Type and trends in outcomes research in breast cancer between 2000 and 2007

M. Valentini1*, A. Milesi2, A. Bettini2, C. Tondini2 & A. Nicolucci1

1Department of Clinical Pharmacology and Epidemiology, Consorzio Mario Negri Sud, S. Maria Imbaro; 2Division of Oncology, Ospedali Riuniti, Bergamo, Italy

Received 30 July 2010; revised 9 November 2010; accepted 10 November 2010

Background: Interest in outcomes research (OR) derives from the need to know the value and the effectiveness of health interventions, especially for oncology. We focused our research on OR in breast cancer, providing an overview of the trend of publications.

Patients and methods: We carried out a Medline search to retrieve all articles in English published from 2000 through 2007. The abstracts were reviewed and classified according to the research topics and the primary design of the trial.

Results: We selected 405 articles: their number remained constant until 2003, rose during 2004–2005 and decreased during the last 2 years. The most common topic was surgery (n = 234), alone or in association with other interventions. The category more investigated was the process. Clinical outcomes, and among them disease-free survival, were more frequent than other outcomes. The median value of 2007 Impact Factor of the journals publishing the selected references was 2.466 (range 0.272–25.547) and the median value of Citation Index was 8 (range 0–143).

Conclusions: Our research showed a decreasing interest in OR during the more recent years. We are hopeful that it will regenerate interest, particularly by the light of the funds allocated to the comparative effectiveness research in the United States.

Key words: breast neoplasms, outcomes and process assessment, outcomes assessment, review, treatment outcome

Introduction

Outcomes research (OR) is an evolving concept of monitoring health care performance. The US Agency for Healthcare Research and Quality defined the OR as follows: ‘Outcomes research seeks to understand the end results of particular health care practices and interventions. End results include effects that people experience and care about, such as change in the ability to function. In particular, for individuals with chronic conditions—where cure is not always possible—end results include quality of life as well as mortality’ [1].

The interest in this field of research derives from the need to know the actual value and the effectiveness of health interventions [2]. In fact, mounting concerns over geographic variability in practice patterns, lack of documented benefit from medical treatments or procedures and skyrocketing medical costs had coalesced into a call for standardization of health care based on empirical data [3] This is especially true for oncology because of its tremendous impact in terms of care delivery, social and economic burden.

We focused our research on breast cancer, the most frequent neoplasm among women, as an interesting field of application for OR. As a matter of fact, in spite of the presence of a lot of efficacious treatment strategies and clear guidelines for the treatment of this disease, there is a great variability in terms of mortality across countries [4].

Moreover, over the last decade, we assisted to the development of innovative diagnostic and surgical new procedures, effective targeted drugs and multigene expression profiles aimed to outdo traditional predictive and prognostic factors [5, 6]. OR could be the instrument to study the impact of these innovations outside clinical trials as well as their benefit in real practice. Therefore, we conducted a structured review of the literature in OR in breast cancer care. Our aim was, first, to provide an overview of the trend of publications during the beginning of this century; second, we sought to examine the interest of the scientific community for these publications.

Methods

Articles selection

In March 2008, we carried out a Medline search to retrieve all articles in English published from 2000 through 2007 with the following MeSH
(Medical Subject Heading) terms: costs and cost analysis; cost–benefit analysis; health services accessibility; health services research; outcomes assessment (health care); quality of health care; quality assurance, health care; quality indicators, health care; guidelines/guidelines as topic; practice guidelines/practice guidelines as topic; decision making; decision support techniques; economics; quality of life. These are the same MeSH terms used by Lee et al. [5] in their review published in 2000; we had to add ‘as topic’ to ‘guidelines’ and ‘practice guidelines’ because these two terms have not existed since 2008.

All these topics were combined using the Boolean operator ‘OR’ and were not exploded. Subsequently, all these terms were combined with the MeSH term ‘breast neoplasms’ with the Boolean operator ‘AND’. Then, we removed all articles classified as ‘comment’, ‘editorial’, ‘letter’, ‘news’ or ‘case reports’ from the publication type. Finally, we eliminated all articles with ‘Mass screening’ as MeSH and ‘Prevention and control’ as SH (subheading). All the research was limited to human subjects and English language.

**articles classification**

The abstracts were independently reviewed by two of the authors (MV and AM) and classified according to the research topics/end points and to the primary design of the trial. Discrepancies were resolved by consensus or mediation by a third party (AN or CT), as necessary. Articles were retrieved as needed for clarification. Articles were considered eligible when evaluating clinical and/or humanistic and/or economic outcomes in diagnosis or therapy of breast cancer. The assessment could be relative to structures, processes or outcomes. When the articles evaluated several neoplasms, we considered only the data relative to breast cancer.

All clinical trials were ineligible, as well as the study nested in clinical trials and feasibility studies. Articles without any intervention were also considered ineligible, as well as studies of validation and studies about innovative interventions, i.e. those interventions whose efficacy was not yet confirmed by large-scale clinical trials. We decided to select only those interventions with proven evidence of efficacy.

We classified the retrieved articles according to the study type (case–control, cohort, record of cases, population studies, administrative database, pharmacoeconomic studies), category (structure, process, outcome), intervention (surgery, chemotherapy, radiotherapy, hormone therapy, palliative care, other) and type of outcome (clinical, humanistic, economic). Each study type was further classified as retrospective or prospective.

With regard to the categories, we referred to the Donabedian’s definitions: ‘structure’ denotes the attributes of the settings in which care occurs … ‘process’ denotes what is actually done in giving and receiving care … ‘outcome’ denotes the effects of care on the health status of patients and populations’ [7].

With regard to the type of outcome, we adopted the following classification, modified by Epstein et al. [8]:

- **Clinical**: clinical events (including therapy complications and toxicity), physiological and metabolic measures (including specificity and sensibility of test or scan), overall survival and disease-free survival (DFS) (including the response to second-line therapies).
- **Humanistic**: symptoms, quality of life, functional status, patient satisfaction; these outcomes should be assessed by specific and validated questionnaires.
- **Economic**: direct medical costs, indirect medical costs.

We evaluated the 2007 Journal Impact Factor (IF) and the subject of the journals publishing the selected references and the Science Citation Index (CI) of the articles. The IF is the ratio of the number of citations of the journal and the number of articles published in the same journal during the previous 2 years [9]. Therefore, the 2007 IF refers to 2005 and 2006. The scientific journals were also grouped according to the classification per subject of Institute of Scientific Information [9]. The CI is the number of times an article has been cited by other articles in Science Citation Index journals [10].

In a previous review on breast cancer quality-of-life outcomes, Mandelblatt et al. [11] excluded studies with <100 patients since there would be insufficient power for the majority of measures; therefore, we decided to divide the retrieved articles in two subgroups according to the study population (i.e. ≤100 versus >100).

**statistical analyses**

The relationships between number of evaluated patients and 2007 IF and CI were estimated using the Spearman’s rank correlation. We explored also the correlation between CI and 2007 IF and CI and year of publication. Results are expressed in terms of Spearman’s correlation coefficient and P value. All analyses were carried out using SAS Statistical Package, version 9.1 (SAS Institute, Cary, NC) [12].

**results**

We carried out the Medline search in March 2008, obtaining 1753 articles. After the abstract review, we selected 405 (23.1%) articles that were classified and analyzed.

Of the remaining 1348 articles, the main causes of exclusion were the following: 315 (23.4%) were related to prognostic or predictive factors, 173 (12.8%) were about innovative interventions or validation of questionnaires/tests, 148 (11%) were randomized clinical trials or reviews, 127 (9.4%) studied the impact of the disease, independent of any intervention.

The number of patients involved was extremely variable because we included studies based both on records of cases and on administrative databases and population studies. Therefore, the median number of patients was 129, and the range was 5–46 687. We divided the studies in two subgroups according to the number of patients: for 164 references (40.5%), the number of patients was ≤100; among them, 94 (57.3%) were based on <50 patients; in two cases, the exact number of patients was not specified. In 239 articles (59.0%), the number of patients was >100. Two references were not classified because the number of patients could not be verified.

The number of articles by years of publications is represented in Figure 1A. The number of references remained constant until 2003; it rose during 2004–2005 and then decreased during the last 2 years. This trend was similar when we split the references by the number of patients. However, when we considered the ratio of selected/retrieved references, this value remained constant during the period 2000–2007 (Figure 1B).

Table 1 shows the type of study: the greatest percentage of articles was based on record of cases. Obviously, population-based studies and those based on administrative databases were more frequent among papers with the higher number of patients.

The distribution of references according to the intervention is presented in Figure 2. Many articles investigated more than one intervention and in several cases the intervention was the entire process of therapy. The most common topic of these publications was surgery (n = 234), alone or in association with other interventions; among surgery procedures, the most frequent intervention was plastic surgery (37.6%).
Interventions classified as ‘Other’ (i.e. information, relation between physician and patient, etc.) were more frequent among studies with more patients. Chemohormonal therapy and supportive care were investigated only by studies with >100 patients involved; on the other hand, studies on palliative care involved predominantly <100 patients.

The category more investigated was the process, followed by outcome and finally structure; some studies investigated more than one category. The distribution was the same in the two subgroups (Table 2). The frequency of the outcomes is shown in Table 3; clinical outcomes, and among them DFS, were more frequent than other outcomes. Among humanistic outcomes, quality of life was the most common one. Among economic outcomes, direct medical costs represented the aspect more often investigated.

The greatest percentage of articles was carried out in the United States (38.8%), followed by UK (7.7%), The Netherlands (6.4%), Germany (5.4%), Italy (4.9%), Japan and Australia (both 4.4%), Canada (4.2%) and France (4.0%). Other countries accounted for <10 articles per country. The same countries published most of the articles with >100 patients.

Articles were published in 127 scientific journals. The greatest number of studies was published in Cancer, with 20 (4.9%) references, followed by Breast Cancer Research and Treatment (19 references, 4.7%), American Journal of Surgery and International Journal of Radiation Oncology, Biology, Physics (17 references, 4.2%), European Journal of Surgical Oncology (16 references, 4.0%), Annals of Surgical Oncology (14 references, 3057%), Journal of Clinical Oncology (13 references, 3.2%) and Plastic and Reconstructive Surgery (12 references, 3.0%). Other scientific journals published <10 articles.

The median value of 2007 IF of the journals publishing the selected references was 2.466 (range 0.272–25.547). The value was not available for 68 (16.8%) references, corresponding to 30 (23.6%) journals. The median value of 2007 IF of the journals publishing studies with up to 100 patients (n = 135) was 2.074 (range 0.272–10.734). The median value of 2007 IF of the journals publishing studies with >100 patients (n = 200) was 3.3180 (range 0.597–25.5470).

The information about the subject of the journals was also missing for the 68 references without 2007 IF. The most common journal subject was oncology (192 references, 51 journals), followed by surgery (98 references, 23 journals), without differences according to study size.

The median value of CI was 8 (range 0–143); this information was missing for 43 references. The median value of CI of studies with up to 100 patients (n = 141) was 6 (range 0–132), while the median value of CI of the studies with >100 patients (n = 220) was 9 (range 0–143).

The evaluation of the influence of the number of patients on the 2007 IF of the scientific journal showed that the higher the number of patients the higher was the 2007 IF (Spearman’s correlation coefficient = 0.31, P < 0.0001). The analysis showed a similar correlation when we evaluated the influence of the number of patients on the CI of the references: the larger was the population study the higher was the CI (Spearman’s correlation coefficient = 0.11, P = 0.03).

Moreover, the CI was significantly correlated to the 2007 IF (Spearman’s correlation coefficient = 0.42, P < 0.0001) and to the year of publication (Spearman’s correlation coefficient = −0.43, P < 0.0001).

**Discussion**

OR explores what happens to the patients outside clinical trials. Therefore, it should be particularly useful in breast cancer where we assist to a proliferation of clinical studies with the introduction of new techniques and treatments everyday in clinical practice. In this contest, OR could answer the question of what is the impact of these innovations in a ‘real’ clinical setting and could help health providers to make informed decisions.

### Table 1. Type of study

<table>
<thead>
<tr>
<th>Study type</th>
<th>All references (No. of references (%)</th>
<th>≤100 patients (No. of references)</th>
<th>&gt;100 patients (No. of references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series of cases</td>
<td>193 (47.7)</td>
<td>95 (57.9)</td>
<td>97 (40.6)</td>
</tr>
<tr>
<td>Longitudinal study</td>
<td>145 (35.8)</td>
<td>54 (32.9)</td>
<td>90 (37.7)</td>
</tr>
<tr>
<td>Case–control study</td>
<td>33 (8.2)</td>
<td>13 (7.9)</td>
<td>20 (8.4)</td>
</tr>
<tr>
<td>Population study</td>
<td>20 (4.9)</td>
<td>0</td>
<td>20 (8.4)</td>
</tr>
<tr>
<td>Administrative database</td>
<td>14 (3.5)</td>
<td>2 (1.2)</td>
<td>12 (5.0)</td>
</tr>
<tr>
<td>Total</td>
<td>405</td>
<td>164</td>
<td>239</td>
</tr>
</tbody>
</table>
decision. Given these premises, we expected an increase in the interest for this field across the years; results showed that our expectations were not met: in fact, the number of the selected articles during the years 2000–2007 was not uniform, and it decreased during the last two observed years, even if the ratio of retrieved/selected articles remained constant during the same period. This phenomenon could be partly related to the use of the MeSH terms: in fact, the meaning of several terms we used is in evolution and it is meaningful that some terms present in the MeSH database until 2007 (‘guidelines’ and ‘practice guidelines’) disappeared in 2008 because they were substituted by other expressions (‘guidelines as topic’ and ‘practice guidelines as topic’). Another reason could be that recently many randomized clinical trials include nested studies on quality of life or preferences of patients or other humanistic outcomes. This probably reduced the interest for OR even if populations included in clinical trial are highly selected and could not be representative of the effectiveness of an intervention in a broader set of patients. The investigators are often more interested in randomized clinical trials, maybe assuming that during the daily practice it is obvious to obtain the same results of an experimental study. Moreover, the randomized controlled trials are frequently sponsored by industries, unlike the OR studies, that are often without adequate fundings.

It could be very useful to recover the role of OR in breast cancer, not limiting it only to the study of humanistic outcomes in randomized clinical trials; some interventions, very promising according to the results of randomized clinical trials, could not obtain the same effects in the daily practice; this is a very interesting field of ‘practical research’ where the OR plays a key role.

At our knowledge, there are no similar findings for other cancer diseases; the reviews collected in one of the monographs of the National Cancer Institute, dedicated to the OR, included also randomized clinical trials; moreover, the temporal trends, when evaluated, were focused on years 1990–2000, so we could not compare our findings with those reviews [13]. It could be interesting to repeat the same search strategy we adopted for breast cancer to the most common neoplasms, such as lung.

![Figure 2. Interventions.](chart)

### Table 2. Categories distribution

<table>
<thead>
<tr>
<th>Categories</th>
<th>All references No. of references (%)</th>
<th>≤100 patients No. of references (%)</th>
<th>&gt;100 patients No. of references (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>231 (57.0)</td>
<td>99 (60.4)</td>
<td>131 (54.8)</td>
</tr>
<tr>
<td>Outcome</td>
<td>159 (39.3)</td>
<td>65 (39.6)</td>
<td>93 (39.7)</td>
</tr>
<tr>
<td>Structure</td>
<td>21 (5.2)</td>
<td>1 (0.6)</td>
<td>20 (8.6)</td>
</tr>
</tbody>
</table>

*The sum of the columns does not correspond to the number of articles because some studies involved more than one category.

### Table 3. Outcomes frequency

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Type of outcome</th>
<th>All references No.</th>
<th>≤100 patients No.</th>
<th>&gt;100 patients No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical</td>
<td>DFS</td>
<td>144</td>
<td>69</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td>124</td>
<td>55</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Physiological and metabolic measures</td>
<td>89</td>
<td>33</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Clinical events</td>
<td>81</td>
<td>42</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>291</td>
<td>123</td>
<td>166</td>
</tr>
<tr>
<td>Humanistic</td>
<td>Quality of life</td>
<td>101</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Patient’s satisfaction</td>
<td>42</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Symptoms</td>
<td>38</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Functional status</td>
<td>16</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>150</td>
<td>58</td>
<td>91</td>
</tr>
<tr>
<td>Economical</td>
<td>Direct medical costs</td>
<td>13</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Indirect medical costs</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>15</td>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

*The sum of the columns does not correspond to the number of articles because some studies involved more than one outcome.
cancer and colorectal cancer, to understand if the patterns of publication are similar.

Our research shows that the majority of the articles are based on retrospective analyses and the interest is mostly focused on process, i.e. all that is done to the patient during her diagnostic–therapeutic course. The emerging data are of studies on clinical outcome looking at DFS and mortality. The scenario seems that OR in breast cancer has not found original ways of development without mimicking the structure of the clinical trials and that could be one of the reasons of the diminishing interest in the field in the last 2 years. OR could be very useful for evaluating the application of target therapies and novel biomarkers (including the multigene expression profiles) in the real clinical context. However, our review selected few references on this argument. Maybe this is due to the recent publications of the large clinical trials that confirmed the importance of this kind of therapy in breast cancer [14, 15]. The emotional impact of positive results in big clinical trials maybe moves the focus away from the possible difficulties of transferring in real practice these innovations.

With regard to the type of intervention, the selected references showed a preference for surgery, and specifically breast reconstruction. On one hand, this matter is very important because reconstructive surgery can help women with breast cancer to avoid a mutilation, thus improving their quality of life; on the other hand, it is essential that this kind of intervention does not overshadow subsequent therapies and the follow-up policy. For surgery is particularly consolidated that the real contest determines the result: expertise of the surgical team, technical supports and so on. This can explain the prevalence of surgical studies on OR.

On the opposite side, our review showed that the attention for advanced disease is very low: in fact, only 19 articles studied palliation and a single article investigated supportive care. This is unexpected because therapies for advanced breast cancer are more variable than treatments of nonmetastatic disease [16]. Maybe the uncertainty about the best available treatment (as the result of maximizing cure and minimizing toxicity) induces the physicians to enroll the great part of these patients in randomized clinical trials, avoiding other kinds of studies. The number of patients involved could support this hypothesis: in fact, all references but three relative to palliative and supportive care involved <100 patients.

Among the types of outcomes, quality of life is the most frequent humanistic outcome. In breast cancer, the evaluation of quality of life is important in every phase of the treatment since the physician has at his/her disposal different types of surgery, hormonal, chemo- and palliative therapy with similar results in terms of survival. Therefore, quality of life or patient’s preferences can affect the choice of the surgical intervention, the hormonal therapy or the chemotherapy regimen. Knowing the impact of treatments in terms of survival and humanistic outcomes in the normal clinical practice, outside the clinical trials, is therefore of primary importance.

We retrieved few articles that studied economic outcomes, probably this was due to the lack of specificity of the MeSH we used for this field of research.

The number of patients involved is a critical matter. We selected the cut-off of 100 patients, but it is noticeable that 23.2% of all articles involved 50 patients or less. Forty percent of the selected references involved 100 patients or less; these articles were published more frequent by journals with a low 2007 IF and were less cited by other authors. Therefore, an article involving few patients seems to be almost totally ignored by the scientific community. OR studies the end results of particular health care practices and interventions in the common clinical practice, so the definition of an adequate sampling is difficult: in fact, in a clinical practice with a lot of patients the study sample should be broad, while in a setting with few patients the study sample could be small. To our knowledge, there are no rules about the number of patients to include in OR studies. Mandelblatt et al. [11] excluded from their review studies with <100 patients ‘since there would be insufficient power for the majority of measures’; however, they included also clinical trials and maybe it could be the reason for using such a cut-off.

Our review has some limitations. First of all, the change of some terms in the MeSH database could have affected our research. On the other hand, these changes are unavoidable when the aim is to investigate a broad range of years. Another limitation could be represented by the use of the IF; not all the scientific world agrees with the use of this instrument to assess the real value of an article. To limit this weakness we used also other indicators, such as the CI. Moreover, we decided to exclude breast cancer screening from our review because our intention was to focus on the treatment of this neoplasm. In the light of our result, it could be interesting to repeat the search strategy also in this field.

In the last years, particularly in breast cancer, tailored strategies are becoming more and more important, i.e. a specific therapy for specific patients based on the tumor and patients characteristics. From this point of view also the results of the OR conducted with few patients could be helpful: for a physician working in a general hospital with few patients and paucity of resources, the clinical experience of a colleague working in a similar setting could be more informative than the results of a clinical trial conducted in high specialized structures.

Maybe the recent American Recovery and Reinvestment Act (ARRA) of 2009, that allotted $1.1 billion to support the comparative effectiveness research (CER) will call the attention to the OR. In fact ‘CER is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition, or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers and policy makers to make informed decisions that will improve health care at both the individual and population levels’. This definition also includes studies on the interventions in patients who are typical of day-to-day clinical care [17]. There is a great debate about how use the funds correctly: according to the Institute of Medicine committee, the CER program will be perceived as successful if it funds research that improves the uptake of new knowledge and its translation into better decision making [17]. The OR could represent one of the answers to this need because it represents the opportunity to link current health care practice to the outcome of care, also in breast cancer setting.
conclusions

Our research showed a decreasing interest in OR during the more recent years under study. We hope this trend will be reversed because OR could improve the knowledge derived from clinical trials, including information from daily clinical practice. In particular, an increase in OR studies could be useful mostly in the contest of the new target therapies and of metastatic disease, in order to integrate the high-quality data derived from randomized studies with the information from the clinical practice. What is the impact of a new treatment or procedure in term of psychological burden for the patient, loss of day of work, community cost and effectiveness of application are only some of the answers that outcome research can be able to give. Considering the today opportunity of easily collecting data thank to the spreading of informatics support, we are hopeful that OR will regenerate interest, particularly by the light of the needs consequent to the recent ARRA of 2009 and the funds allocated to the CER. The possibility of prospectively planning the collection of original data from the everyday practice is fascinating and for sure can open new understanding on the complexity of the cure of breast cancer.

disclosure

The authors declare no conflict of interest.

references

10. Science Citation Index. Thomson Reuters.