Estimation of the Number of Women Living with Metastatic Breast Cancer in the United States

Angela B. Mariotto1, Ruth Etzioni2, Marc Hurlbert3,4, Lynne Penberthy1, and Musa Mayer3

Abstract

Background: Distant metastatic breast cancer (MBC), including metastases found at diagnosis (de novo) and those occurring later (recurrence), represents the most severe form of the disease, when resource utilization is most intensive. Yet, the number of women living with MBC in the United States is unknown. The objective of this article is to use population-based data to estimate the prevalence of MBC.

Methods: We used a back-calculation method to estimate MBC prevalence from U.S. breast cancer mortality and survival from the Surveillance, Epidemiology and End Results (SEER) registries. On the basis of the illness–death process, this method assumes that each observed breast cancer death is the result of MBC, either de novo or a recurrence with metastatic disease.

Results: We estimate that by January 1, 2017, there will be 154,794 women living with MBC in the United States, three in four initially diagnosed with stage I–III breast cancer who later progressed to MBC.

Median survival and 5-year relative survival for de novo MBC increased over the years, especially in younger women. We estimate a two-fold increase in 5-year relative survival rate from 18% to 36%, for women diagnosed with de novo MBC at age 15–49 between 1992–1994 and 2005–2012, respectively.

Conclusions: This study demonstrates an increasing number of women in the United States living with MBC, likely the result of improvements in treatment and aging of the U.S. population.

Impact: The increasing burden of MBC highlights the importance of documenting recurrence to foster more research into the specific needs of this understudied population. Cancer Epidemiol Biomarkers Prev; 26(6): 809–15. ©2017 AACR.

Introduction

In 2016, there are approximately 3.5 million women living with a history of breast cancer in the United States (1). This number includes newly diagnosed women with breast cancer undergoing surgery and adjuvant treatment, long-term survivors who may be cured of the disease, and women who have experienced a recurrence after a disease-free interval. Distant metastatic cancers, including metastases found at diagnosis (de novo) and those occurring later in the disease course (distant recurrence), who represent the majority of cases, constitute the most advanced form of the disease. Many groups, including the Orphan Drug Program of the FDA, health services researchers, and especially the cancer survivorship and advocacy community are increasingly interested in assessing the prevalence of women with metastatic breast cancer (MBC), as these women have significant health care needs when resource utilization tends to be continuous and intensive (2–6).

The prevalence of women initially diagnosed with MBC can be directly estimated (7) using population-based cancer registry data on de novo MBC and vital status at the study cutoff date. However, estimating prevalence of those diagnosed with early-stage breast cancer who later have had a distant recurrence is challenging, as there are no nationally representative data that capture recurrence. Currently, registries in the United States do not routinely collect or report recurrence data.

In the absence of empirical data on the incidence of recurrent MBC, a back-calculation method, Mortality Incidence Approach MODel (MIAMOD; refs. 8, 9), has been used to reconstruct prevalence of recurrent cancer in Australia (10). This method calculates the incidence of MBC (de novo and distant recurrence) based on breast cancer mortality and MBC survival. The method has also been used to estimate the prevalence of breast cancer survivors in states within the United States (11) when cancer incidence data are not available over the long-term.

The objective of this article is to use national data on breast cancer mortality and MBC survival from Surveillance, Epidemiology and End Results (SEER) registries to estimate the prevalence of women living with MBC in the United States, including both women initially diagnosed with MBC and those who have progressed to distant MBC. We also calculate separately the prevalence of women diagnosed with de novo MBC in SEER and the United States (7). The SEER de novo MBC prevalence is compared with an estimate based on the back-calculation method to validate the method and calibrate survival (11).

Materials and Methods

Data sources and definitions

The Surveillance Epidemiology and End Results (SEER) Program collects clinical, demographic, and vital status information.
on all cancer cases diagnosed in defined geographic areas. Data included in this report are from the SEER-9 and SEER-11 registries (November 2015 Submission) obtained using SEER’Stat software version 8.3.2 (www.seer.cancer.gov/seerstat). SEER-9 covers approximately 11% of the U.S. population and includes Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah. For survival analyses, we used data from 1992–2012 from the SEER-11 registries which include SEER-9, Los Angeles and San-Jose Monterey. We only included invasive breast cancers.

Stage at diagnosis was defined using adjusted American Joint Committee on Cancer (AJCC) 6th edition staging classification (12). This stage definition uses extent of disease information for cases diagnosed in 1988–2003 and collaborative staging for cases diagnosed in 2004–2012. De novo MBC was defined as AJCC 6 stage IV which includes only tumors with distant metastasis. Stage IV from previous AJCC editions and distant stage from SEER historical summary staging classification include some locally advanced tumors without distant metastasis, for example, tumors with positive supraclavicular lymph node involvement without distant metastasis. Recurrent MBC was used to designate women initially diagnosed with AJCC stages I–III breast cancer, whose disease later progressed (metastasized) after treatment to distant organs or tissues.

Main inputs to the back-calculation methods are cancer deaths, all cause-deaths, population sizes, and MBC survival. We obtained U.S. female deaths due to breast cancer and all causes, from 1990 to 2012 from the National Center for Health Statistics (NCHS) and U.S. female populations by age, sex, race, and calendar year. Survival associated with cancer diagnosis was assessed via relative survival calculated using SEER’Stat. Relative survival is based on the ratio of overall survival (all causes of death) among cancer cases to the expected survival in individuals without cancer. The expected survival is estimated from U.S. life tables matched to the group of cancer patients by age, sex, race, and calendar year. Relative survival captures all excess mortality among cancer cases including deaths attributable to treatment and as such serves as a proxy for disease-specific survival that accounts for treatment-related mortality. In calculating relative survival, we excluded women diagnosed through death certificate or autopsy because of uncertainties in the diagnosis date. We also excluded cases with no follow-up information.

**Prevalence of de novo MBC using the counting method**

The prevalence of de novo MBC in the SEER-9 areas (counts and proportions) is calculated directly using the SEER’Stat counting method (7), which counts all women alive on December 31, 2013, with a previous diagnosis of stage IV breast cancer (1988–2012) in the SEER-9 areas. The method also adjusts for cases lost to follow-up. To estimate the de novo MBC prevalence counts in the United States, we applied the SEER-9 prevalence proportions by 5-year age group and race to the respective female U.S. populations.

**Modeling survival time from MBC including de novo and recurrence**

To model survival for de novo MBC cases, we estimated relative survival by age and year at diagnosis for women diagnosed with stage IV breast cancer from 1992 to 2012 in the SEER-11 areas. To extrapolate survival beyond the observed data, as required by the back-calculation method, we fit a Weibull mixture cure survival model to de novo MBC relative survival data. The mixture cure survival model assumes that a proportion of patients with cancer is cured of cancer whereas the remaining patients die following a Weibull survival distribution. While most patients with stage IV breast cancer die of their cancer, this model is used because it allows for modeling of long-term survivors and extrapolation of survival beyond the observed data. We fit a separate model to each of the 5 age groups (15–49, 50–64, 65–74, 75–84, 85–99) and used calendar year as a covariate in the model using the CANSURV software (13, 14; https://surveillance.cancer.gov/cansurv/).

Details of the model are provided in the Supplementary Materials.

Because population-level data on survival from MBC recurrence are unavailable, we use an adjustment to the de novo MBC survival based on a University of Texas M.D. Anderson Cancer Center (MDACC) study that included 2,881 and 643 women, retrospectively identified and diagnosed between 1992 and 2007 with recurrent and de novo MBC, respectively (15). The comparison of the overall survival curves for recurrent and de novo MBC, Figure 1 in Dawood and colleagues (15), showed an average risk of death for recurrent relative to de novo MBC of 1.35 [i.e., 1.35 = log(recurrent survival)/log(de novo survival)]. Recurrent MBC survival was estimated by applying the 1.35 relative risk adjustment to each of the modeled de novo MBC survival curves (recurrent MBC survival = de novo MBC survival1.35). We also performed sensitivity analyses and provided prevalence estimates using a lower and a higher relative risk adjustment of 1.2 and 1.5, respectively.

To model survival from de novo or recurrent MBC, we compute a weighted average of the de novo MBC survival and the recurrent MBC survival, that is, MBC survival = w × (de novo MBC survival) + (1 − w) × (recurrent MBC survival), where w is the fraction of breast cancer deaths that are a consequence of de novo MBC, and (1 − w) is the fraction of breast cancer deaths that are a consequence of recurrent MBC. We use incidence-based mortality by stage in SEER to estimate w. Details of the calculation are provided in the Supplementary Materials and Supplementary Fig. S1 which shows that the resulting estimated w is 0.2, implying that 20% of breast cancer deaths in a given year originate from women diagnosed with de novo MBC, whereas 80% are deaths from women diagnosed with earlier stage breast cancer who progressed to recurrent MBC.

**Back-calculation method**

We used MIMOD (8, 9) to estimate incidence and prevalence from breast cancer mortality and MBC survival. The method is based on the illness–death process and 2 equations relating incidence, survival, prevalence, and mortality. The method assumes that each observed breast cancer death is the result of MBC, either de novo or recurrent. The first equation specifies mortality as the sum of prior incidence and survival and back-calculates incidence of MBC (de novo or recurrent), by single-year ages and single calendar years, from breast cancer deaths and
MBC survival. The second equation is used to estimate prevalence from the estimated incidence and survival. The MIAMOD software can be downloaded from [http://www.eurocare.it/Miamod Piamod/tabid/60/Default.aspx](http://www.eurocare.it/Miamod Piamod/tabid/60/Default.aspx), and details of this application are included in the Supplementary Materials. Prevalence projections from 2014 to 2020 assume constant breast cancer mortality rates at 2014 levels and constant survival but use dynamic population size projections for these years.

To adjust for data inconsistencies, such as underreporting of deaths and misclassification of deaths to site of metastasis as found elsewhere (11), we calibrate the back-calculation method by comparing the SEER-9 counting-method prevalence of de novo MBC with the one obtained from the MIAMOD method. The calibration suggests adjusting MBC survival by a factor of 0.92 = exp(–0.08) to correct for 11% underestimation of the observed prevalence, that is, $S_{MBC}(t) = S_{SMBC}(t)^{0.92}$. Results from the calibration are shown in the Supplementary Figs. S2 and S3.

## Results

In 2013, the last year with observed data, we estimate a prevalence of MBC of 138,622; of which, 38,897 (28%) are survivors who were initially diagnosed with de novo stage IV disease and 99,725 (72%) survivors initially diagnosed with stage I–III breast cancer who later progressed to MBC (Table 1). The back-calculation method also estimates 50,344 new diagnoses of MBC in 2013, of which 12,966 (26%) are de novo and 37,378 (74%) recurrences, thus 3 in 4 are undocumented diagnoses of MBC. We project that by January 1, 2017, there will be 154,794 women living with MBC in the United States. Using relative risk adjustment of 1.5 and 1.2, instead of 1.35, we estimate 136,419 women living with MBC in 2013, of which 12,966 (26%) are de novo and 37,378 (72%) survivors initially diagnosed with stage I–III breast cancer who later progressed to MBC. The 5-year relative survival rate had a 2-fold increase from 18% to 36%, for women diagnosed with de novo MBC at age 15–49 between 1992–1994 and 2005–2012, respectively. Despite a poor prognosis, there is a small but meaningful percentage of these cases who survive 10 years or more; more than 11% of women diagnosed between 2000–2004 under the age of 64 years survived 10 years or more. Younger women diagnosed with de novo MBC have higher survival than women diagnosed at older ages (Fig. 3).

Figure 4 compares MBC survival in SEER and in the MDACC study cohort. The MDACC cohort included 2,881 and 643 women with recurrent and de novo MBC, respectively, retrospectively identified and diagnosed between years 1992 and 2007 and ages 17 and 91 years. To be comparable, we selected women diagnosed with de novo MBC in SEER in the same calendar years (1992–2007) and ages 15 through 84. In the MDACC cohort, the median age at diagnosis was 52 and 50 years for de novo and recurrent MBC, respectively, whereas in SEER, the median age at diagnosis was 61 years. Relative survival for women diagnosed with de novo MBC in the SEER areas was lower than overall survival among women in the MDACC cohort. The 4-year relative survival rate of de novo MBC in SEER was 27% compared with 41% and 29% overall survival for de novo and recurrent MBC in the MDACC cohort, respectively. Relative survival is generally higher than overall
survival for de novo MBC (Table 1), thus these results suggest that the MDACC cohort represents a lower risk cohort than the general population. The absolute difference decreased with longer follow-up and 10-year relative survival was 10% in SEER versus 14% in the MDACC for women diagnosed with de novo MBC (Fig. 4).

Discussion

Despite the progressive and incurable nature of almost all MBC, median survival after diagnosis with metastatic disease has been increasing, resulting in a growing number of women living with MBC in the United States. The increased survival is especially noted for women diagnosed at younger ages. We estimate a 2-fold increase in 5-year relative survival rate from 18% to 36%, for women diagnosed with de novo stage IV at age 15–49 between 1992–1994 and 2005–2012, respectively, translating into an increase of approximately one third in the number of women living with MBC, from 105,354 in 1990 to 138,622 in 2013. We further project that by 2017, there will be 154,794 women living with MBC in the United States.

To our knowledge, this is the first time that the number of women living with MBC in the United States has been estimated. These estimates provide a new perspective on the population burden of breast cancer and have great potential significance to the research and advocacy community working on behalf of patients with MBC and their families.

Other studies have also shown improvement in survival for women with de novo distant disease or metastatic recurrence (16–18), attributed to improved treatment. The improvement in MBC survival may also be explained by changes in staging. A study using SEER data (19) has shown that incidence of distant breast cancer has been increasing, especially among young women (Supplementary Fig. S4). Also, the incidence of stage III and unknown stage has been decreasing (Supplementary Fig. S5). Thus, although survival may have increased because of improvements in treatment, part of the increase may be also due to stage migration from stage III or unstaged to stage IV or early detection of stage IV, likely due to increasing availability of better imaging techniques.

Strengths of our study include the large population size, the population-based setting, the long follow-up, and the fact that we used consistent definitions of staging and other variables across time. The calibrated back-calculation method showed a very good agreement with reported incidence and directly estimated prevalence of de novo MBC in the SEER areas. The calibration corrects for possible underreporting and miscategorization of cause of death.

The main limitation of this study is the absence of population-based survival estimates following MBC recurrence. To represent...
survival/mortality associated with MBC recurrence, we used a 1.35 higher risk of cancer death (inflation factor) for recurrent MBC relative to de novo disease based on a single-institution study conducted at MDACC (15). This factor accounts for greater susceptibility to the cancer as well as greater vulnerability to treatment morbidities due to accumulation of cancer treatments.

![Figure 3](https://example.com/figure3.png)

**Figure 3.**
Relative survival by time since diagnosis for women diagnosed with de novo stage IV in the SEER-11 areas between 2005 and 2012 at different age groups.

### Table 2
Number of women, median overall and relative survival in months and 5-year relative survival in percentage (95% confidence interval) for women diagnosed with de novo stage IV breast cancer in the SEER-11 areas by grouped age and year at diagnosis.

<table>
<thead>
<tr>
<th>Year</th>
<th>Age, y</th>
<th>N</th>
<th>Overall</th>
<th>Relative survival</th>
<th>5-y relative survival (95% CI)</th>
<th>10-y relative survival (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992–1994</td>
<td>15–49</td>
<td>430</td>
<td>22.2</td>
<td>22.3</td>
<td>18% (14%–21%)</td>
<td>10% (8%–14%)</td>
</tr>
<tr>
<td>1992–1994</td>
<td>50–64</td>
<td>777</td>
<td>18.4</td>
<td>19.1</td>
<td>15% (13%–18%)</td>
<td>8% (6%–11%)</td>
</tr>
<tr>
<td>1992–1994</td>
<td>65–74</td>
<td>598</td>
<td>16</td>
<td>17.6</td>
<td>15% (12%–18%)</td>
<td>7% (5%–10%)</td>
</tr>
<tr>
<td>1992–1994</td>
<td>75–84</td>
<td>442</td>
<td>10.1</td>
<td>10.9</td>
<td>16% (12%–20%)</td>
<td>7% (4%–11%)</td>
</tr>
<tr>
<td>1992–1994</td>
<td>85+</td>
<td>168</td>
<td>3.8</td>
<td>4.1</td>
<td>6% (2%–13%)</td>
<td>4% (0%–16%)</td>
</tr>
<tr>
<td>1992–1994</td>
<td>All ages</td>
<td>2,415</td>
<td>15.7</td>
<td>16.7</td>
<td>15% (14%–17%)</td>
<td>8% (7%–9%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>15–49</td>
<td>894</td>
<td>24.5</td>
<td>24.7</td>
<td>24% (21%–27%)</td>
<td>11% (9%–15%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>50–64</td>
<td>1,321</td>
<td>20.3</td>
<td>20.6</td>
<td>21% (18%–23%)</td>
<td>10% (8%–12%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>65–74</td>
<td>978</td>
<td>14.4</td>
<td>15.2</td>
<td>17% (15%–20%)</td>
<td>6% (5%–8%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>75–84</td>
<td>799</td>
<td>10.4</td>
<td>11.8</td>
<td>13% (10%–16%)</td>
<td>7% (5%–10%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>85+</td>
<td>292</td>
<td>4.7</td>
<td>5.5</td>
<td>16% (10%–23%)</td>
<td>8% (2%–21%)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>All ages</td>
<td>4,284</td>
<td>16.5</td>
<td>17.7</td>
<td>19% (17%–20%)</td>
<td>8% (8%–9%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>15–49</td>
<td>1,307</td>
<td>29</td>
<td>29.3</td>
<td>29% (26%–31%)</td>
<td>14% (12%–16%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>50–64</td>
<td>2,270</td>
<td>24.6</td>
<td>25.1</td>
<td>24% (23%–26%)</td>
<td>11% (10%–13%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>65–74</td>
<td>1,319</td>
<td>18.9</td>
<td>20.3</td>
<td>20% (18%–23%)</td>
<td>8% (6%–10%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>75–84</td>
<td>1,142</td>
<td>10.3</td>
<td>11.4</td>
<td>15% (13%–18%)</td>
<td>8% (6%–10%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>85+</td>
<td>436</td>
<td>5.7</td>
<td>7.2</td>
<td>14% (9%–20%)</td>
<td>9% (3%–19%)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>All ages</td>
<td>6,474</td>
<td>19.8</td>
<td>21.1</td>
<td>22% (21%–23%)</td>
<td>10% (9%–11%)</td>
</tr>
<tr>
<td>2005–2012</td>
<td>15–49</td>
<td>2,748</td>
<td>38.4</td>
<td>38.7</td>
<td>36% (34%–38%)</td>
<td>—</td>
</tr>
<tr>
<td>2005–2012</td>
<td>50–64</td>
<td>4,861</td>
<td>29</td>
<td>29.7</td>
<td>25% (24%–27%)</td>
<td>—</td>
</tr>
<tr>
<td>2005–2012</td>
<td>65–74</td>
<td>2,468</td>
<td>23.3</td>
<td>24.5</td>
<td>24% (22%–26%)</td>
<td>—</td>
</tr>
<tr>
<td>2005–2012</td>
<td>75–84</td>
<td>1,820</td>
<td>12</td>
<td>14</td>
<td>18% (16%–21%)</td>
<td>—</td>
</tr>
<tr>
<td>2005–2012</td>
<td>85+</td>
<td>865</td>
<td>6</td>
<td>8.2</td>
<td>13% (9%–17%)</td>
<td>—</td>
</tr>
<tr>
<td>2005–2012</td>
<td>All ages</td>
<td>12,762</td>
<td>25.2</td>
<td>26.9</td>
<td>26% (25%–27%)</td>
<td>—</td>
</tr>
</tbody>
</table>

**Abbreviation:** CI, confidence interval.
received before the point of recurrence. Other causes of death, not associated with breast cancer or its treatment, are assumed to be similar between patients with de novo and recurrence MBC. Sensitivity analyses to this assumption showed that U.S. prevalence of MBC estimates would vary from 136,000 to 178,000 in 2017 using a higher relative risk of death (RR = 1.5) or a lower relative risk of death (RR = 1.2) for recurrent MBC survival compared with de novo MBC survival. However, we noted that SEER survival was lower than survival in the MDACC. Possible explanations may be the fact that MDACC patients were younger than SEER patients and that, by definition, they were in treatment at a major cancer center, and therefore more likely to receive optimal care. Given these differences, collection of additional data to estimate recurrent MBC survival would be of value.

We used the adjusted 6th edition stage IV to define MBC to only include tumors that have metastasized to distant sites. If instead, we used SEER historical distant stage definition, prevalence would have been higher, as some tumors without a distant metastasis are included in this definition.

At one time, a diagnosis of distant recurrence or de novo stage IV meant that death from breast cancer was likely to be imminent. Today, with the development of new therapies that target the drivers of breast cancer, and with improved palliative care, MBC is not the immediate death sentence it once was. With optimal care, women with MBC can and often do live for years with reasonable quality of life, albeit undergoing constant treatment to keep their disease under control.

This study demonstrates that there are a large number of women in the United States living with MBC and that this number has increased in more recent years, likely the result of treatment and aging of the U.S. population. This study demonstrates a growing burden of MBC in the United States. It also makes clear that the majority of patients with MBC, the three out of four who are diagnosed with nonmetastatic cancer but progress to distant disease, has never been properly documented. Given the growing burden of MBC, it is critical to collect data on recurrence to foster more research into the specific needs of this understudied population.

In an ideal world, a cancer registry would record the experiences of all patients throughout the entire cycle of disease, enabling researchers, health policy experts and planners, providers, patients, and advocates to understand the full impact of cancer. Finding ways to incorporate information on metastatic disease progression would be an important advance and a key first step toward a comprehensive assessment of the population burden of disease.

Disclosures of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Disclaimer
The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the NIH.

Authors’ Contributions
Conception and design: A.B. Mariotto, M. Hurlbert, L. Penberthy, M. Mayer
Development of methodology: A.B. Mariotto, R. Etzioni
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.B. Mariotto
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.B. Mariotto, R. Etzioni
Writing, review, and/or revision of the manuscript: A.B. Mariotto, R. Etzioni, M. Hurlbert, L. Penberthy, M. Mayer
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A.B. Mariotto, L. Penberthy
Study supervision: A.B. Mariotto

Acknowledgments
The authors would like to thank Joan Warren and Julia Rowland whose comments/suggestions helped improve and clarify this article. The authors acknowledge the members of the Metastatic Breast Cancer Alliance and patients.
living with MBC who inspired this work and encourage readers to learn more at www.mbcalliance.org.

Grant Support
This work was supported by the National Cancer Institute at the NIH.

References


