Use of Routine Health Datasets to Assess the Appropriateness of Diagnostic Tests in the Follow-Up of Breast Cancer Patients: A Population-Based Study on 3930 Patients

Massimo Gion1, Giulia Cardinali2, Stefano Guzzinati3, Paolo Morandi4, Chiara Trevisiol5, Aline SC Fabricio5, Massimo Rugge3,6, Manuel Zorzi3

1Regional Center for Biomarkers, Department of Clinical Pathology, Azienda ULSS 3 Serenissima, Venice, Italy; 2Management Control Unit, Azienda ULSS 3 Serenissima, Venice, Italy; 3Veneto Tumour Registry, Azienda Zero, Padua, Italy; 4Medical Oncology Unit, Azienda ULSS 3 Serenissima, Venice, Italy; 5Veneto Institute of Oncology IOV - IRCCS, Padua, Italy; 6University of Padova, Department of Medicine DIMED, Padua, Italy

Correspondence: Massimo Gion, Regional Center for Biomarkers, Department of Clinical Pathology, Azienda ULSS 3 Serenissima, Ospedale SS. Giovanni e Paolo – Castello, Venezia, 6777 – 30122, Italy, Tel +39 041 5294260, Fax +39 041 5295603, Email massimo.gion@aulss3.veneto.it

Purpose: Clinical practice guidelines (CPGs) recommend against intensive follow-up in asymptomatic women with breast cancer (BC). The present study assessed the adherence to CPGs of diagnostic tests ordering during BC follow-up by exploring routinely collected health data through an algorithm developed to distinguish patients according to their status at follow-up.

Patients and Methods: A retrospective population-based cohort study was performed monitoring the diagnostic tests ordered during 5 years of follow-up in all BC cases incident in 2013 in the Veneto Region, Italy. Data were extracted from the Veneto Tumour Registry, the Hospital Discharge Records and the Outpatients’ Records of Diagnostic and Therapeutic Procedures. The algorithm was developed using information on infusion of anticancer agents, imaging exams ordered, and death.

Results: The algorithm classified patients by status at follow-up in four groups: (i) probably no-evidence-of-disease (NED), (ii) suspicious signs of relapse not confirmed, (iii) increased risk of relapse and (iv) advanced disease at presentation or progressive disease. A total of 3930 consecutive incident cases were followed-up for 5 years, corresponding to 17,184 person-years, 15,345 of which pertaining to NED cases. In NED cases, 32,900 tumour markers and 15,858 imaging exams were ordered. Liver ultrasonography and chest radiography were most frequently ordered.

Conclusion: In contrast with recommendations of CPGs, a substantial overordering of tumour markers and imaging exams occurred in NED BC patients. The developed algorithm can be repeatedly applied to routine health datasets for regular monitoring of the adherence to CPGs and of the impact of interventions to improve appropriateness.

Keywords: tumour markers, imaging exams, adherence to guidelines, routinely collected health data

Introduction

Breast cancer (BC) is the most frequent malignancy among women in both developed and developing countries, with an estimated 2 million new cancer cases diagnosed worldwide in 2018 (23% of all cancers).1 In developed countries, the concomitant increase of incidence and decrease in mortality lead to a progressive increase of prevalent BC cases.2 In Italy, 834,154 prevalent BC cases were registered in 2020, which represent the 43.4% of all cancers in women and the 23.1% of all cancer in both sexes.3

In the frame of BC management, long-term postoperative follow-up involves a significant portion of prevalent cases for a time length ranging from 5 to 10 years, thus leading to a substantial workload of health-care organizations. Several studies showed that follow-up based on imaging other than mammography and/or blood tests does not influence patient survival nor quality of life of patients operated for breast cancer, regardless of location of services.4 7 Systematic reviews of randomized
clinical trials confirmed that follow-up schemes based on regular physical examinations and yearly mammography alone are as effective as more intensive approaches based on regular performance of laboratory and instrumental tests in terms of timeliness of recurrence detection, overall survival and quality of life. On the basis of available evidence, clinical practice guidelines (CPGs) advise that follow-up of asymptomatic patients should consist of routine annual mammographic screening and clinical evaluation every 6 months in the first 5 years after surgery, whereas they recommend against intensive follow-up, including tumour markers (TMs) and/or imaging other than mammography, in asymptomatic women. However, the adherence of clinical practice to CPGs has been reported to be in general low, including inappropriate use of diagnostic exams.

Accordingly, studies performed in Italy at a nationwide scale comparing ordered TMs with epidemiological data showed that ordered CA15.3 are higher than expected if CPGs were respected. A key element of quality monitoring is the availability of standardized measures of care. Some metrics based on evidence from scientific literature and CPGs, which should be routinely assessed with reference to predefined standards, have been developed by professional and health-care provider organizations. As concerns follow-up, indicators have been designed to measure the appropriate utilization of the recommended approach, but no indicators have been proposed to assess the inappropriate overordering of not recommended tests. Most likely, indicators have not been developed as it is not possible to establish whether the use of diagnostics tests during follow-up is appropriate or not, because prevalent cases include relapsed patients, in which the use of diagnostic tests is allowed.

The aim of this study was to analyze the adherence to CPGs of TMs and imaging exams ordering during post-operative follow-up of a cohort of BC women. An ad hoc algorithm built with routinely collected health data was used to define the clinical status of study patients.

Materials and Methods

Study Design and Patients

A retrospective population-based cohort study was performed analysing data from all women with a new diagnosis of infiltrating BC occurred in 2013 in the Veneto Region (Italy) and monitoring in those patients all the diagnostic tests ordered in the first 5 years of follow-up after diagnosis.

Sources of Data

The examined computerized sources of data were (i) the Veneto Tumour Registry (VTR), (ii) the Hospital Discharge Records (HDRs) and (iii) the Outpatients’ Records of Diagnostic and Therapeutic Procedures and prescriptions of drugs reimbursed by the National Health Service (ORDTP). HDRs contain one entry for every episode of hospital admission, discharge and readmission with each entry indicating an anonymous individual code for tracking patients. ORDTP contains codes pertaining to all the diagnostic procedures and therapeutic interventions administered to outpatients within the framework of the Italian National Health Service.

Data Mining and Extraction

Primary BC cases diagnosed from 01/01/2013 to 31/12/2013 were identified by the VTR. Through linkage with HDRs, cases undergoing BC surgery were identified. Each patient was followed from the date of surgery to either the end of the 5th year of follow-up or when the patient was labelled as relapsed or deceased.

Record codes referring to infusion of anticancer agents (IAA), orders of imaging exams, CA15.3 and CEA were acquired from ORDTP from 01/01/2013 (01/01/2012 for IAA) to 31/12/2018. Chest radiography and liver ultrasonography were classified as basic imaging exams, while computed tomography, magnetic resonance other than breast, positron-emission tomography, and whole-body bone scanning were classified as advanced imaging exams. TMs and imaging exams were analysed per year of follow-up and overall, and reported separately in the different patient subgroups as defined by the developed algorithm. The complete study flow is schematically outlined in Box 1.
Algorithm Development

The following specific information registered in the ORDTP were considered to develop the algorithm: IAA in the year before surgery; any new IAA between months 25 and 60 after surgery; number of repeated advanced diagnostic imaging exams (computed tomography and/or magnetic resonance and/or positron-emission tomography and/or whole-body bone scanning) ordered every year; and death. The above information were combined to classify patients by status at follow-up in four different groups: (i) cases suspected to have either advanced disease at presentation or a progressive disease (hereinafter referred to as “Advanced-Relapsed”), (ii) cases with suspicious signs of relapse, not further confirmed (“Suspicious not confirmed”), (iii) cases at increased risk of relapse having had a more advanced or aggressive primary tumour (“no-evidence-of-disease [NED] post neoadjuvant chemotherapy [CHT]”), and (iv) cases probably without evidence of disease (“NED”). The number of cases in the different groups varies during follow-up, as patients can either shift from NED to Advanced-Relapsed and Suspicious not confirmed groups, or deceased. A detailed description of the combination of the selected criteria used to classify patients in the above groups is reported in Supplementary material – Methods.

Assessment of Requested TMs and Imaging Exams

The number of TMs and the type and number of imaging exams ordered over the 5-year follow-up were extracted and examined in the different groups. Advanced-Relapsed patients were excluded because diagnostic tests are required to monitor disease progression and response to therapy. Among patients with at least one TM ordered, the time elapsed from the first TM order per year and the first order of an imaging exam in the same year was analysed.

Criteria to Assess Intensive Follow-Up

According to CPGs, TMs and imaging exams other than mammography ordered during the follow-up of asymptomatic patients were considered inappropriate and the follow-up of these patients was classified as intensive.
Statistical Analysis

Descriptive statistics were used to examine the distribution of TMs and imaging within the entire cohort and in the 4 groups categorized by the algorithm.

The numbers and percentages of ordered diagnostic tests are presented with reference to either the number of patients within each observation year, or the person-years of follow-up for the whole observation period of 5 year.

The comparison of the average number of examinations was tested with the Mann–Whitney test for two groups and with the Kruskal-Wallis test for three or more groups. Non-parametric tests were used due to data not normally distributed. Pearson’s chi-square test was used to compare proportions among groups. The SAS EG v.6.1 (SAS Institute Inc., Cary, NC, USA) statistical package was used for all analyses. All statistical tests were two-tailed. A p-value <0.05 was considered statistically significant.

Ethical Issues/Statement

The study was carried out in accordance with the principles established in the Declaration of Helsinki. The study did not need specific authorisation from an Ethics Committee as it is a retrospective observational study based on cases routinely collected by the Cancer Registry of the Veneto Region, which is authorized by the Italian Guarantor for the Protection of Personal Data to use health-related data for research purposes. All data were anonymised, following the Italian regulations, and were handled for monitoring and quality assurance purposes.

Results

Selection of Study Patients

A total of 4403 new BC cases were diagnosed in women residing in the Veneto Region from 01/01/2013 to 31/12/2013. Fifteen cases were excluded because registered on death certificate only. In the remaining 4388 cases microscopic evidence was available in 4347 of them, while 41 cases were recorded on the basis of non-microscopic clinical information. After linkage with the HDR database, 409 out of 4388 cases were excluded because they did not undergo any BC surgery: 58 (1.3%) matched with codes of BC surgical procedures other than those defined for BC treatment, 183 (4.1%) resulted hospitalized for surgery of diseases not involving breast and 168 (3.8%) were not logged in HDRs. Forty-nine additional cases within the first year were excluded from the follow-up study. Overall, 3930 patients were evaluated and contributed with 17,184 person-years of follow-up.

Estimate of Patients’ Status at Follow-Up

Information on diagnostic and therapeutic procedures available in ORDTP was examined to estimate patients’ status at follow-up. Details are reported in Supplementary material - Methods.

A sample of the flow chart showing the selection algorithm developed on the above assumptions is shown in Figure 1 (the complete flow chart is reported in Supplementary figure S1).

In summary, patients were classified in the four groups according to their status at follow-up as follows:

(A) Advanced-Relapsed. Patients suspected to have either advanced disease at presentation or a progressive disease. These patients were not considered for the assessment of intensive follow-up;

(B) Suspicious not confirmed. Patients in which 2 or more advanced imaging exams were ordered within a year but no additional advanced imaging was ordered in the next year. The lack of further advanced imaging suggests that the possible suspicious of relapse was not confirmed. These patients possibly required a punctual diagnostic work-up;

(C) NED post-neoadjuvant CHT. Women who received neoadjuvant therapy as bearing a larger or more aggressive primary tumour, but did not receive 2 or more advanced imaging exams in close sequence nor IAA during their follow-up. According to current clinical practice in 2013, these patients may have received a more intensive follow-up;
(D) NED. Patients probably without evidence of disease who did not receive 2 or more advanced imaging exams in close sequence nor IAA during their follow-up.

Table 1 shows the number of patients classified with different status at follow-up in each year of follow-up. NED patients represent the most relevant fraction of prevalent cases with 3363/3930 in the first year and 2842/3285 in the fifth year. Monitoring NED patients involved 15,345 out of 17,814 (86.1%) censored person-years of follow-up. Patients with advanced disease at presentation in year 1 or progressive disease from year 1 to year 5 were 466/3930 (11.8%). Patients alive at inclusion in the study and deceased afterwards for any cause over the 5 years of follow-up were 412.
Table 1  Number of Patients with Different Status at Follow-Up in Every Year of the Study

<table>
<thead>
<tr>
<th>Status at Follow-Up</th>
<th>N. of Patients Followed-Up per Year*</th>
<th>Total Person-Years Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year 1</td>
<td>Year 2</td>
</tr>
<tr>
<td>Advanced-Relapsed</td>
<td>203</td>
<td>62</td>
</tr>
<tr>
<td>Suspicious not confirmed</td>
<td>244</td>
<td>62</td>
</tr>
<tr>
<td>NED post-neoadjuvant CHT</td>
<td>120</td>
<td>3.1</td>
</tr>
<tr>
<td>NED</td>
<td>3363</td>
<td>85.6</td>
</tr>
<tr>
<td>Total</td>
<td>3930</td>
<td>100</td>
</tr>
</tbody>
</table>

Notes: *The number of cases in the different groups varies year by year as patients either shift from NED groups to “Relapsed” and “Suspicious not confirmed” or deceased. 
*In the 5th year “Advanced-Relapsed” are possibly underestimated since confirmatory data from the 6th year were not available due to the duration of the study. 
**Abbreviations:** NED, no-evidence-of-disease; CHT, chemotherapy.

Assessment of Putative Appropriateness of Diagnostic Tests Orders

Table 2 summarizes the assumed justification of the possible use of TMs in every group. The putative appropriateness of a possible order of diagnostic tests decreased from probably appropriate in “Advanced-Relapsed” to almost certainly inappropriate in “NED” cases.

Assessment of Ordered TMs According to Status at Follow-Up

CA15.3 and CEA ordered during 5 years are summarized in Table 3 with reference to the total person-years follow-up observed in each group with different status at follow-up.

A total of 39,720 TM tests (20,408 CA15.3 and 19,312 CEA) were ordered in the examined cohort of patients over 5 years in a total of 17,814 person-years follow-up. Most of the TMs (32,900; 82.8% of total. 16,924 CA15.3 and 15,976 CEA) were ordered in patients probably without evidence of disease (NED). The average number of CA15.3 per person-year was 1.15. It was significantly higher in the Advanced-Relapsed group than in Suspicious not confirmed (p<0.0001), NED post-neoadjuvant CHT (p<0.0001) and NED (p<0.0001). It was also significantly higher in Suspicious not confirmed than in NED post-neoadjuvant CHT (p<0.05) and NED (p<0.05), while it was not different between NED post-neoadjuvant CHT and NED. Data of CEA were similar.

TMs measured per year in the four groups over the first 4 years of follow-up are reported in Supplementary Table S1. Data from the 5th year of follow-up were not considered because the classification of patients during the 5th year might be imprecise since confirmatory data from the 6th year were not available due to the duration of the study. The average number of CA15.3 ordered per patient and per year is 3.1 in Advanced-Relapsed, 2.0 in Suspicious not confirmed, 1.7 in NED post-neoadjuvant CHT and 1.8 in NED. No noticeable year-to-year differences occurred in Advanced-Relapsed. Conversely, from year 1 to 4 a significant reduction of the number of tests per patient and per year appeared in Suspicious

Table 2 Criteria to Assess Appropriateness of Tumour Markers and Imaging Exams Ordering According to Status at Follow-Up

<table>
<thead>
<tr>
<th>Status at Follow-Up</th>
<th>TM Request Appropriateness</th>
<th>CPGs Justification for Request</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced-Relapsed</td>
<td>Probably appropriate</td>
<td>Diagnostic work-up of advanced disease or progression</td>
</tr>
<tr>
<td>Suspicious not confirmed</td>
<td>Possibly appropriate</td>
<td>Diagnostic work-up of suspicious signs of progressive disease</td>
</tr>
<tr>
<td>NED post-neoadjuvant CHT</td>
<td>Probably inappropriate but based on a rationale criterion</td>
<td>Monitoring of cases expected to have an increased risk of relapse</td>
</tr>
<tr>
<td>NED</td>
<td>Almost certainly inappropriate</td>
<td>None</td>
</tr>
</tbody>
</table>

Abbreviations: TMs, tumour markers; CPGs, clinical practice guidelines; NED, no-evidence-of-disease; CHT, anticancer chemotherapy.
not confirmed (p<0.0001), NED post-neoadjuvant CHT (p<0.0001) and NED (p<0.0001). Figures of CA15.3 and CEA are superimposable. In the first 4 years of follow-up CA15.3 was ordered in 76.0% (1st year) to 68.4% (4th year) of women and CEA in 71.8% to 63.5%, respectively. It appears that the majority of women operated for primary BC in 2013 in Veneto Region received TMs measurements someway irrespectively of their status at follow-up, as CA15.5 was ordered on average in 71.8% and CEA in 68.0% of NED patients.

### Assessment of Ordered Imaging Exams in NED Patients

Analyses concerning imaging exams were restricted to NED patients, as diagnostic tests can find some justification in the follow-up of patients of the other groups.

A first evaluation considered the total follow-up time as a whole. Table 4 summarizes the number and types of imaging exams ordered in NED patients. If a patient received more exams of the same type in any of the 5 follow-up years, she was counted only once for each imaging exam (“N. patients with imaging” in Table 4). Data were also separately examined in NED patients who have had at least 1 TM test ordered during the 5 years of follow-up and in those who did not.

Overall, 15,858 imaging exams were ordered in NED patients, with 5.3 exams per patient over 5 follow-up years. The most frequent exams were liver ultrasonography (3.0 exams per patient) and chest radiography (2.4 exams per patient). Advanced imaging exams were less frequent, with an average of approximately one exam per type per patient over the 5 years of follow-up. More than 91% of imaging exams occurred in patients who had also had at least 1 TM test ordered during follow-up. The number of exams per patient was significantly higher for liver ultrasonography (3.1 vs 2.2; p<0.0001) and for chest radiography (2.4 vs 1.9; p<0.0001) in women who have had TM ordered than in those who did not.

A second type of evaluation considered the tests ordered in NED patients by year of follow-up. Data were also separately examined in patients with- or without at least 1 TM test ordered in the considered year.

Results are shown in Supplementary Table S2. The frequency of orders of all the evaluated imaging exams significantly decreases over the years (p<0.0001 for liver ultrasonography, whole-body bone scanning and computed tomography, p=0.0154 for chest radiography). The very high percentage of whole-body bone scanning ordered in the 1st year may probably be due to a delayed initial staging procedure. The limited number of magnetic resonance and positron-emission tomography exams precludes any consideration on their figures.

Carrying out a year-by-year evaluation, both imaging exams and TMs were ordered in the same years in 76.9% (whole-body bone scanning) to 84.4% (chest radiography) of cases (Supplementary Table S2). All types of imaging exam were ordered more frequently in patients who have had TMs ordered than in those that did not (p<0.0001).

### Table 3 Tumour Markers Ordered During 5 Years of Follow-Up in Women Operated of Primary Breast Cancer, According to Status at Follow-Up

<table>
<thead>
<tr>
<th>Status at Follow-Up</th>
<th>Person-Years of Follow-Up</th>
<th>CA15.3</th>
<th>CEA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ordered Tests n° (%)</td>
<td>Tests/Person-Year of Follow-Up</td>
<td>Ordered Tests n° (%)</td>
</tr>
<tr>
<td>Advanced-Relapsed</td>
<td>466</td>
<td>1077 (5.3%)</td>
<td>2.31</td>
</tr>
<tr>
<td>Suspicious not confirmed</td>
<td>1478</td>
<td>1889 (9.3%)</td>
<td>1.28</td>
</tr>
<tr>
<td>NED post-neoadjuvant CHT</td>
<td>525</td>
<td>518 (2.5%)</td>
<td>0.99</td>
</tr>
<tr>
<td>NED</td>
<td>15,345</td>
<td>16,924 (82.9%)</td>
<td>1.10</td>
</tr>
<tr>
<td>Total</td>
<td>17,814</td>
<td>20,408 (100%)</td>
<td>1.15</td>
</tr>
</tbody>
</table>

**Notes**: Kruskal–Wallis test. Advanced-relapsed vs suspicious not confirmed, vs NED post-neoadjuvant CHT and vs NED: p<0.0001. Suspicious not confirmed, vs NED post-neoadjuvant CHT and vs NED: p< 0.05.

**Abbreviations**: NED, no-evidence-of-disease; CHT, chemotherapy.
Table 4 Patients Without Evidence of Disease (NED): Imaging Exams Performed During 5 Follow-Up Years

<table>
<thead>
<tr>
<th>Imaging Exam Type</th>
<th>Total</th>
<th>Patients Who Have Had TM Ordered</th>
<th>Patients Who Did Not Have TM Ordered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N. Imaging Exams Ordered</td>
<td>N. Patients with Imaging</td>
<td>N. Exams per Patient</td>
</tr>
<tr>
<td>Liver Ultrasonography</td>
<td>8044</td>
<td>2641</td>
<td>3.0</td>
</tr>
<tr>
<td>Chest Radiography</td>
<td>5429</td>
<td>2278</td>
<td>2.4</td>
</tr>
<tr>
<td>Total body Bone scan</td>
<td>1083</td>
<td>1003</td>
<td>1.1</td>
</tr>
<tr>
<td>Computed Tomography</td>
<td>885</td>
<td>592</td>
<td>1.5</td>
</tr>
<tr>
<td>Magnetic Resonance</td>
<td>294</td>
<td>250</td>
<td>1.2</td>
</tr>
<tr>
<td>Positron-Emission Tomography</td>
<td>123</td>
<td>114</td>
<td>1.1</td>
</tr>
<tr>
<td>Total</td>
<td>15,858</td>
<td>2995(^a)</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Notes: \(^a\)Number of patients who performed at least 1 imaging exam. The number of exams per patient in patients who have had TM ordered than in those who did not was significantly different for liver ultrasonography (p<0.0001) and for chest radiography (p<0.0001).

Abbreviation: TM, tumour marker.
In patients who performed both TMs and imaging exams, we analysed the time elapsed between TM measurement and the imaging ordered in the same year, to explore if TM anticipated imaging ordering. Results are shown in Supplementary Table S3. A positive value indicates that TMs anticipated the imaging exam. During the first year of follow-up, TMs anticipated imaging exams in a limited number of cases: around 50% of chest radiography, but approximately the 25% of liver ultrasonography and chest radiography and less than 10% of whole-body bone scanning. In the following 3 years the anticipation of TMs on imaging occurred in almost 50% of cases. Data for magnetic resonance and positron-emission tomography were not considered due to the limited number of ordered exams.

Discussion

In the present population-based cohort study on data from more than 4300 consecutive women with a BC diagnosed in 2013, we observed that 86% appeared to be free of disease during a five-year follow-up after surgery. Most of them received TMs and/or imaging exams during the five years of follow-up, in disagreement with CPGs. Two seminal randomized clinical trials allocating breast cancer patients to either intensive or clinical follow-up did first demonstrate that no differences occur in 5-year overall mortality between clinical and intensive follow-up protocols. Differently designed randomized clinical trials compared outcome of patients receiving routine follow-up either in hospital or in general practice showing that the majority of recurrences are detected by women as interval events irrespective of continuing hospital follow-up. In addition, intensive follow-up has been shown to have a negative impact on quality of life and a randomized study demonstrated the preference of patients for less frequent follow-up.

The adherence to CPGs is a key issue for health-care systems. The progressive availability of novel diagnostic and therapeutic options and the increasing motivation to reduce inequalities, tend to dramatically increase the costs of health care. In this scenario, the appropriate use of available resources is mandatory to keep under control the sustainability of the health-care systems. The control of appropriateness is more feasible in the case of therapy because the approval pathway of a new drug by Regulatory Agencies is a first barrier to its inappropriate delivery, as indications for the use are clearly established. In addition, post-marketing surveillance is a second control mechanism that facilitates any eventual optimization of appropriate use of drugs. In the case of the diagnostic tests, the approval by regulatory Agencies is far less stringent than for drugs and no post marketing is regularly mandated. Therefore, the risk of inappropriate use is elevated and several studies showed that laboratory tests and imaging are either overused or underused with variable frequencies. Strategies to improve and keep under control appropriateness of diagnostic test are manifold, including dedicated education plans, audit and feedback, or the use of informatics tools. Indicators are the tool to monitor the impact of any intervention focused on improvement of quality of cancer services and quality performance indicator development is a pivotal process in current health-care management. Several groups have proposed quality performance indicators for BC, but a very limited number of them concerns the appropriate use of diagnostic tests in the follow-up. The proposed indicators that we found published are “% of patients testing serum tumour markers (CA15.3) in the 12 months after surgery” and “Receiving chest CT or bone scans or liver US/CT/MR or tumour markers measurement in the year following surgery, excluding patients developing metastasis”.

Calculating indicators related to the follow-up is particularly complex, because some prevalent patients – ie, those in follow-up – have a progressive disease, while others are disease-free. Recommendations concerning the use of diagnostic tests during follow-up of relapsed and not relapsed patients are obviously different and indicators cannot equally fit to all prevalent cases. Data from Cancer Registries are rarely useful to disentangle disease-free prevalent patients from the others, because they cannot include the clinical status at follow-up. Examining the patients’ clinical records, although being reliable and having intrinsic validity, would not respond to other basic indicators’ characteristics such as usability and feasibility that are recommended by the Agency for Healthcare Research and Quality (AHRQ). Therefore, the first issue to solve when building usable and reproducible indicators for the follow-up of cancer patients is to identify with reasonable accuracy the disease-free subgroup among prevalent cases.

Trevisiol et al recently proposed an algorithm to distinguish prevalent BC cases without evidence of disease from those with metastatic spread, based on the ratio between mortality and prevalence and assuming that mortality is a surrogate indicator of metastatic disease. They showed a high overordering rate of CA15.3 in the portion of prevalent
cases classified as possibly free of disease. Although their algorithm can identify the existence of overordering in a patient cohort, it cannot link the ordered tests to individual patients over the follow-up time. Therefore, it remains a valuable approach to check for overordering in a given clinical setting, but it is not suited to monitor adherence to quality indicators.

Several administrative databases applied at regional or national scale in Italy provide standardized data that can be explored through consistent codes. We developed an algorithm capable of extracting, through simple and standardized query procedures, patients’ codes on the basis of their estimated chances of being either free of disease (NED), or at a possible increased risk of relapse (Suspicious not confirmed and NED post-neoadjuvant CHT groups), or actually relapsed. Codes of patients of different groups are easily and reproducibly matched with codes of therapeutic and diagnostic procedures in ORDTP, to check the adherence to CPGs and to eventually develop and use proper indicators.

The number of examined diagnostic tests could be underestimated in the present study as patients could have had exams performed outside public or accredited health systems. However, the Italian National Health System establishes that all patients affected by a malignancy, including breast cancer, are registered with a “disease code” which guarantees that patients receive all services included in a defined core benefit package (LEA) without any charge. In addition, requests of diagnostic tests to outpatients are performed with priority in cancer patients, thus reducing the waiting list time to access the services. Therefore, although we cannot exclude that some patients get tests outside public or accredited health system, we can realistically assume that they should represent a marginal fraction of all breast cancer patients.

With the developed algorithm we showed that TMs and imaging exams are requested in the vast majority of patients in follow-up. The rate of request resulted higher in relapsed patients, but most of the patients apparently without evidence of disease received some TM assays during the follow-up, with a total of 32,900 TMs ordered over 5 years. Also imaging exams were performed in the majority of patients with 13,473 basic and 2385 advanced imaging exams performed in NED patients over 5 years of follow-up. The present study was not designed to perform a cost analysis, but to develop a prototype algorithm to assess the appropriateness of diagnostic test ordering in subgroups of patients with breast cancer according to their status at follow-up. Nevertheless, considering direct costs of the diagnostic test published in the Price List Catalogue Veneto Region Healthcare System, we can estimate that 1,322,944 € for almost certainly not necessary TMs and 1,767,853 € for probably not necessary imaging exams have been wasted over 5 years of follow-up of patients with breast cancer incident in 1 year.

Our findings are comparable to those of some authors, which found an intensive follow-up applied to 79% of patients, but different from others, who found lower rates of TM orders. Differences in the geographic area examined, in the years of recruitment of patients and in the length of the follow-up may possibly explain the discrepancy. However, the major difference between the present and previous studies consists in the fact that we subdivided prevalent cases according to their clinical status at follow-up.

CPGs recommend against intensive follow-up in asymptomatic patients. However, CPGs are a framework that do not intend to exclude the independence of the medical doctor in taking decisions on individual patients. Therefore, we separated cases with possible symptoms requiring imaging to rule out the recurrence, and cases who received neo-adjuvant therapy, in which a higher risk of recurrence may be considered. Results of the latter two groups were reported but not discussed, since we realize that decisions may be responsibility of the medical doctor who assesses the risk case by case. Likewise, we do not comment the findings concerning relapsed patients. Hence, we limit the comment of data to cases that probably remained free of disease over the 5 years of follow-up, receiving 32,900 TMs and 15,858 imaging exams, which according to all available CPGs should be considered inappropriate and imply unnecessary direct costs and improper utilization of health-care structures, as well as indirect costs for women.

Diagnostic tests are generally ordered by oncologists. It cannot be excluded that general practitioners decided to prescribe additional tests on their own, but the order forms in use when the study was performed do not allow for a computerized identification of the physician who requested the test. A manual scrutiny of thousands of requests was neither feasible nor suitable for a general use of the indicator. The Veneto Region started in 2015 moving from the traditional, paper based, to a fully electronic order form that allows for a computerized identification of the physician who ordered the test. Thought it could not be applied in the study, if the algorithm were adopted, it could presently provide information on the physician who orders the test.
A retrospective cohort study conducted in four Canadian provinces has recently shown that quality of breast cancer survivor follow-up care varies among provinces. Therefore, it would have been interesting to assess if differences exist among different care-giver structures included in the study. However, the study included patients operated in 51 surgical structures and followed-up by 39 oncology services. Given the limited number of patients operated and/or followed up in individual structures, a stratification analysis of requests according to the structure would provide possibly unreliable information. However, in Veneto Region public or accredited structures are expected to use shared diagnostic and therapeutic procedures established and regularly updated by a formal body (Rete Oncologica Veneta – ROV – Veneto Region Oncology Network) settled by the Regional government. Therefore, we assumed that major differences in the standard of care provided to women with breast cancer should be limited through Veneto Region oncology network centres.

The present study has some limitations. First, results of TM assays are not available in administrative databases. Therefore, the recognition of positive results is not feasible and the causal relationship between the determination of TMs and a subsequent imaging exam cannot be verified and remains hypothetical.

Second, the available databases used to identify patients and mine their initial data (VTR and HDR) did not provide information regarding the stage of the disease and data could not be stratified according to the disease stage at presentation. Nevertheless, the algorithm was designed to identify cases with increased risks as well as cases in the IV stage at presentation. The HDR has been updated in 2017, and a synthetic version of disease stage is now available. If the algorithm was adopted it could also allow for analyses stratified according to initial disease stage.

Third, the assumption that 2 advanced imaging in Year x and one or more “advanced imaging exams” in the subsequent 12 months defines advanced relapse, settled on the basis of guidelines for the treatment of advanced breast cancer available in 2013\(^{46}\) could easily hide follow-up of incidental findings, ie, lung nodules and a rate of “false positive” classification of relapses can be expected.\(^{47}\) However, this occurrence does not modify – and even possibly attenuate – the conclusions of the study concerning the very high rate of overordering of diagnostic tests in NED patients.

Fourth, given the scope of this study, we searched and examined imaging exams usually performed to screen distant metastases. Therefore, breast imaging was not considered in the analysis and possible underordering was not assessed.

Fifth, HDRs were not found for approximately 10% of BC cases reported by the VTR. Thus, matching data among different databases suggests that some imprecision or time lag possibly exist in logging data. Although the number of unmatched patients is limited and hardly affects the results of the present study, this issue must be considered when using the algorithms in different contexts.

Despite these limitations, the present investigation is a population-based cohort study that does not present the risks of selection bias nor problems of generalizability usually occurring in case series from individual institutions.

### Conclusion

In conclusion, we present an original approach based on an innovative algorithm to identify BC patients who are probably free from disease during their long-term follow-up. In this cohort of patients, we demonstrated a substantial overordering of TMs and imaging exams, in contrast with the recommendations of CPGs. Provided other regions and countries had access to the type of data used in the present study, the developed algorithm can be transferred and repeatedly used over time, thus being suitable for a regular monitoring of the adherence to CPGs and of the impact of interventions to improve appropriateness.

### Abbreviations

BC, breast cancer; BS, whole-body bone scanning; CHT, anticancer chemotherapy; CPG(s), clinical practice guideline(s); HDR, Hospital Discharge Records; IAA, infusion of anticancer agents; NED, no-evidence-of-disease; ORDTP, Outpatient Record of Diagnostic and Therapeutic Procedures; TM(s), tumour maker(s); VTR, Veneto Tumour Registry.

### Ethics Approval

The study did not need specific authorisation from an Ethics Committee as it is a retrospective observational study based on cases routinely collected by the Cancer Registry of the Veneto Region, which is authorized by the Italian Guarantor for the Protection of Personal Data to use health-related data for research purposes.
Acknowledgments
We thank Dr. Antonette Leon for her contribution in editing the manuscript and Mrs. Ornella Scattolin for administrative assistance.

Author Contributions
1. Massimo Gion
Conception and design; Acquisition of data; Analysis and interpretation of data; Drafting of the manuscript; Critical revision of the manuscript for important intellectual content; Obtaining funding; Administrative, technical, or material support; Supervision.

2. Giulia Cardinali
Acquisition of data; Analysis and interpretation of data; Critical revision of the manuscript for important intellectual content; Statistical analysis.

3. Stefano Guzzinati
Acquisition of data; Analysis and interpretation of data; Critical revision of the manuscript for important intellectual content; Statistical analysis.

4. Paolo Morandi
Conception and design; Acquisition of data; Analysis and interpretation of data; Critical revision of the manuscript for important intellectual content; Supervision.

5. Chiara Trevisiol
Conception and design; Drafting of the manuscript; Critical revision of the manuscript for important intellectual content; Administrative, technical, or material support.

6. Aline S.C. Fabricio
Conception and design; Drafting of the manuscript; Critical revision of the manuscript for important intellectual content.

7. Massimo Rugge
Analysis and interpretation of data; Critical revision of the manuscript for important intellectual content; Obtaining funding.

8. Manuel Zorzi
Conception and design; Acquisition of data; Analysis and interpretation of data; Critical revision of the manuscript for important intellectual content; Obtaining funding; Supervision.
All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

**Funding**
This study was supported in part by Veneto Region (IT) through the “Programma Regionale per i Biomarcatori Diagnostici, Prognostici e Predittivi” assigned to Azienda ULSS12 Veneziana (currently named AULSS3 Serenissima). The specific role of the funding organization is Design and conduct of the study.

**Disclosure**
The authors declare no potential conflicts of interest.

**References**


