Incidence and Prevalence Approaches for Estimating Disease Specific Costs

Martin L. Brown, Ph.D.

Health Services and Economics Branch

Applied Research Program

Division of Cancer Control and Population Sciences

AHRQ-NCI Workshop

Standardizing Costs for Healthcare Studies

December 7, 2007

National Cancer Institute

Overview

- Episode of care

- Prevalence, incidence, and costs of care

- Attributable cost

- Comparison of methods for estimating incidence and prevalence

Estimating Cost for An Episode of Care

- Goal: obtain consistent estimates of disease specific cost by comparable episodes of care

Challenges

- Clinically appropriate definition of episode of care may vary by
  - disease or condition
  - severity of disease
  - nature of disease control intervention (e.g. prevention, screening, treatment)
- Flow of cost may not be constant within episodes of care
- Assignment of mutually exclusive and exhaustive costs to disease entities is not obvious
- Medical technology, practice patterns and costs are dynamic, but health cost data is either cross-sectional or longitudinal over a relatively short period of observation

Alternative definitions of episodes of care

National Cancer Institute

- Prevalence
  - Cross-section of individuals with disease
  - Cost per patient
  - Aggregate costs
- Incidence
  - Longitudinal pattern following diagnosis
  - Cost per period or cost per patient
  - Cumulative: from diagnosis to year x (e.g. 5 years)
- Modeled Phase of Care
  - Costs in initial, continuing, and end-of-life phases applied to survival probabilities
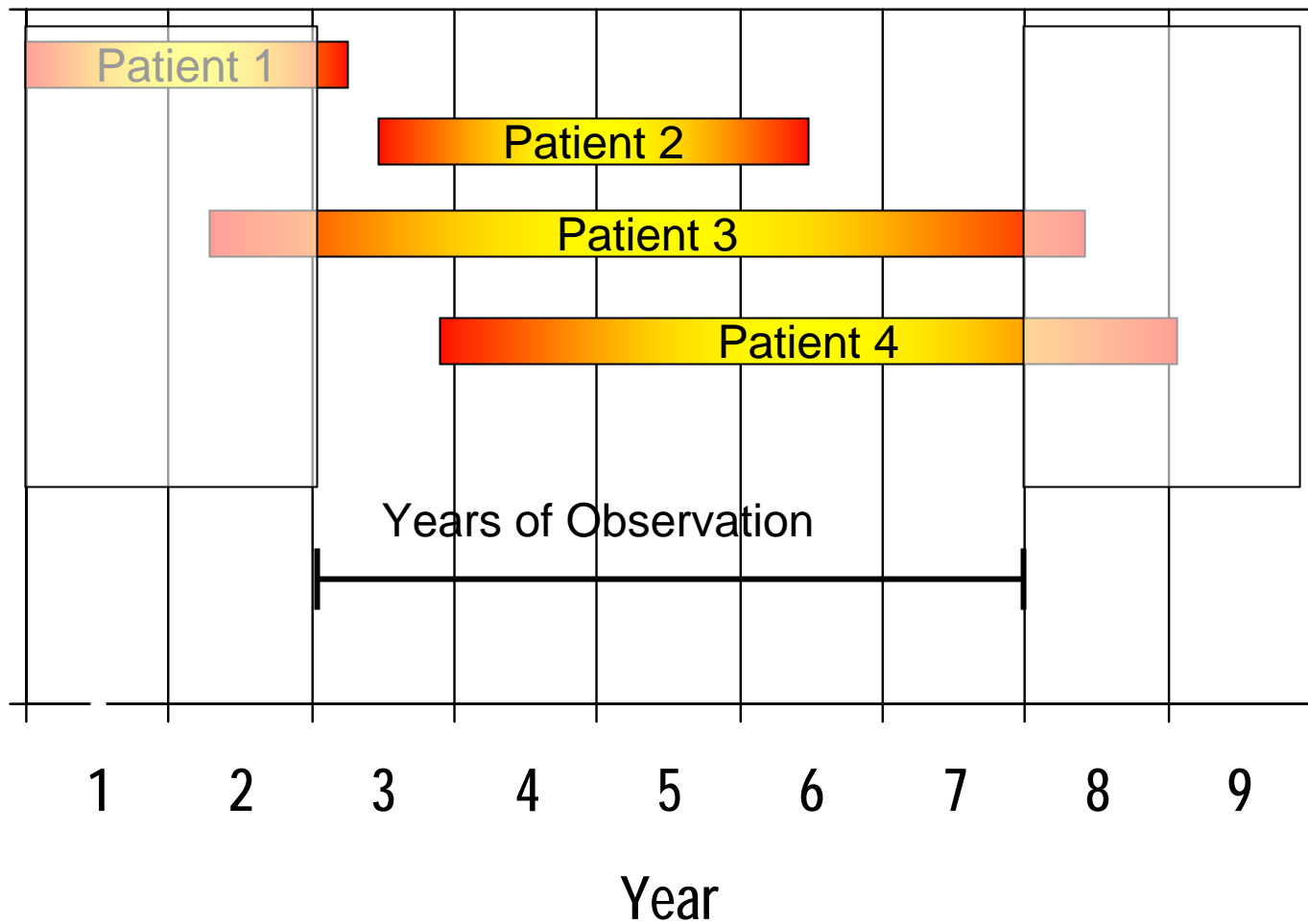 - Long-term estimates from diagnosis to death

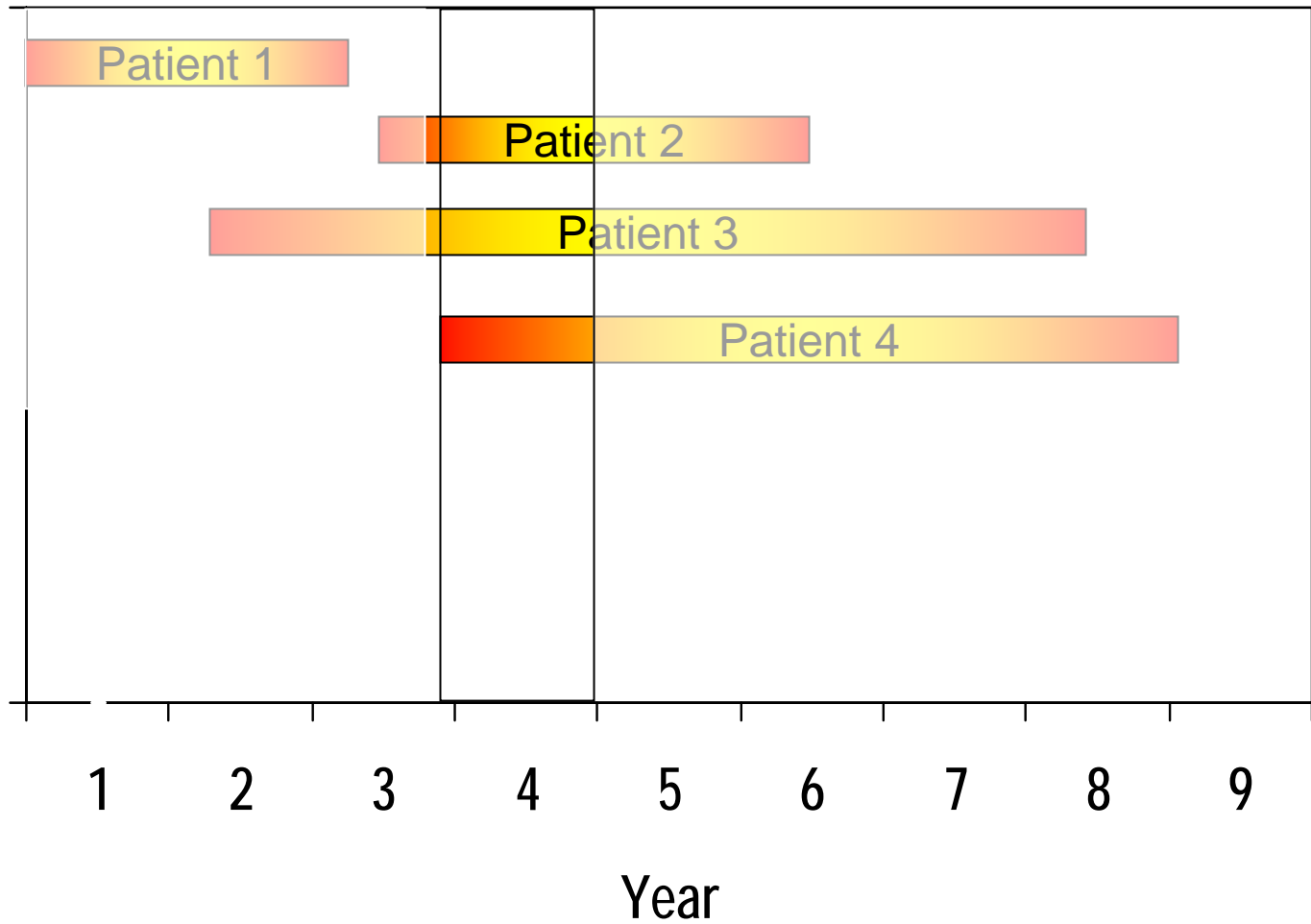# Breast Cancer Costs by Month From Diagnosis

Dollars
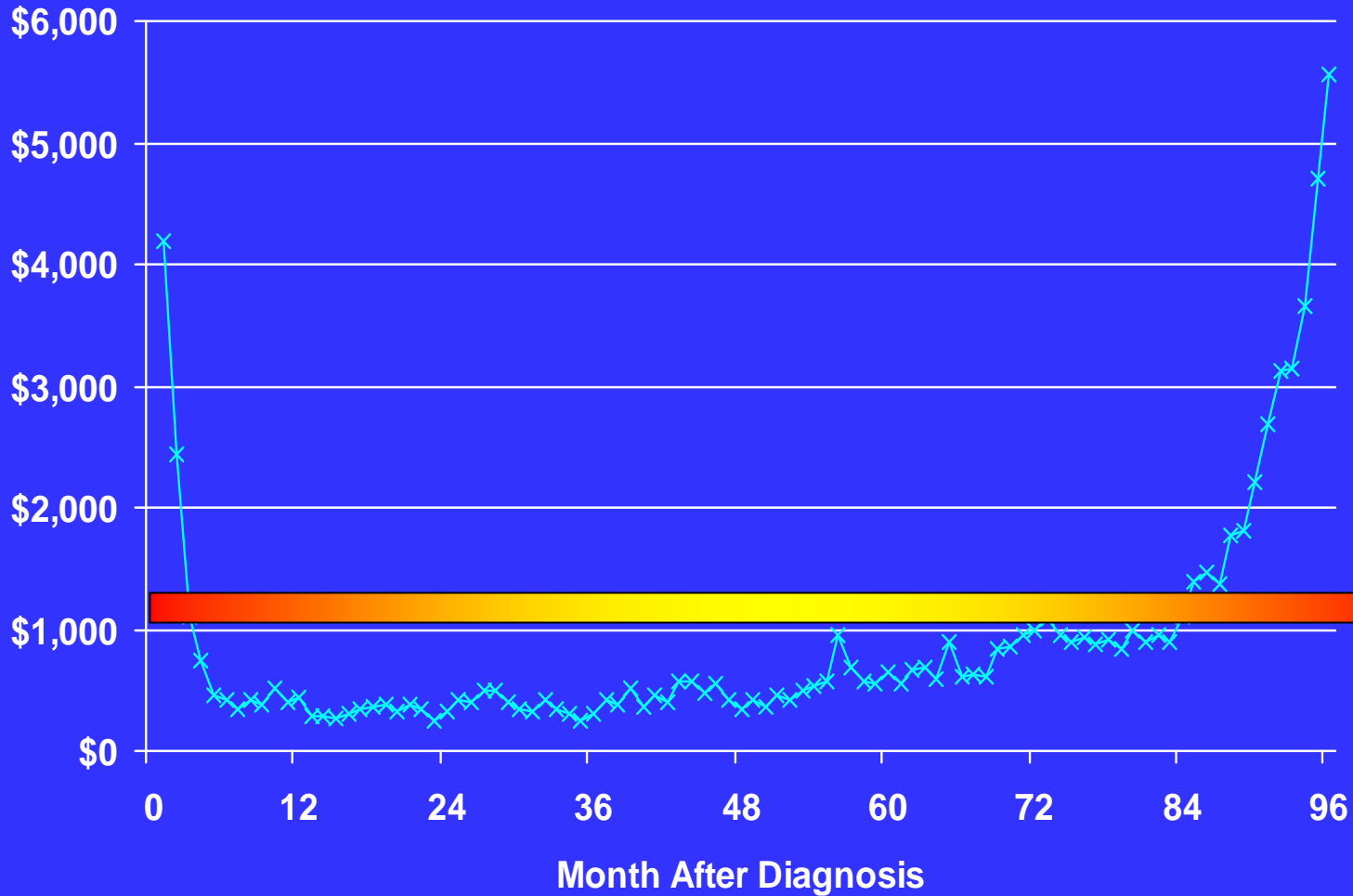
Month After Diagnosis
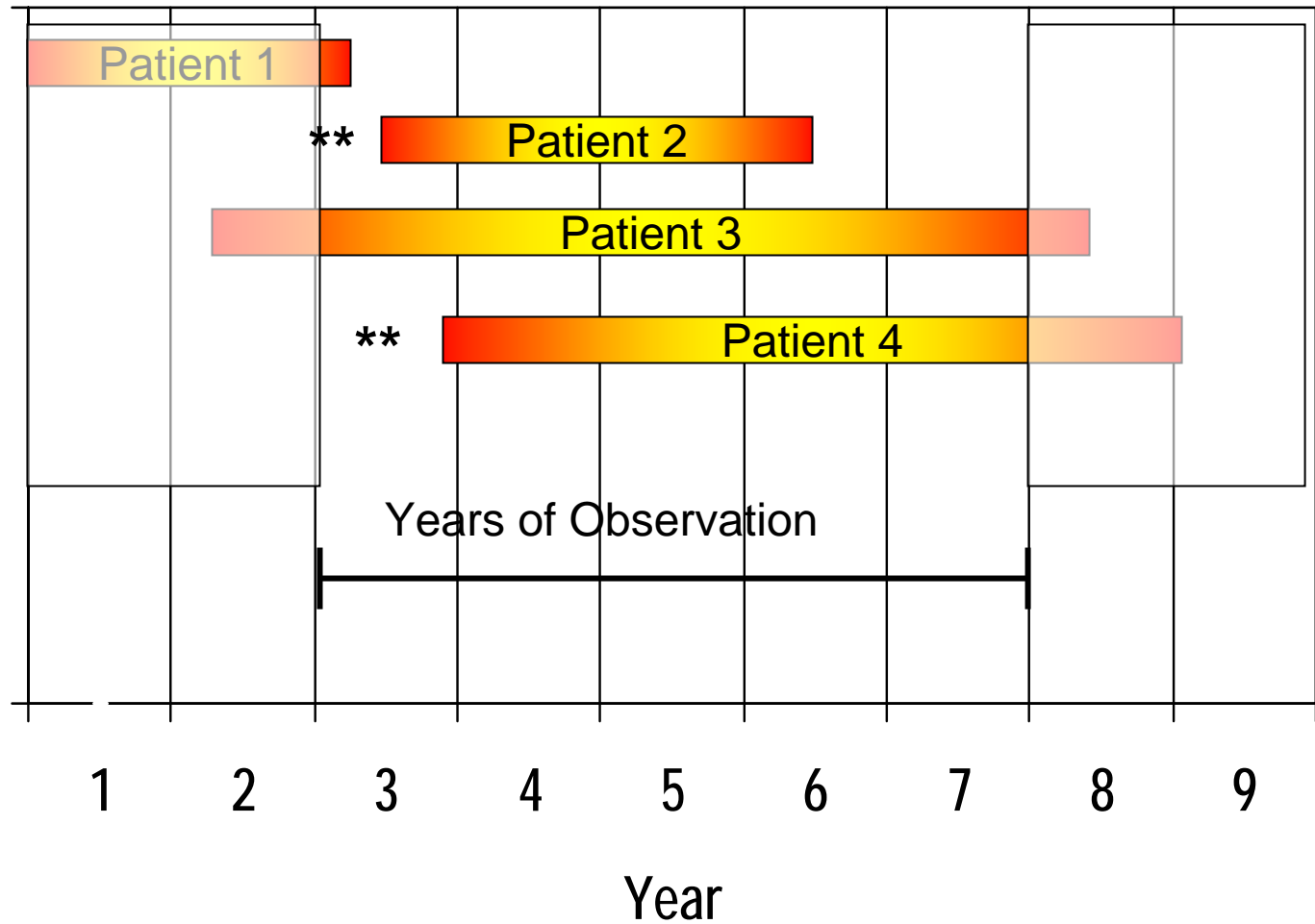
Phase of Care Specific Costs
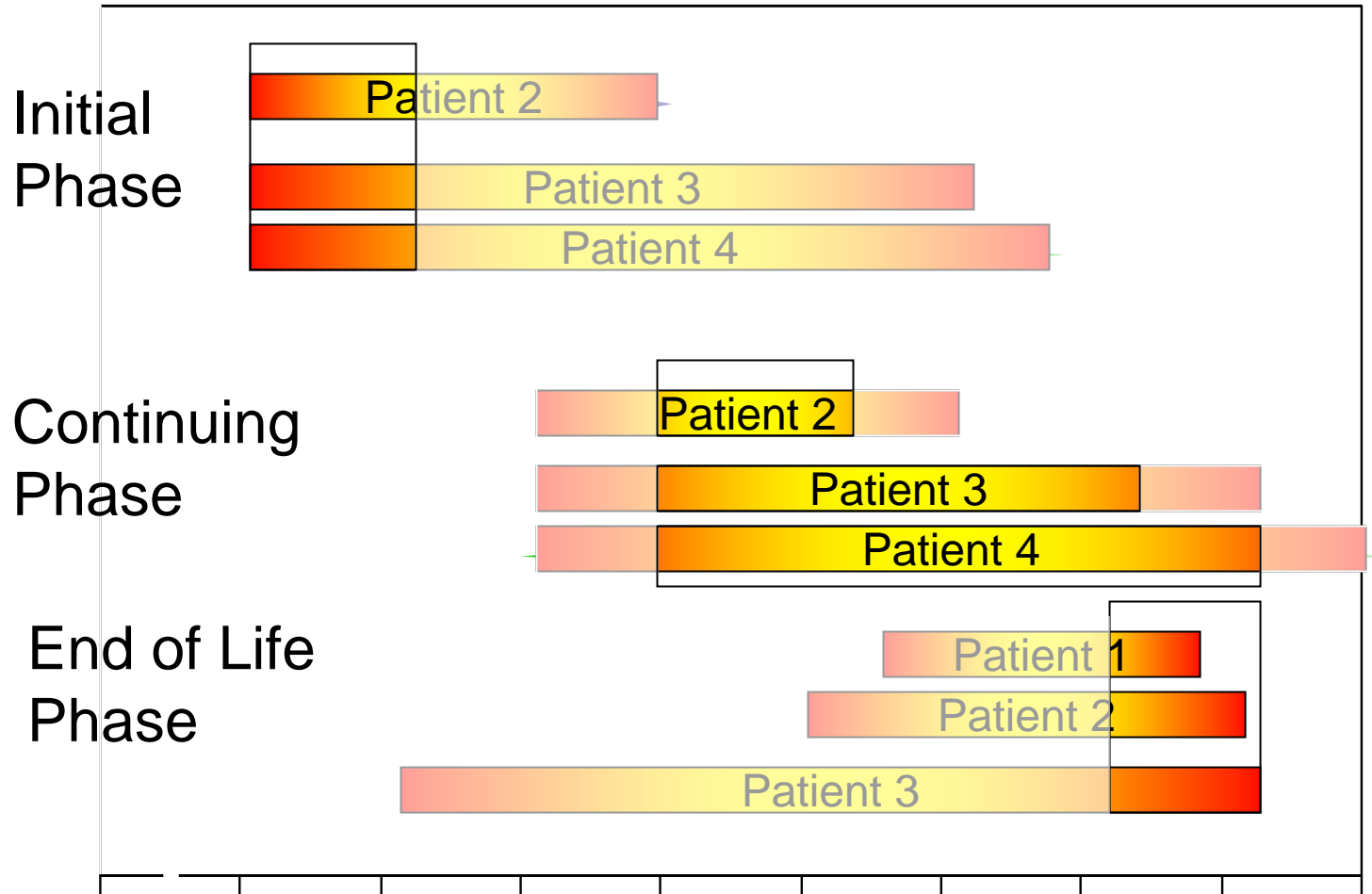
Incidence Cohort and Phase of Care Costs: Observed and Derived Measures

- Directly observed estimates can be compared:
  - Incidence: cost in year 1 since diagnosis
  - Phase of care: cost in initial phase (different from year 1 cost)
  - Cumulative cost to year X (from cohort)

- Derived estimates can be compared, using survival probabilities
  - Phase of care: cost in year 1 since diagnosis (from phase)
  - Cumulative costs to year X (from both)
  - "Life-time", cumulative cost from diagnosis to death (from both)
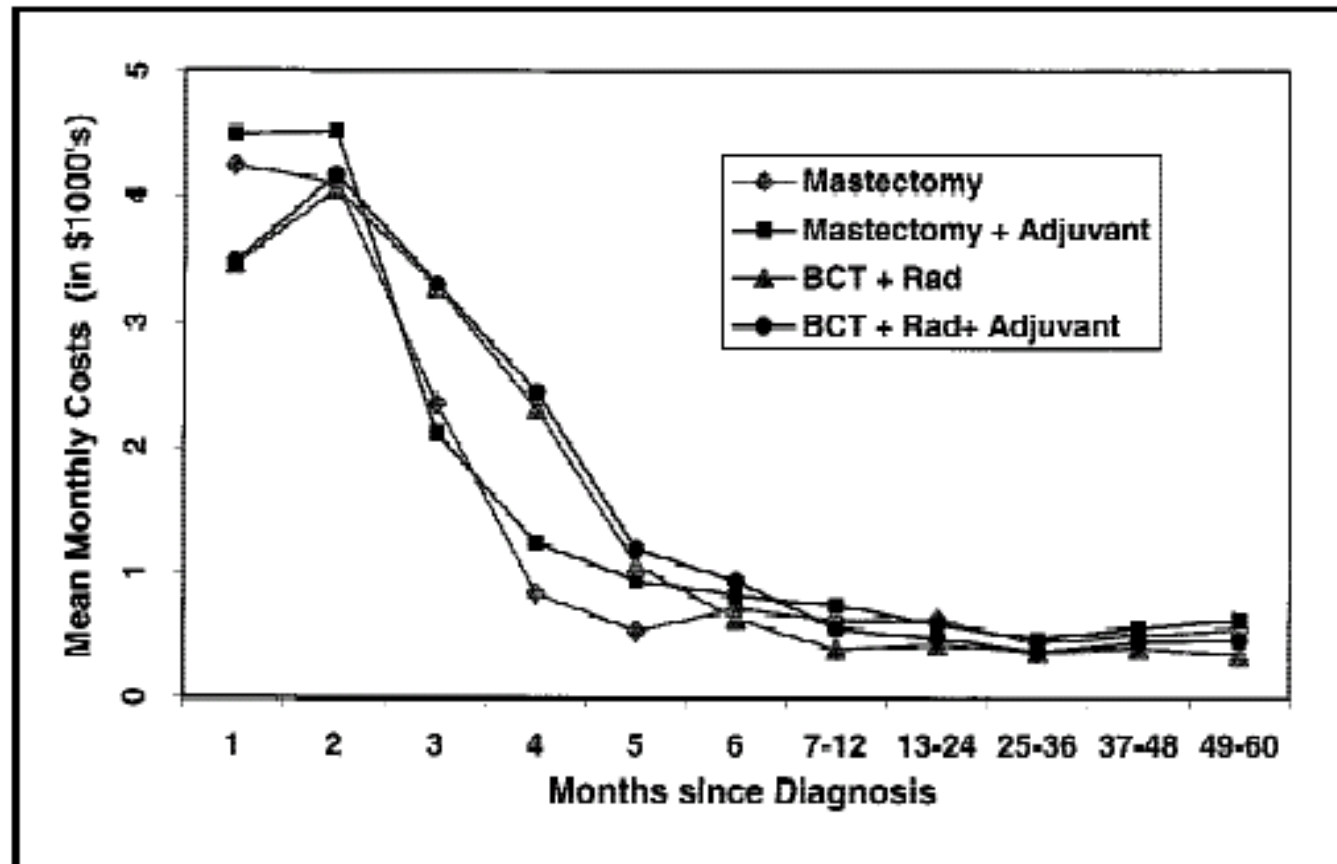
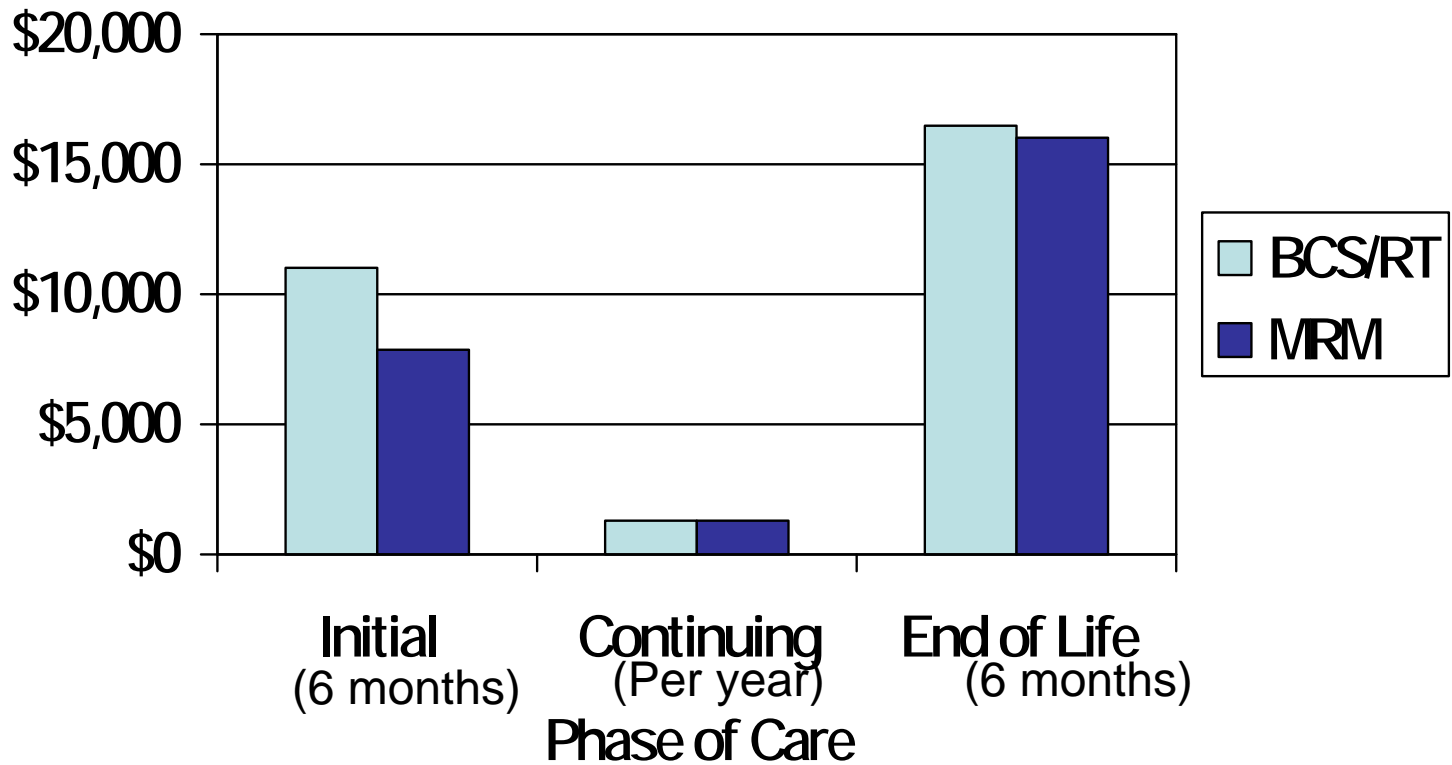# Incidence Cohort Approach: Comparing Treatments

Fig. 3. Adjusted mean total medical care costs by month from diagnosis for each treatment group adjusted for age and tumor stage. Adjuvant = hormonal therapy and/or chemotherapy; BCT = breast-conserving therapy; Rad = radiation therapy.

Source: Barlow et al. J Natl Cancer Inst 2001;93:447–55

Derived Cumulative Cost Estimates

- Incidence cost: Kaplan Meier Sample Average (KMSA)
  - Calculate average cost per month among those still alive at the end of each month
  - Multiply each monthly average by the (crude) survival probability
  - Sum across months (could also apply discounting)
- Phase of Care cost:
  - Analogous to the above, but apply appropriate survival probabilities to estimates from initial, continuing and last year of life phases of care
- When sufficient data is available to apply both methods, the incidence-KMSA and phase-specific approach result in similar estimates of cumulative cost (Etzioni et al. Health Econ 10(3):245-56

Prevalence Cost Estimates

- Observed

OR

- Derived using phase of care approach
  - Estimate phase specific prevalence during observation period using assumptions about incidence and survival (method developed by Angela Mariotto and colleagues)
  - Apply phase specific cost estimates
  - Used to project costs under varying assumptions

Projected Costs of Colorectal Cancer Care in United States, 2000-2020
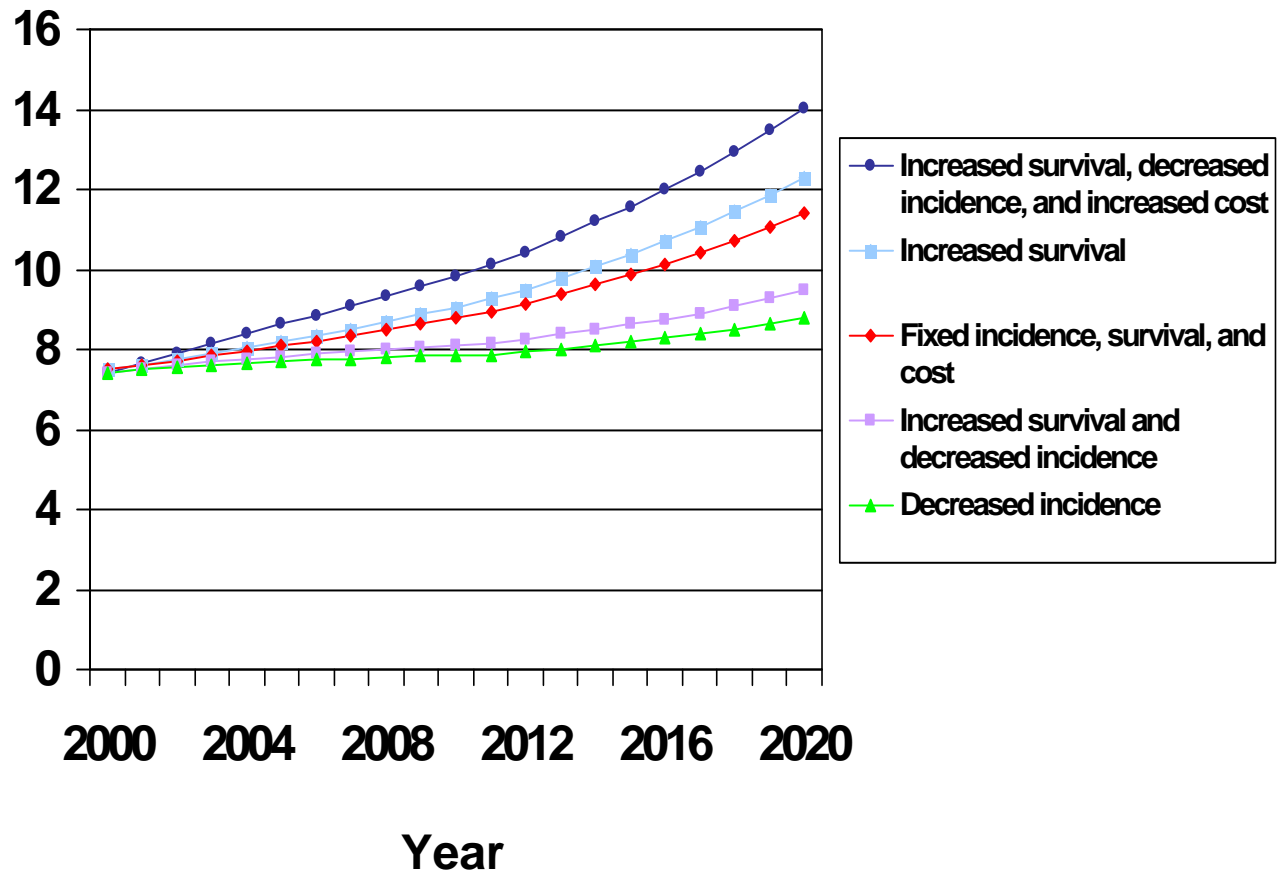
Source: Yabroff KR, et al. *Health Economics* 2007.

Pros and Cons: Prevalence

- – Relative easy to implement from many existing data sources
- – Useful for broad descriptive purposes
- • BUT
  - – Composition ("vintage") of prevalence/incidence cases may vary between data sources or be ill-defined
  - – Influenced by cost trajectory (e.g., u-shape vs. -- -shape)
  - – Not very useful for analytical/evaluative purposes, e.g. cost-effectiveness analysis

Pros and Cons: Incidence

- Useful for analytical/evaluative purposes
- Can be used to construct prevalence estimates
- BUT
  - High requirement for data:
    - Date of diagnosis
    - Survival
    - Comprehensive longitudinal costs
  - Hazard of death differs between disease cases and controls
  - Need large N if death events rare

Pros and Cons: Phase of Care

- Efficient use of data
- Flow of cost is homogeneous within phase
- Can be used to estimate prevalence cost
- BUT
  - High requirement for data
  - Depends on modeling assumptions
  - May not incorporate changes in practice patterns
  - Applicable to cancer, but is it feasible/relevant for other diseases?

Attributable Disease Specific Costs

- Case control approach
    - Match with similar control patients without the case condition (e.g., age, gender, region)
    - Match with same patients prior to diagnosis (pre-post)
- "Cost Driver" approach
- Clinical scenario/algorithm approach (e.g. POHEM)
- Macro-accounting approaches (e.g. regression models)

Comparing Data Sources and Methods for Estimating Costs

- Background Papers (in meeting binder)
- Data Sources
  - Claims
  - Registry-claims
  - Surveys (MEPS)
- Incidence Methods
  - Incidence cohort
  - Phase of care
- Prevalence Methods
  - Years of data to identify cases
  - Sample definition

SEER

- **S**urveillance **E**pidemiology and **E**nd **R**esults
- Started in 1973
- Population-based tumor registry that collects incident cases
- Currently represents 26% of the US population
- Detailed tumor characteristics at time of diagnosis
  - Histology
  - Stage
  - Grade
- Primary treatment within 4 months of diagnosis
- Active follow-up for vital status

SEER Registries

National Cancer Institute
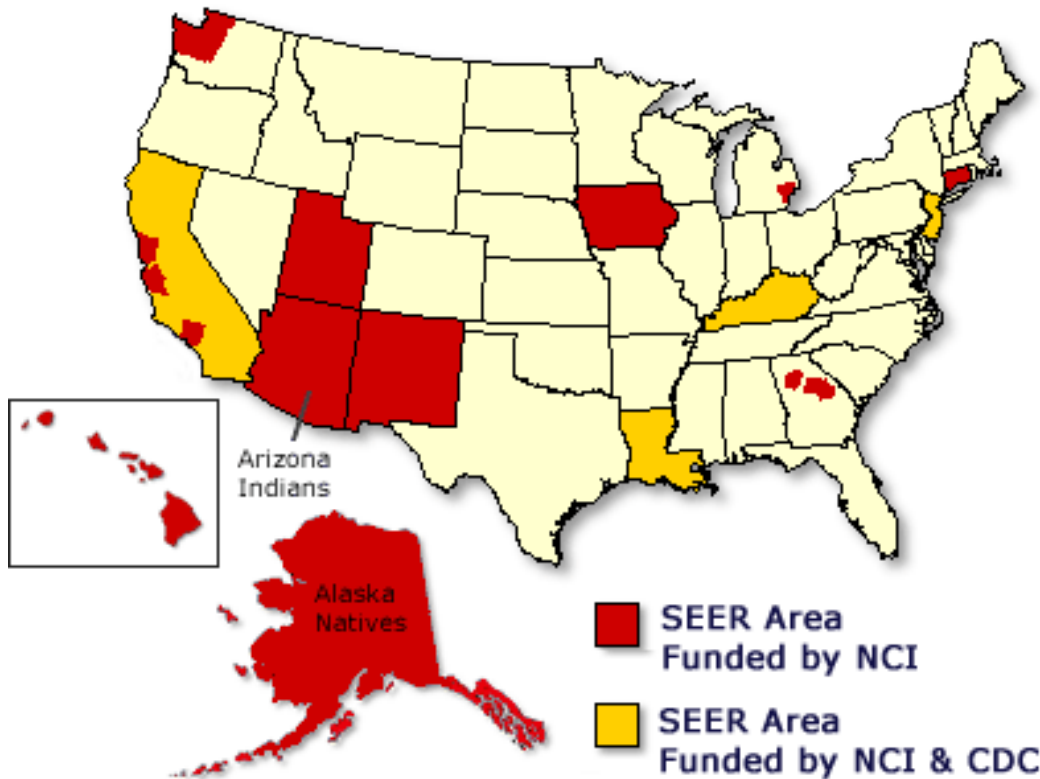
Los Angeles
San Francisco – Oakland
San Jose-Monterey
Greater California
Connecticut
Detroit
Atlanta
Rural Georgia
Hawaii
Iowa
Kentucky
Louisiana
New Jersey
New Mexico
Seattle-Puget Sound
Utah

Linked SEER-Medicare

- Cases identified in SEER registries linked to Medicare claims files

- Among individuals with cancer 65+, 94% linked to Medicare enrollment data

- Over 2.4 million persons with cancer and longitudinal information on health care before, during, and after diagnosis

- Use 5% random sample to identify similar individuals in SEER area without cancer (controls)

- Combines detailed tumor information with longitudinal service use and costs

Comparison of incidence and prevalence colorectal cancer costs

- Observation period 1998-2002
- Colorectal cancer patients aged 65+ at diagnosis/identification
- Controls matched on age, gender, geographic region
- Payments as proxy for cost
- Total and net costs of care

- Incidence
  - SEER-Medicare incidence cohort
  - Medicare claims only incidence cohort
  - SEER-Medicare Phase of Care
- Prevalence
  - SEER-Medicare
  - Medicare claims only
  - Medical Expenditure Panel Survey (MEPS)
  - Vary levels of sample definition (e.g., years, prior cancers)

Comparison of incidence approaches for estimating costs in colorectal cancer patients
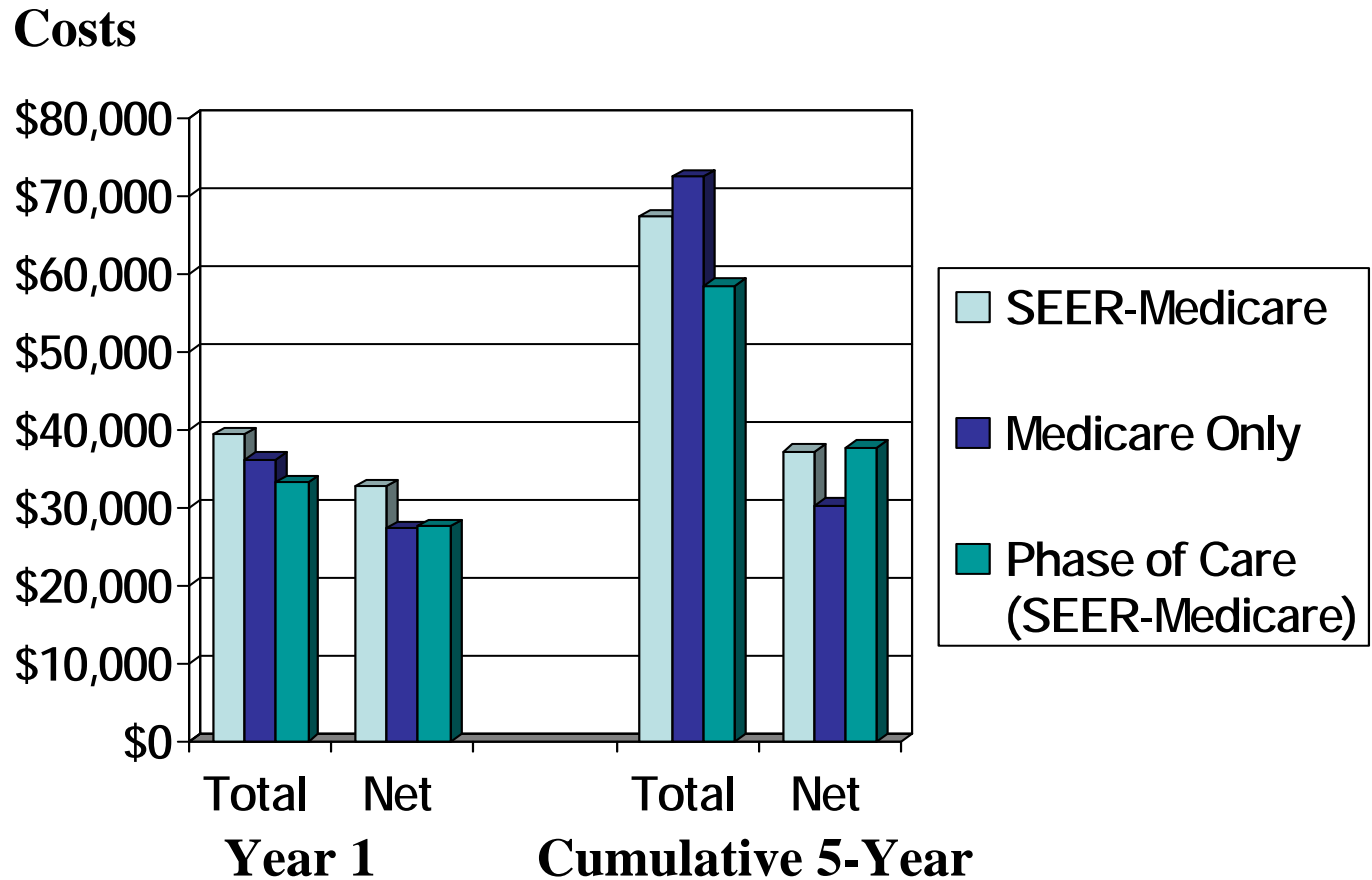
Source: Yabroff et al, Background paper

Comparison of prevalence approaches for estimating costs in colorectal cancer patients

National Cancer Institute
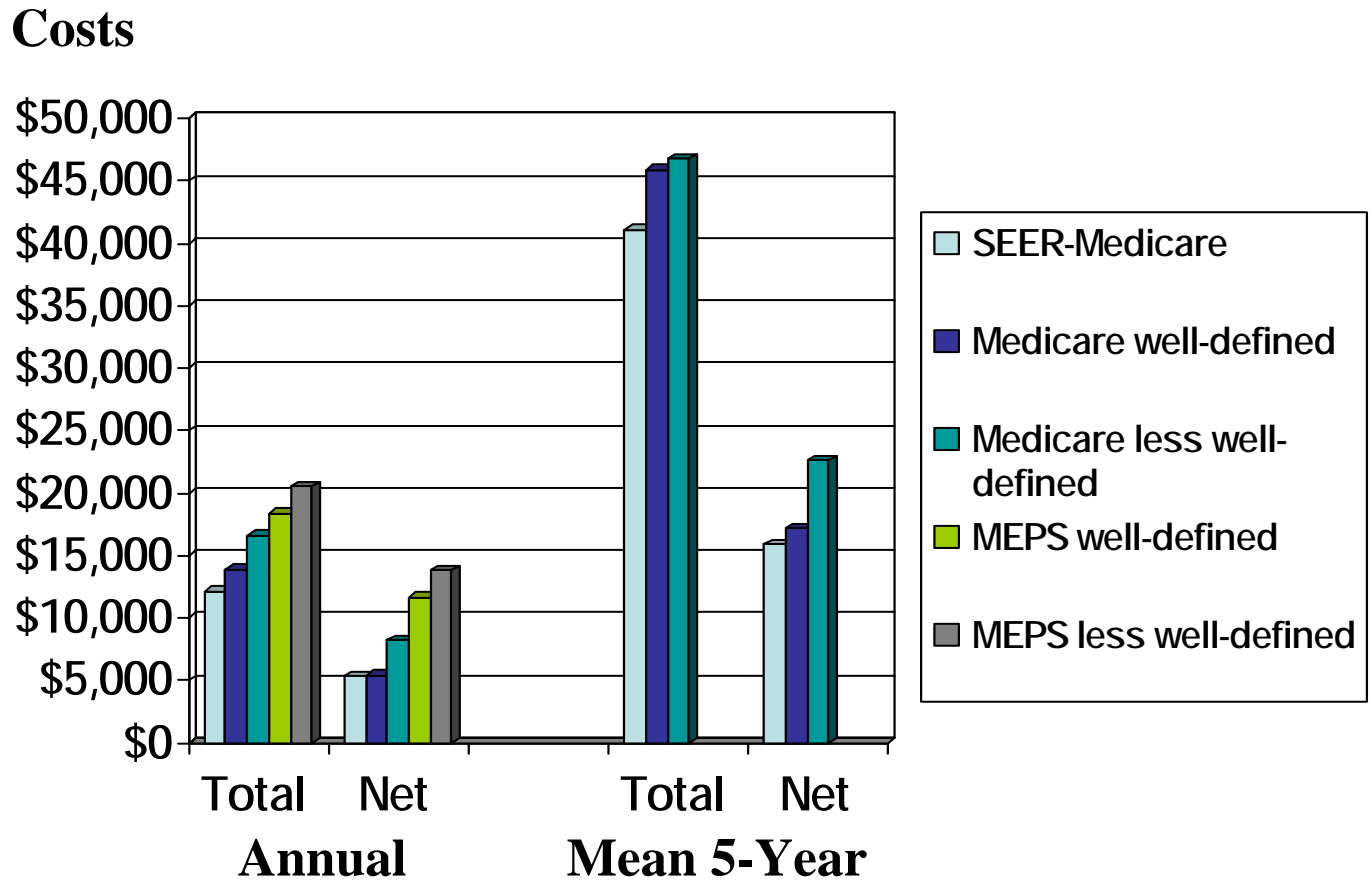
Source: Yabroff et al, Background paper

Summary

- Registry important for incidence AND prevalence
- Different methods, different cost estimates
- Different data sources, different cost estimates
- Attributable cost estimates more similar than total cost estimates
- Incidence
  - Modeled 5-year phase of care estimate similar to SEER-Medicare cohort
  - Misclassification - claims only overstates and understates
- Prevalence
  - Claims only methods overstate costs, effect larger with fewer years of data (fewer long-term survivors)
- Shape of the cost curve likely important (u-shape vs. –shape)
- What are implications for estimating costs of care in diseases without registries?
- How to balance generalizability of costs (e.g. MEPS) against accurate disease identification (e.g. SEER)?