annual screening mammogram reading volume, and the percentage of working time devoted to breast radiology). The survey is particularly topical given that the budget constraints that the Italian National Health Service is facing may lead to increasing flexibility of mammogram-reading teams, as has been reported from the United Kingdom.[14] The data from the survey will have to be evaluated both from a radiological and an epidemiological point of view, because they could support the hypothesis that radiologist's experience-related characteristics are associated with current results of local screening programmes.[10]

■ In the 1980s, the implementation of the new National Health Service, coupled with an aggressive policy of deficit spending, originated a dramatic increase in the number of medical and paramedical staff in the public sector. The imminent retirement of this workforce makes it urgent to assess the professional and training needs of screening centres. The basic role of high-level specific training would suggest that part of the resources currently devoted to mammography screening programmes be allocated to the creation and maintenance of a limited number of multidisciplinary national training centres, following the experience of other European countries. Epidemiologists are responsible, in collaboration with radiologists, for monitoring the effects that a loss of cumulative professional experience in screening centres could cause on their performance.

Conflicts of interests: none declared

References/Bibliografia

1. Ciatto S, Naldoni C, Ponti A et al. Interval cancers as indicators of performance in screening programmes. *Epidemiol Prev* 2008;2:93-98.
2. Broeders M, Nyström L, Ascunce N et al. Epidemiological guidelines for quality assurance in breast cancer screening. In: Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds). *European guidelines for quality assurance in breast cancer screening and diagnosis. 4th edition.* Office for Official Publications of the European Communities, Luxembourg, 2006:15-56.
3. Paci E, Mantellini P, Giorgi Rossi P et al. Tailored Breast Screening Trial (TBST). *Epidemiol Prev* 2013;4-5:317-27.
4. Zorzi M, Guzzinati S, Puliti D, Paci E. A simple method to estimate the episode and programme sensitivity of breast cancer screening programmes. *J Med Screen* 2010;17(3):132-38.
5. Giordano L, Webster P, Segnan N, Austoker J. Guidance on breast screening communication. In: Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds). *European guidelines for quality assurance in breast cancer screening and diagnosis. 4th edition.* Office for Official Publications of the European Communities, Luxembourg, 2006:379-94.
6. Rosselli del Turco M, Hendriks J, Perry N et al. Radiological guidelines for quality assurance in breast cancer screening. In: Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds). *European guidelines for quality assurance in breast cancer screening and diagnosis. 4th edition.* Office for Official Publications of the European Communities, Luxembourg, 2006: 181-95.
7. Ministero della salute, Direzione generale della prevenzione sanitaria. *Programma di screening mammografico: significato e gestione dei cancri d'intervallo. Seconda Edizione.* Roma, 2008. [http://www.senologiasirm.org/index.php?option=com_content&view=article&id=12&Itemid=15].
8. Houssami N, Irwig L, Ciatto S. Radiological surveillance of interval breast cancers in screening programmes. *Lancet Oncol* 2006; 7(3):259-65.
9. Armstrong BK. The role of the cancer registry in cancer control. *Cancer Causes Control* 1992;3(6):569-79.
10. Giorgi D, Giordano L, Ventura L et al. Lo screening mammografico in Italia: survey 2010. *Epidemiol Prev* 2012;6(Suppl.1):8-27.
11. Ciatto S, Bernardi D, Pellegrini M et al. Proportional incidence and radiological review of large (T2+) breast cancers as surrogate indicators of screening programme performance. *Eur Radiol* 2012; 22(6):1250-54.
12. Verbale Gruppo di lavoro area diagnosi, Torino 2013, Convegno Nazionale GISMa. [http://www.gisma.it/index.php?view=article&catid=38%3Agruppi-di-lavoro&id=263%3Averbale-gruppo-di-lavoro-area-diagnosi&option=com_content&Itemid=81].
13. Zorzi M, Fedato C, Baracco S. Confronto tra pT2+ agli esami successivi e cancri intervallo per la stima della sensibilità. Convegno Nazionale GISMa 2013: Sessione poster. [http://www.gisma.it/index.php?option=com_content&view=article &id=258:convegno-nazionale-gisma-2013-sessione-poster-&catid=39:atti-dei-convegni&Itemid=129].
14. Scott HJ, Gale AG. Breast screening: PERFORMS identifies key mammographic training needs. *Br J Radiol* 2006;79(2):S127-33.