Public Meeting on Estimating the Costs of Chronic Diseases
CDC-CDD-NPC Collaboration

November 18, 2004, Atlanta, GA
8am – 4pm
Atlanta Marriott Century Center, 2000 Century Boulevard NE, Atlanta, Georgia 30345 USA,
Phone: 404-325-0000, Fax: 404-325-4920

Organizers:
CDC: Centers for Disease Control
Barbara Bowman, Acting Associate Director for Science, National Center for Chronic Disease Prevention and Health Promotion
Terry Pechacek, Office on Smoking and Health
Paul D. Mowery, Sr. Statistician, Visiting Scientist, Office on Smoking and Health

CDD: Chronic Disease Directors
Christopher Maylahn, Director, New York

NPC: National Pharmaceutical Council
Lou Rossiter, NPC Consultant from the College of William & Mary
Jean Polatsek, Director of Program Development
Patricia Adams, Vice Pres., Business Operations & External Affairs

Participants: (see participant list)

Purpose:
The two purposes of the CDD-CDC-NPC Collaborative are 1) to better understand the ways state chronic disease and Medicaid programs can collaborate; and 2) to develop econometric models for states to estimate and discuss the fiscal impact of chronic diseases.

Outline of the Meeting:
Welcome and Introductions -- Lou Rossiter, NPC
Background and Motivation for Project -- Chris Maylahn, CDD
Methodological Issues in Estimating the Costs of Chronic Diseases -- Eric Finkelstein, RTI
1st Discussant -- Russ Toal, GA State University
2nd Discussant -- Wendy Max, University of California, SF
3rd Discussant -- Bryan Luce, MEDTAP International
4th Discussant -- Steve Cohen, Agency for Health Care Research and Quality
Breakout Group #1: A Research Agenda for Estimating Attributable Costs
Breakout Group #2: Small Area Estimates
Breakout Group #3: Partitioning Costs Across Public and Private Sectors
Summary and Conclusions -- Chris Maylahn, CDD

Welcome and Introductions Lou Rossiter, NPC

Background and Motivation for Project Chris Maylahn, CDD

- Aim to help policymakers sort out Medicaid costs
- Learn what states can do to control Medicaid costs
- Identify what populations to target
- Recommend methods for estimating disease costs in standard manner

Methodological Issues in Estimating the Costs of Chronic Diseases Eric Finkelstein, RTI

1. Background- In 2003 Finkelstein published a paper in Health Affairs titled, "National medical expenditures attributable to overweight and obesity: how much and who's paying?" In the paper, the national estimates for the cost for obesity totaled $93 billion, with half the costs being paid by Medicare/Medicaid. In response to state policy makers and the need to focus on state level estimates, Finkelstein published a second paper in 2004 looking at state-level estimates of annual medical expenditures attributable to obesity.
2. Two common approaches for estimating the medical costs of chronic diseases
    a. Incidence-based methods: estimates the expected lifetime costs for an individual diagnosed with the disease of interest in the base year
        i. Requires data on disease progression and survival probabilities to compute costs
        ii. More appropriate for cost-effectiveness analysis
    b. Prevalence-based methods: estimates disease cost for the prevalent population in a given time period, usually a year
        i. Includes costs for newly diagnosed cases and those in advanced stage of disease (more appropriate for disease burden and policy analysis)
        ii. Two common prevalence-based approaches to estimate cost are:
            1. Diagnosis-based approach
            2. Econometric approach
3. Diagnosis-based approach
    a. Option 1: sum the costs or charges for all events with the disease of interest listed as the primary diagnosis in health insurance claims data or other similar databases
        i. May either understate or overstate costs attributable to the disease of interest
            1. Understate: does not include attributable costs when disease of interest (e.g., diabetes) is listed as a secondary diagnosis
            2. Overstate: may include costs attributable to secondary diagnoses
    b. Option 2: sum the costs or charges for all events with the disease of interest listed as the primary or *secondary* diagnosis
        i. May overstate the costs attributable to chronic diseases
            1. The costs associated with each claim may be applied to multiple conditions and, therefore, double-counted

2. Chronic diseases like diabetes and hypertension frequently appear on the same claim

ii. Total chronic disease costs may exceed aggregate medical spending

4. Econometric approach
    a. Correlate current health care costs with the disease of interest while controlling, to the extent possible, for other observable characteristics that are likely to affect cost
    b. Does not require looking at diagnosis codes, except perhaps to identify the population of interest
    c. A simplified version of the regression equation used to estimate attributable cost based on person-level data is as follows: $\boldsymbol{Costsi = \beta 0 + \beta'(C.D.) + \lambda'(N.C.D) + Y'Xi + ei}$
    d. The dependent variable is equal to each individual's medical costs over a given time period, typically a year.
    e. ***B'***, the vector of key coefficients, represent the increase in costs associated with each chronic disease
    f. Can run model separately by type of cost (e.g., inpatient)
    g. We've used this approach to estimate annual costs of obesity (national and state-level), smoking, injuries, and diabetes
    h. Limitations
        i. Limited to information contained in the data
        ii. Undiagnosed or untreated persons will not be included in disease population
        iii. Relationship between chronic disease and cost is correlational not causal
        iv. Collinearity
5. Recommendation: The Econometric Approach
    a. The econometric approach has several major advantages over the diagnosis-based approach
        i. Regressions control for covariates (e.g., age, gender, comorbidities)
        ii. Allows flexibility in the modeling
        iii. Avoids double-counting of costs for related diseases (e.g., diabetes and hypertension)
6. Potential datasets for generating national and state-level cost estimates: a) Medicaid claims data; b) Medical Expenditure Panel Survey (MEPS); c) Medicare 5% Beneficiary Administrative Data; d) MarketScan (employed population and dependents); e) National Inpatient Sample (NIS) of the Healthcare Cost and Utilization Project (HCUP); f) National Hospital Discharge Survey (NHDS); g) National Nursing Home Survey (NNHS); h) National Home and Hospice Care Survey (NHHCS)
7. Medicaid MAX/SMRF Claims Data
    a. In order to facilitate analyses for the Medicaid population, CMS cleans the claims and eligibility data submitted by the states into Medicaid Analytic Extract (MAX) files that have a uniform format for all states
        i. Prior to 1999 these datasets were known as State Medicaid Research Files (SMRF) and data submission was voluntary
        ii. Includes person-level eligibility records with demographic and claims data
    b. We could construct variables necessary to run regressions, including:

i. Chronic disease flags based on diagnosis codes
ii. Demographic information (e.g., age, gender, race/ethnicity)
iii. Months of eligibility during the year
iv. An indicator for dual eligibility
v. Medicaid costs, in total and broken out by type of service

c. Advantages
i. Includes prescription claims
ii. Single source for state-specific Medicaid prevalence, demographic, and cost data
iii. Large number of observations
iv. Available for all states

d. Disadvantages
i. Estimating costs for each state would be costly and time consuming
ii. Unable to estimate costs for non-Medicaid population
iii. Misses some costs for dual eligibles
iv. Misses costs for non-covered services

8. MEPS-NHIS
a. The Medical Expenditures Panel Survey (MEPS) is a nationally-representative survey of the US civilian non-institutionalized population that quantifies each participant's total annual medical spending (including insurance spending).
b. Includes information on health insurance status and demographic characteristics.
c. The MEPS Conditions File allows for identifying any medical condition for which a participant sought treatment by 3-digit ICD-9-CM code
d. MEPS also links to the National Health Interview Survey, which includes self-reported chronic disease identifiers: "Has a doctor every told you that you have…" (including hypertension, CHD, stroke, diabetes, asthma, multiple cancers)
e. Advantages
i. Nationally-representative dataset
ii. Allows for stratification by payer (in theory)
iii. Data are free and publicly available
f. Disadvantages
i. Data do not allow for state-specific estimates
ii. Even after pooling multiple years, sample size may be inadequate for detailed stratifications (combined, 1998-2001 MEPS include approximately 80,000 adults, but only 5,000 Medicaid recipients)
iii. Data do not include institutionalized population
iv. MEPS estimates of annual medical expenditures are approximately half the corresponding estimates from National Health Accounts

9. Current CDC project
a. RTI is beginning a CDC-funded project to estimate the costs attributable to cardiovascular disease (e.g., CHD, stroke, hypertension, and congestive heart failure) in a few Medicaid populations
i. Primary analysis uses Medicaid MAX/SMRF data to estimate attributable costs for 3 states

ii. Secondary analysis uses MEPS data to generate national cost estimates (both total and Medicaid-specific)
iii. If MEPS estimates prove accurate, then all states will be able to quantify their Medicaid costs associated with cardiovascular diseases solely by multiplying disease prevalence estimates by the MEPS unit cost estimates-- MEPS analysis may provide a viable alternative (though less precise) for states that do not have the capacity to run a full claims-based analysis

10. MEPS-NHIS/BRFSS
    a. The Behavioral Risk Factor Surveillance System (BRFSS) is a state-representative telephone survey of the adult non-institutionalized population that tracks health risks in the United States (BRFSS does not include cost estimates)
    b. We (and others) have used regression results from the MEPS-NHIS national model to predict medical expenditures for each person in the BRFSS sample
        i. This is accomplished by multiplying each BRFSS participant's characteristics (independent variables) by the corresponding coefficients generated from the national regression analysis
    c. RTI has previously used this approach to quantify the state-level costs of obesity and smoking stratified by payer
    d. Because MEPS underestimates annual medical expenditures, state-specific attributable fractions are multiplied by previously published estimates of state (and payer) expenditures from National Health Accounts
    e. Advantages
        i. Allows state-level estimates to vary by state-specific characteristics
    f. Disadvantages
        i. Post 2000, it is no longer possible to stratify BRFSS data by payer (e.g., Medicaid, Medicare, and private insurance)
            1. Can still differentiate between insured and uninsured populations
        ii. BRFSS only includes prevalence data for a handful of chronic diseases, including:
            1. Diabetes
            2. Hypertension
            3. High Cholesterol
            4. Asthma
        iii. Estimates likely have large standard errors
11. Recommendations: use MEPS data perhaps supplemented with BRFSS or other data sources
12. Brief Discussion of Indirect Costs
    a. Indirect costs of disease are economic losses for which no payment is made, but for which an economic effect is still observed
        i. Morbidity losses
            1. Absenteeism (for both employees and caregivers)
            2. Presenteeism (decreased on-the-job productivity)
            3. Household productivity (increased bed days)
        ii. Premature mortality losses
            1. Only applies to incidence-based approach (lifetime costs)
    b. Approach for estimating productivity losses associated with excess morbidity

i. Use National Health Interview Survey (NHIS)
    1. In addition to demographic variables and disease identifiers, NHIS includes data on missed work days and bed days attributable to injury or illness
    2. Use regression analysis with either missed work days or excess bed days as the dependent variable and the disease of interest as the key explanatory variable -- Estimates productivity losses
    3. Multiply by previously published wage data to generate per capita cost of attributable productivity losses

1st Discussant -- Russ Toal, GA State University

1. Background- Former state policy maker who ran Georgia's Medicaid program; in addition to Medicaid, GA provided health care coverage to state employees and state-affiliated university staffs (these three populations comprised about 20% of Georgia's population)
2. The need to focus on the state perspective
    a. What are we trying to do and for whom? We need to focus on tools that will help state policy makers.
    b. What type of cost are we talking about- total cost, cumulative cost? The immediate concern for state policy makers is what they have direct responsibility for (i.e. Medicaid, etc.).
3. State health care programs (premiums, etc.) have become a political issue
4. From the state's perspective, what do states want?
    a. Data tool that has credibility that includes data from their state (ex. GA spent additional money on BRFSS to sample additional populations so that the data would be representative of their state)
    b. Data need to be timely
    c. Data have to be accessible
    d. How much of this cost could be avoided if the states took the appropriate intervention?
5. Most states do not have the capacity to carry out the data collection and address the research questions outlined in #4 above (no funding by legislators, no statisticians/epidemiologists in the department)- money needs to be put on the table (need funding from somewhere); need collaboration with CDC
6. In defining what tool/model to be used: states need to be familiar with and have trust in the organization that is doing the data collection/analysis (CMS needs to be familiar with this)
7. What other organizations or people should be involved? Possibly state employee or university system health plans- made-up of state personnel directors
8. Issues to overcome
    a. Measurement of cost of dual eligibles
    b. Specifying federal share of these costs
    c. Data that is problematic -- The data that goes to CMS is not corrected, and Georgia had a difficult time with the claims processing system
    d. Problems with getting Medicare claims
    e. Eligibility turn-over

   f. How one assesses mental disability or eligibility for mental disability
   g. Cost in state Medicaid programs from SSI -- no one has good data from the application process
   h. Institutionalized populations- no claim-specific data on this group
   i. Claims are not reliable for disease state assessment
   j. Issues related to state comparability – there is a huge variation in state Medicaid programs
   k. Capacity issue- CDC should contract with an independent research organization or some other entity to research questions outlined in #4 b/c states don't have the capacity
9. Problems with estimating costs
   a. MEPS- not state specific, doesn't include institutionalized population
   b. Medicare data is difficult to get (GA wanted Medicare data b/c it helped them look at dual eligibles, retired populations within the state health benefit plan, etc., they wanted to do quality studies)
10. Can costs of chronic disease be put in context for these states?
11. State health benefit plans/ university plans/ school plans are going to vary more than Medicaid; there is better data on these plans
12. CDC needs to play a leadership role and funding needs to be put on the table

2nd Discussant – Wendy Max, University of California, SF

1. Background: Published work on national and state-level estimates of attributable costs due to smoking and other kinds of illnesses. She has produced state and county level smoking attributable costs estimates in California and based her presentation on reactions to Finkelstein's methodology.
2. People want current estimates and want to produce something they can update on their own
3. Reactions to Eric's presentation: looks like a great approach, and bases her reactions and comments knowing the inherent weaknesses of the data
4. Questions
   a. Why the problem is framed this way: the cost of 5 chronic illnesses (cancer, diabetes, stroke, HD, CLRD)?
   b. What is the goal here?
   c. Will separate estimates be produced for each of the 5 conditions? By payer?
5. Definition of Chronic Illness
   a. How is it defined?
   b. Focus on the 5 most prevalent chronic diseases?
   c. Example: All cancer is not chronic and many other diseases are
   d. What proportion of all chronic disease incidence is accounted for by these five chronic diseases?
   e. Nursing home care (institutionalized) is important component for chronic diseases
6. Adding in a possible secondary diagnosis Option 3
   a. Compare healthcare costs for people with the secondary disease of interest (i.e. cervical cancer) to those with no secondary disease by major diagnostic category
   b. Then multiply difference by number of people with the secondary diagnosis
   c. Cost of Secondary Diagnoses: Example

i. Hospitalization cost for those (n=43) with a primary respiratory diagnosis and a secondary diagnosis of cervical cancer = $9415
ii. Hospital cost for those with primary respiratory diagnosis and no secondary diagnosis = $3244
iii. Thus, cost of secondary cervical cancer = $6172 * 43 = $265,396
iv. Do for each major diagnostic category and add them up
v. Then need to subtract the secondary costs of other diseases where the primary disease is cervical cancer
vi. We tried to do this, and the N's were much too small

7. How are secondary diagnoses to be treated here?
    a. Model doesn't really allow for secondary diagnoses, does that mean we will include only primary diagnoses?
    b. MEPS data only include those who are currently treated for something
    c. Will the NHIS question (has a physician ever told you you had…) be used?
    d. Some chronic conditions, by their nature, may not be captured
8. Questions about the Model
    a. Will any interaction terms be included?
    b. Some issues similar to secondary diagnoses (diagnosis of 2 chronic conditions)
    c. What about age-specific costs? Example: being over 80 may have different implications for cancer treatment than for diabetes treatment
    d. Model assumes that treatment patterns don't differ by state (using national coefficients), but we know they do
9. Importance of making sure states don't sum to more than the total
    a. In our work, we always develop the estimates and then adjust by a factor to reflect the total
    b. Example: add up the state totals, compare to the national total, and adjust each state by the ratio of national/state totals
10. Comment on Premature Mortality Losses
    a. "only applies to incidence-based approach"
    b. NO – if assume steady state
    c. That is, prevalence approach means that losses are summed for this year among those who died this year + this year's productivity for someone who died last year + this year's productivity for someone who died 2 years ago +…
    d. Adding up losses over the lifetime of those who died this year could be viewed as a proxy for that approach
11. Inclusion of Mortality Losses
    a. My recommendation would be to include:
        i. Number of lives lost
        ii. Number of years of life lost to life expectancy
        iii. Value of lost productivity associated with those lost years of life

3rd Discussant – Bryan Luce, MEDTAP International

1. Introduction: "…an imprecise estimate of the right concept is superior to a precise estimate of a wrong concept" (Mishan, 1972)
2. Concur with Econometric approach
    a. Define costs broadly to include M/M

   b. "Indirect costs need to include "presenteeism"
   c. "Precision" vs. "accuracy" issue
3. Underlying purpose of effort
   a. To lesson burden
   b. To demonstrate value of interventions
   c. "How much of these costs could be avoided"? (Russ Toal)
4. Challenge: To develop a model which would provide opportunity to meet that challenge
   a. Incorporate a Markov-type disease progression process (incidence-prevalence blend)
      i. Allow modeling of interventions
      ii. Demonstrate value of interventions
         1. Reducing M/M and costs
   b. Jo Mauskopf and Mark Nuijten (PE, ViH?) have published
5. Don't fall into the "an ounce of prevention equals a pound of cure" trap that has been translated to be "prevention saves money".
6. Rather than focusing on saving money, focus on getting "total value" for money
   a. Intervention costs vs. medical cost offsets
   b. Improvement in M/M
   c. Improvement in productivity
7. Focus on ultimate goal of decision-makers (improvements in chronic disease)
8. Opt for "accuracy" at expense of "precision"
9. Opt for "value" rather than costs or cost savings
10. Develop dynamic incidence/prevalence blend model rather than static model
11. Indirect costs might need to be added to model; they could be small estimates, but it will challenge future researchers to improve the estimation of indirect costs
12. Need to argue that intervening in the process will lessen that burden- effort needs to extend past just estimating costs
13. Challenge: develop a model which would provide opportunity to meet that challenge
14. How model needs to be improved: add transition states, add concept of indirect costs; make model broader

4th Discussant -- Steve Cohen, Agency for Health Care Research and Quality

1. Look at the concentration of expenditures- skewed distribution of what population cause most of the diseases- small % population causing large % of cost
2. Next Step after Econometric Model: look at details of the distribution- individuals with top 10% of expenditures and what diseases they have; identify individuals with lower costs, and understand other factors that might be reducing their costs
3. Look at the persistence of expenditures (use data's span of 2-years)- look at where things are stable, where things are going up
4. Quality of care measures, how patient perceives encounter with quality of care- by incorporating this component, we might find out what might lead to more stable expenditures over the years
5. Estimates needed within the year of survey; projection data- we would want to be able to make projections
6. Limitations of various data sets- cross collaboration and talk among which data sets would be of better use (weighing in on limitations and advantages)

7. Need for state level estimates is critical- looking at the existing surveys and their ability to go deeper into state level estimates (some have the ability to make state level estimates); Steve will be using MEPS to be producing estimates for the 10 largest states
8. State-level estimations: current model good- using existing national data, putting in additional state/regional constructs; model/tool to be used for state level estimates; using MEPS and BRFSS data
9. Mean-square error estimations- trying to focus on accuracy as opposed to precision
10. Help develop analytical tools to use various data sets and the information we already have

General Discussion -- Terry Pechacek, Office on Smoking and Health, Moderator

Risk Factors

- Current estimate of risk factor- does it reflect the history of use?
- Differences over time on impact of risk factors
- Platform for cost of risk factors to identify interventions (maybe we shouldn't focus on risk factors since they are more difficult to identify?)
- Behavioral risk factors or other risk factors (hypertension, etc.)? we are mainly looking into behavioral risk factors to reduce costs
- Minimum Data Set (MDS) has useful risk factor data

Improving Data

- Add risk factor collection information to Medicaid information? A first step would be adding the question to the surveys (e.g. adding a Medicaid question (are you covered?) to health risk factor surveys like NHIS, adding risk factor questions to surveys of the Medicaid population). This is getting easier as more states are heading toward electronic data collection, but this might be hard for some states. New database needed? This would be ideal, but might not be attainable.
- Do we have data sets that include behavior from 10 years ago and what their costs are today? The data are not really there- and if they are, there are many limitations
- Troublesome reliance on BRFSS to estimate state-level costs of chronic disease; originally not designed for that; we may want to look into using state employee programs for research (surveying state officials responsible for health care coverage programs who might be interested in the data and learning what they want)
- Do we construct new surveys or do we improve on the ones available? we need a common mission and vision to guide us

Research Goals

- What is the Goal? Research or policy purpose? How do we pitch this as improving the research agenda and how do we get funding, and how do we collect information that will be useful for state policy makers?

Audience needs

- Employers and policy makers need basic, local burden information to support estimates of comparative effectiveness of interventions
- Build models for CDD to use with Med. Dir.

Funding

- Look toward other business, health, and other organizations that might be able to collect data on population at risk (retirees)- to bring in funding and cross-collaboration

Methodology

- Estimation of impact on care-giver? What is the external cost on families as opposed to public?
- Consideration of indirect costs
- Interventions are different to determine impact, especially with multiple risk factors that are not additive
- Possible steps to go forward:
    - Cost Assessment- predict costs and then compare them to what they really are
    - Identify % preventable
    - Calculate cost of prevention
    - What is the net benefit?

Breakout Group #1: A Research Agenda for Estimating Attributable Costs --
Amanda Honeycutt, RTI, Facilitator

Major issues emerged that need to be considered in setting a research agenda for estimating attributable costs which include, but are not limited to: (1) Who is our audience? (2) Who are our partners and collaborators? (3) What is our ultimate goal? (4) What data sources exist? (5) What are the chronic diseases of interest? (6) What type of costs should be included? (7) What is the appropriate methodology?

Who is our Audience
Who are we estimating these costs for? Is our audience state policy makers, CDC, or both? What type of tool is needed for the audience? The intended audience should be an active participant and offer suggestions to guide the research agenda. If state-policy makers are our intended audience, we need feedback from more state-policy makers to determine what type of estimates and tools they want. If the estimates are ultimately for CDC to disseminate, we need to determine what type of tool CDC wants the state-policy makers to have.

Who are our Partners and Collaborators
We need to identify potential collaborators (e.g. universities, managed healthcare organizations, business) who may have similar interests and relevant research that can partner with us and help us obtain research funding.

What is our Ultimate Goal
What is our long-term vision? Do we want to stop at burden of illness estimates? Identifying an ultimate goal will help guide the research agenda.

What Data Sources Exist
More discussion about existing data sets is needed. To examine whether BRFSS estimates are reliable and representative for states, conducting a small study and comparing state-level prevalence estimates to those from BRFSS would be a start. Meeting participants warned of reliance on BRFSS data in making state-level estimates, and voiced concerns about the failure of BRFSS to represent key constituents because it is a telephone survey, and may possibly disappear in the future.

What are the Chronic Diseases of Interest
What diseases are we going to focus on? What factors affect our choice of chronic diseases to be included (i.e. mortality, political)?

What Type of Costs Should be Included
Are we talking about total or cumulative costs? Do we include both direct and indirect costs? State-level estimates should represent the full costs and recognize the distribution of those costs among taxpayers. The group felt that indirect costs need to be included in all models even if they are not of primary interest to state policymakers. CDC should present the full burden of these chronic diseases and reflect the importance of both direct and indirect costs.

What is the Appropriate Methodology

In making state-level estimates, do we opt for precision or accuracy? Opt for accuracy at the expense of precision. Models need to account for double-counting and indirect costs, which could mean creating a multi-stage model to control for the collinearity between variables (i.e. risk factor model).

Breakout Group #2: Small Area Estimates -- Eric Finkelstein, RTI, Facilitator

In defining an approach to making state-level estimates, what will we ultimately provide to states?

What Will We Ultimately Provide to States

Do we want to provide states with a methodology or a tool-kit? States want something simple like a tool-kit where they can obtain their own state-level estimates (e.g. Visual Basic like Finkelstein's Obesity Cost Calculator). The first approach is that we develop a tool-kit for the states. We might take estimates from MEPS, develop an equation to model total costs, and allow them to plug in their state-specific numbers (e.g. state, population, Medicaid population, etc.) to obtain their own estimates. Precision of the estimate would reflect the accuracy of their state-specific numbers. A second approach is that we do the estimations for them, where we make clear that we are only giving them estimates and do some small area corrections. We would want to conduct sensitivity analyses to determine if we were close. In making state-level estimates, we want to achieve 10% Relative Standard Error (RSE) (ex. +/- 20%). Ultimately, states want a tool that is easy to use and that is accurate in predicting state-level estimates.

Breakout Group #3: Partitioning Costs Across Public and Private Sectors --
Russ Toal, GSU, Facilitator
Terry Pechacek, Office on Smoking and Health, Presenter

The discussion focused on two main themes: (1) how the CDC-CDD-NPC Collaborative can work effectively with states to implement the group's findings; and (2) methodologies and data sources for estimating the economic burden of chronic diseases at the state level.

How to Work with States

Discussion of the first topic included the challenges of how to assess what states want and what data states can provide, as well as the issue of packaging the tools (models, data sources and data collection protocols). A suggestion was made that we start with a few states and do a pilot study. Suggestions for approaching states included presentations at annual meetings of associations such as NGA, NASBO, NCSL, and the State Medicaid Director's Association. It was stated that these groups could bring increased credibility to the effort. The group felt that a few states would volunteer to participate in the pilot. Some specific questions to address during the pilot included:

- What cost components are being measured or assessed at present?
- What are states looking for in terms of cost estimates and cost breakouts?
- What questions are policy makers asking that can be addressed by better methodologies and better data sources?
- What data are currently being used?

Data Sources and Methodologies

Suggestions for data sources included the use of administrative data, such as claims data, not collected for research. The task of cleaning and making these data available for research is formidable but probably do-able in some states. The group also recommended trying to obtain data from health insurance companies, company health insurance plans (about 2/3 of health care payments by large company plans is paid to retirees), as well as health insurance plans sponsored by DOD and the VA. Constraints of this approach included the possible lack of demographic identifiers, and limits on use of patient data due to the Health Insurance Portability and Accountability Act (HIPAA). Identifiers deemed essential included age, whether the patient was Medicaid eligible, and whether the respondent was institutionalized. The group also recommended looking at AARP and the American Hospital Association as possible partners who would be willing to share some data. There was an extended discussion of the pros and cons of claims data.

In terms of methodologies, the group recommended a dual track research agenda. Work should continue on analyzing MEPS and other national data sources to estimate econometric models. These models are then applied to a state's specific demographic distribution. The 2nd track is to develop state-specific estimates based on data collected by a state, such as Medicaid and state employee plan administrative data. The 2nd approach would have more credibility with state policy makers because it uses only data from the policy maker's state.

Summary and Conclusions Chris Maylahn, CDD

There is value in producing national and state-level estimates of chronic disease cost. State-policy makers are interested in tools or a set of tools that allow them to measure the public cost of chronic diseases for their specific state. We (members of this working group) want to address questions like:

- What do Medicaid officials want to know and what are they using now?

- The percentage of costs due to chronic disease and its burden, and percent that could have been averted through interventions.

- What fraction of costs are from high cost users of state health care systems?

The estimates need to be reliable and accurate and delivered in a timely manner. We also need confidence or error limits on the estimates.

There are many gaps in current data used to estimate chronic disease cost, and there should be a move toward filling those gaps and improving data collection.

In moving forward, CDC will play a leadership role, and we will seek the help of collaborators (i.e. business organizations, health organizations, universities), and seek input from the audience of Medicaid officials and state-policy makers.

We need to be sensitive to the needs of our audience. We need to understand not only the Medicaid perspective, but also the perspectives of many other groups (e.g., CDC, researchers, etc.) in creating a multi-disciplinary perspective in approaching this task.

Next Steps

CDC-CDD-NPR will move forward in supporting an analysis of MEPS and MAX data sets to produce national estimates for selected states. Next steps include:

- Forming a steering committee to guide the research agenda
- Planning a pilot study using MEPS and MAX data sets in these selected states and at the national level
- Forming three taskforces:
    1. Taskforce #1: To steer the pilot study
    2. Taskforce #2: To select the diseases
    3. Taskforce #3: To facilitate dialogue with partners
- Publishing a paper describing the proposed research agenda to be agreed upon by consensus of this working group