April 2011

An Assessment of the State of the Art for Measuring Burden of Illness

Final Report

Prepared for

Ansalan Stewart
U.S. Department of Health and Human Services
Washington, DC 20201

Prepared by

Amanda A. Honeycutt
Thomas Hoerger
Alex Hardee
Linda Brown
Kevin Smith
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0212050.005.001

RTI Project Number
0212050.005.001

An Assessment of the State of the Art for Measuring Burden of Illness

Final Report

April 2011

Prepared for

Ansalan Stewart
U.S. Department of Health and Human Services
Washington, DC 20201

Prepared by

Amanda A. Honeycutt
Thomas Hoerger
Alex Hardee
Linda Brown
Kevin Smith
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

Contents

Figures

Number **Page**

Table

Number **Page**

1. BACKGROUND

Burden of illness measures provide information about the impacts of diseases and risk factors on individuals, governments, and society as a whole. Burden measures quantify diverse effects of diseases and risk factors using a variety of units, measurement techniques, and levels of aggregation. Two examples of burden measures are healthy life years lost, which summarize the disability and mortality impacts of disease in a single measure; and cost of illness, which measures disease burden in terms of medical costs and productivity losses. Burden of illness estimates may be useful for establishing a population disease burden baseline against which future progress toward achieving disease prevention and health promotion goals may be measured. Additionally, as health care systems respond to increasing demand for and rising costs of medical care, burden measures offer the potential to assess the efficiency of resource allocations to prevent or treat specific diseases and improve health.

The purpose of this project was to collect information on current uses of burden of illness measures, trends in burden of illness measurement, and the methodological and data challenges that affect burden of illness measurement and reporting. We conducted three main activities to collect and compile information about burden of illness measurement: a literature review, an environmental scan, and a roundtable meeting with experts. The literature review summarizes key burden of illness measures and measurement approaches and challenges for implementing each measure. The literature review is provided in its entirety in Appendix A. The environmental scan describes current efforts, innovative initiatives, and gaps in measures of the disease burden in the United States. To conduct the environmental scan, we held telephone interviews with 13 burden of illness experts from across the United States with expertise in various types of burden of illness measures. We also searched federal health agency and private foundation Web sites to identify new initiatives and new or recent grant awards focused on burden of illness measures or measurement. The resulting environmental scan is provided in Appendix B of this report.

Our final activity was to conduct a daylong roundtable meeting of experts on November 10, 2010, in Washington, DC. Fifteen government and nongovernment policy and burden of illness experts participated in the meeting, during which we discussed key challenges for burden of illness measurement (presentations from that meeting are available upon request). At the meeting's conclusion, all participants contributed to a discussion of burden of illness areas of consensus and areas needing further consideration. A summary of the roundtable meeting presentations and discussions is provided in Appendix C.

This report compiles and summarizes key findings from all three project activities: the literature review, the environmental scan, and the roundtable meeting. The purpose of the

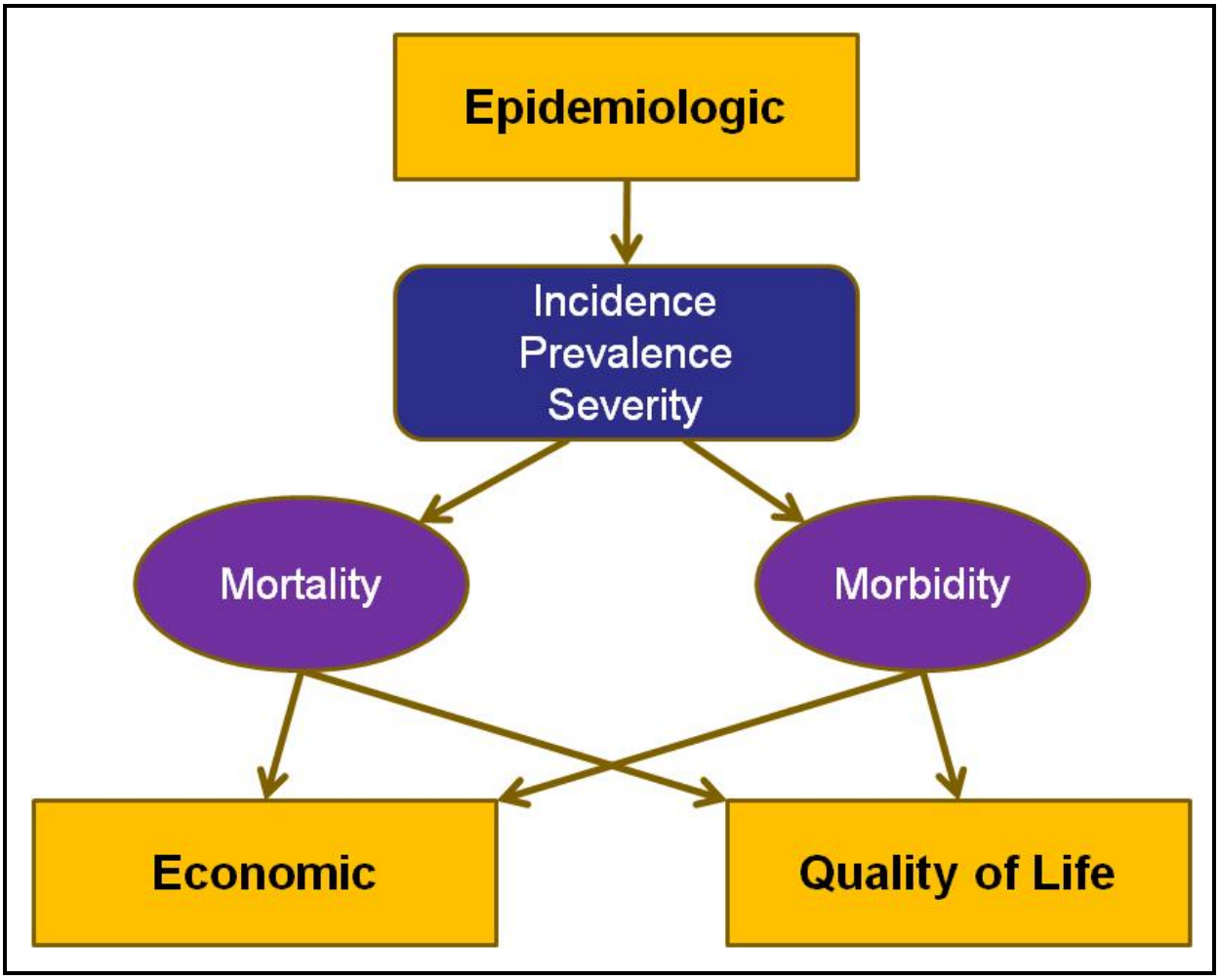
report is to provide policy makers with an understanding of the current landscape regarding metrics, methods, and data for quantifying the burden of illness. The report is therefore expected to serve as a primer on burden of illness for Office of the Assistant Secretary for Planning and Evaluation (ASPE) staff and other federal policy makers so that burden estimates can be better utilized in policy development. In Section 2, we provide a brief overview of commonly used burden of illness measures. Section 3 describes key methodological trends and needs in burden of illness measurement, and Section 4 describes the main data trends and needs in burden of illness measurement. Section 5 discusses issues surrounding the use of burden of illness data for making policy decisions, and Section 6 summarizes lessons learned from this project about burden of illness and discusses issues for further consideration.

2. OVERVIEW OF BURDEN OF ILLNESS MEASURES

Burden of illness measures provide information about the impact of disease or specific diseases on society, government, and the individuals affected by disease. Some measures capture the number of people affected by a given disease or risk factor, whereas others capture the impact of disease on longevity, costs, and quality of life. Burden of illness measures are especially useful for comparing differences in burden. For example, they may be used to assess differences in the burden of a specific disease or of all illnesses and injuries across groups of people, such as comparing disease burden between racial groups and between groups of people of different income levels. Burden measures are also useful for considering differences across diseases by addressing such questions as "Does heart disease cause more deaths than cancer?" or "Which costs more—heart disease or cancer"? Finally, burden measures may be used to consider differences across time in the same or similar population groups, by assessing whether the U.S. population lives longer now than at some time in the past and analyzing whether increased spending on health care is associated with improvements in health outcomes.

In reviewing burden of illness measures for this project, we found it useful to group measures into three broad categories: epidemiologic, economic, and quality of life. As shown in Figure 2-1, epidemiologic measures, such as incidence, prevalence, mortality, and life expectancy, provide the building blocks for economic cost and health-related quality of life (HRQoL) measures. The latter two measures generally place a value on epidemiologic measures of burden. Examples are mortality costs, which place a monetary value on the number of deaths attributable to a disease; and the disutility of being in a disease state, which assigns a preference value to disease incidence or prevalence. Although no single measure of burden of illness has emerged as the preferred measure for informing health policy decisions, metrics from different categories are frequently combined for use in analytical studies. For example, summary measures of population health, such as the quality-adjusted life year (QALY) and disability-adjusted life year (DALY), combine information on mortality and nonfatal health outcomes to present population health in a single numerical index (Parrish, 2010). QALYs are further combined with economic costs to perform cost-effectiveness analyses of health interventions. A cost-effectiveness ratio describes the additional price to obtain a health improvement when compared with a baseline intervention. Different measures thus capture different aspects of the burden of illness and should be selected for use according to the research or policy question of interest. In the next three subsections, we provide an overview of each of the three main types of burden of illness measures: epidemiologic burden, economic burden, and quality of life.

Figure 2-1. Three Categories of Burden Measures



2.1 Epidemiologic Burden

Epidemiology is the study of the occurrence and distribution of diseases, causes of death, and behaviors and their determinants in populations, as well as the application of the knowledge obtained to control health problems (Last, 2001). Two main types of epidemiological studies—descriptive and analytic—are used to collect, analyze, and interpret information on the distribution and determinants of disease, respectively (Dicker et al., 2006).

Important descriptive epidemiologic measures of burden of illness are incidence, prevalence, mortality, and life expectancy. For each of these, age adjustment is often useful for making comparisons across groups. Incidence is the number or rate of new cases of disease arising in a given period of time in a population. Prevalence is the proportion of individuals with a disease in a population at a specific point in time. Together, these measures form the basis for measuring disease occurrence and enable epidemiologists to estimate the overall

magnitude of a health problem or the short-term trends within a population. Mortality, most commonly expressed as a rate, is the total number of deaths in a population or deaths due to a specific disease, scaled to the size of the population and per unit of time. Life expectancy is defined as the average number of years an individual at a given age is expected to live if current mortality rates continue. It is most commonly used as an indicator of overall population health. For example, women in Japan have the highest life expectancy at birth of any subpopulation in the world—85.9 years—whereas women in Zimbabwe have the lowest life expectancy at birth—42.7 years (World Health Organization [WHO], 2010).

Recent studies have compared life expectancy across countries (Preston and Ho, 2009) and across geographically and ethnically defined populations within the United States (Murray et al., 2006) and have linked risk factors, such as obesity, to life expectancy over time within the United States (Olshansky et al., 2005). Murray et al. (2006) described health outcomes for eight different geographic and ethnicity groupings (i.e., "eight Americas") and found a life expectancy gap at birth between the highest and lowest life expectancy groups (Asian females and high-risk urban black males) of 20.7 years in 2001. Olshansky et al. (2005) estimated that growth in obesity in the United States could lead to reductions in life expectancy and that weight loss could result in life expectancy increases at birth of 0.21 to 1.08 years across sex and race groups.

Analytical epidemiology searches for causes and effects; examples are attempts to quantify the association between health exposures and outcomes or to test a hypothesized causal relationship developed through descriptive studies. Although epidemiology by itself cannot prove that a particular exposure was the exact cause of a certain health outcome, it often provides sufficient evidence for public health officials to take the appropriate control and prevention measures. Analytical epidemiologic measures of association include excess risk, relative risk, and the odds ratio. Measures of potential impact, such as population attributable risk, assume the observed association is causal. See text box for brief definitions of these measures. Additional detail is provided in the literature review (see Appendix A).

A challenge when estimating the epidemiological burden of illness is the difficulty of attributing health outcomes to any single disease because of the high degree of comorbidities among people with chronic illnesses. Some estimates double count utilization, death, costs, or other outcomes (e.g., attributing an outcome to both diabetes and kidney disease when an individual had both). In cases where individuals have multiple illnesses or risk factors, it is very difficult to predict what the health outcomes would have been in the absence of the disease or risk factor because of competing risks. For these analyses, researchers should clearly state their assumptions and indicate the degree to which uncertainty in the parameters affects estimates by providing confidence intervals.

Of the three burden of illness categories we have specified, epidemiologic measures are the most specific. They also provide the foundation for developing economic and health status measures of disease burden (Spasoff, 1999). For example, estimates of disease prevalence are used to estimate the aggregate cost and quality of life impacts of disease. However, epidemiologic measures cannot by themselves show an overall change in health status. If some indicators, such as the crude death rate, show better health, whereas others, such as chronic disease prevalence, show worse health, it cannot be said whether the population as a whole is better off or not. To make this assessment, an assessment of preferences for the different health states is needed. Consequently, many descriptive epidemiologic measures serve as health indicators, but they also form the basis of population health measures with a higher level of aggregation, such as the Quality of Well-Being (QWB) index or the EuroQol-5D index (EQ-5D).

Selected Analytical Epidemiology Measures Defined

A. Measures of Association

Absolute Measure

Excess risk: The rate of disease in an exposed population minus the rate of disease in the unexposed population.

Relative Measures

Relative risk: The rate of disease among the exposed divided by the rate of disease among the unexposed.

Odds ratio: Used when the risk of disease cannot be directly calculated because the population at risk is not known (e.g., case-control study design). The disease odds ratio is the odds of having the disease among the exposed divided by the odds of having the disease among the unexposed. This is equivalent to the exposure odds ratio, the odds of exposure among the disease cases divided by the odds of exposure among the controls.

B. Measure of Potential Impact

Population attributable risk: The proportion of disease among the total population that would be eliminated if the exposure were eliminated (assumes that the association is causal).

(Please see Appendix A for additional summaries of analytical epidemiology measures.)

2.2 Economic Burden

Economic burden of illness is often described in terms of health care spending, both at a single point and across time. In the United States, per capita spending on health care rose from an inflation-adjusted $1,080 in 1960 to almost $7,700 in 2008 (Centers for Medicare & Medicaid Services [CMS], 2011). The government's share of spending on health care has increased from 24.5% in 1960 to 47.3% in 2008, and health care spending, which represented only 5.2% of the U.S. gross domestic product (GDP) in 1960, accounted for 16.2% of GDP in 2008.

Although health care spending represents a large and growing share of the full economic burden of illness, other important components are nonmedical spending, lost work days resulting from increased morbidity or early mortality, and the impact on family members' employment or patients' psychological well-being. Estimates of the economic burden of illness attempt to capture the full economic costs, or "opportunity cost" of a disease, where opportunity costs are the value of health and non-health outcomes foregone as a result of the disease. For example, part of the opportunity cost of multiple sclerosis is the value of lost productivity for the patient and family members who switch jobs or quit working as a

result of the patient's illness. Opportunity costs differ from accounting, or financial, costs in that opportunity costs value even those costs or losses for which no financial cost is incurred, such as productivity losses or unpaid caregiving provided by family members.

The economic burden of illness is typically valued by applying cost of illness methods. Hodgson and Meiners (1982) discussed recommendations of a 1978 Public Health Services task force on how to perform cost of illness studies to ensure that future studies use consistent approaches, thereby improving the comparability of studies. They noted that the approach most commonly used to form values for illness, disease, and health care services was to identify the cost-generating components and to attribute a monetary value to them. They recommended including both direct and indirect costs in cost of illness analyses. Direct costs consist of medical and nonmedical spending to diagnose, treat, manage, and live with an illness (e.g., doctor visits, transportation costs, family spending for household help). Indirect costs capture the productivity losses that arise when people are unable to work and psychosocial costs, such as the costs of financial strain or uncertainty over a person's future health and well-being.

After describing all possible costs of illness, Hodgson and Meiners (1982) acknowledged the limitations of attempting to include all costs associated with an illness and provide recommendations for cost of illness studies. Key among these recommendations are that cost of illness researchers (1) specify the costs to be included in a study; (2) clearly describe the methods and data; (3) use a range of discount rates from 2.6% to 10% to discount foregone future benefits (e.g., mortality-related productivity losses); (4) avoid double-counting of costs by excluding transfer payments, for example; (5) include nonmedical and psychosocial costs whenever relevant; and (6) conduct sensitivity analyses to examine the impact on disease cost estimates of uncertainty in key parameter values.

Hundreds of cost of illness analyses have since been conducted to characterize the full or partial economic burden of specific diseases. In response to a request from the Senate Committee on Appropriations, the National Institutes of Health (NIH) compiled a comprehensive set of direct and indirect cost estimates for many of the diseases for which NIH conducts and supports research (Kirschstein, 2000). The first set of these estimates was provided in September 1995; Committee members were especially interested in costs for the top 15 causes of mortality as identified by the Centers for Disease Control and Prevention (CDC). In response to the initial and subsequent requests, NIH has developed tables and reports that show disease-specific estimates of the direct and indirect costs of illness alongside the level of NIH support for each disease. When updated by RTI in 2006, the list included direct and indirect cost estimates for almost 75 diseases or risk factors (e.g., smoking, obesity). This updated table is included as an attachment to Appendix A.

In the past few years, the need to provide guidance on a consistent set of methods for economic burden of illness estimation has once again emerged. Several studies have shown that cost of illness estimates for a given disease vary widely, even when the same data are used (Akobundu et al., 2006; Honeycutt et al., 2009). Moreover, estimated medical costs from cost of illness studies often attribute the same costs to multiple diseases, with the implication that estimated costs may exceed aggregate health care spending when summed across diseases (Kirschstein, 2000; Trogdon et al., 2008).

To establish guidelines for estimating health care costs of disease and to identify specific areas for future research, disease costing researchers met in December 2007 at a conference co-sponsored by the National Cancer Institute and the U.S. Agency for Healthcare Research and Quality. The recommendations from this workshop for standardizing disease cost estimation and future research needs are provided in the July 2009 supplement to *Medical Care*, "Health Care Costing: Data, Methods, Future Directions" (Yabroff et al., 2009). Approaches to valuing the economic burden of disease often use three broad economic burden of illness categories: direct medical spending, direct nonmedical spending, and indirect costs.

Direct medical costs of an illness capture the costs for all medical services, including hospital inpatient, physician inpatient, physician outpatient, emergency department outpatient, nursing home care, hospice care, rehabilitation care, specialists' and other health professionals' care, diagnostic tests, prescription drugs and drug sundries, and medical supplies (Segel, 2006). Direct nonmedical costs include transportation and relocation expenses and the costs of modifications required to the home, automobile, or diet. Nonmedical costs also capture the value of informal caregiving.

Indirect costs capture the productivity (labor and household) losses resulting from excess morbidity and from early mortality. Broader measures of indirect costs also capture the psychosocial costs of illness.

The human capital approach is generally used to value productivity losses by multiplying time lost from labor by average earnings. A key advantage of the human capital approach for valuing morbidity-related productivity losses is its ease of implementation, as data are often available on workdays lost and wages. Moreover, the resulting monetary values are easily interpretable as economic losses to employers and employees (Hodgson and Meiners, 1982). Unfortunately, wage data are not available for unemployed workers, retired workers, children, or homemakers. Although the human capital approach assigns value to household productivity for adults, the approach does not provide a comprehensive valuation of life because labor and household productivity losses do not capture personal discomfort, pain, and suffering, which may be significant. The human capital approach is therefore often used to provide a lower-bound estimate of the full economic costs of illness (Tarricone, 2006).

The willingness-to-pay (WTP) approach is an alternative method that values the prevention of illness and death from disease as the sum of what people are willing to pay to reduce their own risks plus the sum of the additional amounts that people are willing to pay to prevent illness and death in others (Freeman, 1993). The main method for estimating WTP uses wage differentials between high- and low-risk occupations (Freeman, 1993). Another method for estimating WTP is to ask people directly about how much they value specific reductions in illness and death risks (Freeman, 1993). A key advantage of the WTP approach is that it can capture in a single measure all of the benefits of disease prevention, including the value of productivity losses, pain and suffering, and even out-of-pocket medical spending. In addition, if the disease impacts are limited to short-term impacts that do not include death, valuations of those impacts can also be performed using WTP (Johnson et al., 1997).

Although WTP measures capture both direct and indirect costs of illness in a single measure, approaches to estimate WTP tend to be more costly and time-consuming than direct estimation of health care costs and productivity losses. Most measures of WTP require primary data collection (i.e., asking people how much they value not being exposed to a disease risk, such as asbestos) and extensive pretesting of survey instruments to capture the expected impact of disease characteristics on utility. Nonetheless, they are often preferred by economists because they are consistent with the theory that individuals maximize utility (Bayoumi, 2004).

The value of a statistical life (VSL) approach is a specific example of WTP that is sometimes used to estimate the mortality impacts of a disease. Rather than representing the specific value of an individual life, the VSL method measures the aggregate of what individuals are willing to pay for small reductions in their risk of dying (EPA, 2000). A strength of the VSL method is that it considers both earnings and the intrinsic value of being alive, whereas the human capital approach considers only earnings losses (Miller et al., 2004).

2.3 Quality of Life

HRQoL refers to a person or group's perceived physical and mental health at a point in time. It is important for measuring the impact of both chronic diseases and medical interventions on patient well-being, because it is a direct measure of quality of life, rather than a proxy, such as range of joint motion or level of pain. HRQoL is also needed to generate other measures of burden, such as QALYs, that summarize morbidity and mortality impacts of disease in a single measure. Many steps are involved in developing a measure of HRQoL. First, a standardized questionnaire is used to collect information about the impacts of a disease or intervention on a number of physical and mental health functioning domains (e.g., mobility, cognitive functioning, pain). The most commonly used questionnaires are generic instruments, but some instruments focus on a single disease state or organ system

to capture more specific aspects of functioning impaired by that disease. An example of an organ-specific scale is the National Eye Institute Visual Functioning Questionnaire-25, which asks 25 questions to evaluate different aspects of eyesight. A problem with specific scales is that health domains not thought to be related to the specific system are not included in the questionnaire, so respondents with much different overall health can appear very similar. A generic health status profile seeks to go a step further by assessing multiple health domains. The most commonly used health status profile is the SF-36, which uses 36 questions to rate eight scales of a respondent's health, including physical function, bodily pain, social functioning, and vitality. The EQ-5D creates index scores for five areas of functioning: ambulation, pain, self-care, usual activities, and anxiety-depression (Fryback, 2010).

The second step involves attaching preference weights to each of the functional states that are tracked in the questionnaires. Preference weights have been obtained for several generic HRQoL indexes, but each index has used different approaches to elicit preferences in different subpopulations and different time periods (e.g., QWB preference values were obtained from a sample residing in San Diego in 1974–75; EQ-5D preference values have been obtained from subpopulations of 15 countries by the EuroQol Group). The indexes score health using a standardized weighting representing community preferences for health states. Accurate representation of community preferences is important for cost-utility analysis, because only with appropriate stated preferences between different health states can researchers estimate the effectiveness of potential health interventions for a population (Fryback et al., 2007). Utility scores are generated using preference weights from a sample of the general population in a specific area. Evidence suggests that valuations of different health states could differ for people in different countries due to differences in demographic backgrounds, social-cultural values, and political or economic systems; therefore, it is advisable to use country-specific preference weights when valuing health with a generic index (Huang et al., 2007).

The third step is to generate a summary measure of the impact of a disease or intervention on HRQoL by using a scoring algorithm to combine the preference-weighted index values. The scoring algorithm for the health utilities index (HUI) is multiplicative, whereas the QWB scoring algorithm is additive. Researchers who use any of these generic HRQoL estimates to characterize disease burden should be aware of differences across indexes in the elicitation of preferences and the scoring algorithm assumptions and understand how these differences may affect the resulting HRQoL estimate. The final HRQoL estimate is a single measure that typically lies between 0 (death) and 1 (perfect health) (Fryback, 2010).

A related HRQoL measure attempts to characterize the amount of time people are unhealthy. Developed by CDC for use in the ongoing Behavioral Risk Factor Surveillance System (BRFSS), the "Healthy Days Measures" consist of four items eliciting information

about the number of days in the past month with poor physical or mental health (http://www.cdc.gov/hrqol/hrqol14_measure.htm) (Moriarty et al., 2003). Another HRQoL measure that is sometimes used in the United States is the Health and Activity Limitation index (HALex) (Tengs and Wallace, 2000). The HALex was constructed from data in the National Health Interview Survey and is based on a self-assessment of health ("excellent," "very good," "good," "fair," or "poor") and on limitations on five activities of daily living.

A single measure that combines both morbidity and mortality impacts of a disease is often desired to make comparisons across diseases or across alternative treatment approaches. Measures of health-adjusted life years (HALYs) combine information on HRQoL and life expectancy to yield such an estimate. HALYs are summary population health measures that may be used to make comparisons across a range of illnesses, interventions, and populations (Gold et al., 2002). HALY is an umbrella term that refers to a generic measure of health-weighted life years. It includes subclasses such as QALYs and DALYs. HALYs are of significant and varied use in public health research, including for comparing the health of one population to the health of another, comparing the health of the same population at different points in time, quantifying health inequalities within populations, providing appropriate attention on non-fatal health outcomes, and analyzing the benefits of health interventions for use in cost-effectiveness analysis (Gold et al., 2002).

QALYs and DALYs both permit morbidity and mortality to be simultaneously described within a single number. Their original purposes were different, however; thus, it is important to understand their similarities and differences. QALYs were developed in the late 1960s for use in cost-effectiveness analysis and represented a breakthrough in quantifying the health outcome in a cost-effectiveness ratio. The QALY represents a year of life weighted along a continuum of 0 (death) to 1 (ideal health). The benefits of an intervention are thus maximized by increasing both the "utility" and the longevity of individuals and populations. A preference-weighted HRQoL measure, such as one deriving from the HUI or the EuroQol (EQ-5D), is needed for the utility portion of the QALY (Gold et al., 2002; Fryback, 2010).

In large part because of their widespread use and ease of calculation, QALYs were recommended as the preferred measure for HALY assessment by the Institute of Medicine (IOM) Committee to Evaluate Measures of Health Benefits for Environmental, Health, and Safety Regulation (IOM, 2006). QALY estimation uses the HRQoL index measures for the possible health states associated with a particular disease, combined with the time spent in each health state (duration) over the remaining life span. QALYs are useful for decision making because the measure "reflects trade-offs [that people would be willing to make] between survival and quality of life" (IOM, 2006, p. 87).

Two of the more commonly used alternatives to QALYs for measuring disease burden are healthy year equivalents (HYEs) and DALYs. HYEs reflect the number of years in optimal

health that would yield the same level of utility as a given lifetime health profile (e.g., series of health states over time). The HYE measure allows for quality of life to be affected by length of time spent in a particular health state, in contrast to QALYs. However, the HYE approach requires extensive data collection on people's preferences across a large number of health profiles (Hauber, 2009).

DALYs were first presented in the World Development Report 1993 by the World Bank. They express the number of healthy years of life lost due to death or disability. It is significant to note that while QALYs assign a score to health states, DALYS assign disability scores to diseases. Particularly for global burden of disease studies, self-assessments of health status can be problematic because different populations have different attitudes about desired health (WHO, 2008). WHO researchers give the example of aboriginal populations of Australia, who experience higher mortality than the population at-large, yet are less likely to rate their health as "poor" or "fair" (Gold et al., 2002). This concern also led the creators of the DALY to establish disability weights based on expert opinion and secondary data, which indicate a disease's average impacts in a population, rather than preferences drawn from population samples (WHO, 2004). QALYs, by contrast, use HRQoL scores to capture morbidity impacts of disease. Such HRQoL scores rely on patient-reported assessments of health states or functioning.

The DALYs attributable to a specific disease, and for a particular population, equal the sum of years of life lost and years of life lived with disability (DALY = YLL + YLD). Years of life lost (YLLs) are calculated by multiplying the number of deaths for a given cause (N) by the average life expectancy at the age of death (L), [YLL = N x L] (Murray and Lopez, 1996; WHO, 2004). Years of life in disability multiplies the number of disability cases (I) by the average duration of the disease (L) and by a weighting factor (DW) that reflects the severity of the disease (YLD = I x L x DW) (WHO, 2004). Because both mortality and morbidity are captured, the DALY allows population-based comparison of total burden from very different diseases and conditions, such as infectious disease, which inflicts high mortality, and depression, a leading cause of disability (WHO, 2004).

3. METHODOLOGICAL TRENDS AND NEEDS IN BURDEN OF ILLNESS MEASUREMENTS

In this section, we describe three broad issues in burden of illness methodologies and discuss trends and needs related to those issues. We first discuss methods for attributing burden to a specific disease or risk factor, so that a given cost or life year lost or quality of life impact is not assigned to more than one disease (i.e., double-counted). We then discuss trends in developing and reporting on summary measures of population health. We also discuss the importance of valuing time lost to disease and disability—both for patients and caregivers. Finally, we discuss inconsistencies in quality of life measures of burden.

3.1 Attribution of Burden to a Specific Disease or Risk Factor: Avoiding Double-Counting in Burden Estimates

3.1.1 Global Burden of Disease Approaches

The Global Burden of Disease (GBD) framework is the principal attempt to integrate, validate, analyze, and disseminate information on population health worldwide. In 1992, the World Bank commissioned the initial GBD study to provide a comprehensive analysis of the global disease burden in 1990. This study developed and utilized the DALY (Murray et al., 2002).

The principal goal of the GBD study is to use the DALY approach to quantify the loss of health in any of the following health domains: mobility, self-care, pain, and cognitive impairment. Diseases and injuries and risk factors and environmental determinants are all considered to contribute to loss of health. Constructing summary measures requires a method to attribute a health problem to a cause, such as a disease or risk factor. The GBD project uses two types of causal inference: counterfactual analysis and categorical attribution. Counterfactual analysis attempts to estimate what the burden would have been in the absence of the disease or risk factor. For example, disease cost analyses often estimate the marginal impact of a disease on medical costs, then use the estimates to calculate what costs would have been in the absence of the disease. The difference between actual costs and what they would have been in the absence of the disease or risk factor is then used as an estimate of the costs attributable to the disease or risk factor. In contrast, categorical attribution is used to attribute the fatal and nonfatal burden of diseases and injuries to a mutually exclusive and exhaustive set of disease and injury categories (World Health Organization [WHO], 2004). This is commonly known as the "one death, one cause" premise and can be somewhat arbitrary in the event of comorbidities or multicausality. For example, a death that followed hospitalization for both ischemic heart disease and cerebrovascular disease must be attributed to just one of the two conditions. According to Christa Fischer Walker (2010), the categorical attribution in GBD obscures the importance of

multi-disease causality and may lead to incorrect estimates of the possible effects of interventions on mortality.

A comprehensive revision of the GBD framework and approach is underway. These revisions focus on updating epidemiologic estimates, disability weights, and improving methodologies. Estimates of incidence and prevalence have not been fully revised since the original study was performed nearly 20 years ago, and there has not been a comprehensive review of disability weights since the original study. Methodological improvements over time mean that disease burden estimates from the original 1990 study and follow-ups in 2000, 2001, and 2002 are not comparable. The newest study, GBD 2010, was funded by the Bill and Melinda Gates Foundation and will be released in 2011. It will make use of new and improved data sources, such as country-level vital statistics, and new methods of modeling missing data, such as the DisMod III model (Institute for Health Metrics and Evaluation [IHME], 2011).

Researchers in resource-poor settings frequently encounter the problem of incomplete disease registry or administrative records, upon which disease prevalence estimates are based. DisMod III is designed to produce consistent estimates in such circumstances, using a generic model of disease dynamics. IHME is also creating user-friendly DisMod III software that will allow researchers to perform sensitivity analyses for important parameters. By applying a simple model of how disease moves through a population, researchers can combine data from multiple sources to reduce errors in the available data and impute the data that are completely absent. Initially, it will be used internally by IHME researchers; as it is developed, it will be available from IHME to epidemiologists, doctors, and public health researchers (IHME, 2011).

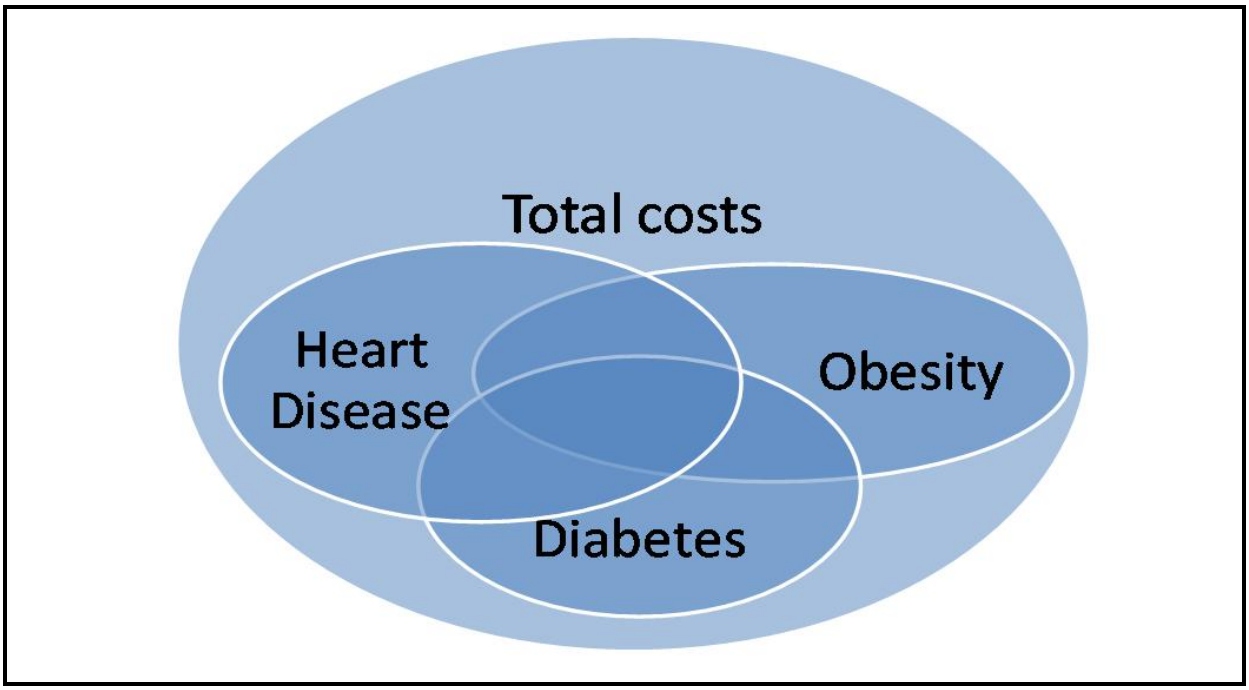
GBD 2010 researchers are reviewing epidemiological evidence to find the most current information by disease and injury on incidence, prevalence, case-fatality, and mortality. They are investigating approximately 300 conditions and 40 separate risk factors across 21 global regions. These epidemiological reviews will be analyzed by GBD researchers using the DisMod III modeling tool. Results for each condition will be critiqued by the expert groups to ensure consistency across the final burden estimates (IHME, 2011).

3.1.2 Approaches for Attributing Medical Costs to Specific Diseases

When people have multiple diseases, attributing burden (whether the measure of interest is deaths, costs, or quality of life) to any one of the diseases can lead to double-counting, or assigning burden to more than one of the diseases. Figure 3-1 is a Venn diagram illustrating costs for selected diseases and conditions. Diabetes is a costly disease that accounts for a significant share of medical care costs. Diabetes is also a risk factor for heart disease, another costly disease. Therefore, some of the costs of heart disease will also be included in a measure of the costs of diabetes. In addition, obesity is a risk factor for both diabetes and

heart disease, so the total cost of obesity also includes some of the costs of diabetes and some of the costs of heart disease. The diagram clearly shows the overlap in costs between diabetes, heart disease, and obesity. Some costs are double-counted, and others are even triple-counted.

Figure 3-1. Double-Counting Costs



Different approaches for estimating the cost of illness can lead to different cost estimates, even when the same data are used. Disease-specific cost of illness studies measure the cost of a single or limited number of diseases, whereas general cost of illness studies allocate a population's total expenditures to a large group of diseases (Rosen, 2010). Most published studies on the cost of a given disease have used a disease-based approach. Yet, because no methodological standards have been issued for cost of illness, and because different approaches lead to different cost estimates, wide variation exists in published disease cost estimates. In particular, these studies do not typically constrain estimates to sum to total medical spending. One study found that the cost of illness for 80 diseases from published cost of illness studies totaled 211% of national health expenditures (Bloom et al., 2001). General cost of illness studies start with total health sector costs and allocate some portion of spending to each disease, which ensures that the total does not exceed the sum of the cost of all diseases. Still, it is unclear what the best approach is to attribute or allocate medical spending to each disease or risk factor of interest.

Double-counting costs can be troubling for policy makers trying to make comparisons between diseases because most cost of illness studies currently focus on individual diseases. Because of double-counting, the sum of the reported individual disease costs may greatly exceed total medical care expenditures. This would seem to violate the basic principle of cost accounting whereby the sum of the individual components of costs should equal total costs. In addition, from a practical standpoint, policy makers may lose confidence in the individual cost studies if each disease's cost is "big" and the total from the studies is higher than total costs.

In this section, we review recent efforts to avoid double-counting when the costs of multiple diseases are estimated simultaneously. Although this is an important methodological issue, from a policy standpoint individual disease cost studies and simultaneous disease cost studies may provide valuable and complementary information.

Alternative Methods for Estimating the Cost of Multiple Diseases Simultaneously

In research supported by the National Institutes of Health's (NIH's) National Institute of Aging, Allison Rosen, David Cutler, and colleagues at the National Bureau of Economic Research are examining the strengths and weaknesses of three alternative approaches to estimating the cost of multiple diseases simultaneously. All three approaches use data from the Medical Expenditures Panel Survey (MEPS) and take a top-down approach starting with total health care spending and allocating it to individual diseases. The encounter-based approach assigns individual patient encounters (e.g., office visits, emergency room visits, hospital stays, prescriptions) to a specific disease based on primary diagnosis codes submitted for insurance purchases. This approach has the advantage that each encounter is assigned to at most one disease, thereby avoiding the double-counting problem. However, many encounters (accounting for about 19% of costs) have no associated diagnoses, and the approach does not account for comorbidities or secondary diagnoses. The episode-based approach uses commercial software to group related encounters into episodes associated with underlying diseases or conditions. This method groups almost all spending into episodes associated with specific diseases, but lack of transparency within the grouping algorithms makes the software something of a black box. The person-based (or regression-based) approach uses regression analysis to estimate a person's total health spending as a function of disease variables. Total health expenditures are regressed on personal characteristics (e.g., age, sex, race) and disease variables that equal 1 if the person has the disease and 0 otherwise. This approach accounts for comorbidities but may miss out on other factors that affect spending.

Table 3-1 shows the annual cost per person for selected diseases by method in 2003. The methods produce similar costs for some diseases, such as chronic renal failure and osteoarthritis. The person-based method produces much higher cost estimates than the other two methods for other diseases, including dementia and congestive heart failure.

Table 3-1. Annual Per Person Cost for Selected Diseases by Method, 2003

Disease	Encounter	Episode	Person
Colon cancer	$8,100	$4,458	$10,475
Lung cancer	$12,082	$14,213	$23,895
Dementia	$596	$1,111	$9,231
Depression and bipolar disease	$616	$984	$1,070
Hypertension	$225	$522	$376
Coronary atherosclerosis	$3,415	$4,342	$3,303
Congestive heart failure	$2,869	$2,476	$12,645
Cerebrovascular disease	$2,563	$2,818	$5,759
Asthma	$348	$639	$519
Chronic renal failure	$11,105	$11,433	$11,964
Osteoarthritis	$1,184	$1,726	$1,450

Source: Rosen, 2010.

According to Rosen (2010), there is not yet a best method for estimating the costs of multiple diseases simultaneously. Instead, the best approach may depend on the available data, the purpose of the analysis, and the resolution of ongoing research into the methods.

BEA Research

The U.S. Department of Commerce's Bureau of Economic Analysis (BEA) has been studying burden of illness measures and has developed plans to develop a prototype medical care satellite account. The BEA is responsible for accurately measuring the nation's output, productivity, and inflation. Because health care expenditures account for such a large share of gross domestic product (GDP), it is necessary to accurately measure output and prices in the health sector. Although total health care spending appears to be measured accurately, there is concern that price is overstated and quality is understated. As a result, GDP growth may be understated by as much as 0.2 percentage points per year and inflation may be similarly overstated (Aizcorbe et al., 2011). Consequently, measured productivity is understated and future budget projections overstate inflation.

Therefore, BEA is developing prototype medical care accounts. These will focus on new price indexes for medical care spending and estimates of spending by disease. The BEA has developed alternative medical price indexes for 2001–2005 using MEPS data. Although all price indexes moved upward during the period, they did so at varying rates (Aizcorbe et al., 2011). The BEA will publish an article on developing the prototype accounts in a September 2011 *Survey of Current Business* article.

The CDC Chronic Disease Cost Calculator

The Centers for Disease Control and Prevention's (CDC's) Chronic Disease Cost Calculator (http://www.cdc.gov/chronicdisease/resources/calculator/index.htm) provides person-based estimates of the cost for six chronic diseases (heart disease, congestive heart failure, hypertension, stroke, diabetes, and cancer). The calculator uses MEPS data and regression analysis to estimate the costs for the six diseases simultaneously while also controlling for age, sex, race, education, income, and a series of other disease indicators. The current version of the calculator only estimates costs for Medicaid beneficiaries; however, the calculator is being expanded to include patients insured by private insurers or Medicare.

By estimating costs for multiple diseases simultaneously, the person-based approach used in the calculator partially resolves the double-counting problem. However, the regression specification underlying the calculator raises a more subtle double-counting issue. Health care cost data have special characteristics (many patients have zero expenditures, and expenditures are skewed to the right) that make ordinary least square estimates inappropriate. Therefore, the calculator employs a nonlinear estimator that is more appropriate for the data. However, the nonlinear framework complicates the allocation of costs to individual diseases when a patient has more than one disease; this complication can lead to double-counting if the coefficients are not interpreted carefully. To avoid double-counting, the calculator uses a cost allocation developed by Trogdon et al. (2008). When a patient has two (or more) comorbidities, this cost allocation assigns a share of the costs attributed to the comorbidities to each of the individual conditions. The sum of the shares equals 1, thereby avoiding double-counting.

3.1.3 Novel Epidemiological Approaches to Detect Causal Relationships between Risk Factor Exposure and Disease Outcomes

Methodological advances in approaches to detect causal relationships between risk factors and health outcomes may lead to improvements in burden of disease measurement. New methods have focused on using observational data, such as data from health insurance claims or electronic medical records, to make causal inferences. Researchers at the Harvard School of Public Health and the MD Anderson Cancer Center have been active in developing novel statistical approaches for estimating the impacts of risk factors or specific treatments on health outcomes. In an ideal world, policy and clinical decisions would be based on experiments that compare the effectiveness of several randomly assigned interventions in large samples of people that adhere to the study parameters. Unfortunately, such ideal randomized experiments are uncommon due to concerns about the time and resources needed to implement these approaches as well as concerns about the ethics of social experimentation. Research at Harvard School of Public Health is directed toward emulating hypothetical experiments by combining observational data, assumptions, and statistical methods. The research focuses on using analytic approaches where assumptions do not

conflict with current subject matter knowledge. Research at MD Anderson Cancer Center has used Bayesian statistics to develop innovative clinical trials, laboratory experiments, and observational studies with a focus on cancer.

3.2 Summary Measures of Population Health

In this section, we briefly describe ongoing efforts to create consistent summary measures of population health that can be used to assess U.S. population health over time or across subgroups or that can be used to value health outcomes in analyses of the effectiveness and cost-effectiveness of clinical and health promotion interventions. We first describe the Patient-Reported Outcomes Measurement Information System (PROMIS) initiative, an effort to develop comprehensive questions for use in assessing the impact of disease or intervention on a broad range of health functioning outcomes. We then describe recent efforts to standardize and improve on health outcome assessment using QALYs. Next, we describe recent efforts to use the CDC Healthy Days Measures. Finally, we discuss the Canadian experience with developing and using HRQoL measures to monitor population health and to conduct research using standardized measures.

3.2.1 PROMIS

In 2004, NIH launched an initiative to address major gaps in patient-reported clinical outcomes. PROMIS is focused on building standardized and accessible item banks to measure patient perceptions of common chronic conditions. A better understanding of patient-reported outcomes, including fatigue, emotional health, pain, and mobility, will enable a significantly higher understanding of the burden of chronic conditions and will positively affect medical care by improving clinical trial research and advancing clinical practice guidelines. The goal of PROMIS is to develop psychometrically robust item banks and computerized adaptive tests that the clinical research community can administer from a Web-based repository. Clinical researchers would be able to efficiently collect patient-reported outcome measurements from their subjects on key symptoms and domains of interest and report instant health assessments that can be compared on a common metric (NIH, 2011).

The PROMIS questionnaires generate function scores similar to generic health status profiles such as the SF-36, although their calibration and adaptation yield higher precision than is possible with a fixed form. During the roundtable meeting, presenter Ron Hays elaborated on the statistical precision of PROMIS: it uses the item response theory testing methodology, which allows new items to be cross-linked with existing items and is critical in obtaining desired precision using the minimum possible number of questions. PROMIS provides an opportunity to improve health care outcomes by giving decision makers quantitative data on how conditions and care affect what patients are able to do and how they actually feel (Hays, 2010).

Cherepanov and Hays (in press) indicated that the future importance of patient-reported outcome measures, such as those provided by PROMIS, is enhanced with the recent passage of HR 3590, the Patient Protection and Affordable Care Act, which established a nonprofit corporation, the Patient-Centered Outcomes Research Institute, to advance the quality and relevance of evidence on disease prevention, diagnosis, monitoring, and treatments. Patient-reported outcome measures greatly enhance health researchers' ability to track wellness among the U.S. population. As such, Cherepanov and Hays expect patient-reported outcomes to gain particular significance in the future, as the focus of health care delivery shifts from diagnosis and treatment of health problems to wellness and prevention of disease.

3.2.2 Quality-Adjusted Life Years (QALYs) and Health-Related Quality of Life (HRQoL)

A great deal of recent efforts to improve on QALY measurement has focused on assessing differences in the HRQoL indexes used to generate QALYs. The National Health Measurement Study was one of the most ambitious efforts to date to compare HRQoL measures across different generic indexes. The study administered six indexes using the same data collection approach to the same group of individuals in 2005–2006 to derive comparable results across all six (Fryback et al., 2007). The indexes were the EQ-5D, the HUI2 and HUI3, the QWB, the SF-6D, and the HALex. Fryback et al. (2007) found significant differences in the mean age- and gender-specific scores across the six indexes, suggesting that the QALY estimates that result from using, for example, the QWB would be far different from those estimated using the EQ-5D. More recently, Fryback and colleagues (2010) have attempted to develop a crosswalk between five health indexes. They found that such crosswalks may work fairly well for lower health states, but do not work well for better states of health (Fryback et al., 2010).

3.2.3 CDC Healthy Days Measure

The United States has not adopted a standardized measure of HRQoL, such as the QWB or EQ-5D, for use in population health comparisons. CDC produces a different measure of HRQoL—the Healthy Days Measures—which consist of four components of physical and mental health. Questionnaire responses are summed to generate a score of total physically and mentally unhealthy days in a given month. Recent work has used data from National Health Measurement Study to create a crosswalk between the CDC Healthy Days Measures and preference-based utility scores (Fryback et al., 2010). Dr. Matthew Zack of CDC related at the roundtable meeting that HRQoL information such as the Healthy Days Measures can be useful for communicating a summary health measure to health practitioners and members of the public. He emphasized that although preference-weighted HRQoL measures, such as the EQ-5D and the HUI, are necessary for evaluating the effectiveness and cost-effectiveness of clinical and prevention interventions, they may not be meaningful

to health practitioners and public health policy makers. Indicators such as the CDC Healthy Days Measures may be more easily understood and acted upon to improve the public's health (Zack, 2010).

3.2.4 Statistics Canada Efforts to Standardize HRQoL Measures

In a presentation at the roundtable meeting, Keiko Asakawa, a health economist with Statistics Canada, described efforts to standardize HRQoL measures in Canada and how Canada's experience differs from that of the United States (Asakawa, 2010). Dr. Asakawa first discussed the sources of possible differences in HRQoL measures across generic HRQoL indexes. First, there are notable differences in possible health states across generic HRQoL indexes, from 243 in the EQ-5D to 972,000 in the HUI3. There are inherent tradeoffs in selecting one measure over another for use in national surveys. Psychometric properties vary across instruments, and due to domains captured, certain instruments are favorable for specific diseases. The generic indexes also differ in the approaches used to elicit preferences across health states and in the final scoring algorithm, and these differences imply that one index may be better for measuring burden for certain diseases, but worse for other diseases. None of the available generic indexes provides the best burden measure for all diseases. For example, there is a large ceiling effect for the EQ-5D, meaning it is difficult to judge differences in respondents with very good health; 36% of EQ-5D respondents are reported in "perfect" health. The SF-6D, conversely, is subject to a floor effect, suggesting that it may not be useful for detecting differences in utility among the very sick. Furthermore, as research by Fryback and others has shown, the preference-based HRQoL measures generated by different instruments are not interchangeable, and a lack of understanding of their differences can lead to misuse or misinterpretation.

Canada has standardized collection of the HUI3 in national health surveys, including the National Population Health Survey (cross-sectional and longitudinal), the Canadian Community Health Survey (cross-sectional), and the Canadian Health Measures Survey (cross-sectional). The Canadian health agencies chose the HUI3 because of its ability to rate diverse and independent health attributes, to detect small changes in health status, and to provide preference-based utility scores for use in cost-effectiveness analysis. Although no measure is perfect, standardization enabled Statistics Canada to link population health surveys to administrative and census data at the individual level (e.g., a hospital discharge database). Much debate in the United States has arisen from discussions of whether one specific health index should be selected as the "best" for use in HRQoL measurement. Because this debate continues, with some even calling for the development of a new HRQoL summary measure using the PROMIS item banks (Fryback, 2010), it is unlikely that a single measure will be selected or recommended anytime soon for use in U.S. cost-effectiveness analyses.

3.3 Valuation of Time Lost to Disease and Disability

Economic measures of burden often include both direct and indirect costs, where indirect costs capture the value of time lost to disease, disability, or death. A great deal of recent economics research has focused on how to value those time losses. Probably the most commonly used method is the human capital approach, which assigns earnings losses to people's time lost to early mortality and excess morbidity. However, no accepted guidelines exist to value time losses for children, and the time costs captured for elderly people or others who do not work reflect valuations of household productivity only. This is a critical limitation for health policy purposes because many health interventions are targeted at either children or the elderly, and these groups may account for as much as 50% of health care costs.

Despite its potential utility for valuing time losses in both working and nonworking populations, use of willingness-to-pay (WTP) in health policy has been somewhat limited. In contrast, WTP estimates of the value of life are already used in environmental and safety regulation. For many diseases, implementing WTP approaches can be challenging. For example, it is often difficult to identify economic situations, such as occupational choice, that easily lend themselves to the estimation of reduced injury and death to provide relevant estimates for the disease of interest. Furthermore, survey approaches, although they may be tailored to address the specific features of the disease of interest, are expensive to conduct, and problems with over- and under-estimation of WTP values have been widely noted (Portney, 1994). Future research to generalize and refine WTP methods is warranted to overcome these challenges.

Economists are also beginning to address the time costs of preventing, treating, or managing a disease. These non-monetary costs may represent large components of a disease's overall burden, but they are frequently overlooked in cost analyses or economic evaluations of treatment interventions (Russell, 2009; Freeman and Loewe, 2000). Russell (2009) emphasizes the need to collect and report on non-monetary costs, such as the value of a patient's time spent obtaining care or managing a disease, as well as the time costs for unpaid caregivers, such as a family member who accompanies a patient for doctor visits.

3.4 Inconsistencies in Quality of Life Measures of Burden and Sources of Those Inconsistencies

As discussed in Sections 3.2.2 and 3.2.4, inconsistencies have been identified in QALY losses and QALY changes for the same set of respondents, depending on which generic HRQoL measures are used to generate the QALY estimates. Environmental scan interviewees provided us with in-depth descriptions of the primary strengths and weaknesses of HRQoL and HALY measures. Several described the importance of considering the nature of the disease or group of diseases under study to select the most appropriate

measures of HRQoL, a recommendation that was also made in the IOM (2006) report, *Valuing Health for Regulatory Cost-Effectiveness Analysis*. Below we describe the main issues concerning generic HRQoL indexes that users should take into account when selecting an index for a given analysis or when critically reviewing published studies. No single generic HRQoL index is best in all situations. Researchers should therefore select the HRQoL measure that most closely matches the illness and population being studied.

- **Domains measured**. The domains included in the index should correspond to the functional domains or health states affected by the disease or intervention. For example, if a disease has impacts on cognition, the researcher should select a generic index that captures cognitive impacts, such as the HUI, or the appropriate disease-specific index.
- **Population surveyed for preference elicitation**. The generic HRQoL indexes have used various populations for preference elicitation. Researchers should consider the extent to which the population under study (e.g., institutionalized versus community-dwelling adults) is represented in the HRQoL measurements when selecting a given index. For example, the HUI2 collected preference information for children from parents. Consequently, the HUI2 may be the best generic index for valuing the burden of a disease in children. The EQ-5D is the only index that has recently collected preferences from a sample of American adults. Asakawa and Feeny (2009) recently found that the determinants of health differ between people living in institutions and those residing in the community.
- **Nature of the disease or intervention—healthy versus sick people**. Some of the most widely used HRQoL indexes exhibit floor effects (i.e., cannot fully capture the impacts of disease for people with poor quality of life), whereas others exhibit ceiling effects (i.e., do not adequately distinguish between health states for people who are generally healthy). Therefore, it is important to select an index that will best capture the full range of HRQoL impacts of a disease or intervention. For example, if the disease or illness of interest tends to affect people in overall good health, such as injury among runners, then the SF-36 or SF-6D should be selected over the EQ-5D or HUI, because the latter two exhibit ceiling effects for relatively healthy populations.
- **Nature of the disease or intervention—single domain versus multiple domains affected**. The approaches used to combine the preference-weighted HRQoL scores into a single measure differ across the widely used generic indexes. The QWB uses a straight linear additive model, and the EQ-5D, SF-6D, and SF-36 use linear additive models with additional terms. The HUI2 and HUI3 use multiplicative models that allow for interactions between preferences across domains. If a disease or intervention affects multiple domains, these interactions across domains may be important to capture. For example, individuals with diabetes experience physical impairment that may be causally related to mental health effects of diabetes, such as depression. This issue is unlikely to be important for illnesses that affect only one health state or domain. See the literature review (Appendix A) for links to the questionnaires that form the basis of some of the generic health indexes.

This page intentionally left blank.

4. DATA TRENDS AND NEEDS IN BURDEN OF ILLNESS MEASUREMENT

Many U.S. data sources provide information related to burden of illness. Some measures that can be obtained from U.S. data sources can provide information about one of the three broad categories of epidemiology, economics, and health-related quality of life (HRQoL); for example, disease prevalence (from the National Health and Nutrition Examination Survey [NHANES]), direct expenditure (from the Medical Expenditure Panel Survey [MEPS]), and health utility weights (from the 2005–2006 National Health Measurement Study). Others, such as summary measures of population health, are hierarchically dependent on other measures. For example, the disability-adjusted life year (DALY) relies on mortality data from country-level vital registration systems and the number of incident disability-causing cases for a particular condition, which is either determined from population-level surveys or imputed if no data are available.

4.1 State and Local Data on Burden of Illness—More Local Data Are Needed and Desired

Quantifying disease burden at the local level is important because it provides targets for local public health intervention. The need for more complete and uniform burden data at state and local levels was the primary data need described by environmental scan and roundtable meeting participants. Environmental scan interviewees indicated that local data collection on burden of illness is important because the burden for any particular disease may vary a great deal from one place to another, and decisions about public health priorities and how to allocate public health resources tend to be made at state and local levels in the United States. State and local public health leaders need information about disease burden that is specific to their populations to make informed decisions.

Currently, few local areas consistently collect data on disease burden within the community, and data collection approaches are not uniform across those communities that do collect data. The Behavioral Risk Factor Surveillance System (BRFSS) allows for uniform data collection at the state level and for some communities within the United States, but many of the BRFSS modules are optional and therefore are not collected across all states or communities.

Some efforts are currently underway to estimate burden at local levels within the United States. For example, Ezzati et al. (2008) estimated the mortality impacts of uncontrolled hypertension at the state level using BRFSS data and applying estimated relationships between self-reports of hypertension and clinical findings on hypertension prevalence from NHANES.

4.1.1 Mobilizing Action Towards Community Health (MATCH) Initiative

The Mobilizing Action Towards Community Health (MATCH) Initiative, facilitated by the University of Wisconsin and funded by the Robert Wood Johnson Foundation, is a recent example of aggregating local measures of population health data to create new insights. MATCH measures and ranks community health for every county in the United States using health outcomes, which are measures of current health, and health factors, which are considered to be measures of future health. The project's summary health outcomes metric gives equal weight to premature death (a measure of years of life lost), unhealthy days, and low birthweight (both measures of morbidity). MATCH uses data collected in the BRFSS. Health factors are considered to be an indication of a county's future health, due to their time-lagged effects on public health. They include health behaviors, such as smoking rate; socioeconomic factors, such as percentage of children in poverty and homicide rate; and environmental factors, such as access to healthy foods and particulate matter concentration. MATCH researchers have publicized the study (countyhealthrankings.org) and its public policy implications: that policy affects health factors, which in turn affect health outcomes. They have published state-level rankings of health by county, so that citizens and policy makers can gauge their county's health relative to the rest of the state. This comparison is designed to mobilize community- and state-level action toward improving health factors and current and future health outcomes.

4.2 Burden of Rare Diseases

Quantifying the burden of rare diseases is another data challenge. Roundtable meeting participants noted that the National Institutes of Health (NIH) conducts research for thousands of diseases, but we have good burden data for only about 120 diseases. This raises the question of how to quantify burden for rare diseases—those for which population-based studies cannot be used to assess disease prevalence or impact (e.g., end-stage renal disease, schizophrenia). Disease registries are one way to quantify burden, but they are expensive to start and maintain. Moreover, as roundtable meeting participants noted, advocacy groups often support the development and maintenance of disease registries, but because it is infeasible to create disease registries for all rare diseases, policy makers need input on how to assess the need for a registry for any particular disease and the usefulness of registry data for generating burden measures.

4.3 Quality of Burden Data from Health Insurance Claims and Other Administrative Data Sources

Data from health insurance claims and other administrative databases are often used to estimate the prevalence or cost of a given disease. However, because these data are collected for the purpose of billing and are not drawn from population-based surveys, they are not generally representative of the U.S. population and may result in biased estimates of disease burden. For example, the early stages of chronic kidney disease are rarely

reported as diagnoses in administrative records because people with chronic kidney disease often have other chronic conditions (e.g., diabetes or heart disease) that are reported instead. Consequently, using health care claims data to estimate the prevalence of chronic kidney disease would likely result in an underestimate of the number of people with the condition. Similarly, estimates of per-person health care spending for chronic kidney disease are likely to be overstated in health care claims, because those patients for whom a chronic kidney disease diagnosis is entered are likely to be the sickest and most costly patients. Roundtable meeting participants noted that, to ensure that burden estimates are accurate, it is important to collect data expressly for the purpose of measuring disease burden.

4.4 New Methods to Estimate Deaths Attributable to Disease When Limited Data Are Available

Sophisticated statistical methods are being applied in new ways to more accurately estimate epidemiologic burden measures for which limited data are available. For example, Rajaratnam et al. (2010) and Obermeyer et al. (2010) have developed new methods for estimating adult mortality for countries without a central registry in place to track deaths. These include using Gaussian process regression to estimate the annual probability of death between 15 and 60 years and development of the Corrected Sibling Survival method. This is used for adjusting sibling reports of deaths to account for the fact that in families with high mortality, all siblings may have died and be unable to provide survey data on sibling deaths. It can also account for lack of reporting due to forgetting some sibling deaths or being unsure of sibling status. Researchers have also recently developed new methods for estimating child mortality when incomplete birth history data are available. These approaches may prove most useful for estimating child and adult mortality in developing countries, where vital registration data are often not available or are invalid or contradictory.

Other methodological advances that have implications for burden of disease measurement are efforts to improve approaches for detecting causal relationships between exposures to risk factors (e.g., smoking), or interventions to reduce the prevalence of risk factors, and disease outcomes. In particular, methods have focused on using observational data, such as data from health insurance claims or electronic medical records, to make causal inferences. Researchers are currently developing statistical approaches for estimating the impacts of risk factors or specific treatments on health outcomes.

This page intentionally left blank.

5. USING BURDEN OF ILLNESS TO INFORM POLICY DECISION MAKING

Figure 5-1 lists potential policy uses for burden of illness measures. Burden measures can play a large role in priority setting by identifying diseases that are likely to have large needs for clinical services, risk factors that may be preventable, and diseases that need further research to better understand the disease and how it can be prevented or treated. Although understanding the burden associated with a disease is a necessary first step in developing priorities, burden measures alone seldom provide enough information to make informed resource allocation decisions. Additional information is usually needed. For example, are there currently available interventions that can treat the disease? How much of the burden can be reduced by the available interventions? How much will the interventions cost? Thus, burden measures should be viewed as the starting point for priority setting and resource allocation, to be supplemented with additional information.

Figure 5-1. Potential Uses of Burden of Illness Measures

Setting Priorities for Resource Allocation
- Health services
- Prevention
- Research

Monitoring and Evaluating Policy
- Cost-effectiveness analysis of individual interventions
- Monitoring performance of health systems

Monitoring Population Health

Identifying Disparities

Understanding the Burden of Morbidity and Assessing Morbidity/Mortality Tradeoffs

Budgeting for Future Health Expenditures

Improving Measurement of National Income and Inflation

Identifying Opportunities for International Development
- Technology transfer to low-income nations
- Neglected diseases
- Reductions in burden may increase human capital formation

Although priority setting for health care services may focus on treatments for specific diseases, priority setting for prevention often focuses on specific, modifiable risk factors, such as smoking, poor diet, physical inactivity, lack of immunization, and drug and alcohol use. This creates the need for burden measures related to the risk factors. Methodologically, this means that the risk factors must be linked to specific diseases and their consequences using attributable fractions or other methods. As with priority setting for disease treatment, resource allocation decisions for prevention may start with burden measures but will also require information on available interventions, intervention outcomes, and costs.

Even by themselves, burden measures may provide a strong starting point for setting research priorities. Although research allocation will also be affected by opportunities for breakthroughs and the opportunity to translate previous research into practice, one of the main objectives of research is to develop new breakthroughs that improve health. Diseases with a large burden represent an obvious target for breakthroughs. A strong correlation has been observed between National Institutes of Health (NIH) spending and the burden of disease, as measured by disability-adjusted life years (DALYs) (McKenna and Zohrabian, 2009), suggesting that information about disease burden may be guiding, to some extent, U.S. decisions about research priorities.

Burden of illness measures can also be used to evaluate and monitor health policies. As suggested above, burden measures may be combined with information on interventions, costs, and outcomes to evaluate individual interventions in cost-effectiveness analyses. Changes in burden measures at the aggregate level or for particular diseases may also be tracked over time to determine whether the overall performance of the health care system is improving. For example, is life expectancy improving and infant mortality falling? Is the burden from heart disease, cancer, and other leading diseases decreasing? How does the change in burden compare with changes in health care spending on these diseases? In a similar way, burden measures can be used to monitor changes in population health over time.

Eliminating health disparities was one of the two overarching goals behind the Healthy People 2010 framework for health promotion and disease prevention in the United States. For Healthy People 2020, the goal has been expanded to "achieve health equity, eliminate disparities, and improve the health of all groups" (U.S. Department of Health and Human Services, 2009). Burden of illness data are essential for identifying disparities and for monitoring progress toward the Healthy People 2020 goals. Epidemiological data have long been collected to measure gaps in life expectancy at the national level between blacks and whites and between males and females. More recently, burden studies have looked at disparities in life expectancy at the local level (Murray et al., 2006; Kindig et al., 2010). The studies show wide disparities, and some of the disparities have widened over time. Local burden data may support targeted policies to reduce disparities.

Quality-of-life measures play an important role in quantifying the role of morbidity, providing perspective on the tradeoff between morbidity and mortality in policy decisions. One of the most important insights from the Global Burden of Disease (GBD) project is that unipolar depressive diseases are the third leading cause of the burden of disease worldwide, as measured by DALYs in 2004 (World Health Organization [WHO], 2008). This cause did not rank among the top 20 causes for years of life lost, but ranked first in years lived with disability.

In addition to their other roles, economic measures of burden—combined with epidemiologic data on trends in risk factors and demographics—are useful in predicting future health expenditures, both at the national level and for public health insurance programs including Medicare and Medicaid. These programs account for a large and growing share of federal spending; therefore, predicting their future expenditures is essential for efforts to control the federal budget.

On a more esoteric, but still important, policy level, the Bureau of Economic Analysis (BEA) efforts to develop satellite health accounts have the potential to improve measurements of gross domestic product (GDP), health care productivity, and inflation. Inflation measures directly affect cost-of-living increases for Social Security and other programs, and better measurement of health care productivity would feed into Medicare fee adjustments for physicians. As a by-product, the BEA health accounts may provide more standardized estimates of individual diseases.

Finally, measures of the burden of disease may inform efforts to target foreign aid and promote international development. Projects such as WHO's GBD may identify diseases where technology can be transferred from high-income to low-income nations at relatively low cost to fight diseases that are common to both types of nations. The transfer of HIV/AIDS drugs from high-income nations to Africa provides a good example. Burden measures may also identify diseases that are rare in high-income nations but common in low-income nations. Policies to subsidize development of treatments for these neglected diseases may be desirable. There is also evidence that reductions in disease burden for children encourage education and other forms of human capital formation. Reducing burden may therefore increase international development.

Although burden measures provide useful information, policy makers should also be aware of the limitations of burden measures when making decisions. First, different measures capture different dimensions of burden. Therefore, looking at multiple measures for the same disease may provide a more complete picture of the disease's burden than relying on a single measure of burden. For example, accidents are the fifth leading cause of death in the United States, but because many accident victims die at early ages they account for the third most years of life lost.

Second, approaches to measure disease burden have not been fully standardized. Different approaches to valuing disease costs or quality of life impacts can lead to vastly different estimates of burden, and different values for the same burden measure may create confusion among policy makers as they try to select the best burden estimates for a given disease. Efforts are underway to improve consistency in burden measurement for health care costs and for measuring patient-reported health outcomes. These efforts are likely to lead to more consistency in future estimates of disease burden, which will ultimately benefit

policy makers and the population as a whole, as better estimates of disease burden may contribute to well-informed decisions about public health resource allocation.

Third, policy makers should be aware that most burden measures apply implicit or explicit values on health states. For example, the years of life lost before age 75 ignores (i.e., assigns a value of 0 to) deaths occurring after age 75, and seemingly objective measures like deaths place an implicit value of 0 on morbidity. These values may not be shared by all members of society.

Fourth, when using burden measures to set priorities for resource allocation, there may be a tendency to focus only on common diseases with high prevalence and large burdens. This is an oversimplification: collectively, rare diseases may have a large burden and individually there may be strong opportunities for addressing specific diseases. A more nuanced approach would ensure that rare diseases still receive resources, but at a rate that is approximately equal to their relative burden.

6. SUMMARY OF LESSONS LEARNED ABOUT BURDEN OF ILLNESS AND ISSUES FOR FURTHER CONSIDERATION

This study has compiled information on burden of illness measures and measurement approaches from the literature; recent and ongoing federal initiatives; and experts in the fields of epidemiology, economics, health services research, medicine, and health-related quality of life. Burden of illness was defined very broadly for our purposes as any impact of disease or disease risk factors on individuals, society, or government. Our efforts to compile information on burden of illness and to identify trends and challenges in burden of illness measurement culminated in a roundtable meeting of experts from various fields to discuss the usefulness of different measures of burden and to identify and discuss approaches to address remaining challenges in burden of illness measurement. Working with these experts and with those who provided input on the environmental scan and drawing from the published literature, we have identified a few important messages about burden of illness measures and measurement to help guide policy makers in making use of burden of illness measures. We also highlight remaining methodological and data issues in measuring the burden of illness.

6.1 A Single Summary Measure of U.S. and Local Disease Burden May Not Be Needed or Desired

Each burden of illness measure conveys different information about the type of impact of a disease or risk factor and, in some cases, about the particular individuals affected. For example, a measure of years of life lost (YLL) at birth provides information about the impact of a disease on mortality over the lifetime, whereas a measure of days of work lost describes the impact of a disease on disability that causes job or work loss primarily among working-aged adults. Appendix D contains a table that summarizes the impact of 20 common diseases or risk factors in the United States on seven different measures of disease burden in 1996: disability-adjusted life years (DALYs), years lived with disability (YLD), YLL, deaths, costs, hospital days, and days of work lost. The most burdensome illness in terms of DALYs, YLL, and costs is ischemic heart disease. However, depression is the most burdensome illness based on measures of YLD and days of work lost, and it is one of the most burdensome based on a measure of hospital bed days.

All of these different burden measures may be useful to policy makers for deciding on resource allocation among the many different diseases that affect people in the United States and worldwide. However, no one measure is perfect; each has limitations. For example, deaths entirely ignore disability impacts of disease, and DALYs may not reflect patient preferences for having different illnesses. Because each measure provides a slightly different description of disease burden, it is important to consider multiple measures, rather than make policy decisions based on a single summary measure, such as the DALY or

quality-adjusted life years (QALYs). Moreover, different measures are suited to different purposes. It may be most relevant to use one measure to guide certain policy decisions (e.g., policies affecting working-aged adults) and another measure to guide other decisions (e.g., decisions related to costs). Given the limitations of focusing on a single measure of disease burden, policy makers may find it most useful to be presented with multiple burden measures and to integrate them pragmatically for decision making.

6.2 More Work Is Needed on the Allocation of Burden to Specific Diseases

Work is underway, but more work is needed, on the allocation of burden to specific diseases. Much of the current work on allocating burden to specific diseases focuses on economic measures of burden, specifically addressing the problem of double-counting of costs when people have more than one disease or risk factor. Double-counting arises when the same hospitalization cost, for example, is assigned to diabetes, heart disease, and obesity for an individual who has all three conditions. The challenge is to develop an approach for assigning costs to each individual disease, while constraining total costs not to exceed national health care spending. Although several efforts are underway to address the problem, there is not yet a best method for estimating the burden of multiple diseases simultaneously. Instead, the best approach may depend on the available data, the purpose of the analysis, and recommendations that may emerge from ongoing efforts (e.g., National Health Expenditure Accounts, Bureau of Economic Analysis, Global Burden of Disease) to advance the methods for assigning burden to specific diseases and risk factors.

6.3 More Work Is Needed on Preference-Based Measurement of Quality of Life

Measures of health-related quality of life (HRQoL) can vary widely depending on the underlying health functioning domains included in the study and how preferences for those domains are elicited and valued. Such methodological differences in HRQoL measures can lead to very different estimates of QALY losses for a disease, even from the same set of respondents. Although recent research has attempted to develop algorithms to derive one HRQoL measure from another, efforts are ongoing and no consensus exists for using any particular measure of HRQoL in cost-effectiveness and other analyses. The Institute of Medicine (IOM) (2006) has recommended that researchers consider the characteristics of the disease and patients when determining which HRQoL measure to use in a given analysis of the effectiveness or cost-effectiveness of an intervention.

6.4 Local Estimates of Disease Burden Are Needed to Inform Local Public Health Decision Making

Within the United States, large differences exist in the types of conditions that create the most burden. In some areas, alcohol-related illnesses impose the greatest burden, whereas

in others, obesity-related illnesses create high levels of burden. Because of these differences across local areas, local estimates of disease burden are desired and needed to inform local public health decision making. In the United States, local public health departments are tasked with responding to the health needs of their local communities, and data on local burden could help these agencies best allocate their public health resources.

6.5 Burden Measures Are Important but Are Not Sufficient to Inform All Policy Decisions

Although burden measures are an important first step in quantifying the impact of a disease or risk factor, additional information is nonetheless required to make fully informed decisions about allocating resources across health care, prevention efforts, and research opportunities. For example, when deciding how to invest resources for optimizing population health, information is also needed on the effectiveness and costs of interventions or research strategies to diagnose, treat, or prevent disease. These data can help inform policy decisions about where and how much to invest to achieve improved health outcomes at both local and national levels.

Finally, our project has focused strictly on health outcomes, but there may be non-health outcomes related to illness and injury that policy makers should consider in decision making. For example, social and financial well-being are additional measures of disease impact that may be useful for policy makers to consider in addition to health-related measures of burden.

This page intentionally left blank.

REFERENCES

Aizcorbe, A. (2010). *BEA's medical care satellite account: Relevance of burden of illness measures and plans for a prototype account*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington, DC.

Aizcorbe, A., Bradley, R., Greenaway-McGrevy, R., Herauf, B., Kane, R., Liebman, E., et al. (2011). *Alternative price indexes for medical care: Evidence from the MEPS Survey.* Retrieved from http://www.bea.gov/papers/pdf/MEPS_Draft_012011.pdf.

Akobundu, E., Ju, J., Blatt, L., & Mullins, C. D. (2006). Cost-of-illness studies: A review of current methods. *Pharmacoeconomics, 24*(9), 869–890.

Asakawa, K. (2010). *Generic preference-based measures: Fundamentals, comparisons and applications*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington, DC.

Asakawa, K., & Feeny, D. (2009). Do the determinants of health differ between people living in the community and in institutions? *Social Science & Medicine, 69*(3), 345–353.

Bayoumi, A. M. (2004). The measurement of contingent valuation for health economics. *Pharmacoeconomics*, *22*(11), 691–700.

Bloom, B. S., Bruno, D. J., Maman, D. Y, & Jayadevappa, R. (2001). Usefulness of U.S. cost-of-illness studies in healthcare decision making. *Pharmacoeconomics, 19*(2), 207–213.

Centers for Medicare & Medicaid Services (CMS). (2011). National Health Expenditure Data Fact Sheet. Retrieved from https://www.cms.gov/NationalHealthExpendData/25_NHE_Fact_Sheet.asp.

Cherepanov, D., & Hays, R. D. (in press). Health and quality of life outcomes: The role of patient-reported measures. In J. L. Magnabosco & R. W. Manderscheid (Eds.), *Outcomes measurement in the human services: Cross-cutting issues and methods* (2nd ed.). Washington, DC: National Association of Social Workers Press.

Dicker, R., Coronado, F., Koo, D., & Parrish, R. G. (2006). Principles of epidemiology in public health practice, 3rd ed. [SS–1000]. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Office of Workforce and Career Development.

Druss, B. G., Marcus, S. C., Olfson, M., & Pincus, H. A. (2002). The most expensive medical conditions in America. *Health Affairs (Millwood)*, *21*(4), 105−111.

Environmental Protection Agency (EPA). (2000). *Guidelines for preparing economic analysis.* Chapters 6 and 7. Report 240-R-00-003. Washington, DC: Office of the Administrator, USEPA.

Ezzati, M., Oza, S., Danaei, G., & Murray, C. J. L. (2008). Trends and cardiovascular mortality effects of state-level blood pressure and uncontrolled hypertension in the United States. *Circulation, 117*(7), 905–914.

Freeman, A. M. (1993). *The measurement of environmental and resource values: Theory and methods* (1st ed.). Washington, DC: Resources for the Future.

Freeman, J., & Loewe, R. (2000). Barriers to communication about diabetes mellitus: Patients' and physicians' different views of the disease. *The Journal of Family Practice*, *49*(6), 507–513.

Fryback, D. G. (2010). *Measuring health-related quality of life*. Monograph presented at the Workshop on Advancing Social Science Theory: The Importance of Common Metrics. Washington, DC: The National Academies, Division of Behavioral and Social Sciences and Education.

Fryback, D. G., Palta, M., Cherapanov, D., Bolt, D., & Kim, J. (2010). Comparison of 5 health-related quality of life indexes using item response theory. *Medical Decision Making*, *30*(1), 5–15.

Fryback, D. G., Dunham, N., Palta, M., Hanmer, J., Buechner, J., Cherepanov, D., et al. (2007). U.S. norms for 6 generic health-related quality-of-life indexes from the national health measurement study. *Medical Care*, *45*(12), 1162–1170.

Gold, M. R., Stevenson, D., & Fryback, D. G. (2002). HALYs and QALYs and DALYs, oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health, 23*(1), 115–134.

Hauber, A. B. (2009). Healthy-years equivalent: Wounded but not yet dead. *Expert Review of Pharmacoeconomics and Outcomes Research, 9*(3), 265–269.

Hays, R. (2010). *A 15-minute introduction to PROMIS*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington, DC.

Hodgson, T. A., & Meiners, M. R. (1982). Cost-of-illness methodology: A guide to current practices and procedures. *Milbank Memorial Fund Quarterly Health and Society, 60*(3), 429–462.

Honeycutt, A. A., Segel, J. E., Hoerger, T. J., & Finkelstein, E. A. (2009). Comparing cost-of-illness estimates from alternative approaches: An application to diabetes. *Health Services Research*, *44*(1), 303–320.

Huang, I. C., Willke, R. J., Atkinson, M. J., Lenderking, W. R., Frengakis, C., & Wu, A. W. (2007). US and UK versions of the EQ-5D preference weights: Does choice of preference weights make a difference? *Quality of Life Research, 16*(6), 1065–1072.

Institute for Health Metrics and Evaluation. (2011). Global burden of diseases, injuries, and risk factors study 2010. Retrieved from http://www.healthmetricsandevaluation.org/research/project/global-burden-diseases-injuries-and-risk-factors-study-2010.

Institute of Medicine. (2006). Valuing health for regulatory cost-effectiveness analysis. In W. Miller, L. A. Robinson, & R. S. Lawrence (Eds.), *Committee to evaluate measures of health benefits for environmental, health, and safety regulation.* Washington, DC: Institute of Medicine.

Johnson, F. R., Fries, E. E., & Banzhas, H. S. (1997). Valuing morbidity: An integration of the willingness-to-pay and health-status index literatures. *Journal of Health Economics, 16*(6), 641–665.

Kindig, D. A. (2010). *Burden of illness and MATCH county health rankings*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington, DC.

Kindig, D. A., Booske, B. C., & Remington, P. L. (2010). Mobilizing Action Toward Community Health (MATCH): Metrics, incentives, and partnerships for population health. *Preventing Chronic Disease, 7*(4).

Kirschstein, R. (February 2000). *Disease-specific estimates of direct and indirect costs of illness.* Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health.

Last, J. (2001). *A dictionary of epidemiology*. Oxford: Oxford University Press.

McKenna, M. T., & Zohrabian, A. (2009). U.S. burden of disease—Past, present and future. *Annals of Epidemiology, 19*(3), 212–219.

Michaud, C. M., McKenna, M. T., Begg, S., Tomijima, N., Majmudar, M., Bulzacchelli, M. T., et al. (2006). The burden of disease and injury in the United States 1996. *Population Health Metrics*, *4*, 11.

Miller, W., Richardson Vigdor, E., & Manning, W. G. (2004). Covering the uninsured: What is it worth? *Health Affairs*, W-4, 157–167.

Moriarty, D. G., Zack, M. M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures—population tracking of perceived physical and mental health over time. *Health and Quality of Life Outcomes*, *1*(1), 37.

Murray, C. J. L., Kulkarni, S. C., Michaud, C., Tomijima, N., Bulzacchelli, M. R., Iandiorio, T. J., & Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, *3*(9).

Murray, C. J. L., & Lopez, A. D. (1996). *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard School of Public Health (Global Burden of Disease and Injury Series, vol. I).

Murray, C. J. L., Salomon, J. A., Mathers, C. D., & Lopez, A. D. (2002). *Summary measures of population health: Concepts, ethics, measurement and applications.* Geneva, Switzerland: WHO Press.

National Institutes of Health (NIH). (2011). Patient-Reported Outcomes Measurement Information System. Retrieved from http://www.nihpromis.org/default.aspx.

Obermeyer, Z., Rajaratnam, J. K., Park, C. H., Gakidou, E., Hogan, M. C., Lopez, A. D., & Murray, C. J. L. (2010). Measuring adult mortality using sibling survival: A new analytical method and new results for 44 countries, 1974–2006. *PLoS Med, 7*(e1000260), 1–17.

Olshansky, S. J., Passaro, D. J., Hershow, R. C., Layden, J., Carnes, B. A., Brody, J., Hayflick, L., Butler, R. N., Allison, D. B., & Ludwig, D. S. (2005). A potential decline in life expectancy in the United States in the 21st century. *New England Journal of Medicine*, *352*(11), 1138–1145.

Parrish, R. G. (2010). Measuring population health outcomes. *Preventing Chronic Disease*, *7*(4), A71.

Portney, P. R. (1994). The contingent valuation debate: Why economists should care. *The Journal of Economic Perspectives, 8*(4), 3–17.

Preston, S. H., & Ho, J. Y. (2009). Low life expectancy in the United States: Is the health care system at fault? *NBER Working Paper Number 15213*.

Rajaratnam, J. K., Marcus, J. R., Levin-Rector, A., Chalupka, A. N., Wang, H., Dwyer, L., Costa, M., Lopez, A. D., & Murray, C. J. L. (2010). Worldwide mortality in men and women aged 15–59 years from 1970 to 2010: A systematic analysis. *The Lancet*, *375*, 1704–1720.

Roehrig, C., Miller, G., Lake, C., & Bryant, J. (2009). National health spending by medical condition, 1996−2005. *Health Affairs (Millwood)*, *28*(2), w358−367.

Rosen, A. B. (2010). *Challenges allocating national health care spending to diseases*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington DC.

Russell, L. (2009). Completing costs: Patients' time. *Medical Care*, *47*(7 Suppl 1), S89–S93.

Schmidt, W. P., Cairncross, S., Barreto, M. L., Clasen, T., & Genser, B. (2009). Recent diarrhoeal illness and risk of lower respiratory infections in children under the age of 5 years. *International Journal of Epidemiology, 38*(3), 766–772.

Segel, J. E. (2006). Cost-of-illness studies—A primer. Research Triangle Park, NC: RTI International. Retrieved August 2, 2009, from http://www.rti.org/pubs/COI_Primer.pdf.

Spasoff, R. A. (1999). *Epidemiological methods for health policy*. Oxford: Oxford University Press.

Tarricone, R. (2006). Cost-of-illness analysis: What room in health economics? *Health Policy, 77*(1), 51–63.

Tengs, T. O., & Wallace, A. (2000). One thousand health-related quality-of-life estimates. *Medical Care, 38*(6), 583–637.

Trogdon, J. G., Finkelstein, E. A., & Hoerger, T. J. (2008). Use of econometric models to estimate expenditure shares. *Health Services Research, 43*(4), 1442–1452.

U.S. Department of Health and Human Services. Healthy People 2020: Framework. Retrieved August 2, 2009, from http://www.healthypeople.gov/2020/Consortium/HP2020Framework.pdf.

Walker, C. F. (2010). *Estimating the global burden of disease in populations with limited data*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington DC.

World Health Organization (WHO). (2008). *The Global Burden of Disease: 2004 update*. Geneva, Switzerland: WHO Press.

World Health Organization (WHO). (2004). *The Global Burden of Disease concept*. Geneva, Switzerland: WHO Press.

World Health Organization (WHO). (2010). *World health statistics 2010*. Geneva, Switzerland: WHO Press.

Yabroff, K. R., Brown, M. L., Lawrence, W. F., Barnett, P. G., Lipscomb, J., eds. (July 2009). Health care costing: Data, methods, future directions. *Medical Care, 47*(7 Suppl 1), S1–S142.

Zack, M. (2010). *Quality of life: Recent developments and remaining challenges*. Presented at ASPE/RTI Burden of Illness Roundtable, Washington, DC.