Using Claims to Study Cancer Survivors
Objectives

• Characterize the burden of cancer survivorship on patients and the healthcare system.

• Quantify and predict excess risks of:
  – Functional impairment
  – Hospitalization for comorbid disease
  – Cost
  – In survivors versus non-cancer controls
Importance

• Preliminary data portend high morbidity and cost burden from cancer survivors.

• Current data are insufficient for informing policy and clinical interventions.
Importance

• Findings seek to impact 18 million cancer survivors expected by 2022.

• Vulnerable group and priority patient population to study.
Background

• The number of cancer survivors is rising.

Background

• Cancer survivors exert
  – Disproportionate healthcare costs
  – Decreased economic productivity
• $4,500 excess per patient per year

• Our pilot data: $3,500-$5,000 per survivor per year excluding treatment costs.
• Total: ~$40 billion excess, 75% hospitalizations
ASCO has sounded a call...
Obstacles to Research

• Cancer survivors are difficult to study:
  – Heterogeneous cancer diagnoses
  – Heterogeneous treatment histories
  – Non-cancer health characteristics

• Post-treatment trajectory requires longitudinal study
  – Oncologist’s insights on history
“Lost in Transition”

- Non-oncology, usual care
- Survivors understood based on non-oncologic paradigm
- Oncologists lack longitudinal feedback
- Prospective strategies to risk-stratify survivors and prevent adverse outcomes
Current Gaps & Limitations

• Limited conceptual approach

• Focuses on *treatment* to predict subsequent adverse outcomes

• Uni-dimensional approaches result in ineffectual risk-prediction models
An expanded conceptual model...
Uni-Dimensional Model...

PATIENT

Got Adriamycin

Got Radiation

Cardiac event

Cancer Relapse
Improved model...

PATIENT

Cancer-Related

Demographic Through treatment Education Behaviors

Non-Cancer Longitudinal

Before treatment

Beyond treatment
Data source

- Population-based cancer survivor cohort
- Derived using the SEER-Medicare data
- Propensity score matched 1:1 with non-cancer controls.

- SEER-Medicare-Health Outcomes Survey (MHOS) linked data?
Aims: Compare Cancer vs. Controls

Aim #1: What is the excess risk of hospitalization for commonest, costliest comorbid diagnoses?

Myocardial infarction, heart failure, stroke, infection
Aims: Compare Cancer vs. Controls

Aim #2: What are excess sources of healthcare resource use?

Total costs, inpatient costs, other measures such as hospital readmission after index, LOS, use of resources such as intensive care and SNF
Aims: Compare Cancer vs. Controls

Aim #3: What is the excess risk of functional impairment in cancer survivors, including physical, (cognitive and social function)?
Conceptual Issues

• What is a cancer “survivor”?

• What may be the unique value of claims studies in providing insights about survivors?
Conceptual Issues

• What is a cancer “survivor”?  
  – Acute: at the time of diagnosis and treatment
  – Extended: immediately after tx completed (usually months)
  – Permanent: after a longer period (usually yrs) 

From cancer.net (ASCO)
Conceptual Issues

• What is “survivorship”? 
  – “The process of living with, through, and beyond cancer” 
  – Survivorship begins at diagnosis 
  – Can include recurrence 
  – Can include management of cancer as a chronic disease 
  – Can include life after cancer

From cancer.net (ASCO)
Survivorship Paradigm

Analytic perspective:

Cancer diagnosis
A period of follow-up
Unique value of claims? (S-M)

- Nationally diverse, population-based cohort
- Relatively detailed Cancer Dx, Ca Tx
- Windows: Pre-cancer, Ca Tx, Post-Tx, End of life
- Cost
• Can morbidity and/or cost be predicted by variables just before, during, or after cancer tx?
• Are these preventable outcomes?
Are subsequent morbidities and cost(s) in excess of that predicted for never cancer controls?
Methodological Issues

• Patient selection:
  – cancer survivors
  – matched non-cancer controls

• Measures
  – Functional status
  – Comorbidity
  – Morbidity
  – Cost
Patient selection

• How to select cancer survivors?

Obstacles to Research

• Cancer survivors are difficult to study:
  – Heterogeneous cancer diagnoses
  – Heterogeneous treatment histories
  – Non-cancer health characteristics

• Include multiple, common cancer diagnoses to reflect heterogeneity
Patient selection: Pilot group

- Breast, prostate, lung included
- Future: lymphoma, others?

- Current selection creates a patient subgroup by cancer diagnosis and matches accordingly.

  Vs

- Select all cancer patients and match as one large group
Cancer Survivors ("cases")

Pre-cancer Hx → Cancer Dx

Enrolled in Medicare any time btw 1991-2009 AND Cancer Dx

Require continuous FFS, Part A & B for 1 yr after cancer dx date to document treatment details

Ca Tx → C

NC

Survivor selection paradigm

Require continuous coverage 30 days after MI

Cont coverage during this interval
Sample cancer denominator: Prostate

<table>
<thead>
<tr>
<th>Condition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosed with prostate cancer, from 1991 to 2009</td>
<td>SEER</td>
</tr>
<tr>
<td>Exclude if any second cancer diagnosis</td>
<td>SEER</td>
</tr>
<tr>
<td>Exclude if AJCC stage IV at diagnosis</td>
<td>SEER</td>
</tr>
<tr>
<td>Age 66+ at diagnosis</td>
<td>SEER</td>
</tr>
<tr>
<td>Exclude no pathological confirmation</td>
<td>SEER</td>
</tr>
<tr>
<td>Treated cancer: prostatectomy or radiation or hormone therapy within 1 year of diagnosis date</td>
<td>Medicare</td>
</tr>
<tr>
<td>Define cancer-free interval after 1 year (primary interval); No chemo, XRT, ?further surgery; ICD diagnosis code for mets</td>
<td>Medicare</td>
</tr>
</tbody>
</table>
Sample control denominator:

<table>
<thead>
<tr>
<th>Condition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-cancer patients</td>
<td>SEER SUMDENOM</td>
</tr>
<tr>
<td>Men</td>
<td>SEER SUMDENOM</td>
</tr>
<tr>
<td>Medicare enrollment year from 1991 to 2009</td>
<td>Medicare</td>
</tr>
<tr>
<td>Reason for entitlement is age (not ESRD, disability)</td>
<td>Medicare</td>
</tr>
<tr>
<td>Exclude if patient found in cancer sample</td>
<td>SEER</td>
</tr>
</tbody>
</table>
Propensity Score Match

- Year of entitlement
- Age
- [Sex]
- Race
- SEER Region
- Dual Eligibility
## Propensity Score Match

<table>
<thead>
<tr>
<th></th>
<th>Cancer: 111864</th>
<th>Control: 192622</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Before Match</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. After Match</td>
<td>Cancer: 106536</td>
<td>Control: 106536</td>
</tr>
</tbody>
</table>
Analytic perspective: Cancer diagnosis
A period of follow-up
How to match index dates?

Cancer Survivors ("cases")

Ca Tx

1 year after diagnosis (arbitrary window)

H

cont coverage during this interval

Never Cancer ("controls")

H

Matched DUMMY TIME ZERO
A few alternatives
The last year of life

Cancer Survivors ("cases")

Never Cancer ("controls")

Follow backwards for the year before death
Breast Cancer Survivors ("cases")

Enrolled in Medicare any time btw 1991-2009 AND Had breast cancer in that period

Never Cancer ("controls")

Enrolled in Medicare any time btw 1991-2009

Ca Tx

MI

Require continuous FFS, Part A & B for 1 yr after cancer dx date to document treatment details

± cont coverage during this interval

Identify 1st MI admission in claims. Time ZERO

Require continuous coverage 1 yr prior to MI and 30 days after MI

± cont coverage during intervals

Identify all MI admissions ever claimed in the patient’s entire claims. All are potential sources of match.

If there was an MI within the last 30 days (or 1 yr? or whatever window seems clinically relevant?) exclude it, because it could represent a “readmission” as opposed to “primary event”.

Create the best match at the time of MI for the Case and Control: Age, Year of Event, Race, and SEER registry (essentially city/state).

*In this case, the 2nd MI was excluded from the match pool because it occurred in close proximity to the initial event.

Scenario #1: Each Case Contributes 1 Event, Each Control Contributes 1 Event
Scenario #2: Each Case Contributes 1 Event, Each Control Contributes Up to Multiple Events

Breast Cancer Survivors ("cases")

Never Cancer ("controls")

Identify 1st MI admission in claims. Time ZERO for Case 1

Identify 1st MI admission in claims. Time ZERO for Case 2

Require a minimum period between MI #1 and MI #2, for example, 1 year, or 3 years, or 5 years between events to say that these two MI's, though they come from the same person, can be independently matched to two different Cases’ MIs.
Breast Cancer Survivors ("cases")

Never Cancer ("controls")

Scenario #3: Cases Or Controls Can Contribute Multiple Events

Minimum interval between MI #1 and MI #2, for example, of 1 year, or 3 years, or 5 years would be required here.
Feasibility?
### Cancer

#### a. Five years for all patients

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2702</td>
<td>2544</td>
<td>2491</td>
<td>2421</td>
<td>2424</td>
</tr>
<tr>
<td>Inpatient</td>
<td>1842</td>
<td>1874</td>
<td>1888</td>
<td>1874</td>
<td>1905</td>
</tr>
<tr>
<td>Percent</td>
<td>68.2%</td>
<td>73.7%</td>
<td>75.8%</td>
<td>77.4%</td>
<td>78.6%</td>
</tr>
</tbody>
</table>

#### b. Three years, excluding patients who died within 3 years

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>19794.7</td>
<td>20386.96</td>
<td>22678.41</td>
</tr>
<tr>
<td>Inpatient</td>
<td>14536.9</td>
<td>15150.84</td>
<td>16983.34</td>
</tr>
<tr>
<td>Percent</td>
<td>73.4%</td>
<td>74.3%</td>
<td>74.9%</td>
</tr>
</tbody>
</table>

#### c. Three years, excluding patients who died due to cancer.

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>19741.62</td>
<td>20254.26</td>
<td>21939.96</td>
</tr>
<tr>
<td>Inpatient</td>
<td>14556.31</td>
<td>15231.06</td>
<td>16879.41</td>
</tr>
<tr>
<td>Percent</td>
<td>73.7%</td>
<td>75.2%</td>
<td>76.9%</td>
</tr>
</tbody>
</table>

### Control

#### a. Five years for all patients

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1869</td>
<td>1971</td>
<td>1959</td>
<td>2053</td>
<td>2063</td>
</tr>
<tr>
<td>Inpatient</td>
<td>1568</td>
<td>1669</td>
<td>1644</td>
<td>1725</td>
<td>1720</td>
</tr>
<tr>
<td>Percent</td>
<td>83.9%</td>
<td>84.7%</td>
<td>83.9%</td>
<td>84.0%</td>
<td>83.4%</td>
</tr>
</tbody>
</table>

#### b. Three years, excluding patients who died within 3 years

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>16103.41</td>
<td>17098.71</td>
<td>18407.19</td>
</tr>
<tr>
<td>Inpatient</td>
<td>13338.26</td>
<td>14261.95</td>
<td>15329.33</td>
</tr>
<tr>
<td>Percent</td>
<td>82.8%</td>
<td>83.4%</td>
<td>83.3%</td>
</tr>
<tr>
<td></td>
<td>Breast cancer</td>
<td>Prostate cancer</td>
<td>Lung cancer</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>N(PS) Cancer/control</td>
<td>N(PS) Cancer/control</td>
<td>N(PS) Cancer/control</td>
</tr>
<tr>
<td>Any readmission after MI</td>
<td>22.00% 22.70%</td>
<td>21.80% 22.50%</td>
<td>27.00% 24.90%</td>
</tr>
<tr>
<td>Matched with age, race, state</td>
<td>110352 (248/1128) (137/603)</td>
<td>200748 (710/3261) (359/1595)</td>
<td>44616 (151/560) (75/301)</td>
</tr>
<tr>
<td></td>
<td>22.00% 21.60%</td>
<td>21.80% 22.10%</td>
<td>27.00% 21.50%</td>
</tr>
<tr>
<td>Matched with age, race, state, enrol_year, buyin**</td>
<td>110352 (248/1128) (221/1021)</td>
<td>197452 (695/3183) (537/2426)</td>
<td>44336 (150/556) (107/497)</td>
</tr>
</tbody>
</table>
Conclusions

• Cancer survivors are a impactful patient population, as they are growing in number and healthcare resource utilization.

• Little studied.

• Claims data may provide unique insights on this patient group.

• Comparing an observational cohort of survivors with matched controls is feasible.